The Quantization of Linear Dynamical Systems I: (Mostly!) Finite Systems

Butterfield, Caulton and Roberts

Philosophical Aspects of QFT, Michaelmas 2021: version for Tues 2 November 2021

Sections 1, 2 and 4 prepare for Chap. 8.2 of Roberts 'Reversing Time's Arrow", on 9 Nov. 2021. So we cut the document at the end of Section 4.

This document, and its successor (on the Quantization of Linear Dynamical Systems with Infinitely many degrees of freedom), expound a rigorous quantization procedure developed by Segal and Mackey (1963) and Segal (1967). This document stresses (in order): complex structures, symplectic structure, and thereby the definitions of one-particle structure and of Fock space. We will end with the Stone-von Neumann Theorem, which sums up the "good control" we have of finite systems in euclidean space (by the way: not just linear ones). This means we postpone to the successor document: more detailed coverage of infinite systems, i.e. fields; and ideas from algebraic quantum theory; which will include e.g. inequivalent representations, 'getting out of Fock space', Haag's theorem etc.; for which (cf. e.g. Emch 1972). But the present material:

(i) gives a strong grip on the first (forbiddingly concise!) third of Wald (1994), which is the basis for the rest of that book on QFT in curved spacetime and thus e.g. the Unruh effect;

(ii) is of intrinsic interest... though please be warned that here you will find: no Lagrangian, no path integrals, no renormalization, no gauge theory, no curved spacetime, no gravitation; indeed, no interactions, and overall, not much physics ... we will focus on the harmonic oscillator (!),the free KG field and spin-chains (and without putting a Hamiltonian on the chain...). Nor will you find much straight-up philosophy ... but perhaps the light here shed on field/wave vs. particle counts as philosophy, since wave vs.particle is, like continuum vs. discrete, a perennial dichotomy of *natural philosophy*...

In this document, we consider only finitely many degrees of freedom, and lead up to the Stone-von Neumann Theorem, which essentially guarantees that the quantization of point particles in $\mathbb{R}^n$ is unique. We begin by introducing the Weyl form of the CCRs; and posing the quest for its representations (Section 1). Then we present the complexification and realification of vector spaces, complex structures etc. (Section 2). Then in Section 3, we review the symplectic and Poisson bracket structures of classical mechanics. We specialize, for the most part, to symplectic vector spaces and linear systems. So this will include an "advanced look" at the harmonic oscillator. But we will also glimpse classical linear *fields*. With all this in hand, we can then see the task of quantization as "unitarizing" a Hamiltonian evolution in a symplectic space so as to give an evolution in a complex Hilbert space.

For this task, the main ideas will be a *one particle structure* and—a more familiar idea!—the Fock space built out of it. We review both of these in Section 4. We start with the new idea of one particle structure, and illustrate its application to the classical harmonic oscillator. This delivers us as the quantum state space—not the familiar quantum harmonic oscillator, with (in one spatial dimension) Hilbert space $L^2(\mathbb{R})$!—but 'merely' the world's simplest complex Hilbert space, viz. $\mathbb{C}$ i.e. the complex plane. To get the familiar quantum harmonic oscillator, i.e. $L^2(\mathbb{R})$ (equipped with the quantum harmonic oscillator Hamiltonian), we need to take the Fock space built from $\mathbb{C}$. Thus we will *factorize* the usual understanding of canonical quantization—viz. (for the 1-dimensional harmonic oscillator) "replace the two-dimensional classical phase space $\mathbb{R}^2 \ni (q, p)$, with $L^2(\mathbb{R})$, i.e. L^2 functions on the configuration space $\mathbb{R}$—into: first, build a 1-particle structure; second, build the Fock space on that. To explain this second step, i.e. Fock space, we will again state the general idea. Then we illustrate with the harmonic oscillator. Then we can glimpse how this plays out for (linear) fields.

Underlying this quantization scheme (both in general and for the harmonic oscillator as an example) is a *two out of three property* of the unitary group: which concerns its relation to certain orthogonal and symplectic groups (Section 5). Finally, we state (i) the Stone-von Neumann Theorem; and (ii) an analogous theorem (the Jordan-Wigner theorem) about the uniqueness of the representation of the CARs (as against CCRs) of a *finite* system, such as a spin chain (Section 6).

> "Let us try to introduce a quantum P.B. [Poisson Bracket] which shall be the analogue of the classical one....we are thus led to the following *definition for the quantum P.B. of any two variables u and v:* $uv - vu = i\hbar[u, v]$." — P.A.M. Dirac (1947, pp.86-87)

> "There is thus a complete harmony between the wave and light-quantum descriptions of the interaction." — P.A.M. Dirac (1927, p.245)

> "First quantization is a mystery, but second quantization is a functor." — attributed to Edward Nelson

> "Probably all these connections would have been clarified long ago, if quantum physicists had not been hampered by a prejudice in favour of complex and against real numbers." — Freeman Dyson (1996, p.1200)

> "The career of a young theoretical physicist consists of treating the harmonic oscillator in ever-increasing levels of abstraction." — attributed to Sydney Coleman

Contents

1 Canonical quantization introduced 4
1.1 Commutation relations: from Heisenberg to Weyl 4
1.2 The Weyl algebra 8
2 Complexification, complex structures—and all that 9
2.1 Complexification 10
2.2 Complex structures 12
2.3 Compatibility of a complex structure with bilinear forms 13
2.4 A compatible J is not unique—and encodes some dynamics 15
2.5 Complex conjugation of spaces 17
3 Symplectic mechanics 18
3.1 From Hamilton's equations to symplectic forms 19
3.2 Poisson brackets 25
3.3 Looking back to the Weyl algebra 27
3.4 A geometrical perspective 29
3.5 Symplectic vector fields from time-translation invariance 35
3.6 Linear solution spaces 38
3.7 Example: The simple (classical) harmonic oscillator 40
4 One-particle structures and Fock space 40
4.1 'Particle' and 'field' 41
4.2 The general idea of one-particle structure 44
4.3 Example: the simple harmonic oscillator 45
4.4 The free bosonic field on any one-particle structure: the idea 47
4.5 Example: the simple harmonic oscillator again 48

4.6 (Apparently) rival quantizations 50
5 "Unitarization": complex structures, metrics, and the 2-out-of-3 property 51
6 The Stone-von Neumann and Jordan-Wigner uniqueness theorems 53
6.1 The two theorems 53
6.2 Proving the theorems 58
6.3 Going beyond these theorems 58
7 Further Reading 59

References **59**

1 Canonical quantization introduced

1.1 Commutation relations: from Heisenberg to Weyl

> "The Problem of finding quantum conditions is not of such a general Character as those we have been concerned with up to the present. It is instead a special Problem which presents itself with each particular dynamical System one is called upon to study. There is, however, a fairly general method of obtaining quantum conditions, applicable to a very large class of dynamical Systems. This is the method of *classical analogy*" (Dirac 1947, Section 21, pg.84)

The idea of *canonical quantization* is familiar from elementary quantum mechanics: to "promote" the classical Poisson bracket relations

$$\{q^i, q^j\} = \{p_i, p_j\} = 0; \qquad \{q^i, p_j\} = \delta^i_j, \tag{1.1}$$

where $i, j \in \{1, 2, \ldots n\}$, to the *Heisenberg canonical commutation relations* (CCRs)

$$[Q^i, Q^j] = [P_i, P_j] = 0; \qquad [Q^i, P_j] = i\hbar\delta^i_j \mathbb{1}; \tag{1.2}$$

(we will usually choose units of $\hbar = 1$). This Poisson bracket-commutator correspondence originated with Dirac, and can be found in the 1947 Third Edition[1] of his *Principles of Quantum Mechanics*:

> "The Problem of finding quantum conditions now reduces to the Problem of determining P.B.'s [Poisson Brackets] in quantum mechanics." (Dirac 1947, p.87)

The standard representation of Equation (1.2) is the familiar irreducible Schroedinger representation: namely, for n configurational degrees of freedom $q_1, \ldots, q_n$, e.g. a spinless particle in Euclidean n-space, or n such particles on a line, the Heisenberg CCRs are satisfied if,

$$Q^i\psi := q_i\psi, \qquad P_j\psi := -\frac{ih}{2\pi}\frac{\partial\psi}{\partial q_j} \qquad \text{for } \psi \in L^2(\mathbb{R}^n, d\mathbf{q}) \text{ and } i = 1, \ldots, n. \tag{1.3}$$

This prompts four main topics. They are of increasing scope, and we will consider only the first.

(a): To examine canonical quantization as just described for position and momentum in $\mathbb{R}^n$. The big positive result here is the Stone-von Neumann theorem, stating (roughly) that for $\mathbb{R}^n$ as the configuration space, the Schroedinger representation of (2) is unique up to unitary equivalence. Cf Section 6. But so as to set the scene for quantum field theory, and more generally so as to get materials useful for contexts other than $\mathbb{R}^n$, we will lead up to this slowly. This will mean expounding some ideas of *Segal quantization*, which is the most straightforward generalization of the above ideas. In short: it replaces $\mathbb{R}^n$ as the classical configuration space, by an arbitrary n-dimensional manifold.

(b): To extend quantization to other quantities, in particular functions (polynomial, or even "arbitrary", functions) of position and momentum.

(c): To consider other methods of quantization.

(d) To pursue the *pure mathematical* interest of quantization. For a glimpse of this, cf. Folland (2008, p.49) and Vogan (1987) cited there. In short: the interest lies in how it helps one find all the irreducible unitary representations of a connected Lie group G: i.e. in physical language, finding all quantum systems in which G acts irreducibly as a symmetry

[1] The spirit of this statement appears in the First Edition of Dirac (1930), though not the clear presentation of the problem of quantization stated here.

group. The corresponding classical problem is to find all symplectic manifolds on which G acts transitively as a group of canonical transformations (symplectomorphisms), i.e. all symplectic homogeneous G-spaces. But this classical problem is "under good control". For the orbits of the co-adjoint action of G on $\mathfrak{g}^*$ are symplectic homogeneous G-spaces; and furthermore, all symplectic homogeneous G-spaces can be, more or less, built from orbits of such co-adjoint action. (Here, "more or less" signals issues about central extensions and covering spaces). Thus a "good" quantization procedure for such spaces is likely to be illuminating for the task of finding all the irreducible unitary representations of G.

Of course, we foreswear (d); and for the most part, we foreswear (b) and (c). For an introduction to both, and of course (a), we recommend:

- Landsman (2007) 'Between Classical and quantum', especially Section 3, preprint available at `http://philsci-archive.pitt.edu/2328/`
- Ali and Engliš (2005) 'Quantization methods: a guide for physicists and analysts', `https://arxiv.org/abs/math-ph/0405065`

In particular, as to (b): Ali and Engliš (2005, Section 1) review the obstructions confronting quantization of (even just a "handful" of polynomial) functions of position and momentum. These obstructions concern ambiguities of operator-ordering. That is: natural general constraints on the quantization map Q ("adding a hat") that sends a classical (real-scalar) quantity $f : \mathbb{R}^{2n} \to \mathbb{R}$ to a quantum quantity, i.e. to a self adjoint operator $Q(f) : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$, lead to *contradictions*. This topic originates in papers by Groenewold and van Hove. Recent developments include: Gotay et al. (1996) and Gotay (1999).

As to (c): Ali and Engliš (2005, Section 3f) review geometric quantization, deformation quantization etc. But even their Section 2 gives details of e.g. the inequivalent quantizations involved in the Aharonov-Bohm effect.

But the four topics are of course closely related. For example, these obstructions mean that a main motivation to pursue (c)'s other methods of quantization is to extend quantization to as many quantities as possible.

For us, concentrating on (a): the main point about (b), i.e. the obstructions, will be that (cf. Wald 1994, Section 2.2, pp. 17-18): Segal quantization "works" for:

(i) a classical configuration space that is an arbitrary n-dimensional manifold M (so that classical quantities are real functions of the cotangent bundle T^*M); *provided that*

(ii) we restrict consideration to quantities that are at most linear in the momenta (i.e. the momenta canonically conjugate to arbitrary configurational coordinates q on M).

Here, the word "works" means that the quantization map Q maps Poisson brackets into commutators, divided by $i\hbar$. (In more formal jargon: "Q respects Lie algebra structure"). That is: Q obeys, for classical quantities $f, g : T^*M \to \mathbb{R}$ that are appropriately restricted by condition (ii) above:

$$[Q(f), Q(g)] = i\hbar\, Q(\{f, g\}) \tag{1.4}$$

In this sense, Segal quantization is a good framework for the quantization of finite-dimensional systems.

And Segal quantization has other merits. We will also see that for *linear* classical systems, it "respects" the dynamics. That is: the Segal quantization of the classical Hamiltonian of such a system (which is essentially like that of a harmonic oscillator: "$p^2 + q^2$") is the "correct" quantum Hamiltonian. Besides, we will eventually see that it works for (some!) quantum

field theories. Specifically, it works for the quantization of the free bose field (e.g. De Faria and De Melo 2010, Section 6.3). Furthermore, it does this in a manner that generalizes readily to constructing quantum field theories on *curved* spacetimes (Wald 1994, p.31 and Section 3.2).

So much by way of preamble. For our main topic, i.e. (a) above, the first job is to pass from the Heisenberg CCRs to the *Weyl form of the CCRs*. The point here is that since the classical position and momentum quantities, for a phase space $\mathbb{R}^{2n}$, are unbounded, we expect the quantum position and momentum Q^i, P_j to also be unbounded, indeed to have all of $\mathbb{R}$ as their spectra—so that, if they are to be self-adjoint, they cannot be defined on all of $L^2(\mathbb{R}^n)$.

Indeed, setting aside the physical desideratum that the spectra should be unbounded: there is a simple theorem that if two *bounded* self-adjoint operators Q, P have a commutator that is proportional to the identity, they must *commute*. That is: If $[Q, P] = \alpha I$ for some $\alpha \in \mathbb{C}$, then[2] $\alpha = 0$.

In short: we face issues of domains. We remedy this by formulating to the *Weyl form* of the CCRs. These govern unitary exponentiations of linear combinations of the position, and similarly, of the momentum operators. We will define these unitary exponentiations, and deduce their CCRs (i.e. the Weyl form) from the Heisenberg CCRs, in (B) below. But first, we will in (A) take a more general perspective, so as to use some of the tools of the Hilbert Space Review (viz. the spectral theorem and Stone's theorem) and introduce the jargon of a *transitive system of imprimitivity*.

(A): The Weyl CCRs can be viewed as arising from a choice of a quantity (observable), subject to a continuous group of symmetries, in the following way.

Let $\Delta \mapsto E_\Delta$ be a projection-valued measure (PVM) on Borel sets $\Delta \subseteq \mathbb{R}^n$. If we think of this PVM as representing spatial position in euclidean n-space, then each E_Δ would be interpreted as the experimental outcome, 'The system is in the spatial region Δ'. (Of course, the standard examples are: $n = 1$ for a particle on a line; or $n = 3$; or $n = 3N$ for N distinguishable particles in three-dimensional space, in which last example 'the system being in the spatial region Δ' really means 'the system configuration being in $\Delta \subset \mathbb{R}^{3N}$'.) However, we need not give the PVM this sort of spatial interpretation: 'being in Δ' can viewed as a non-spatial 'mark' or 'score'. Recall our Philosophical Remarks in the Hilbert Space Review. Given this PVM, let us write Q^i for the associated self-adjoint operators, defined by the spectral theorem: $Q^i := \int_{-\infty}^{\infty} \lambda \, dE^i_\lambda$, for each $i = 1, \ldots, n$. And we write $\mathbf{Q}$ for the vector-operator $\langle Q^1, Q^2, ..., Q^n \rangle$: where 'vector-operator' means as usual that the components transform as expected under rotations.

Roughly speaking, the symmetry we are concerned is the statement that the statistical outcomes of this quantity (observable) remain unchanged when the set Δ is translated in $\mathbb{R}^n$; and correspondingly, for the PVM:

$$E_\Delta \mapsto E_{\Delta'} = E_{\Delta - \mathbf{a}} \, ; \tag{1.5}$$

where for any $\mathbf{a} \in \mathbb{R}^n$, we use the shorthand $\Delta - \mathbf{a} := \{\mathbf{x}' \in \mathbb{R}^n \mid \mathbf{x}' = \mathbf{x} - \mathbf{a} \text{ for some } \mathbf{x} \in \Delta\}$. For example, if Δ were a region in space, then the symmetry captures *spatial homogeneity*, i.e. that the statistical outcomes of an experiment are the same no matter where it is set up in space: as discussed by Jauch (1968, Section 12-2).

Motivated by Wigner's theorem, we interpret preservation of statistical outcomes as

[2] If $[Q, P] = \alpha I$, then setting $\beta = \alpha/i$ we have $[Q^n, P] = ni\beta Q^{n-1}$ for all n. Thus, $\beta n |Q^{n-1}| = |ni\beta Q^{n-1}| = |Q^n P - PQ^n| \leq |Q^n P| + |PQ^n| \leq 2|Q^n||P|$, so $n \leq 2|Q||P|/\beta$ for all n. Since n can be arbitrarily large, this means that if Q and P are both bounded, then $\beta = 0$ and hence $\alpha = 0$. See also De Faria and De Melo (2010, Lemma 2.11) and Jauch (1968, p.205 Problem 4).

implying that the transformation 1.5 is unitary.[3] So, more precisely now: let us assume that together with the PVM E_Δ, there is a strongly continuous unitary group $U(\mathbf{a})$ indexed by $\mathbf{a} \in \mathbb{R}^n$ such that: (1) $U(\mathbf{0}) = I$, $U(\mathbf{a} + \mathbf{a}') = U(\mathbf{a})U(\mathbf{a}')$ for $\mathbf{a}, \mathbf{a}' \in \mathbb{R}^n$; with the additivity and strong continuity being of course understood component-wise, $i = 1, 2, ...n$; and also (2):

$$U(\mathbf{a})E_\Delta U(\mathbf{a})^* = E_{\Delta - \mathbf{a}}\,. \tag{1.6}$$

The pair $(\Delta \mapsto E_\Delta, \mathbf{a} \mapsto U(\mathbf{a}))$ satisfying (1) and (2) is called a *transitive system of imprimitivity* following Mackey (1976, Section 3.7).

It is easy to check[4] that it follows that this unitary $U(\mathbf{a})$ 'translates' the 'position' vector operator $\mathbf{Q}$ as expected. That is: $U(\mathbf{a})\mathbf{Q}U(\mathbf{a})^* = \mathbf{Q} + \mathbf{a}I$; again, of course understood component-wise, $i = 1, 2, ...n$.

Moreover, by Stone's theorem there is a unique self-adjoint vector-operator $P \equiv \langle P_1, P_2, ..., P_n \rangle$ such that $U(\mathbf{a}) = e^{-i\mathbf{a}\cdot\mathbf{P}}$ for all $\mathbf{a} \in \mathbb{R}^n$. A simple calculation[5] then shows that there is a dense domain of vectors $D_{QP} \subseteq \mathcal{H}$ on which both $\mathbf{Q}$ and $\mathbf{P}$ are defined, and such that,

$$[Q^j, P_k]\psi = i\delta^j_k\psi \text{ for all } \psi \in D_{QP} \tag{1.7}$$

To sum up: the familiar Heisenberg form of the canonical commutation relations can be viewed as arising from a transitive system of imprimitivity—as is natural to postulate for the quantity position, for a homogeneous space.

(B): We return to assuming *ab initio* the usual position and momentum operators. So we assume we are given Q^i and P_j as in eq. 1.3. Now let us define for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$,

$$U(\mathbf{a}) := e^{-i\mathbf{a}\cdot\mathbf{P}/\hbar} \;\; ; \qquad V(\mathbf{b}) := e^{-i\mathbf{b}\cdot\mathbf{Q}/\hbar}; \tag{1.8}$$

Since the Us and Vs are each unitary, they are bounded, and so are defined everywhere in $L^2(\mathbb{R}^n)$. We have:

$$(U(\mathbf{a})\psi)(\mathbf{x}) = \psi(\mathbf{x} - \mathbf{a}) \;\; ; \;\; (V(\mathbf{b})\psi)(\mathbf{x}) = e^{-i\mathbf{b}\cdot\mathbf{x}/\hbar}\psi(\mathbf{x}) \tag{1.9}$$

so that each $U(\mathbf{a})$ represents a translation by $\mathbf{a}$ in euclidean n-space by $\mathbf{a}$; and each $V(\mathbf{b})$ represents a translation in momentum-space by $\mathbf{b}$.

We have, of course, commutation within each family of the Us and Vs:

$$U(\mathbf{a})U(\mathbf{a}') = U(\mathbf{a}')U(\mathbf{a}) = U(\mathbf{a} + \mathbf{a}') \qquad V(\mathbf{b})V(\mathbf{b}') = V(\mathbf{b}')V(\mathbf{b}) = V(\mathbf{b} + \mathbf{b}') \tag{1.10}$$

To deduce the commutation relations between U and V operators, we need the *Campbell-Baker-Hausdorff formula* for products of exponentials of non-commuting operators. This goes as follows.

Given a self-adjoint operator A, we say that a vector $\psi \in \mathcal{H}$ is *analytic* if for all n, $A^n(\psi)$ is defined, and so is $e^A\psi$. Then the version of the Campbell-Baker-Hausdorff formula which is appropriate here (De Faria and De Melo, Lemma 2.12) says that if:

(i) A, B and $A + B$ have a common dense domain D of analytic vectors, and

(ii) [A,B] commutes with A and with B:

[3] Strictly speaking it could also be antiunitary; but since there is no strongly continuous group of antiunitary operators we must here treat these operators as unitary.

[4] ToDo: ADD THIS.

[5] ToDo: ADD THIS.

then in D:

$$e^A e^B = e^{A+B+\frac{1}{2}[A,B]} \equiv e^{A+B} e^{\frac{1}{2}[A,B]} \tag{1.11}$$

To apply (1.11) to (1.8), we of course need to be assured that Q^i, P_i and $Q^i + P_i$ have a common dense domain D of analytic vectors. But taking this for granted here: we set $A := -i\mathbf{a}.\mathbf{P}/\hbar$ and $B := -i\mathbf{b}.\mathbf{Q}/\hbar$, to deduce that

$$U(\mathbf{a})V(\mathbf{b}) = \exp(\frac{1}{2}i(\mathbf{a}\cdot\mathbf{b})/\hbar).\exp(-i(\mathbf{a}.\mathbf{P}/\hbar + \mathbf{b}.\mathbf{Q}/\hbar)) \; ; \tag{1.12}$$

and *mutatis mutandis*, we set $A := -i\mathbf{b}.\mathbf{Q}/\hbar$ and $B := -i\mathbf{a}.\mathbf{P}/\hbar$, to deduce that

$$V(\mathbf{b})U(\mathbf{a}) = \exp(-\frac{1}{2}i(\mathbf{a}\cdot\mathbf{b})/\hbar).\exp(-i(\mathbf{a}.\mathbf{P}/\hbar + \mathbf{b}.\mathbf{Q}/\hbar)). \tag{1.13}$$

Combining these immediately gives the *Weyl commutation relations*:[6]

$$U(\mathbf{a})V(\mathbf{b}) = e^{i\mathbf{a}.\mathbf{b}/\hbar} V(\mathbf{b})U(\mathbf{a}). \tag{1.14}$$

1.2 The Weyl algebra

So from now on, we take as our CCRs, not the Heisenberg form (1.2), but (1.14) together with the trivial commutations of Us and Vs alone i.e. (1.10).

At the end of the last Section (especially (B)) we built the Us and Vs concretely from given $\mathbf{Q}, \mathbf{P}$. But in the usual tradition of physics, we can:

(i) consider an abstract algebra of Us and Vs subject to the relations (1.14) and (1.10); any such algebra is called *the Weyl algebra*; (later, we will discuss the algebraic and topological conditions satisfied by this algebra—in short, it is a *C*-algebra*); and then,

(ii) try to classify the representations of this algebra, especially the unitary representations on some Hilbert space.

As already announced at the start of Section 1.1, the main result about (ii), for finite-dimensional systems, will be the Stone-von Neumann uniqueness theorem. But as that discussion also suggested: the Weyl algebra, and Segal quantization, will also be centre-stage for quantizing fields (including on curved spacetime) and for the pure mathematical topic (d) of Section 1.1.

Now, we first make two comments, (1) and (2), about this endeavour (in order of increasing importance for us); and then, in (3), develop a more abstract formulation of the Weyl relations, which will be central in all that follows.

(1): *The relation between the Heisenberg and Weyl forms*:– The Weyl form of the CCRs implies the Heisenberg form, and so a representation of the Weyl form is also a representation of the Heisenberg form. But uniqueness (up to unitary equivalence) of a representation of the

[6]Beware: (i) many authors 'flip' the notation of U and V, so that V represents translations in space; and (ii) some authors (even rigorous ones e.g. Prugovecki 1981, Chapter IV, Sections 6.2, 6.4!) also put the $\hbar$ in the numerator of the exponent, so that the exponent is in dire danger of having dimension action-squared! Besides, (iii): various texts also get the sign of the exponent in (1.14) wrong. (See later for discussion of different choices of sign in the two definitions of (1.8).) We are following S. Summers (2001: in *John von Neumann and the Foundations of quantum mechanics*, ed. M. Redei and M. Stoeltzner). Summers puts the $\hbar$ in the denominator of the exponent, is perfectionist about signs; and his use of U for translation in space, is like Weyl himself (1932, Chapter IV, Section 14, building on Chapter II, Section 11): this last text being no doubt correct, but—with all due respect!—incomprehensible.

Weyl form does not imply uniqueness of the implied representation of the Heisenberg form. The reason lies in the simple theorem above, that two *bounded* self-adjoint operators Q, P cannot obey the Heisenberg form. In fact, the Heisenberg form does not imply the Weyl form, even if Q and P are essentially self-adjoint on their respective domains; though conditions can be added that make the implication go through, e.g. the Dixmier (1958) condition (in French!) discussed by Jauch (1968, pp.204-205).

(2): *Allowing for projective unitary representations*:— Of course, the quantum state is *non-redundantly* represented by a *ray* rather than a unit vector. This motivates considering *projective* representations of groups, rather than "true" representations. Such representations allow a phase to occur in equations stating the group composition law for the representing operators. Indeed, we see this even for elementary abelian groups, like the phase-space translation groups we are concerned with: cf. the phase in (1.14), and in (1.16) below.

(3): *A more abstract formulation*:— Equation (1.14) can be given a more abstract formulation, which both:

(i) brings out the role being played by the symplectic structure in the underlying framework of Hamiltonian mechanics, and

(ii) underpins how Segal quantization succeeds in quantizing linear classical systems, both finite-dimensional and infinite-dimensional.

Setting $z := (\mathbf{a}, \mathbf{b}) \in \mathbb{R}^{2n}$, we define the family of operators

$$W(z) := e^{\frac{1}{2}i\mathbf{a}\cdot\mathbf{b}} U(\mathbf{a}) V(\mathbf{b}). \tag{1.15}$$

Then the Weyl form of the CCRs, i.e. (1.14) and (1.10), are equivalent to the following, which is thus also called the *Weyl algebra*: for all $z, z_1, z_2 \in \mathbb{R}^{2n}$,

$$\begin{array}{rcl} W(z_1)W(z_2) & = & e^{\frac{1}{2}i\Omega(z_1,z_2)} W(z_1 + z_2); \\ W^\dagger(z) & = & W(-z); \end{array} \tag{1.16}$$

where Ω is the *symplectic product*:

$$\Omega(z_1, z_2) := \mathbf{a}_2 \cdot \mathbf{b}_1 - \mathbf{a}_1 \cdot \mathbf{b}_2, \tag{1.17}$$

The symplectic meaning of Ω will be explained in Section 3. But as a preliminary to that, we spell out in Section 2 some elementary ideas and results about complexification and complex structures: which are often treated very concisely if at all (e.g. Wald 1994, p.190).

2 Complexification, complex structures—and all that

There is a circle of ideas which can be traversed starting from almost any point... We begin with complexification, then describe complex structures, then the compatibility of a complex structure with a bilinear form, such as an inner product or symplectic form. This will give us a glimpse of how we can "go back and forth" between certain classical phase spaces (viz. symplectic vector spaces) and Hilbert spaces. It will also give us a glimpse of (i) Kahler manifolds, and (ii) how in a quantum theory different choices of a complex structure are associated with different splittings of positive and negative frequencies, and thereby (iii) the Unruh effect.The Section ends with discussion of the complex conjugation of spaces.

2.1 Complexification

2.1.A Complexification as tensor product:— The *complexification* $V^{\mathbb{C}}$ of a real vector space V is defined as the tensor product of V with the complex numbers $\mathbb{C}$

$$V^{\mathbb{C}} := V \otimes \mathbb{C} \, . \tag{2.1}$$

Here we think of $\mathbb{C}$ as a copy of $\mathbb{R}^2$, with a basis $\{(1,0),(0,i)\}$. So far, this is just a real vector space. Every vector in $V^{\mathbb{C}}$ can be written uniquely as

$$v = v_1 \otimes 1 + v_2 \otimes i \tag{2.2}$$

and the (real) dimension of $V^{\mathbb{C}}$ is twice the dimension of V. But we make it into a complex vector space, by defining complex scalar multiplication by

$$\alpha(v \otimes \beta) = v \otimes (\alpha\beta) \;\text{ for all } v \in V \text{ and } \alpha, \beta \in \mathbb{C} \, ; \tag{2.3}$$

where we also of course require scalar multiplication to distribute over addition, i.e. we 'extend by linearity':

$$\alpha(v \otimes \beta + u \otimes \gamma) := \alpha(v \otimes \beta) + \alpha(u \otimes \gamma) \equiv v \otimes (\alpha\beta) + u \otimes (\alpha\gamma) \, . \tag{2.4}$$

Since every vector in $V^{\mathbb{C}}$ can be written uniquely as $v = v_1 \otimes 1 + v_2 \otimes i$, it is usual to drop the tensor product symbol and just write

$$v = v_1 + iv_2. \tag{2.5}$$

One then checks that the definition eq. 2.1, equivalently eq. 2.2, implies that the complex scalar multiplication defined by eq. 2.3, can be written in the usual-looking form. Namely: for a complex number $\alpha = a + ib$ with $a, b \in \mathbb{R}$

$$(a + ib)(v_1 + iv_2) = (av_1 - bv_2) + i(bv_1 + av_2). \tag{2.6}$$

So we regard $V^{\mathbb{C}}$ as the direct sum of two copies of V, equipped with a complex scalar multiplication defined by eq. 2.6 .

There is a natural embedding of V in to $V^{\mathbb{C}}$ given by

$$v \mapsto v \otimes 1 \, . \tag{2.7}$$

V may thus be regarded as a *real* subspace of $V^{\mathbb{C}}$. If V has a basis $\{e_i\}$ over $\mathbb{R}$ then a correponding basis for $V^{\mathbb{C}}$ is given by $\{e_i \otimes 1\}$ over $\mathbb{C}$. The *complex* dimension of $V^{\mathbb{C}}$ is therefore equal to the *real* dimension of V:

$$\dim_{\mathbb{C}} V^{\mathbb{C}} = \dim_{\mathbb{R}} V. \tag{2.8}$$

2.1.B Complexification as direct sum:— Alternatively, we can *define* the complexification of V as the direct sum

$$V^{\mathbb{C}} := V \oplus V \tag{2.9}$$

equipped with a *complex structure* (cf. below for details) given by the operator $J : V^{\mathbb{C}} \to V^{\mathbb{C}}$, where J is defined by

$$J(v, w) := (-w, v) \, . \tag{2.10}$$

Here J encodes multiplication by i in the sense that setting $a = 0, b = 1$ in eq. 2.6 yields

$$i(v_1 + iv_2) = -v_2 + iv_1 = -v_2 \otimes 1 + v_1 \otimes i \tag{2.11}$$

where the last expression on the right is in the notation of eq. 2.2.

Let $\dim_{\mathbb{R}} V = n$. Then in matrix form, J is given by a $2n \times 2n$ matrix J, viz.

$$J = \begin{pmatrix} \mathbf{0} & -\mathbb{1}_{\mathbf{V}} \\ \mathbb{1}_{\mathbf{V}} & \mathbf{0} \end{pmatrix} \ . \tag{2.12}$$

where $-\mathbb{1}_{\mathbf{V}}$ is the identity map on V. Thus $V^{\mathbb{C}}$ can be written as $V \oplus JV$ or as $V \oplus iV$, so as (i) to avoid the tensor product notation, and (ii) to signal the fact that the direct sum in eq. 2.9 is endowed with J. J swaps the summands in the sense that $J(v, 0) = (0, v)$.

Examples: (i) the complexification of $\mathbb{R}^n$ is $\mathbb{C}^n$; (ii) if V is the $m \times n$ matrices with real entries, then $V^{\mathbb{C}}$ is the $m \times n$ matrices with complex entries.

Again we have (cf. eq. 2.8): the *complex* dimension of $V^{\mathbb{C}}$ is equal to the *real* dimension of V, which is half the *real* dimension of $V \oplus V$:

$$\dim_{\mathbb{C}} V^{\mathbb{C}} = \dim_{\mathbb{R}} V = \frac{1}{2} \dim_{\mathbb{R}} (V \oplus V) \ . \tag{2.13}$$

2.1.C A matter of convention:— The above discussion (in 2.1.A and 2.1.B) has an obviously conventional aspect. Suppose that in 2.1.A, we had taken the basis of $\mathbb{C}$ as a copy of $\mathbb{R}^2$, to be in the opposite order, i.e. $\{(0, i), (1, 0)\}$. Then eq. 2.2 would become

$$v = v_1 \otimes i + v_2 \otimes 1 \tag{2.14}$$

Then the definition of complex scalar multiplication, eq. 2.3 and 2.4, remain as they are. But the notation that drops the tensor product, i.e. eq. 2.5, becomes

$$v = iv_1 + v_2 \ ; \tag{2.15}$$

and the usual-looking form of the complex scalar multiplication that we now deduce is the following analogue of eq. 2.6: for a complex number $\alpha = a + ib$ with $a, b \in \mathbb{R}$

$$(a + ib)(iv_1 + v_2) = (av_2 - bv_1) + i(av_1 + bv_2). \tag{2.16}$$

Similarly, for the alternative direct sum approach of 2.1.B. Instead of eq. 2.10, we define the complex structure J on the direct sum $V \oplus V$ by

$$J(v, w) := (w, -v) \ . \tag{2.17}$$

Then, setting $a = 0, b = 1$ in eq. 2.16 yields

$$i(iv_1 + v_2) = -v_1 + iv_2 = iv_2 - v_1 = v_2 \otimes i - v_1 \otimes 1 \tag{2.18}$$

where the last expression on the right is in the notation of eq. 2.14. This J as defined by eq. 2.17 is of course just minus the J defined by eq. 2.10. The matrix form of J as defined by eq. 2.17 is thus the negative of eq. 2.12. That is:

$$J = \begin{pmatrix} \mathbf{0} & \mathbb{1}_{\mathbf{V}} \\ -\mathbb{1}_{\mathbf{V}} & \mathbf{0} \end{pmatrix} \ . \tag{2.19}$$

This last equation will give us, shortly, an obvious comparison with the matrix expression of a symplectic form.

2.2 Complex structures

2.2.A Basics:— A *complex structure* on a real vector space V is an automorphism J of V that squares to minus the identity map, $-\mathbb{1}$. That is: $J^2 = -\mathbb{1}$. Such a structure on V allows one to define multiplication by complex scalars in a canonical fashion so as to regard V as a complex vector space. Namely:

$$(x + iy)v := xv + yJ(v) \text{ for all } v \in V \text{ and } x, y \in \mathbb{R} ; \tag{2.20}$$

which (check!) makes V into a complex vector space, denoted V_J.

If V is any real vector space, there is a canonical complex structure J on the direct sum $V \oplus V$: namely, the complex structure on the complexification $V^{\mathbb{C}}$ of V, i.e. on the tensor product $V \otimes \mathbb{C}$, written as $V \oplus JV$ or as $V \oplus iV$. That is, J is given by $J(v, w) := (-w, v)$, i.e. by eq. 2.10, ; and the matrix form of J is as in eq. 2.12. In this notation for complexification—i.e. the notation, $V \oplus JV$ or $V \oplus iV$—we can write: $V \oplus JV = (V \oplus V)_J$ or similarly $V \oplus iV = (V \oplus V)_J$.

One can go in the other direction. Any complex vector space W is also a real vector space, with the same vector addition and real scalar multiplication. On this underlying real vector space, one defines a complex structure J by $J(w) := iw$ for all $w \in W$; where the right-hand-side is given us by W being a complex vector space. With this complex structure defined, we of course get back the original complex vector space W.

In fact, if V_J has complex dimension n, then V must have real dimension $2n$. That is, a finite-dimensional real space V admits a complex structure only if it is even-dimensional. If $\{v_1, ..., v_m\}$ is a basis of the complex vector space V_J, then $\{v_1, J(v_1)..., v_m, J(v_m)\}$ is a basis of the underlying real vector space V.

Every even-dimensional real vector space V admits a complex structure. Indeed, many. For any basis $\{e_1, e_2, \dots, e_{2n}\}$ of V can be divided in to n pairs, say $\{e_1, e_2\}, \dots, \{e_{2n-1}, e_{2n}\}$, and then one can define J as the 'swap with a minus' on each such pair, i.e. $J(e_1) := e_2, J(e_2) := -e_1, \dots, J(e_{2n-1}) := e_{2n}, J(e_{2n}) := -e_{2n-1}$, and then one extends by linearity to all of V. So $J^2 = -\mathbb{1}$.

Suppose that we are given a real linear transformation $A : V \to V$ on a real vector space V, and that V admits a complex structure J. Then A defines a complex linear transformation of the complex space V_J if and only if A commutes with J, i.e. if and only if $AJ = JA$: (trivial check, cf. eq. 2.20).

Likewise, a real subspace U of V is a complex subspace of V_J (i.e. is closed under complex-linear combinations) if and only if J preserves U, i.e. if and only if $J(U) < U$; (trivial check).

2.2.B: Basic example:— Obviously, the main example of a complex structure is the structure on $\mathbb{R}^{2n}$ coming from the complex structure on $\mathbb{C}^n$. That is, the complex n-dimensional space $\mathbb{C}^n$ is also a real $2n$-dimensional space. Here, one uses the same vector addition and real scalar multiplication: while multiplication by the complex number i is not only a *complex* linear transform of the space, thought of as a complex vector space, but also a *real* linear transform of the space, thought of as a real vector space. This is just because scalar multiplication by i:

(a) commutes with scalar multiplication by real numbers, i.e. $i(\lambda v) = (i\lambda)v = (\lambda i)v = \lambda(iv)$, and

(b) distributes across vector addition.

As a complex $n \times n$ matrix, this complex structure is simply the diagonal matrix with i on the diagonal. The corresponding real $2n \times 2n$ matrix is denoted J. What this matrix J looks like will depend on how we order the basis: cf. eq. 2.22 and 2.23 in (1) and (2) below.

Again, there is the general equation that counts dimensions, with $V^{\mathbb{C}} = (V \oplus V)_J$

(cf. eq. 2.13):

$$\frac{1}{2}\dim_{\mathbb{R}}(V\oplus V)_J = \dim_{\mathbb{C}}(V\oplus V)_J = \dim_{\mathbb{R}}V = \frac{1}{2}\dim_{\mathbb{R}}(V\oplus V)\ . \tag{2.21}$$

And in this example, with $V=\mathbb{R}^n$: these numbers are all n.

2.2.C: The "look" of J:— Suppose given a complex vector space, of complex dimension n, and a basis $\{e_1, e_2, \ldots, e_n\}$. This set, together with these vectors multiplied by i, namely $\{ie_1, ie_2, \ldots, ie_n\}$, form a basis for the underlying real vector space. (Cf. 2.2.A, paragraph 4, above.) There are two natural ways to order this basis.

(1): If one orders the basis as $\{e_1, ie_1, e_2, ie_2, \ldots, e_n, ie_n\}$, then the matrix for J takes the following block-diagonal form, where the blocks are the 2×2 matrix $J_2 := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. That is: J is (with subscript $2n$ added, so as to indicate dimension):

$$J_{2n} := \begin{pmatrix} J_2 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & J_2 & \ldots & \mathbf{0} \\ & & \ddots & \\ \mathbf{0} & \mathbf{0} & \ldots & J_2 \end{pmatrix}\ . \tag{2.22}$$

(2): If one orders the basis as $\{e_1, e_2, \ldots, e_n, ie_1, , ie_2, \ldots, ie_n\}$, then the matrix for J is block-antidiagonal:

$$J_{2n} := \begin{pmatrix} \mathbf{0} & -\mathbb{1}_{\mathbf{n}} \\ \mathbb{1}_{\mathbf{n}} & \mathbf{0} \end{pmatrix}\ : \tag{2.23}$$

This is more natural when one thinks of the real space as a direct sum of real spaces, as in the second, alternative, approach to complexification at the end of Section 2.1. Thus eq. 2.23 is the same as eq. 2.12.

2.3 Compatibility of a complex structure with bilinear forms

2.3.A: Basics:— Later we will be much concerned with vector spaces that have: either an inner product (like a Hilbert space) or a symplectic product (as in Hamiltonian mechanics; cf. Section 3). So we here consider, in general, the "meshing" of a complex structure with bilinear forms. This will lead, in 2.3.B and 2.3.C, to "building a Hilbert space", and to the construction in the reverse direction, from a Hilbert space to a symplectic space.

If B is a bilinear form on a real vector space V, i.e. $B: V\times V\to\mathbb{R}$, then we say that J *preserves* B if for all $u, v\in V$

$$B(Ju, Jv) = B(u, v)\ . \tag{2.24}$$

Recall that since J is an automorphism with $J^2=-1$, we have $J^{-1}=-J$. This implies that eq. 2.24 is equivalent to J being skew-adjoint with respect to B. That is:

$$B(Ju, v) = -B(u, Jv)\ . \tag{2.25}$$

Examples of bilinear forms are inner products and symplectic products. If g is an inner product on V then J preserves g if and only if J is an orthogonal transformation. Likewise, J preserves a non-degenerate, skew-symmetric form ω, i.e. a symplectic product, if and only if J is a symplectic transformation, i.e. $\omega(Ju, Jv) = \omega(u, v)$. If ω and J obey, for all non-zero $u\in V$, $\omega(u, Ju) > 0$, we say that J *tames* ω.

2.3.B: From symplectic form and compatible J to real-valued inner product:— A symplectic form ω on a real vector space V, together with a complex structure J that preserves ω, define: a symmetric bilinear form g_J on the complex vector space V_J. Namely, by:

$$g_J(u,v) := \omega(u, Jv) \ . \tag{2.26}$$

This is called *the Kähler condition*. We note that g_J is symmetric because J being skew-adjoint with respect to ω, i.e. eq. 2.25, implies that the rhs of eq. 2.26, i.e. $\omega(u,Jv) = -\omega(Ju,v) \equiv \omega(v,Ju) =: g_J(v,u)$. One similarly checks trivially that: (i) J preserves g_J; (ii) if J *tames* ω, then g_J is positive-definite, i.e. an inner product.

One also checks trivially that on the complex vector space V_J: g_j is *complex-linear*, even though g_J is *real-valued*. Thus, applying the initial definition of complex scalar multiplication for V_J, eq. 2.20, we write:

$$\begin{aligned} g_J((x+iy)u,v) &:= \omega((x+iy)u, J(v)) \equiv \omega((xu + yJ(u)), J(v)) \\ \equiv \omega(xu, J(v)) + \omega(yJ(u), J(v)) &\equiv x\omega((u, J(v)) + y\omega(J(u), J(v)) \\ &\equiv x g_J(u,v) + y g_J(Ju,v) \ . \end{aligned} \tag{2.27}$$

2.3.C: Defining a complex-valued inner product:— From 2.3.B, we assume we are given: (i) a real vector space V with (ii) a symplectic form ω, and (iii) a complex structure J that preserves and tames ω; and thereby (iv), on the complex vector space V_J, a positive-definite real-valued inner product g_J: namely as defined by the Kähler condition, eq. 2.26.

Now let us define a *complex-valued* function on $V \times V$ in terms of g_J and ω by

$$\langle u, v\rangle \equiv \langle u, v\rangle_{\omega,J} := g_J(u,v) + i\omega(u,v) \tag{2.28}$$

where the subscript shows the dependence on the given ω and J. It is trivial that this function is additive in each argument, i.e. $\langle u+w, v\rangle = \langle u, v\rangle + \langle w, v\rangle$ and similarly for additivity of the second argument. One checks (exercise!) that it is *sesquilinear*. That is: it is complex-linear in the second argument, but antiinear in the first argument. That is, with $x, y \in \mathbb{R}$:

$$\langle (x+iy)u, v\rangle = x\langle u, v\rangle - iy\langle u, v\rangle \ \text{ and } \ \langle u, (x+iy)v\rangle = x\langle u, v\rangle + iy\langle u, v\rangle \ . \tag{2.29}$$

The check of eq. 2.29 uses most of the properties we have postulated. Namely: the definition eq. 2.26 of g_J in terms of ω and J; the antisymmetry of ω and the symmetry of g_J; and the fact that J preserves ω.

Besides, recall that we assumed that J tames ω, so that g_J is positive-definite, i.e. a real-valued inner product (cf. 2.3.B). Then since ω is also non-degenerate, one checks (exercise!) that $\langle \cdot, \cdot \rangle$ is positive-definite. To conclude: $\langle \cdot, \cdot \rangle$ is a complex inner product in the usual sense: sesquilinear and positive-definite.

We recall that a (complex) *Hilbert space* is a a complex inner product space, that is complete in the norm induced by the inner-product. That is: Cauchy sequences, in the norm, converge to a vector in the space. This completeness does not follow from the above assumptions, unless the given real vector space V is finite-dimensional. (Thus a complex inner product space is often called a *pre-Hilbert space*.) But even if V is infinite-dimensional, and not complete in the norm, there is a canonical construction of a Hilbert space from it. This is like the canonical construction, for an arbitrary metric space (X,d), of a complete metric space $(\bar{X}, \bar{d})$, into which (X,d) can be isometrically embedded. Namely, the points of $\bar{X}$ are appropriately defined equivalence classes of Cauchy sequences in X. For details, see e.g. Prugovečki (1971, Section 3.3).

2.4 A compatible J is not unique—and encodes some dynamics

There are three remarks to make at this point, about this development from 2.2.A to 2.3.C. They all concern the non-uniqueness of J, and they give a glimpse of further, more physical, developments, including *the Unruh effect*—glimpsed in 2.4.C below.

2.4.A: On the non-uniqueness of J:— We stressed in 2.2.A that an even-dimensional real vector space V admits many complex structures. For any basis $\{e_1, e_2, ..., e_{2n}\}$ can be divided into n pairs (in many ways), with J can then defined as the 'swap with a minus' on each such pair, extended by linearity.

But in this Subsection, since 2.3.B, we have assumed that a symplectic form ω is given, and that J is compatible with it. So does fixing ω and requiring compatibility still leave freedom in the definition of J? In fact, it does.

Fixing ω defines (by an analogue of the Gram-Schmidt diagonalization of a bilinear form) bases such that ω's matrix form is that of J in eq. 2.12 (cf. Section 2.1, and Section 3 below, about symplectic structure). This is best understood in terms of how Hamiltonian mechanics defines a symplectic form on the phase space "of qs and ps", i..e. of positions and momenta. This naturally associates each q one-to-one with a p, and so the basis of $2n$ vectors breaks down in to n pairs. We might write the basis as $\{q_1, q_2, ..., q_n, p_1, p_2, ...p_n\}$, with each (q_i, p_i) forming a pair that J is to "swap with a minus". (Cf. the discussions above about the direct-sum way of thinking about complexification and complex structure.) The J thus defined will be, by construction, compatible with the given ω. So does compatibility with this fixed ω also fix, i.e. determine, J?

No. For we must remember that our vector space has no concept of *length* of vectors: it has only a concept of area given by the symplectic form (cf. the discussion in Section 3). So for each $i = 1, ..., n$, and each q_i in the basis yielding the matrix form in eq. 2.12, there is a positive-real-parameter family of vectors p_i, any one of which can be chosen while preserving ω's form in eq. 2.12. So with $\dim(V) = 2n$, there is an entire $(\mathbb{R}^+)^n$ 'hyperquadrant' in $\mathbb{R}^n$ of choices of the n vectors p_i. (Note that this freedom in J is not just a choice of sign, as discussed for complexifications in Section 2.1.C.)

We will see later a physical rationale for this: elegant and helpful, since it concerns the simple harmonic oscillator (SHO). In one spatial dimension, the SHO has a phase space $\mathbb{R}^2 \ni (q_1, p_1)$, with the system's possible trajectories (histories) being ellipses. But this copy of $\mathbb{R}^2$ has no concept of length, but only of area. A choice of J will thus encode facts about the eccentricity of the ellipses, and thus about the dynamics (the Hamiltonian). (The image $J((q_1, p_1))$ of a point (q_1, p_1) under the action of J will lie on the same ellipse as (q_1, p_1).)

The idea that J—and a closely associated map K that "maps from the (complex!) classical solution space to the quantum Hilbert space"—encode facts about the dynamics will be important in the sequel: also for understanding the *Unruh effect*. Cf. 2.4.C below.

2.4.B: From vector space to manifold:— In Hamiltonian mechanics, the phase space is in general a manifold, not a vector space. Namely, a symplectic manifold. Usually, this is the cotangent bundle of the configuration space. But if it is not, Darboux' theorem secures that locally it can be written as a cotangent bundle, and so has a canonical decomposition in to qs and ps, that associates each q one-to-one with a p.

However, in the sequel, we will be mostly concerned with the "happy" case of a phase space that is a vector space. It may be infinite-dimensional, as for classical fields; or it may be finite-dimensional, as for n uncoupled SHOs. In either case, a linear combination of solutions is itself a solution. For classical fields on a spatial manifold, e.g. $\mathbb{R}^3$, we add—or more generally, linearly combine—the field configurations and the momenta pointwise. For n uncoupled SHOs,

we add (linearly combine) for each SHO independently. If we are given two solutions for the ith SHO (with a frequency ω_i say), labelled by their amplitude and phase (i.e. amplitude at time $t = 0$), we just add the two amplitudes and the two phases.

For any symplectic manifold M, we can of course rehearse for the tangent space T_pM at each point $p \in M$, and for its dual space T_p^*M, the development above from 2.2.A to 2.3.C. This means that given a symplectic form ω that smoothly varies across a local neighbourhood $U \subset M$, the bases it defines as in (1) above, i.e. the bases of T_pM at each point $p \in U$ such that ω's matrix is as in eq. 2.12 (cf. Section 2.1), also vary smoothly. And so the expression of J varies smoothly. In short, the local constructions presented above, from 2.2.A to 2.3.C, can be smoothly meshed with each other at the points in a local neighbourhood $U \subset M$.

But this still leaves open the question of *global existence* of a smooth J compatible with the global smooth ω. There can be obstructions to global existence. (The exposition of Wald (1994) assumes there are none.) So when we do the local construction of J at each point $p \in M$, as above, we say there is an *almost complex structure*. For details of this, see e.g. Da Silva (2001, Section V).

2.4.C: Complexifying the classical solution space; and then splitting the frequencies in different ways:— When we study linear systems (Section 3.7), we will see that a complex structure J corresponds to a splitting of the frequencies of complex classical solutions into positive and negative frequencies; and we will later see that having more than one complex structure J underlies the *Unruh effect*. The idea will be that in the Unruh effect, there are two different notions of time-evolution (two different Killing fields, two different Hamiltonians), that determine different *one particle structures* (cf. Section 4), and so different complex structures J. The general ideas are as follows.[7]

We first take the complexification of the solution space of the classical linear system. Here, we identify the solutions with the initial states, thanks to the determinism of the classical equations of motion. So writing S for the real symplectic vector space of solutions, the complexification is $S^{\mathbb{C}}$ (cf. Section 2.1).

We then define a 'positive frequency'/'positive energy' Hilbert space $\mathcal{H}$ by its being spanned by (as the span of) the complex classical solutions that oscillate with purely positive frequency (NB: also written ω!). For the simple harmonic oscillator, this means the complex classical solutions: $q(t) = \alpha \exp(-i\omega t)$, α a constant in $\mathbb{C}$. (Think of the momentum information being in the imaginary part.) For n uncoupled simple harmonic oscillators with frequencies $\omega_1, ..., \omega_n$, this means: $q_j(t) = \alpha_j \exp(-i\omega_j t)$ with $j = 1, ...n$. So for the latter case, $\mathcal{H}$ has complex dimension n.

Then the 'negative frequency'/'negative energy' Hilbert space $\bar{\mathcal{H}}$ is the span of the complex classical solutions that oscillate with purely negative frequency. In Section 2.5, just below, we will see that $\bar{\mathcal{H}}$ can be taken as the *complex conjugate of* $\mathcal{H}$, as defined there.

$S^{\mathbb{C}}$ is then the direct sum of the positive and negative frequency Hilbert spaces: $S^{\mathbb{C}} = \mathcal{H} \oplus \bar{\mathcal{H}}$. This direct sum structure means that there is a real-linear one-to-one onto "projection map" $K : S \to \mathcal{H}$ that extracts the positive frequency part of any real classical solution. This map K "maps from the (complex!) classical solution space to the quantum Hilbert space". It is the (main part of the definition of) *one particle structure*, which wil be central in the sequel, both for quantization in general (obviously!) and for e.g. the Unruh effect. Cf. Section 4.

The Unruh effect then arises in a scenario (defined on Minkowski spacetime!) in which two different notions of time-evolution (two different Killing fields, two different Hamiltonians) yield: two different frequency-splittings in (two different direct sum decompositions of) $S^{\mathbb{C}}$, and so two different Js; and so two different maps K; and thus two different vacua (ground states),

[7] For further reference, see Wald (1994): (i) pp.24-29, for finite systems; and (ii) pp.35-43, especially 39-41, for infinite systems, i.e. the Klein-Gordon field.

and two different Fock spaces built from these vacua.

Besides: the failure of the Stone von Neumann theorem for infinite systems, means that here, 'different' means 'unitarily inequivalent'. That is: the two different Fock spaces built from the two vacua give unitarily inequivalent representations of the Weyl algebra.

Incidentally, Wald (1994, p.29 par.2) points out that also for finite systems, e.g. n uncoupled time-independent simple harmonic oscillators, one can choose a different frequency-splitting than the usual one, and so define a different vacuum (ground) state, which is usually called a *squeezed vacuum*. But here, there *is* unitary equivalence of representations.

There is a general philosophico-mathematical theme hereabouts: *singular limits*. That is: for every finite n, we have unitary equivalence; but for $n = \infty$, there is unitary inequivalence. We will see exactly the same for *spin-chains*. There, the canonical *anti*-commutation relations (CARs)—rather than CCRs—have for finite spin chains a unique representation upto unitary equivalence (the Jordan-Wigner theorem). But for infinite spin chains there are countless unitarily inequivalent representations.

2.5 Complex conjugation of spaces

2.5.A: Basics:— The *complex conjugate of complex vector space* W is the complex vector space $\overline{W}$ that has the same elements and additive group structure as $\overline{W}$, but whose scalar multiplication involves conjugation. That is: we define the scalar multiplication $*$ in $\overline{W}$ in terms of the scalar multiplication $\cdot$ in W by:

$$\alpha * w := \overline{\alpha} \cdot w \ , \quad \text{for all } \alpha \in \mathbb{C}, w \in \overline{W} \tag{2.30}$$

Various properties and results ensue!

(1) $\overline{\overline{W}} = W$.

(2) W and $\overline{W}$ have the same complex dimension. Note that the identity map $id : W \to \overline{W}$ is an antilinear map, since,

$$id(\alpha \cdot w) = \alpha \cdot w \equiv \overline{\alpha} * w = \overline{\alpha} * id(w) \tag{2.31}$$

and id maps any basis of W into a basis of $\overline{W}$. So id is an *anti-isomorphism* from W to $\overline{W}$. It is a "canonical" one in the sense that its definition needs no choice of basis. That is: it is defined in terms of the underlying identity of vectors.

But of course, there are countless anti-isomorphisms defined in terms of such bases (just like there are countless isomorphisms!). For given any two bases, $\{e_i\}$ and $\{f_i\}$, of W and $\overline{W}$ respectively, the map $\Theta : e_i \to f_i$ can be extended by *antilinearity* to be an antilinear map, an *anti-isomorphism*, from W to $\overline{W}$.

(3) If W and U are complex vector spaces, an antilinear map $f : W \to U$ can be regarded as an ordinary linear map $f : \overline{W} \to U$, since:

$$f(\alpha * w) = f(\overline{\alpha} \cdot w) = \overline{\overline{\alpha}} \cdot f(w) = \alpha \cdot f(w) \ ; \tag{2.32}$$

where in the last two expressions, $\overline{\overline{\alpha}} \cdot f(w)$ and $\alpha \cdot f(w)$, the $\cdot$ is of course scalar multiplication in the codomain space U.

Conversely, any linear map g defined on $\overline{W}$, $g : \overline{W} \to U$, gives rise to an antilinear map from W to U, which again we write with a g. That is, we write: $g : W \to U$. For if we write the scalar multiplication in W as $\cdot$ (as before) and the scalar multiplication in U as $\cdot\cdot$, then the map $g : W \to U$ obeys:

$$g(\alpha \cdot w) \equiv g(\overline{\alpha} * w) = \overline{\alpha} \cdot \cdot g(w) \ , \tag{2.33}$$

since $g : \overline{W} \to U$ is linear. So the defined map $g : W \to U$ is antilinear.

(4) A linear map between complex vector spaces, $f : W \to U$, gives rise to a corresponding *also!* linear map $\overline{f} : \overline{W} \to \overline{U}$ which has the same action as f. For $\overline{f}$ preserves scalar multiplication, since

$$\overline{f}(\alpha * w) := f(\overline{\alpha} \cdot w) = \overline{\alpha} \cdot f(w) = \alpha * \overline{f}(w) \ . \tag{2.34}$$

If W, U are finite-dimensional, and the matrix of f with respect to bases $\{e_i\}$ of W and $\{g_j\}$ of U is (c_{ij}), i.e. $f(e_i) = c_{ij} g_j$, then the matrix of the linear map $\overline{f} : \overline{W} \to \overline{U}$ with respect to the *same* (as regards the underlying identity of vectors!) bases, i.e. $\{e_i\}$ of $\overline{W}$ and $\{g_j\}$ of $\overline{U}$, is the matrix whose entries are the complex conjugates of the c_{ij}. For in U, $c_{ij} g_j$ is short for $c_{ij} \cdot g_j$. But $c_{ij} \cdot g_j = \overline{c_{ij}} * g_j$. In short: to get the matrix of $\overline{f}$ from the matrix of f, we take complex conjugates of entries—but we do not transpose!

(5) *The complex conjugate of a Hilbert space.* That a Hilbert space $\mathcal{H}$ has extra structure additional to being a vector space, viz. the inner product, implies that there *is* a canonical aka natural, i..e basis-independent, isomorphism between $\mathcal{H}$ and $\overline{\mathcal{H}}$.

Indeed, recall *Riesz' theorem*: for a separable Hilbert space $\mathcal{H}$, every continuous linear functional $F : \mathcal{H} \to \mathbb{C}$ is given by taking the inner product with a unique vector $\psi_F \in \mathcal{H}$. That is: $F(\cdot) = (\psi_F, \cdot)$. Since this inner product is *sesquilinear*, i.e. $(\alpha\psi, \beta\phi) = \overline{\alpha}\beta(\psi, \phi)$, there is natural *antilinear* bijection between continuous linear functionals and vectors in $\mathcal{H}$: $F \mapsto \psi_F$. This is antilinear because $(\alpha F) \mapsto \psi_{(\alpha F)} \equiv \overline{\alpha}.\psi_F$. (Here, the . is good old scalar multiplication in $\mathcal{H}$!).

So there is natural *linear* bijection—i.e. an isomorphism!—between continuous linear functionals and vectors in the complex conjugate Hilbert space $\overline{\mathcal{H}}$. That is the dual space of linear functionals, $\mathcal{H}^*$ can be identified with $\overline{\mathcal{H}}$. It then follows that if we identify $\mathcal{H}^{**}$ with $\mathcal{H}$, there is natural isomorphism between $\mathcal{H}^{**} \equiv \mathcal{H}$ and $(\overline{\mathcal{H}})^*$.

Exercise! : Is there a natural isomorphism between $(\overline{\mathcal{H}})^*$ and $\overline{\mathcal{H}^*}$?

(6) *The relation of complexifications to complex structures.*

YET TO DO (a) general ideas then (b) physics, i.e. about the complexification of classical solutions as direct sum of positive-frequency and negative frequency subspaces, with the J thus encoding a choice of positive-frequency.

3 Symplectic mechanics

The choice of complex structure discussed above plays a remarkable role in the passage from classical physics to quantum field theory. So: let us now return to classical systems! Not the mechanics of Newton and Leibniz, but the symplectic mechanics of Élie Cartan and Vladimir Arnold, in which Segal quantization is formulated.

We first review the general mathematical structure in which Hamilton's equations are formulated, in four Subsections. In the first two, we treat the phase space Γ informally. In the third, we look back at the Weyl algebra in the light of our treatment of Poisson brackets. In the fourth Subsection, Γ is a manifold, i.e. a cotangent bundle. In more detail, the plan is as follows.

In the first Subsection, starting from Hamilton's equations we begin to develop the idea of symplectic structure (Section 3.1). Then we write the classical Poisson brackets in terms of the symplectic product (Section 3.2). Then Section 3.3 will look back to the ideas of the Weyl algebra, given in Section 1.2, in the form using operators W—which combine the translations

in position and in momentum that were given separately by the operators U and V. Then we give the modern geometric formulation using a manifold, i.e. a cotangent bundle (Section 3.4). Though long, this development: (i) sets us in good stead for later Sections; (ii) is anyway just the tip of the iceberg of symplectic mechanics!

To give one example of (ii), Section 3.5 will discuss how time-translation invariance implies the local Hamiltonian form of time-evolution. (Philosophically, time-translation invariance is an analogue of the spatial homogeneity that we invoked in (A) in Section 1.1.)

Finally in the last two Subsections, we: (i) turn to Hamiltonian systems that admit a linear structure (Section 3.6) and (ii) illustrate this with the harmonic oscillator—and briefly, with linear fields (Section 3.7). This will set us up for the next Section's discussion of a *one-particle structure* that is the central object of quantization on the Segal approach.

3.1 From Hamilton's equations to symplectic forms

(1): Time evolution from the gradient of H:—
We begin with *Hamilton's equations*

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q^i} \quad ; \quad \frac{dq^i}{dt} = \frac{\partial H}{\partial p_i} \, , \tag{3.1}$$

on the phase space Γ of qs and ps. As mentioned in this Section's preamble, we will in this Seubsection treat the phase space Γ informally. Defining

$$\xi^\alpha = q^\alpha, \;\; \alpha = 1, ..., n \quad ; \quad \xi^\alpha = p_{\alpha - n}, \;\; \alpha = n+1, ..., 2n \tag{3.2}$$

Hamilton's equations become

$$\dot{\xi}^\alpha = \frac{\partial H}{\partial \xi^{\alpha+n}}, \;\; \alpha = 1, ..., n \quad ; \quad \dot{\xi}^\alpha = -\frac{\partial H}{\partial \xi^{\alpha-n}}, \;\; \alpha = n+1, ..., 2n \;\; . \tag{3.3}$$

Writing $\mathbf{1}$ and $\mathbf{0}$ for the $n \times n$ identity and zero matrices respectively, we define the $2n \times 2n$ *symplectic matrix* ω by

$$\omega := \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix} \;\; . \tag{3.4}$$

The matrix ω is antisymmetric, and has the properties, writing $\tilde{}$ for the transpose of a matrix, that

$$\tilde{\omega} = -\omega = \omega^{-1} \quad \text{so that} \quad \omega^2 = -\mathbf{1} \quad ; \quad \text{also} \quad \det \omega = 1. \tag{3.5}$$

Using ω, Hamilton's equations eq. 3.3 get the more symmetric form, in matrix notation

$$\dot{\xi} = \omega \frac{\partial H}{\partial \xi} \;\; . \tag{3.6}$$

In terms of components, writing $\omega^{\alpha\beta}$ for the matrix elements of ω, and $\partial_\alpha := \partial \, / \partial \xi^\alpha$, eq. 3.3 become

$$\dot{\xi}^\alpha = \omega^{\alpha\beta} \partial_\beta H. \tag{3.7}$$

Eq. 3.6 and 3.7 show how ω forms, from the naive gradient (column vector) ∇H of H on the phase space Γ of qs and ps, the vector field on Γ that gives the system's evolution: the *Hamiltonian vector field*, often written X_H. At a point $z = (q, p) \in \Gamma$, eq. 3.6 can be written

$$X_H(z) = \omega \nabla H(z). \tag{3.8}$$

The vector field X_H is also written as D (for 'dynamics').

In Section 3.4, we will see how this definition of a *vector* field from a gradient, i.e. a *covector* or 1-form field, arises from Γ's being a cotangent bundle. More precisely, we will see that any cotangent bundle has an intrinsic symplectic structure that provides, at each point of the base-manifold, a natural i.e. basis-independent isomorphism between the tangent space and the cotangent space. For the moment, we will in **(2)** and **(3)** below:

note a geometric interpretation of ω in terms of area; and then

generalize the above discussion of ω into the definition of a symplectic form for a fixed *vector space*.

(2): Interpretation in terms of areas:—
Let us begin with the simplest possible case: $\mathbb{R}^2 \ni (q, p)$, representing the phase space of a particle constrained to one spatial dimension. Here, the 2×2 matrix

$$\omega := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{3.9}$$

defines the antisymmetric bilinear form on $\mathbb{R}^2$:

$$A : ((q^1, p_1), (q^2, p_2)) \in \mathbb{R}^2 \times \mathbb{R}^2 \mapsto q^1 p_2 - q^2 p_1 \in \mathbb{R} \tag{3.10}$$

since

$$q^1 p_2 - q^2 p_1 = \begin{pmatrix} q^1 & p_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} q^2 \\ p_2 \end{pmatrix} = \det \begin{pmatrix} q^1 & q^2 \\ p_1 & p_2 \end{pmatrix} \ . \tag{3.11}$$

It is easy to prove that $A((q^1, p_1), (q^2, p_2)) \equiv q^1 p_2 - q^2 p_1$ is the signed area of the parallelogram spanned by $(q^1, p_1), (q^2, p_2)$, where the sign is positive (negative) if the shortest rotation from (q^1, p_1) to (q^2, p_2) is anti-clockwise (clockwise).

Similarly in $\mathbb{R}^{2n}$: the matrix ω of eq. 3.4 defines an antisymmetric bilinear form on $\mathbb{R}^{2n}$ whose value on a pair $(q, p) \equiv (q^1, ...q^n; p_1, ..., p_n), (q', p') \equiv (q'^1, ...q'^n; p'_1, ..., p'_n)$ is the sum of the signed areas of the n parallelograms formed by the projections of the vectors $(q, p), (q', p')$ onto the n coordinate planes labelled by pairs of axes, $i = 1, 2, ..., n$. That is to say, the value is:

$$\Sigma_{i=1}^n \ q^i p'_i - q'^i p_i \ . \tag{3.12}$$

This induction of bilinear forms from antisymmetric matrices can be generalized: there is a one-to-one correspondence between forms and matrices. In more detail: there is a one-to-one correspondence between antisymmetric bilinear forms on $\mathbb{R}^2$ and antisymmetric 2×2 matrices. It is easy to check that any such form, ω say, is given, for any basis v, w of $\mathbb{R}^2$, by the matrix $\begin{pmatrix} 0 & \omega(v,w) \\ -\omega(v,w) & 0 \end{pmatrix}$. Similarly for any integer n: one easily shows that there is a one-to-one correspondence between antisymmetric bilinear forms on $\mathbb{R}^n$ and antisymmetric $n \times n$ matrices. (In Hamiltonian mechanics as usually formulated, we consider the case where n is even and the matrix is non-singular, as in eq. 3.4.)

This geometric interpretation of ω is important for two reasons.

(i): The first reason is that the idea of an antisymmetric bilinear form on a copy of $\mathbb{R}^{2n}$ is the main part of the definition of a *symplectic form*, which is the central notion in the usual geometric formulation of Hamiltonian mechanics. More details shortly, for a fixed copy of $\mathbb{R}^{2n}$; and in Section 3.4, where the form is defined on many copies of $\mathbb{R}^{2n}$, each copy being the tangent space at a point in the cotangent bundle T^*Q.

(ii): The second reason is that the idea of (signed) area underpins the theory of forms (1-forms, 2-forms etc.): i.e. antisymmetric multilinear functions on products of copies of $\mathbb{R}^n$. And when these copies of $\mathbb{R}^n$ are copies of the tangent space at (one and the same) point in a manifold, these forms lead to the whole theory of integration on manifolds. One needs

this theory in order to make rigorous sense of any integration on a manifold beyond the most elementary (i.e. line-integrals); so it is crucial for almost any mathematical or physical theory using manifolds. In particular, it is crucial for Hamiltonian mechanics. So no wonder the *maestro* says that 'Hamiltonian mechanics cannot be understood without differential forms' (Arnold 1989, p. 163).

However, it turns out that we will not need many details about forms and the theory of integration. This is essentially because we focus only on the elementary idea of solving a mechanical problem by giving the time-evolution (a trajectory through the phase space Γ). This means we will focus on line-integrals: viz. integrating with respect to time the equations of motion; or equivalently, integrating the dynamical vector field on the state space. We have already seen this vector field as X_H in eq. 3.8; and we will see it again, for example in terms of Poisson brackets (eq. 3.49), and in geometric terms (Section 3.4). But throughout, the main idea will be as suggested by eq. 3.8: the vector field is determined by the symplectic matrix, "at" each point in the manifold Γ, acting on the gradient of the Hamiltonian function H. So in short: focussing on line-integrals enables us to side-step most of the theory of forms.[8]

(3): Bilinear forms and associated linear maps:—
We now generalize from the symplectic matrix ω to a *symplectic form*; in five extended comments.

(1): *Preliminaries*:—
Let V be a (real finite-dimensional) vector space, with basis $e_1, ..., e_i, ...e_n$. We write V^* for the dual space, and $e^1, ..., e^i, ...e^n$ for the dual basis: $e^i(e_j) := \delta^i_j$.

We recall that the isomorphism $e_i \mapsto e^i$ is basis-dependent: for a different basis, the corresponding isomorphism would be a different map. Only with the provision of appropriate extra structure would this isomorphism be basis-independent.

For physicists, the most familiar example of such a structure is the spacetime metric $\mathbf{g}$ in relativity theory. In terms of components, this basis-independence shows up in the way that $\mathbf{g}$ and its inverse lower and raise indices. As we will see in a moment, the underlying mathematical point is that because $\mathbf{g}$ is a bilinear form on a vector space V, i.e. $\mathbf{g} : V \times V \to \mathbb{R}$, and is non-degenerate, any $v \in V$ defines, independently of any choice of basis, an element of V^*: viz. the map $u \in V \mapsto \mathbf{g}(u, v)$. (In fact, V is the tangent space at a spacetime point; but this physical interpretation is irrelevant to the mathematical argument.) We will also see that Hamiltonian mechanics has a non-degenerate bilinear form, viz. a symplectic form, that similarly gives a basis-independent isomorphism between a vector space and its dual. (Roughly speaking, this vector space will be the $2n$-dimensional space of the qs and ps.)

On the other hand: for any vector space V, the isomorphism between V and V^{**} given by

$$e_i \mapsto [e_i] \in V^{**} : e^j \in V^* \mapsto e^j(e_i) = \delta^j_i \tag{3.13}$$

is basis-independent, and so we identify e_i with $[e_i]$, and V with V^{**}. We will write $< ; >$ (also written $< , >$) for the natural pairing (in either order) of V and V^*: e.g. $< e_i \ ; \ e^j > = < e^j \ ; \ e_i > = \ \delta^j_i$.

A linear map $A : V \to W$ induces (basis-independently) a *transpose* (aka: dual), written $\tilde{A}$ (or A^T or A^*), $\tilde{A} : W^* \to V^*$ by

$$\forall \alpha \in W^*, \forall v \in V : \ \ \tilde{A}(\alpha)(v) \ \equiv < \tilde{A}(\alpha) \ ; \ v > := \ \alpha(A(v)) \ \equiv \ (\alpha \circ A)(v) \ . \tag{3.14}$$

[8] But forms are essential for understanding integration over surfaces of dimension two or more: which one needs for the integral invariants approach to Hamiltonian mechanics, and its deep connection with Stokes' theorem.

If $A : V \to W$ is a linear map between real finite-dimensional vector spaces, its matrix with respect to bases $e_1, ..., e_i, ...e_n$ and $f_1, ..., f_j, ...f_m$ of V and W is given by:

$$A(e_i) = A_i^j f_j \ ; \quad \text{i.e. with } \ v = v^i e_i, \ \ (A(v))^j = A_i^j v^i \ . \tag{3.15}$$

So the upper index labels rows, and the lower index labels columns. Similarly, if $A : V \times W \to \mathbb{R}$ is a bilinear form, its matrix for these bases is defined as

$$A_{ij} := A(e_i, f_j) \tag{3.16}$$

so that on vectors $v = v^i e_i, w = w^j f_j$, we have: $A(v, w) = v^i A_{ij} w^j$.

(2): *Associated maps and forms*:—
Given a bilinear form $A : V \times W \to \mathbb{R}$, we define the *associated linear map* $A^\flat : V \to W^*$ by

$$A^\flat(v)(w) \ := \ A(v, w) \ . \tag{3.17}$$

Then $A^\flat(e_i) = A_{ij} f^j$: for both sides send any $w = w^j f_j$ to $A_{ij} w^j$. That is: the matrix of $A^\flat$ in the bases e_i, f^j of V and W^* is A_{ij}:

$$[A^\flat]_{ij} = A_{ij}. \tag{3.18}$$

On the other hand, we can proceed from linear maps to associated bilinear forms. Given a linear map $B : V \to W^*$, we define the *associated bilinear form* $B^\sharp$ on $V \times W^{**} \cong V \times W$ by

$$B^\sharp(v, w) \ = \ < B(v) \ ; \ w > \ . \tag{3.19}$$

If we put $A^\flat$ for B in eq. 3.19, its associated bilinear form, acting on vectors $v = v^i e_i, w = w^j f_j$, yields, by eq. 3.17:

$$(A^\flat)^\sharp(v, w) \ = \ < A^\flat(v) \ ; \ w > \ = \ A(v, w) \ . \tag{3.20}$$

One similarly shows that if $B : V \to W^*$, then $\forall w \in W$:

$$(B^\sharp)^\flat(v)(w) \equiv < (B^\sharp)^\flat(v) \ ; \ w > = \ B(v)(w) \equiv < B(v) \ ; \ w > \ \text{ so that } \ (B^\sharp)^\flat = B \ . \tag{3.21}$$

So the flat and sharp operations, $^\flat$ and $^\sharp$, are inverses.

(3): *Tensor products*:—
It will sometimes be helpful to put the above ideas in terms of *tensor products*. If $v \in V, w \in W$, we can think of v and w as elements of V^{**}, W^{**} respectively. So we define their tensor product as a bilinear form on $V^* \times W^*$ by requiring for all $\alpha \in V^*, \beta \in W^*$:

$$(v \otimes w)(\alpha, \beta) \ := \ v(\alpha) w(\beta) \ \equiv < v \ ; \ \alpha >< w \ ; \ \beta > \ . \tag{3.22}$$

Similarly for other choices of vector spaces or their duals. Given $\alpha \in V^*, \beta \in W^*$, their tensor product is a bilinear form on $V \times W$:

$$(\alpha \otimes \beta)(v, w) \ := \ \alpha(v) \beta(w) \ \equiv < v \ ; \ \alpha >< w \ ; \ \beta > \ . \tag{3.23}$$

Similarly, we can think of $\alpha \in V^*, w \in W$ as elements of V^* and W^{**} respectively, and so define their tensor product as a bilinear form on $V \times W^*$:

$$(\alpha \otimes w)(v, \beta) \ := \ \alpha(v) w(\beta) \ \equiv < v \ ; \ \alpha >< w \ ; \ \beta > \ . \tag{3.24}$$

In this way we can express the linear map $A : V \to W$ in terms of tensor products. Since

$$A(e_i) = A_i^j f_j \quad \text{iff} \ \ < A(e_i); f^j > \ = \ A_i^j \tag{3.25}$$

eq. 3.24 implies that

$$A \;=\; A_i^j \; e^i \otimes f_j \;\; . \tag{3.26}$$

Similarly, a bilinear form $A : V \times W \to \mathbb{R}$ with matrix $A_{ij} := A(e_i, f_j)$ (cf. eq. 3.16) is:

$$A \;=\; A_{ij} \; e^i \otimes f^j \tag{3.27}$$

The definitions of tensor product eq. 3.22, 3.23 and 3.24 generalize to higher-rank tensors (i.e. multilinear maps whose domains have more than two factors). But we will not need these generalizations.

(4): *Antisymmetric and non-degenerate forms*:—
We now specialize to the forms and maps of central interest in Hamiltonian mechanics. We take $W = V$, $\dim(V)$=n, and define a bilinear form $\omega : V \times V \to \mathbb{R}$ to be:

(i): *antisymmetric* iff: $\omega(v, v') = -\omega(v, v')$;

(ii): *non-degenerate* iff: if $\omega(v, v') = 0 \;\; \forall v' \in V$, then $v = 0$.

The form ω and its associated linear map $\omega^\flat : V \to V^*$ now have a square matrix ω_{ij} (cf. eq. 3.18). We define the *rank* of ω to be the rank of this matrix: equivalently, the dimension of the range $\omega^\flat(V)$.

We will also need the antisymmetrized version of eq. 3.23 that is definable when $W = V$. Namely, we define the *wedge-product* of $\alpha, \beta \in V^*$ to be the antisymmetric bilinear form on V, given by

$$\alpha \wedge \beta : (v, w) \in V \times V \mapsto (\alpha(v))(\beta(w)) - (\alpha(w))(\beta(v)) \in \mathbb{R} \; . \tag{3.28}$$

(The connection with the interpretation of the symplectic matrix in terms of areas, especially eq. 3.12, will become clear in a moment; and will be developed in **(2)** of Section 3.4.)

It is easy to show that for any bilinear form $\omega : V \times V \to \mathbb{R}$: ω is non-degenerate iff the matrix ω_{ij} is non-singular iff $\omega^\flat : V \to V^*$ is an isomorphism.

So a non-degenerate bilinear form establishes a basis-independent isomorphism between V and V^*; cf. the discussion of the spacetime metric $\mathbf{g}$ in (1) at the start of this Subsection.

Besides, this isomorphism $\omega^\flat$ has an inverse, suggesting another use of the sharp notation, viz. $\omega^\sharp$ is defined to be $(\omega^\flat)^{-1} : V^* \to V$. The isomorphism $\omega^\sharp : V^* \to V$ corresponds to ω's role, emphasised in **(1)** of Section 3.1, of defining a vector field X_H from dH. (But we will see in a moment that the space V implicitly considered in **(1)** of Section 3.1 really has more structure than being just any finite-dimensional real vector space. Namely, it is of the form $W \times W^*$.)

NB: This definition of $^\sharp$ is of course *not* equivalent to our previous definition, in eq. 3.19, since:

(i): on our previous definition, $^\sharp$ carried a linear map to a bilinear form, which reversed the passage by $^\flat$ from bilinear form to linear map, in the sense that for a bilinear form ω, we had $(\omega^\flat)^\sharp = \omega$; cf. eq. 3.20;

(ii): on the present definition, $^\sharp$ carries a bilinear form $\omega : V \times V \to \mathbb{R}$ to a linear map $\omega^\sharp : V^* \to V$, which inverts $^\flat$ in the sense (*different* from (i)) that

$$\omega^\sharp \circ \omega^\flat = id_V \;\; \text{and} \;\; \omega^\flat \circ \omega^\sharp = id_{V^*} \; . \tag{3.29}$$

So beware: though not equivalent, both definitions are used! But it is a natural ambiguity, in so far as the definitions "mesh". For example, one easily shows that our second definition, i.e. eq. 3.29, is equivalent to a natural expression:

$$\forall \alpha, \beta \in V^* : \; < \omega^\sharp(\alpha), \beta > := \; \omega((\omega^\flat)^{-1}(\alpha), (\omega^\flat)^{-1}(\beta)) \; . \tag{3.30}$$

It is also straightforward to show that for any bilinear form $\omega : V \times V \to \mathbb{R}$: if ω is antisymmetric of rank $r \leq n \equiv \dim(V)$, then r is even. That is: $r = 2s$ for some integer s, and there is a basis $e_1, ..., e_i, ..., e_n$ of V for which ω has a simple expansion as wedge-products

$$\omega \ = \ \Sigma_{i=1}^{s} \ e^i \wedge e^{i+s} \ ; \tag{3.31}$$

equivalently, ω has the $n \times n$ matrix

$$\omega \ = \ \begin{pmatrix} \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \ . \tag{3.32}$$

where $\mathbf{1}$ is the $s \times s$ identity matrix, and similarly for the zero matrices of various sizes. This *normal form* of antisymmetric bilinear forms is an analogue of the Gram-Schmidt theorem that an inner product space has an orthonormal basis, and is proved by an analogous argument.

(5): *Symplectic forms*:—
As usually formulated, Hamiltonian mechanics uses a non-degenerate antisymmetric bilinear form: i.e. $r = n$. So eq. 3.32 loses its bottom row and right column consisting of zero matrices, and reduces to the form of the naive symplectic matrix, eq. 3.4, in **(1)** of Section 3.1. Equivalently: eq. 3.31 reduces to eq. 3.12.

Accordingly, we define: a *symplectic form* on a (real finite-dimensional) vector space Z is a non-degenerate antisymmetric bilinear form ω on Z: $\omega : Z \times Z \to \mathbb{R}$. Z is then called a *symplectic vector space*. It follows that Z is of even dimension.

Besides, in Hamiltonian mechanics (as usually formulated) the vector space Z is a product $V \times V^*$ of a vector space and its dual. Indeed, this is already suggested by:

(i) the fact in Lagrangian mechanics, that the canonical momenta $p_i := \frac{\partial L}{\partial \dot{q}^i}$ transform as a 1-form; (a fact to which we will return in Section 3.4); and

(ii) the discussion in **(1)** of Section 3.1 about the one-form field ∇H determining a vector field X_H.

Thus we define the *canonical symplectic form* ω on $Z := V \times V^*$ by

$$\omega((v_1, \alpha_1), (v_2, \alpha_2)) \ := \ \alpha_2(v_1) - \alpha_1(v_2) \ . \tag{3.33}$$

So defined, ω is by construction a symplectic form, and so has the normal form given by eq. 3.4.

Given a symplectic vector space (Z, ω), the natural question arises which linear maps $A : Z \to Z$ preserve the normal form given by eq. 3.4. It is straightforward to show that this is equivalent to A preserving the form of Hamilton's equations (for any Hamiltonian); so that these maps A are called *canonical* (or *symplectic*, or *Poisson*). But since we do not need details about the theory of canonical transformations, we will not go into details about this. Suffice it to say here the following.

$A : Z \to Z$ is symplectic iff, writing $\tilde{}$ for the transpose (eq. 3.14) and using the second definition eq. 3.29 of $\sharp$, the following maps (both from Z^* to Z) are equal:

$$A \circ \omega^\sharp \circ \tilde{A} = \omega^\sharp \ ; \tag{3.34}$$

or in matrix notation, with the *matrix* ω given by eq. 3.4, and again writing $\tilde{}$ for the transpose of a matrix

$$A \omega \tilde{A} = \omega \ . \tag{3.35}$$

(Equivalent formulas are got by taking inverses. We get, respectively: $\tilde{A} \circ \omega^\flat \circ A = \omega^\flat$ and $\tilde{A} \omega A = \omega$.)

The set of all such linear symplectic maps $A : Z \to Z$ form a group, *the symplectic group*, written $\mathrm{Sp}(Z, \omega)$.

To sum up this Subsection:— We have, for a vector space V, $\dim(V) = n$, and $Z := V \times V^*$:

(i): the canonical symplectic form $\omega : Z \times Z \to \mathbb{R}$; with normal form given by eq. 3.4;

(ii): the associated linear map $\omega^\flat : Z \to Z^*$; which is an isomorphism, since ω is non-degenerate;

(iii): the associated linear map $\omega^\sharp : Z^* \to Z$; which is an isomorphism, since ω is non-degenerate; and is the inverse of $\omega^\flat$; (cf. eq. 3.29).

We will see shortly that Hamiltonian mechanics takes V to be the tangent space T_q at a point $q \in Q$, so that Z is $T_q \times T_q^*$, i.e. the tangent space to the space Γ of the qs and ps.

3.2 Poisson brackets

We have seen how a single scalar function H on phase space Γ determines the evolution of the system via a combination of partial differentiation (the gradient of H) with the symplectic matrix. We now express these ideas in terms of Poisson brackets. This is not just because of their central role in canonical quantization. Within classical mechanics, they give a very neat expression for the rate of change of any dynamical variable; it arises from how the Poisson bracket encodes the way that a scalar function determines a (certain kind of) vector field.

(1): Poisson brackets introduced:—
The rate of change of any dynamical variable f, taken as a scalar function on phase space Γ, $f(q, p) \in \mathbb{R}$, is given (with summation convention) by

$$\frac{df}{dt} = \dot{q}^i \frac{\partial f}{\partial q^i} + \dot{p}_i \frac{\partial f}{\partial p_i} \ . \tag{3.36}$$

(If f is time-dependent, $f : (q, p, t) \in \Gamma \times \mathbb{R} \mapsto f(q, p, t) \in \mathbb{R}$, the right-hand-side includes a term $\frac{\partial f}{\partial t}$. But we here set aside the time-dependent case.) Applying Hamilton's equations, this is

$$\frac{df}{dt} = \frac{\partial H}{\partial p_i}\frac{\partial f}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial f}{\partial p_i} \ . \tag{3.37}$$

This suggests that we define the Poisson bracket of any two such functions $f(q, p), g(q, p)$ by

$$\{f, g\} := \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} \ ; \tag{3.38}$$

so that the rate of change of f is given by

$$\frac{df}{dt} = \{f, H\} \ . \tag{3.39}$$

In terms of the $2n$ coordinates ξ^α (eq. 3.2) and the matrix elements $\omega^{\alpha\beta}$ of ω (eq. 3.7), we can write eq. 3.37 as

$$\frac{df}{dt} = (\partial_\alpha f)\dot{\xi}^\alpha = (\partial_\alpha f)\omega^{\alpha\beta}(\partial_\beta H) \ ; \tag{3.40}$$

and so we can define the Poisson bracket by

$$\{f, g\} := (\partial_\alpha f)\omega^{\alpha\beta}(\partial_\beta g) \equiv \frac{\partial f}{\partial \xi^\alpha}\omega^{\alpha\beta}\frac{\partial g}{\partial \xi^\beta} \ . \tag{3.41}$$

In matrix notation: writing the naive gradients of f and of g as column vectors ∇f and ∇g, and writing $\tilde{}$ for transpose, we have at any point $z = (q, p) \in \Gamma$:

$$\{f, g\}(z) = \tilde{\nabla} f(z).\omega.\nabla g(z). \tag{3.42}$$

With these definitions of the Poisson bracket, we readily infer the following five results. (Later discussion will bring out the significance of some of these.)

(1): Since the Poisson bracket is antisymmetric, H itself is a constant of the motion:

$$\frac{dH}{dt} = \{H, H\} \equiv 0 \ . \tag{3.43}$$

(2): The Poisson bracket of a product is given by "Leibniz's rule": i.e. for any three functions f, g, h, we have

$$\{f, h \cdot g\} = \{f, h\} \cdot g + h \cdot \{f, g\} \ . \tag{3.44}$$

(3): Taking the Poisson bracket as itself a dynamical variable, its time-derivative is given by a "Leibniz rule"; i.e. the Poisson bracket behaves like a product:

$$\frac{d}{dt}\{f, g\} = \{\frac{df}{dt}, g\} + \{f, \frac{dg}{dt}\} \ . \tag{3.45}$$

(4): The Jacobi identity (easily deduced from (3)):

$$\{\{f, h\}, g\} + \{\{g, f\}, h\} + \{\{h, g\}, f\} = 0 \ \ . \tag{3.46}$$

(5): The Poisson brackets for the qs, ps and ξs are:

$$\{\xi^\alpha, \xi^\beta\} = \omega^{\alpha\beta} \ \ ; \ \text{ i.e.} \tag{3.47}$$

$$\{q^i, p_j\} = \delta^i_j \ \ , \quad \{q^i, q^j\} = \{p_i, p_j\} = 0 \ . \tag{3.48}$$

Eq. 3.48 is very important, both for general theory and for problem-solving. The reason is that preservation of these Poisson brackets, by a smooth transformation of the $2n$ variables $(q, p) \to (Q(q, p), P(q, p))$, is necessary and sufficient for the transformation being canonical. Besides, in this equivalence 'canonical' can be understood both: in the usual elementary sense of preserving the form of Hamilton's equations, for any Hamiltonian function; and in the geometric sense of preserving the symplectic form (as explained (a) in (5), at the end of **(3)** of Section 3.1, and (b) for manifolds in Section 3.4).

Note here that, as the phrase 'for any Hamiltonian function' brings out, the notion of a canonical transformation is independent of the forces on the system as encoded in the Hamiltonian. That is: the notion is a matter of Γ's geometry—as we will emphasise in Section 3.4.

But we will not need to go into many details about canonical transformations: for we do not aim to survey the whole of Hamiltonian mechanics!

(2): Hamiltonian vector fields:—
We earlier described how the symplectic matrix enabled the scalar function H on Γ to determine a vector field X_H. The previous Subsection showed how the Poisson bracket expressed any dynamical variable's rate of change along X_H. We now bring these ideas together, and generalize.

Recall that a vector X at a point x of a manifold M can be identified with a directional derivative operator at x assigning to each smooth function f defined on a neighbourhood of x its directional derivative along any curve that has X as its tangent vector. Similarly here: the dynamical vector field $X_H =: D$ is a derivative operator on scalar functions, which can be written in terms the Poisson bracket:

$$D := X_H = \frac{d}{dt} = \dot{q}^i \frac{\partial}{\partial q^i} + \dot{p}_i \frac{\partial}{\partial p_i} = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i} = \{\cdot, H\} \ . \tag{3.49}$$

But this point applies to any smooth scalar, f say, on Γ. That is: although we think of H as the energy that determines the real physical evolution, the mathematics is of course the same for such an f. So any such function determines a vector field, X_f say, on Γ that generates what the evolution "would be if f was the Hamiltonian". Thinking of the integral curves as parametrized by s, we have

$$X_f = \frac{d}{ds} = \{\cdot, f\} \ . \tag{3.50}$$

X_f is called the *Hamiltonian vector field* of (for) f; just as, for the physical Hamiltonian, $f \equiv H$, Section 3.1 called X_H '*the* Hamiltonian vector field'.

The notion of a Hamiltonian vector field will be crucial for what follows. We begin with two remarks which we will need later.

(1): So every scalar f determines a Hamiltonian vector field X_f. But note that the converse is false: not every vector field X on Γ is the Hamiltonian vector field of some scalar. For a vector field (equations of motion) X, with components X^α in the coordinates ξ^α defined by eq. 3.2

$$\dot{\xi}^\alpha = X^\alpha(\xi) \ , \tag{3.51}$$

there need be no scalar $H : \Gamma \to \mathbb{R}$ such that, as required by eq. 3.7,

$$X^\alpha = \omega^{\alpha\beta}\partial_\beta H \ . \tag{3.52}$$

Thus Hamilton's equations have the special feature that all the right hand sides are, up to a sign, partial derivatives of a *single* function H. (In fact, this feature underpins the possibility of expressing the equations of motion by variational principles.)

We also note under what condition is a vector field X Hamiltonian. The answer is: X is locally Hamiltonian, i.e. there is locally a scalar f such that $X = X_f$, iff X generates a one-parameter family of canonical transformations. We will give a modern geometric proof of this at the end of Section 3.4. For the moment, we only need to note, as at the end of **(1)** above, that here 'canonical transformation' can be understood in the usual elementary sense as a transformation of Γ that preserves the form of Hamilton's equations (for any Hamiltonian); or equivalently, as preserving the Poisson bracket; or equivalently, as preserving the symplectic form (to be defined for manifolds, in Section 3.4).

3.3 Looking back to the Weyl algebra

We have now developed symplectic mechanics sufficiently that it is useful to look back at the Weyl algebra, given in Section 1.2, in the form using operators W—which combine the translations in position and in momentum that were given separately by the operators U and V.

If (we are lucky enough that!) the classical phase space is a vector space (e.g. $\mathbb{R}^{2n}$), then—as we have seen (in (5) at the end of **(2)** in Section 3.1)—we can make it a *symplectic vector space*: i.e. a pair (Z, Ω), where Z is the phase space—i.e. a real vector space—and Ω is a symplectic product. (Here, we writeΩ instead of ω.) The symplectic product $\Omega : Z \times Z \to \mathbb{R}$ is, by definition, anti-symmetric, linear and non-degenerate (i.e. if $\Omega(z_1, z_2) = 0$ for all z_2, then $z_1 = \mathbf{0}$).

We define *the symplectic product* Ω on $Z = \mathbb{R}^{2n} \ni z_1, z_2$ as in (1.17): which we repeat here:

$$\Omega(z_1, z_2) := \mathbf{a}_2.\mathbf{b}_1 - \mathbf{a}_1.\mathbf{b}_2, \tag{3.53}$$

Then—now looking briefly at the topic of quantization—the Weyl form of the CCRs, i.e. (1.14) and (1.10), are equivalent, as we saw in Section 1.2, to the following: for all $z, z_1, z_2 \in \mathbb{R}^{2n}$,

$$\begin{array}{rcl} W(z_1)W(z_2) & = & e^{\frac{1}{2}i\Omega(z_1,z_2)}W(z_1 + z_2); \\ W^\dagger(z) & = & W(-z)\ . \end{array} \tag{3.54}$$

Note that $\Omega(z, \cdot) : Z \to \mathbb{R}$ is a real-valued function on Z, and so a classical observable. In particular, $\Omega(z, \cdot) = q^i$ iff z has $(n + i)$th component $b_i = 1$ and the rest 0, and $\Omega(z, \cdot) = p_i$ iff z has ith component $a^i = -1$ and the rest 0. In general, $\Omega(z, \cdot)$ is some linear combination of p_is and q^is.

In this formulation, the classical Poisson bracket relations (Equation 1.1) may be written

$$\{\Omega(z_1, \cdot), \Omega(z_2, \cdot)\} = -\Omega(z_1, z_2)\ . \tag{3.55}$$

So the corresponding Heisenberg form of the CCRs are

$$[\hat{\Omega}(z_1, \cdot), \hat{\Omega}(z_2, \cdot)] = -i\Omega(z_1, z_2)\mathbb{1}\ . \tag{3.56}$$

Thus we seek a representation in which the map $z \mapsto \hat{\Omega}(z, \cdot)$ takes elements of Z to self-adjoint operators, and in which the Weyl unitaries defined by

$$W(z) := e^{i\hat{\Omega}(z,\cdot)}. \tag{3.57}$$

obey the Weyl algebra, eq. 1.16.

This is Wald's presentation: see Wald (1994, Ch. 2). Later we will use field operators Φ, for which $\Phi(Jz) = \hat{\Omega}(z, \cdot)$, or $\Phi(z) = -\hat{\Omega}(Jz, \cdot) = \hat{\Omega}(\cdot, Jz)$.

(1): ... And looking forward to symplectic manifolds ...:—
In the case where the classical phase space Γ is not a vector space, we must develop more tools in order to quantize—in particular, in order to define the Weyl algebra. Details are in Section 3.4. But the basic idea will be as follows.

In this case, we seek a group whose action on Γ is *transitive* and preserves the symplectic form $\omega := \sum_i \mathrm{d}p_i \wedge \mathrm{d}q^i$. (In the case that Γ is a vector space, this group is just the (abelian) additive group of translations in Γ, which is isomorphic to Γ. That is what allowed us to treat Γ as a symplectic vector space above.) For illustration, taking the case $\Gamma = \mathbb{R}^{2n}$, the group action is a $2n$-parameter family of diffeomorphisms associated with the vector fields (with constant coefficients)

$$X_z = \sum_{i=1}^{n} b_i \frac{\partial}{\partial q^i} - a^i \frac{\partial}{\partial p_i}, \tag{3.58}$$

for any $z := (\mathbf{a}, \mathbf{b}) \in \mathbb{R}^{2n}$. We may now act on any two such vector fields with the *symplectic form* ω with which Γ—being a symplectic manifold (cf. Section 3.4)—is equipped. This yields

$$\omega(X_{z_1}, X_{z_2}) = \mathbf{a}_2.\mathbf{b}_1 - \mathbf{a}_1.\mathbf{b}_2. \tag{3.59}$$

Our quantization problem then becomes the search for continuous families of unitaries $z \mapsto W(z)$ which respect this symplectic structure, as expressed in the Weyl algebra (1.16), setting $e^{\frac{1}{2}i\Omega(z_1,z_2)} = e^{\frac{1}{2}i\omega(X_{z_1},X_{z_2})}$. Since the Weyl algebra (1.16) is unitary up to the phase factor $e^{\frac{1}{2}i\omega(X_{z_1},X_{z_2})}$, it is a *projective unitary representation* of the group of *symplectomorphisms* on Γ.

3.4 A geometrical perspective

Now we develop the modern geometric description of Hamiltonian mechanics. We will build especially on Section 3.1. There will be four Subsections. First, we introduce the cotangent bundle T^*Q. Then we collect what we will need about forms. Then we can show that any cotangent bundle is a symplectic manifold. This enables us to formulate Hamilton's equations geometrically.

(1): Canonical momenta are one-forms: Γ as T^*Q:—
So far we have treated the phase space Γ informally: saying just that it is a $2n$-dimensional space coordinatized by the qs, a smooth coordinate system on the configuration manifold Q, and the ps. But in the Lagrangian framework—which we have been silent about—the ps are canonical momenta $\frac{\partial L}{\partial \dot{q}^i}$; and one shows that at each point $q \in Q$, the p_i transform as a 1-form. Accordingly we now take the physical state of the system to be a point in the cotangent bundle T^*Q, the $2n$-dimensional manifold whose points are pairs (q, p) with $q \in Q, p \in T^*_q$.

We stress that from now on, the symbol p has a (fruitful!) ambiguity, between "dynamics" and "kinematics/geometry". For p represents both:

(A) the conjugate momentum $\frac{\partial L}{\partial \dot{q}}$, which of course depends on the choice of L; and

(B) a point in a fibre T^*_q of the cotangent bundle T^*Q (i.e. a 1-form or covector); or relatedly: the components p_i of such a 1-form: notions that are independent of any choice of a Lagrangian or Hamiltonian.

In more detail:—

(A): Recall that in the Lagrangian framework, the Lagrange equations being second-order in time prompts us to take the initial q and $\dot{q}$ as chosen independently, with L (encoding the forces on the system) then determining the evolution (the Lagrangian dynamical vector field D)—and so also determining the actual "realized" value of $\dot{q}$ at other times as a function of q, and so ultimately, of t. Similarly here: Newton's second law being second-order in time prompts us to take the initial q and p as independent, with H (encoding the forces on the system) then determining the evolution (the Hamiltonian dynamical vector field D)—and so also determining the actual value of p at other times as a function of q, and so ultimately, of t. Besides, by passing via the Legendre transformation back to the Lagrangian framework, one can check that the later actual value of p is determined to equal $\frac{\partial L}{\partial \dot{q}}$.

(B): But p also represents any 1-form (so that p_i represents the 1-form's coordinates). Here, we need to recall three points:—

(i): A local coordinate system (a chart) on Q defines a basis in the tangent space T_q at any point q in the chart's domain. As usual, we write the chart's coordinate functions as q^i. So we shall temporarily denote the chart by $[q]$, so that there are coordinate functions $q^i : \mathrm{dom}([q]) \to \mathbb{R}$. We write elements of the coordinate basis as usual, as $\frac{\partial}{\partial q^i}$.

(ii): The chart $[q]$ thereby also defines a dual basis dq^i in the cotangent space T^*_q at any $q \in \mathrm{dom}([q])$. (Here we recall, *en passant*, that the isomorphism at each q between T_q and T^*_q, that maps the basis element $\frac{\partial}{\partial q^i} \in T_q$ to the one-form dq^i in the dual basis, is basis-*dependent*. A different basis $\frac{\partial}{\partial q'^i}$ would give a different isomorphism. Cf. the discussion in (1) of **(3)** of Section 3.1.)

(iii): Putting (i) and (ii) together: the chart $[q]$ thereby also induces a local coordinate

system on a neighbourhood of the cotangent bundle around any point $(q,p) \in T^*Q$ with $q \in \mathrm{dom}([q])$ and $p \in T^*_q$.

Putting (i)-(iii) together: the coordinates of any point (q,p) in T^*Q in such a coordinate system are usually also written as (q,p). That is: p is used for the components of *any* 1-form, in the basis dq^i dual to a coordinate basis $\frac{\partial}{\partial q^i}$. So, similarly to (i) above: we will write this induced chart on T^*Q as $[q,p]$.

(2): Forms, wedge-products and exterior derivatives:—
As we said at the end of the discussion of interpreting ω in terms of areas (**(2)** of Section 3.1): we can largely avoid the theory of forms. For what follows, we need to recall only:

(i) the idea of forms of various degrees, together comprising the exterior algebra, and equipped with operations of wedge-product and contraction;

(ii) the ideas of differential forms, the exterior derivative, and of exact and closed forms.

(1): *The exterior algebra; wedge-products and contractions*:—
We begin by recalling some ideas from **(2)** and **(3)** of Section 3.1. Let us again begin with the simplest possible case, $\mathbb{R}^2$, considered as a vector space: not as a manifold with a copy of itself as tangent space at each point.

If α, β are covectors, i.e. elements of $(\mathbb{R}^2)^*$, we define their *wedge-product*, an antisymmetric bilinear form on $\mathbb{R}^2$, by

$$\alpha \wedge \beta : (v,w) \in \mathbb{R}^2 \times \mathbb{R}^2 \mapsto (\alpha(v))(\beta(w)) - (\alpha(w))(\beta(v)) \in \mathbb{R} \ . \tag{3.60}$$

Let us write the standard basis elements of $\mathbb{R}^2$ as $\frac{\partial}{\partial q}$ and $\frac{\partial}{\partial p}$, with elements of $\mathbb{R}^2$ having components (q,p) in this basis; and let us write the elements of the dual basis as dq, dp. Recalling the definition of the area form A, eq. 3.10, we deduce that A is $dq \wedge dp$.

Similarly for $\mathbb{R}^{2n}$. Recall that the symplectic matrix defines an antisymmetric bilinear form on $\mathbb{R}^{2n}$ by eq. 3.12. The value on a pair $(q,p) \equiv (q^1, ...q^n; p_1, ..., p_n), (q', p') \equiv (q'^1, ...q'^n; p'_1, ..., p'_n)$ is the sum of the signed areas of the n parallelograms formed by the projections of the vectors $(q,p), (q',p')$ onto the n coordinate planes formed by pairs of axes. This is a sum of n wedge-products. That is to say: if we write the standard basis elements as $\frac{\partial}{\partial q^i}$ and $\frac{\partial}{\partial p_i}$, this form is $\omega := \Sigma_i \, dq^i \wedge dp_i$. It has the action on $\mathbb{R}^n \times \mathbb{R}^n$:

$$(q^i \frac{\partial}{\partial q^i} + p_i \frac{\partial}{\partial p_i}, q'^i \frac{\partial}{\partial q^i} + p'_i \frac{\partial}{\partial p_i}) \mapsto \Sigma_{i=1}^n \, q^i p'_i - q'^i p_i \ . \tag{3.61}$$

In general, if V, W are two (real finite-dimensional) vector spaces, we define: $L(V,W)$ to be the vector space of linear maps from V to W; $L^k(V,W)$ to be the vector space of k-multilinear maps from $V \times V \times .... \times V$ (k copies) to W; and $L^k_a(V,W)$ to be the subspace of $L^k(V,W)$ consisting of (wholly) antisymmetric maps.

We then define $\Omega^k(V) := L^k_a(V, \mathbb{R})$ for $k = 1, 2, ..., \dim(V)$, so that $\Omega^1(V) = V^*$. We also set $\Omega^0(V) := \mathbb{R}$. $\Omega^k(V)$ is called the space of *(exterior) k-forms* on V. If $\dim(V) = n$, then $\dim(\Omega^k(V)) = \begin{pmatrix} n \\ k \end{pmatrix}$.

The wedge-product, as defined above, can be extended to be an operation that defines, for $\alpha \in \Omega^k(V), \beta \in \Omega^l(V)$, an element $\alpha \wedge \beta \in \Omega^{k+l}(V)$. We can skip the details: suffice it to say that the idea is to take tensor products, as in (3) of **(3)** of Section 3.1, and anti-symmetrize.

But we will need the definition of the *contraction*, (also known as: *interior product*), of a k-form $\alpha \in \Omega^k(V)$ with a vector $v \in V$. We shall write this as $\mathbf{i}_v \alpha$. (It is also written with a hook notation.) We define the contraction $\mathbf{i}_v \alpha$ to be the $(k-1)$-form given by:

$$\mathbf{i}_v \alpha(v_2, ..., v_k) := \alpha(v, v_2, ..., v_k) \ . \tag{3.62}$$

It follows, for example, that contraction distributes over the wedge-product *modulo* a sign, in the following sense. If α is a k-form, and β a 1-form, then

$$\mathbf{i}_v(\alpha \wedge \beta) = (\mathbf{i}_v\alpha) \wedge \beta + (-1)^k \alpha \wedge (\mathbf{i}_v\beta) \ . \tag{3.63}$$

The direct sum of the vector spaces $\Omega^k(V), k = 0, 1, 2, ..., \dim(V) =: n$, has dimension 2^n. When this direct sum is considered as equipped with the wedge-product $\wedge$ and contraction $\mathbf{i}$, it is called the *exterior algebra* of V, written $\Omega(V)$.

(2): *Differential forms; the exterior derivative; the Poincaré Lemma*:—
We extend the discussion just given in (1) to a manifold M of dimension n, taking all the tangent spaces T_x at $x \in M$ as copies of the vector space V, and requiring fields of forms to be suitably smooth.

We begin by saying that a (smooth) scalar function $f : M \to \mathbb{R}$ is a 0-form field. Its *differential* or *gradient*, df, as defined by its action on all vector fields X, viz. mapping them to f's directional derivative along X

$$df(X) := X(f) \tag{3.64}$$

is a 1-form (covector) field, called a *differential 1-form*.

The set $\mathcal{F}(M)$ of all smooth scalar functions forms an (infinite-dimensional) vector space, indeed a ring, under pointwise operations. We write the set of vector fields on M as $\mathcal{X}(M)$, or as $\mathcal{T}^1_0(M)$; and the set of covector fields, i.e. differential 1-forms, on M as $\mathcal{X}^*(M)$, or as $\mathcal{T}^0_1(M)$. (So superscripts indicate the contravariant order, and subscripts the covariant order.)

Accordingly, we define: $\Omega^0(M) := \mathcal{F}(M)$; $\Omega^1(M) = \mathcal{T}^0_1(M)$; and so on. In short: $\Omega^k(M)$ is the set of smooth fields of exterior k-forms on the tangent spaces of M.

The wedge-product, as defined in Equation (3.60), can be extended to the various $\Omega^k(M)$. We form the direct sum of the (infinite-dimensional) vector spaces $\Omega^k(M), k = 0, 1, 2, ..., \dim(V) =: n$, and consider it as equipped with this extended wedge-product. We call it the *algebra of exterior differential forms* on M, written $\Omega(M)$.

Similarly, contraction, as defined in Equation (3.62), can be extended to $\Omega(M)$. On analogy with eq. 3.62, we define, for α a k-form field on M, and X a vector field on M, the contraction $\mathbf{i}_X\alpha$ to be the $(k-1)$-form given, at each point $x \in M$, by:

$$\mathbf{i}_X\alpha(x) : (v_2, ..., v_k) \mapsto \alpha(x)(X(x), v_2, ..., v_k) \in \mathbb{R} \ . \tag{3.65}$$

The *exterior derivative* is a differential operator on $\Omega(M)$ that maps a k-form field to a $(k+1)$-form field. In particular, it maps a scalar f to its differential (gradient) df. Indeed, it is the *unique* map from the k-form fields to the $(k+1)$-form fields $(k = 1, 2, ..., n)$ that generalizes the elementary notion of gradient $f \mapsto df$, subject to certain natural conditions.

To be precise: one can show that there is a unique family of maps $d^k : \Omega^k(M) \to \Omega^{k+1}(M)$, all of which, for simplicity, we write as $\mathbf{d}$, such that:

(a): If $f \in \mathcal{F}(M)$, $\mathbf{d}(f) = df$.

(b): $\mathbf{d}$ is $\mathbb{R}$-linear; and distributes across the wedge-product, *modulo* a sign. That is: for $\alpha \in \Omega^k(M), \beta \in \Omega^l(M)$, $\mathbf{d}(\alpha \wedge \beta) = (\mathbf{d}\alpha) \wedge \beta + (-1)^k \alpha \wedge (\mathbf{d}\beta)$. (Cf. eq. 3.63.)

(c): $\mathbf{d}^2 := \mathbf{d} \circ \mathbf{d} \equiv 0$; i.e. for all $\alpha \in \Omega^k(M)$ $d^{k+1} \circ d^k(\alpha) \equiv 0$. (This condition looks strong, but is in fact natural. For its motivation, it must here suffice to say that it generalizes the fact in elementary vector calculus, that the curl of any gradient is zero: $\nabla \wedge (\nabla f) \equiv 0$.)

(d): $\mathbf{d}$ is a *local operator*; i.e. for any $x \in M$ and any k-form α, $\mathbf{d}\alpha(x)$ depends only on α's restriction to any open neighbourhood of x; more precisely, we define for any open set U of M, the vector space $\Omega^k(U)$ of k-form fields on U, and then require that

$$\mathbf{d}(\alpha \mid_U) = (\mathbf{d}\alpha) \mid_U \quad . \tag{3.66}$$

To express $\mathbf{d}$ in terms of coordinates: if $\alpha \in \Omega^k(M)$, i.e. α is a k-form on M, given in coordinates by

$$\alpha = \alpha_{i_1 \ldots i_k}\, dx^{i_1} \wedge \cdots \wedge dx^{i_k} \quad (\text{sum on } \; i_1 < i_2 < \cdots < i_k), \tag{3.67}$$

then one proves that the exterior derivative is

$$\mathbf{d}\alpha \;\; = \;\; \frac{\partial \alpha_{i_1 \ldots i_k}}{\partial x^j}\, dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \quad (\text{sum on all } j \text{ and } i_1, \cdots < i_k), \tag{3.68}$$

We define $\alpha \in \Omega^k(M)$ to be:

exact if there is a $\beta \in \Omega^{k-1}(M)$ such that $\alpha = \mathbf{d}\beta$; (cf. the elementary definition of an exact differential);

closed if $\mathbf{d}\alpha = 0$.

It is immediate from condition (c) above, $\mathbf{d}^2 = 0$, that every exact form is closed. The converse is "locally true". This result, called the *Poincaré Lemma*, is important. For example, we will use it in **(4)** below to characterize which vector fields preserve a symplectic manifold's symplectic form.

To be precise: for any open set U of M, we define (as in condition (d) above) the vector space $\Omega^k(U)$ of k-form fields on U. Then the *Poincaré Lemma* states that if $\alpha \in \Omega^k(M)$ is closed, then at every $x \in M$ there is a neighbourhood U such that $\alpha \mid_U \in \Omega^k(U)$ is exact.

We will also need (again, for **(4)**'s characterization of which vector fields preserve the symplectic form) a useful formula relating the Lie derivative, contraction and the exterior derivative. Namely: *Cartan's magic formula*, which says that if X is a vector field and α a k-form on a manifold M, then the Lie derivative of α with respect to X (i.e. along the flow of X) is

$$\mathcal{L}_X \alpha = \mathbf{d}\mathbf{i}_X \alpha + \mathbf{i}_X \mathbf{d}\alpha \;\; . \tag{3.69}$$

This is proved by straightforward calculation.

(3): Symplectic manifolds; the cotangent bundle as a symplectic manifold:— Any cotangent bundle T^*Q has a natural *symplectic structure*, which is the geometric structure on manifolds corresponding to the symplectic matrix ω introduced by eq. 3.4, and to the symplectic forms on vector spaces defined in (5) at the end of Section 3.1. (Here 'natural' means intrinsic, and in particular, independent of a choice of coordinates or bases.) It is this structure that enables a scalar function to determine a dynamics. That is: the symplectic structure implies that any scalar function $H : T^*Q \to \mathbb{R}$ defines a vector field X_H on T^*Q.

We first describe this structure (in (1)); and then in (2), show that any cotangent bundle has it.

(1): *Symplectic manifolds*:—

A *symplectic structure* or *symplectic form* on a manifold M is defined to be a differential 2-form ω on M that is closed (i.e. $\mathbf{d}\omega = 0$) and non-degenerate. That is: for any $x \in M$, and any two tangent vectors at x, $\sigma, \tau \in T_x$:

$$\mathbf{d}\omega = 0 \quad \text{and} \quad \forall\, \tau \neq 0, \;\; \exists \sigma : \quad \omega(\tau, \sigma) \neq 0 \;\; . \tag{3.70}$$

Such a pair (M, ω) is called a *symplectic manifold*.

There is a rich theory of symplectic manifolds; but we shall only need a small fragment of it, building on our discussion since Section 3.1. (In particular, the fact that we mostly avoid the theory of canonical transformations means we will not need the theory of Lagrangian submanifolds.)

First, it follows from the non-degeneracy of ω that M is even-dimensional; (cf. eq. 3.32).

It also follows that at any $x \in M$, there is a basis-independent isomorphism $\omega^\flat$ from the tangent space T_x to its dual T_x^*. We saw this in (2) and (4) of **(3)** of Section 3.1, especially eq. 3.17. Namely: for any $x \in M$ and $\tau \in T_x$, the value of the 1-form $\omega^\flat(\tau) \in T_x^*$ is defined by

$$\omega^\flat(\tau)(\sigma) := \omega(\sigma, \tau) \quad \forall \sigma \in T_x \ . \tag{3.71}$$

Here we return to the main idea emphasised already in **(1)** of Section 3.1: that symplectic structure enables a covector field, i.e. a differential one-form, to determine a vector field. Thus for any function $H : M \to \mathbb{R}$, so that dH is a differential 1-form on M, the inverse of $\omega^\flat$ (which we might write as $\omega^\sharp$), carries dH to a vector field on M, written X_H. Cf. eq. 3.8.

So far, we have noted some implications of ω being non-degenerate. The other part of the definition of a symplectic form (for a manifold), viz. ω being closed, $\mathbf{d}\omega = 0$, is also important. We shall see that it underlies the characterization in **(4)** below of which vector fields preserve the symplectic form.

So much by way of introducing symplectic manifolds. We turn to showing that any cotangent bundle T^*Q is such a manifold.

(2): *The cotangent bundle*:—
Choose any local coordinates q on Q ($\dim(Q)$=n), and the natural local coordinates q, p thereby induced on T^*Q; (cf. (B) of **(1)** above). We define the 2-form

$$dp \wedge dq := dp_i \wedge dq^i := \Sigma_{i=1}^n dp_i \wedge dq^i \ . \tag{3.72}$$

To show that eq. 3.72 defines the same 2-form, whatever choice we make of the chart q on Q, it suffices to show that $dp \wedge dq$ is the exterior derivative of a 1-form on T^*Q which is defined naturally (i.e. independently of coordinates or bases) from the derivative (also known as: tangent) map of the projection

$$\pi : (q, p) \in T^*Q \mapsto q \in Q. \tag{3.73}$$

Thus consider a tangent vector τ (not to Q, but) to the cotangent bundle T^*Q at a point $\eta = (q, p) \in T^*Q$, i.e. $q \in Q$ and $p \in T_q^*$. Let us write this as: $\tau \in T_\eta(T^*Q) \equiv T_{(q,p)}(T^*Q)$. The derivative map, $D\pi$ say, of the natural projection π applies to τ:

$$D\pi : \tau \in T_{(q,p)}(T^*Q) \mapsto (D\pi(\tau)) \in T_q \ . \tag{3.74}$$

Now define a 1-form θ_H on T^*Q by

$$\theta_H : \tau \in T_{(q,p)}(T^*Q) \mapsto p(D\pi(\tau)) \in \mathbb{R} \ ; \tag{3.75}$$

where in this definition of θ_H, p is defined to be the second component of τ's base-point $(q, p) \in T^*Q$; i.e. $\tau \in T_{(q,p)}(T^*Q)$ and $p \in T_q^*$.

This 1-form is called the *canonical 1-form* on T^*Q. It is the "Hamiltonian cousin" of a 1-form defined in the Lagrangian framework (and also there called the 'canonical 1-form'.) But our discussion of the "fruitful ambiguity" of the symbol p brings out a contrast with the Lagrangian case. While the Lagrangian 'canonical 1-form' clearly depends on the Lagrangian function L, the definition of θ_H, eq. 3.75, does *not* depend on any function H. θ_H is given just by the cotangent bundle structure. Hence the subscript H here just indicates "Hamiltonian (as against Lagrangian) version"—*not* dependence on a function H.

So much by way of a natural definition of a 1-form. One now checks that in any natural local coordinates q, p, θ_H is given by

$$\theta_H = p_i dq^i. \tag{3.76}$$

Finally, we define a 2-form by taking the exterior derivative of θ_H:

$$\mathbf{d}(\theta_H) := \mathbf{d}(p_i dq^i) \equiv dp_i \wedge dq^i \ . \tag{3.77}$$

where the last equation follows immediately from eq. 3.68. One checks that this 2-form is closed (since $\mathbf{d}^2 = 0$) and non-degenerate. So $(T^*Q, \mathbf{d}(\theta_H))$ is a symplectic manifold.

Referring to eq. 3.12 or eq. 3.33, both in Section 3.1, or eq. 3.61 of (1) in **(2)** above, we see that at each point $(q, p) \in T^*Q$, this symplectic form is, upto a sign, our familiar "sum of signed areas"—first seen as induced by the matrix ω of eq. 3.4.

Accordingly, Section 3.1's definition of a canonical symplectic form is extended to the present case: $\mathbf{d}(\theta_H)$, or its negative $-\mathbf{d}(\theta_H)$, is called the *canonical symplectic form*, or *canonical 2-form*. (The difference from Section 3.1's definition is that on a manifold, the symplectic form is required to be closed.)

(The difference by a sign is of course conventional: it arises from our taking the qs, not the ps, as the first n out of the $2n$ coordinates. For if we had instead taken the ps, the matrix occurring in eq. 3.6 would have been $-\omega \equiv \omega^{-1}$: exactly matching the cotangent bundle's intrinsic 2-form $\mathbf{d}(\theta_H)$.)

A famous theorem (Darboux's theorem) says that locally, any symplectic manifold "looks like" a cotangent bundle: or in other words, a cotangent bundle is locally a "universal" example of symplectic structure. But we turn to giving a geometric perspective on Hamilton's equations.

(4): Geometric formulations of Hamilton's equations:—
We already emphasised in Sections 3.1 and 3.2 the main geometric idea behind Hamilton's equations: that a gradient, i.e. covector, field dH determines a vector field X_H. We first saw this determination via the symplectic matrix, in eq. 3.8 of **(1)** of Section 3.1, viz.

$$X_H(z) = \omega \nabla H(z) \ ; \tag{3.78}$$

and then via the Poisson bracket, in eq. 3.49 of Section 3.2, viz.

$$D := X_H = \frac{d}{dt} = \dot{q}^i \frac{\partial}{\partial q^i} + \dot{p}_i \frac{\partial}{\partial p_i} = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i} = \{\cdot, H\} \ . \tag{3.79}$$

The symplectic structure and Poisson bracket were related by eq. 3.42, viz.

$$\{f, g\}(z) = \tilde{\nabla} f(z).\omega.\nabla g(z). \tag{3.80}$$

And to this earlier discussion, **(3)** above has added the identification of the canonical symplectic form of a cotangent bundle, eq. 3.77.

Let us sum up these discussions by giving some geometric formulations of Hamilton's equations at a point $z = (q, p)$ in a cotangent bundle T^*Q. Let us write $\omega^\sharp$ for the (basis-independent) isomorphism from the cotangent space to the tangent space, $T^*_z \to T_z$, induced by $\omega := -\mathbf{d}(\theta_H) = dq^i \wedge dp_i$ (cf. eq. 3.29 and 3.71). Then Hamilton's equations, eq. 3.8 or 3.78, may be written as:

$$\dot{z} = X_H(z) = \omega^\sharp(\mathbf{d}H(z)) = \omega^\sharp(dH(z)) \ . \tag{3.81}$$

Applying $\omega^\flat$, the inverse isomorphism $T_z \to T^*_z$, to both sides, we get

$$\omega^\flat X_H(z) = dH(z) \ . \tag{3.82}$$

In terms of the symplectic form ω at z, this is (cf. eq. 3.17): for all vectors $\tau \in T_z$

$$\omega(X_H(z), \tau) = dH(z) \cdot \tau \ ; \tag{3.83}$$

or in terms of the contraction defined by eq. 3.62, with $\cdot$ marking the argument place of $\tau \in T_z$:

$$\mathbf{i}_{X_H}\omega := \omega(X_H(z), \cdot) = dH(z)(\cdot) \ . \tag{3.84}$$

More briefly, and now for any function f, it is:

$$\mathbf{i}_{X_f}\omega = df \ . \tag{3.85}$$

Here is a final example. Recall the relation between the Poisson bracket and the directional derivative (or the Lie derivative $\mathcal{L}$) of a function, eq. 3.50 and 3.79: viz.

$$\mathcal{L}_{X_f} g = dg(X_f) = X_f(g) = \{g, f\} \ . \tag{3.86}$$

Combining this with eq. 3.85, we can reformulate the relation between the symplectic form and Poisson bracket, eq. 3.80, in the form:

$$\{g, f\} = dg(X_f) = \mathbf{i}_{X_f} dg = \mathbf{i}_{X_f}(\mathbf{i}_{X_g}\omega) = \omega(X_g, X_f) \ . \tag{3.87}$$

(1): *Which vector fields preserve the symplectic form?*:—
We turn to the promised answer to this question. Namely: A vector field X on a symplectic manifold M preserves the symplectic form ω (i.e. in more physical jargon: generates (a one-parameter family of) canonical transformations) iff X is Hamiltonian in the sense of **(2)** of Section 3.2; i.e. there is a scalar function f such that $X = X_f \equiv \omega^\sharp(df)$. Or in terms of the Poisson bracket, with $\cdot$ representing the argument place for a scalar function: $X(\cdot) = X_f(\cdot) \equiv \{\cdot, f\}$. In summary: a vector field on any symplectic manifold (M, ω)—it need not be a cotangent bundle—generates a one-parameter family of canonical transformations iff it is a Hamiltonian vector field.

Cartan's magic formula and the Poincaré Lemma make it easy to prove this.

We define a vector field X on a symplectic manifold (M, ω) to be *symplectic* (also known as: *canonical*) iff the Lie-derivative along X of the symplectic form vanishes, i.e. $\mathcal{L}_X\omega = 0$.[9]

Since ω is closed, i.e. $\mathbf{d}\omega = 0$, Cartan's magic formula, eq. 3.69, applied to ω becomes

$$\mathcal{L}_X\omega \equiv \mathbf{d}\mathbf{i}_X\omega + \mathbf{i}_X\mathbf{d}\omega = \mathbf{d}\mathbf{i}_X\omega \ . \tag{3.88}$$

So for X to be symplectic is for $\mathbf{i}_X\omega$ to be closed. But by the Poincaré Lemma, if $\mathbf{i}_X\omega$ is closed, it is locally exact. That is: there locally exists a scalar function $f : M \to \mathbb{R}$ such that

$$\mathbf{i}_X\omega = df \ \text{ i.e. } \ X = X_f \ . \tag{3.89}$$

So for X to be symplectic is equivalent to X being *locally Hamiltonian.*

3.5 Symplectic vector fields from time-translation invariance

This Section is deliberately written without prerequisites drawn from previous Subsections. So it serves as a snappy refresher of those Subsections' ideas—in the form of an "exercise" in the philosophy of time-translation invariance.

[9]Here, we assume the notion of the Lie-derivative, in particular the Lie-derivative of a 2-form. Suffice it to say, as a sketch, that the flow of X defines a map on M which induces a map on curves, and so on vectors, and so on co-vectors, and so on 2-forms such as ω. Nor will we go into details about the equivalence between this definition of X's being symplectic, and X's generating (active) canonical transformations, or preserving the Poisson bracket. For as we have emphasised, we will not need to develop the theory of canonical transformations.

In analytic mechanics, the state of a physical system is a point in a $2n$-dimensional manifold M, standardly expressed in terms of local coordinates $(\mathbf{q}, \mathbf{p}) := (q_1, \ldots, q_n, p_1, \ldots, p_n)$. A smooth function $H : M \to \mathbb{R}$ called the *Hamiltonian* then determines how the point $(\mathbf{q}, \mathbf{p})$ changes over time, by the postulate that it is given by the curve $(\mathbf{q}(t), \mathbf{p}(t))$ with $(\mathbf{q}(0), \mathbf{p}(0) = (\mathbf{q}, \mathbf{p})$ satisfying the system of ordinary differential equations known as *Hamilton's equations:*

$$\frac{d}{dt}q_i(t) = \frac{\partial H}{\partial p_i} \qquad\qquad \frac{d}{dt}p_i(t) = -\frac{\partial H}{\partial q_i} \tag{3.90}$$

for each $i = 1, \ldots, n$ and for all $t \in \mathbb{R}$.

Hamilton's equations are invariant under an arbitrary time translation $t_0 \in \mathbb{R}$, in that if $(\mathbf{q}(t), \mathbf{p}(t))$ is a solution, then so is $(\mathbf{q}(t + t_0), \mathbf{p}(t + t_0))$. This property captures an essential aspect of local physics, that experiments can be repeated at different times and produce the same results. We will now sketch how symplectic mechanics treats time translation invariance.

Given a manifold M, a *symplectic form* ω is a closed, non-degenerate, bilinear two-form on M; by this we mean that it is a smoothly defined tensor field at each point $p \in M$, where it is a bilinear function taking pairs of vectors at p to a real number, $\omega : v \times w \mapsto r \in \mathbb{R}$, and which satisfies the following properties.

(i) (*skew-symmetry*) $\omega(v, w) = \omega(w, v)$;

(ii) (*non-degeneracy*) if $\omega(v, w) = 0$ for all vectors w then $v = 0$;

(iii) (*closure*) $d\omega = 0$, where d is the exterior derivative.[10]

The pair (M, ω) is called a *symplectic manifold*, and is the arena in which symplectic mechanics is formulated. To express the axiom of time translation invariance, we say that a possible motion of a system is a smooth vector field X along which the symplectic form ω is invariant; this holds iff its Lie derivative satisfies $\mathcal{L}_X\omega = 0$, or equivalently[11] iff $d(\iota_X d\omega) = 0$, where $\iota_X\omega$ is the interior product[12] of X with ω. Such a vector field is called *symplectic*. The central postulate of symplectic mechanics can then be expressed, *if the vector field X describes a possible dynamical evolution, then X is symplectic, or equivalently $\iota_X\omega$ is closed.*

This postulate — although it may not look like it! — captures the essential structure of Hamilton's equations. One can see this in two steps. The first step is a deep fact of differential geometry, the 'Poincaré lemma': that if α is a closed k-form, then around every point $p \in M$ there is a neighbourhood in which α is *exact:* $\alpha = d\beta$ for some $(k-1)$-form β. One says for short: 'Every closed form is locally exact'. It is easy to check that the converse is globally true: every exact form is closed. So, the central postulate of symplectic mechanics is in fact equivalent to the statement: *if the vector field X describes a possible dynamical evolution, then $\iota_X\omega$ is locally exact:*

$$\iota_X\omega = dH \tag{3.91}$$

meaning that this equation holds in some neighbourhood of a point, and where the smooth function H is unique up to the addition of a constant function. A vector field X on (M, ω) that satisfies Equation (3.91) is called a *Hamiltonian vector field*, and is said to be 'locally generated'

[10] The *exterior derivative* d is the unique local linear mapping from k-forms to $(k+1)$-forms of M such that (i) if f is a smooth function, then df is its differential; (ii) d satisfies the product rule $d(\alpha \wedge \beta) = d\alpha \wedge (-1)^k \alpha \wedge d\beta$ (where α is a k-form); and (iii) $d^2 = 0$ (cf. Marsden and Ratiu 2010, Prop.4.2.4).

[11] This latter equivalence is an application of Cartan's 'magic formula', $\mathcal{L}_X\omega = d\iota_X\omega + \iota_X d\omega$, together with the fact that the exterior derivative satisfies $d^2 = 0$.

[12] The *interior product* of a vector and an n-form can be defined as the contraction of the vector into the first 'slot' of the n-form; in abstract index notation, it would be written $X^a\omega_{ab}$. For a reference, see Marsden and Ratiu (2010, Section 2.4).

by the Hamiltonian H, the value of which is invariant along X. Equation (3.91) is often called the 'coordinate-free' form of Hamilton's equations.

The second step reveals why this is so: an elementary argument shows how, if ω is a symplectic form on a $2n$-dimensional manifold M, then every point $p \in M$ admits a neighbourhood in which ω can be expressed in terms of some coordinate system $(\mathbf{q}, \mathbf{p})$ (called *local Darboux coordinates*) as,

$$\omega := \begin{pmatrix} & -\mathbb{1} \\ \mathbb{1} & \end{pmatrix} \tag{3.92}$$

(Marsden and Ratiu 2010, Theorem 5.1.2). That is: ω swaps the groups of qs and ps, but it otherwise acts as the identity on the 'q_i entries' of a vector, and as the negative identity on the 'p_i entries'. Now: writing the vector field X along one of its integral curves $(\mathbf{q}(t), \mathbf{p}(t))$ as $X = \left(\frac{d}{dt}q_1(t), \ldots, \frac{d}{dt}p_n(t)\right)$, and writing our one-form as $dH = \left(\frac{\partial H}{\partial q_1}, \ldots, \frac{\partial H}{\partial p_n}\right)$, we immediately conclude that Equation (3.91) is just the familiar form of Hamilton's equations — it is just Equation (3.90) in disguise!

This exercise also provides some physical intuition into how to view the symplectic form. For example, recall that a (anti-clockwise) rotation matrix in two dimensions can be written,

$$R_\theta := \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{3.93}$$

Comparing this to the matrix expression of ω for $n = 2$ in Equation (3.92) when $\dim M = 2$, one finds that the symplectic form is just a rotation through $\pi/2$, and so the *inverse* of the symplectic form is a rotation through $-\pi/2$. Since dH is the differential of H, we can view this matrix as 'rotating' dH through $-\pi/2$ into a level surface of h. This just repeats what we have already seen in more technical terms above: a Hamiltonian vector field X is one along which the Hamiltonian h (or 'energy') is conserved.

A yet more general interpretation of ω is in terms of areas. Viewing two vectors $(u, v), (u', v') \in \mathbb{R}^2$ at a point as defining a parallelogram, the matrix ω returns the area of the parallelogram, $(u, v)\omega(u', v')^\intercal = uv' - u'v$, as shown in Figure 1. The sign of the area is positive if the anti-clockwise rotation from (u, v) to (u', v') is less than π, and negative if it is greater. This generalises to arbitrary dimensions: for a pair of vectors ($\mathbf{u}, \mathbf{v}$ and $(\mathbf{u}', \mathbf{v}')$), the symplectic form ω returns the sum of the signed areas of each parallelogram, $\sum_{i=1}^{n} u_i v_i' - v_i' u_i$.

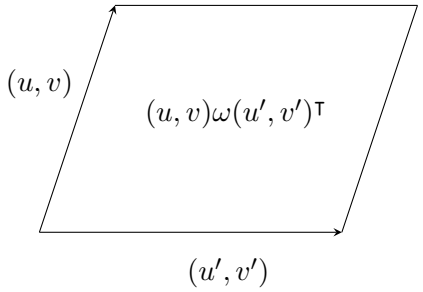

Figure 1: The symplectic form returns the signed area of a parallelogram.

3.6 Linear solution spaces

As announced in the preamble to this Section (Section 3), we end with two Subsections about *linear* systems (again in the Hamiltonian framework). This subsection covers generalities; the next discusses the harmonic oscillator—and briefly, linear fields.

Looking back after this review of symplectic mechanics ... It clearly helps us to fix on the essential structures that we wish to somehow quantize. But there is the difficulty that our classical system has a real manifold M with a symplectic form, but no linear structure: there is no non-conventional way to 'add' two distinct points in M. And yet, to quantize, we must somehow pass to a complex vector space. Section 2 shows how, if only we had a *linear* symplectic manifold — a symplectic manifold M that is also a real vector space — then our path forward would be obvious: choose a complex structure J! More precisely: we should choose a complex structure that preserves and tames the symplectic form, and thereby complexify the real vector space and define a Hilbert space; (recall Section 2.3). (As noted in Section 2.4, we need to beware that such a complex structure J is not unique.)

So we now specialize to this "happy" case. Indeed, there are two aspects here, about (i) kinematics, i.e. the structure of the state space, and (ii) dynamics, i.e. equations of motion. We will assume that :

1. M is a symplectic vector space, as discussed in detail in **(3)** (especially its last part (5)) of Section 3.1; and

2. the equations of motion preserve the linear structure of states: which will be a matter of the Hamiltonian being quadratic (or constant) in the q_i and p_i.

For various reasons, this restriction is less limiting than it might first seem. We can already state two reasons.

(A): In a classical theory of fields, configurations *superpose*: think of the interference, i.e. superposition, of water waves, electromagnetic waves etc. Thus we can hope our discussion will apply to infinite-dimensional linear configuration manifolds (and their cotangent bundles). This comment is about Condition 1. But also classical field theories often use linear wave equations, i.e. equations of motion that preserve the sum-structure of solutions ("the waves do not interact, they just add together"); illustrating Condition 2.

(B): Condition 2 means that, strictly speaking, the only finite-dimensional systems we will examine will be sets of coupled harmonic oscillators. But this is not so limiting: as the practised physicist knows, nearly everything can be modelled as a set of coupled harmonic oscillators! And indeed, the reason for this lies in the fact that if the Hamiltonian is quadratic in momentum, to which is added a potential $V(\mathbf{q})$ that is analytic, then the second-order approximation of its Taylor series is a harmonic oscillator. That is: we are second-order approximating a potential that is 'locally defined' (in the sense that its values everywhere are determined by its derivatives at a single point). Indeed, in even more elementary terms: since the force on a classical particle in a potential V is $-\nabla V$, the simplest, but not spatially constant, form for the force is that it should be linear in displacement—and such a form is given by a quadratic V. Besides, the usual "small oscillations", or "normal coordinates" analysis (diagonalizing the real symmetric matrix of the second partial derivatives of V) can be applied so as to describe a system of coupled oscillators as a set of independent, i.e. uncoupled, harmonic oscillators.

In the rest of this subsection, we make three comments. The first is about Condition 2, the second about Condition 2, and the third is about both Conditions (and combines the first two comments). All three commentsdraw on ideas in preceding Subsections; they are elementary but underpin what follows.

(1): From symplectic manifold to symplectic vector space:—
We recall that a real vector space V is itself a manifold in a trivial way: it has a global coordinate chart, and the tangent space T_xV at each point $x \in V$ is just (isomorphic to) V. So suppose we are given V as: both a symplectic manifold (V, ω), as defined (most officially!) at the start of **(3)** of Section 3.4; and also as a vector space. Then (cf. **(3)** (especially its last part (5)) of Section 3.1) we infer that V is even dimensional, with dimension $2n$ say; and we can write its points suggestively as pairs $(q, p) \equiv (q^1, \ldots, q^n, p_1, \ldots, p_n)$. Then we can view the symplectic form ω as a map on the vector space itself. Let us write it as Ω, following eq. 1.17 and 3.53. That is: we have a bilinear map $\Omega : V \times V \to \mathbb{R}$ that is skew-symmetric and non-degenerate—not any longer on pairs of tangent space vectors to the manifold V but—on pairs of points (q, p) in V. In short, we have a *symplectic vector space*.

(2): Quadratic Hamiltonians:—
We now spell out Condition 2: i.e. how a Hamiltonian being quadratic implies that time-evolution preserves linear structure. So let the phase space Γ be a vector space with global coordinates (q, p). As in **(1)** of Section 3.1, we write ξ^α, with α running form 1 to $2n$: cf. Eq. 3.2.

We now *define* a linear system as one in which the Hamiltonian is a quadratic form $H_{\alpha\beta}$ in the ξs. That is: the energy $= H = (\xi^\alpha)^T[H_{\alpha\beta}\xi^\beta]$. Then taking partial derivatives of the energy H with respect to any ξ^α (holding all other ξ^α constant of course) will give: a linear combination of the various ξ^β, i.e. a linear combination with constant coefficients. Call it $a_\alpha\xi^\alpha$ (with summation convention).Then ∇H is the column of these partial derivatives. Multiplying ∇H by the symplectic matrix (cf. e.g. eq. 3.6 or 3.7) keeps it a linear combination: (for one just flips qs and ps, and adds a – sign!). So the Hamiltonian vector field is a linear combination of the various ξ^β with constant coefficients. Call it $b_\alpha\xi^\alpha$ (with summation convention)

So at each point $\xi = (q, p) \in \Gamma$, the infinitesimal flow is: $b_\alpha\xi^\alpha$. Then it is trivial that the time-evolution preserves the linear structure of solutions. For take two points: $\xi_1 = (q_1, p_1)$ and $\xi_2 = (q_2, p_2)$. At the sum-state got by superposing these states, $\xi_{1+2} := (q_1 + q_2, p_1 + p_2)$, the infinitesimal flow is by definition: $b_\alpha\xi^\alpha_{1+2}$. But this is: $b_\alpha(\xi^\alpha_1 + \xi^\alpha_2)$: and since the b_α are just numbers this is: $b_\alpha(\xi^\alpha_1) + b_\alpha(\xi^\alpha_2)$.

In short: The sum of two instantaneous states has as its infinitesimal Hamiltonian flow (tangent vector in phase space) the sum of the two states' individual Hamiltonian flows (tangent vectors).

(3): A symplectic structure on the set of solutions:—
We recall from (1) at the end of **(4)** of Section 3.4 that a vector field X on a symplectic manifold M preserves the symplectic form iff it is Hamiltonian i.e. induced by a scalar field $f : M \to \mathbb{R}$ i.e. $X = X_f \equiv \omega^\sharp(df)$. This means that if M is a symplectic vector space V, as we are now assuming, then time-evolution under any Hamiltonian f (not necessarily a quadratic one) preserves the symplectic product. That is: given any two points $(q, p), (q', p') \in V$, the value of $\Omega((q, p), (q', p'))$ is invariant along a solution to Hamilton's equations (because the symplectic form is invariant along it):

$$\Omega((q(t), p(t)), (q'(t), p'(t))) = \text{constant}. \tag{3.94}$$

This means that it is possible to define the symplectic structure Ω not just on points in V, but on the set of entire solutions $\mathcal{S}$ to Hamilton's equations. Besides, Conditions 1-2 imply this set of solutions $\mathcal{S}$ is a vector space: it is closed under taking linear combinations. (Here, we are using Condition 2: the Hamiltonian being quadratic.) Thus it follows that $(\mathcal{S}, \Omega)$ is a symplectic vector space. This structure will be the focus of our attention in the theory of quantization.

3.7 Example: The simple (classical) harmonic oscillator

Let's step down from the level of abstraction just described, and remind ourselves about the simple harmonic oscillator in one dimension. Let the phase space $M = \mathbb{R}^2$ be written in global Cartesian coordinates (q, p) and viewed as a vector space, written V above. The symplectic form on the vectors at a point is $\omega = \left(\begin{smallmatrix} & -1 \\ 1 & \end{smallmatrix}\right)$, which gives rise to Hamilton's equations in their standard form. The Hamiltonian for the harmonic oscillator is the familiar one (now writing ν for the frequency, since ω represents the symplectic form):

$$H(q, p) = \tfrac{1}{2}p^2 + \tfrac{1}{2}\nu^2 q^2 \tag{3.95}$$

for some non-zero $\nu \in \mathbb{R}$.

Plugging this in to Hamilton's equations we find that $(d/dt)q(t) = p$ and $(d/dt)p(t) = -\nu^2 q$. Solutions to these equations are easily checked to have the form,

$$\begin{aligned} q(t) &= a\cos\nu t + b\sin\nu t \\ p(t) &= \nu b\cos\nu t - \nu a\sin\nu t \end{aligned} \tag{3.96}$$

for some $a, b \in \mathbb{R}$. That is, each solution is associated with a pair of constants $(a, b) \in \mathbb{R}^2$. One can now quickly confirm that the space $\mathcal{S}$ of solutions forms a vector space: given the solution of Equation (3.96) with constants (a, b) and another solution with constants (a', b'), their sum is obviously also a solution.

We can define Ω in terms of ω to get a symplectic structure $\Omega((q, p), (q', p')) := qp' - q'p$ on the vector space M or V. Its values do not change over time. So with $\mathcal{S}$ as the solution space for the simple harmonic oscillator: $(\mathcal{S}, \Omega)$ is a symplectic vector space.

4 One-particle structures and Fock space

This Section states the two core ideas of the Segal quantization of a linear classical system. *First*: there is a map K from the solution space of the classical system to a Hilbert space. K is required to satisfy conditions that combine the ideas of complex structures (Section 2) and symplectic structures (Section 3), in such a way that the Hilbert space is determined. Besides, it is determined as having a unitary dynamics that is the "unitary cousin" of the classical system's dynamics. K, or the Hilbert space to which it leads, is called a *one-particle structure*. *Second*: there is the usual Fock space construction, which will be applied to the one-particle structure's Hilbert space (i.e. after the first idea has been implemented). So here, the phrase 'one-particle' signals that the Hilbert space is the first (non-zero, i.e. non-vacuum) summand of the usual Fock space sum of ever larger tensor powers.

As we announced in the Preamble to this document: we will see this illustrated for the harmonic oscillator (in one spatial dimension). Starting with classical harmonic oscillator, the first idea delivers us as the quantum state space—not the familiar quantum harmonic oscillator, with (in one spatial dimension) Hilbert space $L^2(\mathbb{R})$!—but 'merely' the world's simplest complex Hilbert space, viz. $\mathbb{C}$ i.e. the complex plane. To get the familiar quantum harmonic oscillator, i.e. $L^2(\mathbb{R})$ (equipped with the quantum harmonic oscillator Hamiltonian), we need to take the Fock space built from $\mathbb{C}$. That Fock space will "be" (i.e. be a Hilbert space isomorphic to) $L^2(\mathbb{R})$. So we will in effect *factorize* the usual understanding of canonical quantization—viz. (for the 1-dimensional harmonic oscillator) "replace the two-dimensional classical phase space $\mathbb{R}^2 \ni (q, p)$, with $L^2(\mathbb{R})$, i.e. L^2 functions on the configuration space $\mathbb{R}$—into: *first*, build a 1-particle structure; *second*, build the Fock space.

The plan is as follows. First, we orient ourselves to Fock space, i.e. variable particle number, with general remarks about 'particle' and 'field' (Section 4.1). Then we review the idea of a one-particle structure in Section 4.2. Then Section 4.3 illustrates it with the simple

harmonic oscillator in one spatial dimension: getting as the Hilbert space "just" $\mathbb{C}$ i.e. the complex plane. Then we move to the "second step". We present the free bosonic field on any one-particle structure, in Section 4.4. We illustrate this with the harmonic oscillator, in Section 4.5. We are then in a position to remark en passant on an instance of the Stone-von Neumann theorem: roughly speaking, that finite systems have a unique quantization. (Details in Section 6.) Thus we describe in Section 4.6 how by changing the frequency of the harmonic oscillator, we can write down an apparently different quantization. But it is in fact unitarily equivalent to the first one.

4.1 'Particle' and 'field'

First, the plan just announced prompts a couple of warning remarks about the word 'quantization'.

(1): Beware: You might think that the one particle structure is about 'first quantization', with the Fock space construction being 'second quantization'. But that is not quite right, for three reasons. .

(1a): As just noted, what the one particle structure delivers us for a standard case, viz. the harmonic oscillator, is *not* the system that is usually called the 'first quantized' system. .

(1b): The classical linear system with which we begin can be infinite-dimensional, e.g. the solution space of a linear wave equation. So to the extent that 'second quantization' means just 'quantization of a field' (as it does for some people): in such a case, the construction of the one particle structure is 'second (i.e. field) quantization'. .

(1c): (Like (1b) but "in the other direction"): in the two-step quantization of the harmonic oscillator, the eventual Fock space is the state-space of a *finite* quantum system, viz. the state-space $L^2(\mathbb{R})$. So the second step can hardly be called 'field quantization'. .

(2): Beware: Some books say: a): first quantization is about particles behaving like waves; (b) Second quantization is about waves behaving like particles (e.g. Blundell and Lancaster *Quantum field theory for the Gifted Amateur*, p. 20). We have no quarrel with (a). After all: think of deBroglie (1924), Schroedinger's December 1925 fundamental idea, to replace the classical Hamiltonian state space T^*Q for a finite classical system with configuration space Q, by $L^2(Q)$; and at the experimental level: electron diffraction, and the two-slit experiment. But (b) is wrong, or at least misleading. For it suggests we need an infinite classical system ("waves"), or a Fock space construction, in order that a quantum theory (or a quantization) yield us particle like features such as: (i) an operator whose spectrum is the non-negative integers, i.e. what we can call a number operator $\hat{N}$; (ii) a position operator, so that we can talk about localization. But we don't! We have (ii) in elementary wave mechanics (Schroedinger picture/position representation). And as to (i) any orthobasis of any denumerable-dimensional Hilbert space can of course have the non-negative integers attached to its elements as eigenvalues, and ladder operators defined with respect to that basis. (And if you demand also that these eigenvalues count energy, then again ... the simple harmonic oscillator is a finite system that delivers (i).)

More positively, we will see that our discussion *does* bear on, indeed illuminate, the particle-wave duality (or better: particle-field duality). Spoiler alert: not in a revolutionary way! Two specific examples are:—

(1): The recovery of the "fieldy-wavy" state-space $L^2(\mathbb{R})$ for the quantum harmonic oscillator from the "particle-ish" Fock space which "counts excitations" will proceed by the "time-honoured" (i.e. historically significant) "equivalence" of matrix mechanics with wave mechanics": that is, the Hilbert space isomorphism from $l^2(\mathbb{N})$ to $L^2(\mathbb{R})$. (By the way: the equivalence is conceptually, and historically subtler than the textbooks suggest. Cf. F. Muller, The Equivalence Myth, *Studies in History and Philosophy of Modern Physics* 1995.)

(2): A deeper example lies in the unitary equivalence of:

(i) a particle (Fock-space- or number-operator-based) "way of thinking" of a quantum field theory; (in effect: the quantities in the *commutant of the number operator*); and:

(ii) a field or wave way of thinking; (in effect: the quantities in the *commutant of the field operator*).

This unitary equivalence is made precise and proven for free bosonic fields, that are obtained by complexifyig the solution space of a classical linear system, in Theorem 1.10 of Baez et al (1992: p. 49, and Section 1.8, 1.9). (Schweber (1962, Chapter 7d and 7e is a heuristic discussion of this equivalence.). Cf. Section 4.4 below: which will define properly what Baez et al call the particle and real wave *pictures* (or *representations*). They also discuss *coherent states*, calling them the *complex wave* picture/representation.

In the rest of this Subsection, we give a general discussion of 'particle' and 'field'. There are two overall themes.

(A): The first (and longer) theme will to beware of a false dichotomy between particle and field, in quantum field theory. For the basic object (individual, "thing") in such a theory is the quantum system itself: which behaves in some regards (especially: in some states) like a particle or a collection of them, and in other regards or states like a field.

(B): The second theme will be that, after all, fields are primary in that the system is defined by an operator-valued field on space or spacetime, i.e. an assignment to each point of space or spacetime of a linear operator on a Hilbert space (whose vectors and density matrices then give states)—and this is a quantum analogue of, for example, the electric field assigning to each point of space or spacetime an electric field vector.

(A): *Beware of a false dichotomy*:— What, after all, do we mean by 'particle' and 'field'? Clearly, the concepts get changed as we pass from classical physics, to elementary ('first quantized') quantum theory, to quantum field theory. So interpreting these theories, especially the last, is in part a matter of plotting those changes. And even the most cursory attempt to do that shows there are many different particle-like, and many different field-like, attributes that one can consider; as follows.

We might list, as attributes of classical particles that quantum particles lack in elementary quantum theory: a continuous spacetime trajectory, impenetrability. We can similarly list new attributes of quantum particles: the Fourier-transformation between position and momentum, quantum statistics. In the transition to quantum field theory, definiteness and conservation of particle number go; creation and annihilation come in.

Turning to fields, we might list, as attributes of classical fields that the wave- functions of elementary quantum theory lack: energy-momentum, being real. Yet a wave-function is like a classical field in that it represents the state ('configuration') of the system concerned, albeit in a mathematically particular representation, namely the position representation; instantaneously or throughout time, depending on one's definitions of wave-function and field. (Here, and in what follows, we use 'representation' to mean an orthonormal basis of state-vectors.) In quantum field theory, this attribute goes: the state of the system is *not* represented by the quantum field, i.e. by the assignment of operators to each (spatial or spacetime) point. The state is represented, as always in quantum theories, by a state-vector.

With this great variety of attributes (and no doubt more) to be considered, and related to one another, there is certainly plenty to do in plotting the changes in the two notions of particle and field. But this variety should also make one wary of loose talk about a conflict between particle and field interpretations of quantum field theory. There is probably no essence in each of these two notions, one essence contradicting the other, allowing one to then try and judge which has the upper hand in interpreting quantum field theory.

This point is supported by a long-standing (and apparently 'essentialist') proposal for how to understand 'particle' within quantum field theory: namely that a quantity is 'particle-like' iff it commutes with all particle number operators (is in the commutants of all the number operators). This proposal is certainly attractive: for instance, it makes position, momentum and spin particle-like. It also suggests that we call a representation that diagonalizes these number operators 'a particle representation'. Such a representation will then be invaluable for describing phenomena in which particle-like quantities are important. Typically, the phenomenon will involve simultaneous eigenstates of number and some particle-like quantity, and we then choose the representation by simultaneously diagonalizing number and that quantity. The paradigm case is of course scattering theory's use of eigenstates of number and momentum.

But of course, there are other (mutually non-commuting) representations, or more generally sets of states, that are invaluable for describing different phenomena, where other quantities, not number, are important (in particular: definite in value). There is the field-operator itself; and related quantities like phase and associated ideas like coherent states; both of which are invaluable for the quantum description of the electromagnetic field; (cf. e.g. Loudon (1973, Chapter 7)). Besides, since as we noted there are various connotations of 'field', there need be no unique best choice amongst these non-commuting representations (sets of states), for a corresponding definition of 'field-like quantity' and 'field representation'—.

Whether or not there is such a choice, the important point is this: it is wrong to ask which of particle and field has the upper hand in interpreting quantum field theory. For the individual described by the theory is the underlying quantum system, with its Hilbert space of states. 'Particle' and 'field' are now both matters of a representation, of a selected set of states. The glory of quantum field theory is that it allows and uses all these representations, variously appropriate for describing particle-like or field-like phenomena—where the 'like' signals due allowance for ambiguities and changes in the concepts, as sketchily plotted above. Surely this is what Dirac (taciturn as always) meant by his brief remark that his quantization of the electromagnetic field gave 'a complete harmony between the wave and light-quantum descriptions'; (Dirac 'The Quantum Theory of the Emission and Absorption of Radiation', *Proceedings of the Royal Society of London* **A** 114 (1927), 243-265, see p. 245.)

(B): *... And yet...*:— Fields are primary in that the system has infinitely many degrees of freedom, and is defined by an operator-valued field on space or spacetime. More precisely: its algebra of quantities is generated by an assignment to each point (or allowing for smearing: to each small region) of space or spacetime, of a linear operator, or a set of linear operators, on a Hilbert space: whose vectors and density matrices then give states of the system. (We will see this in more detail for algebraic quantum field theory.) So this is a quantum analogue of, for example, the classical electric field assigning to each point of space or spacetime an electric field vector—though with the difference, mentioned in (A) above, that classically this assignment is a *state* of the field concerned (or part of the state, since time derivatives may also be needed to get sufficient initial data); while here the assignment presents only the system, not a specific state..

There is a further point here, that is well made by Wald (1994, p. 2-6)—and we believe that he speaks for the majority of physicists. Wald notes that a curved spacetime lacks the symmetries of a flat spacetime (Galilean or Poincaré) that underpin the construction of a representation of the canonical commutation relations, and its uniqueness up to unitary equivalence (the Stone-von Neumann theorem)—so that we must consider various representations. The benefit of the algebraic approach (which he advocates) is that it allows one to consider all these representations on an equal footing. This reinforces the idea that fields that are primary, and that particle notions are derived, and often approximate or phenomenological. (Of course, this is not to deny the interest of foundational studies of eg rival schemes for localization within

quantum field theory: cf. for example, Halvorson (2001), which uses one particle structures and the Fock spaces built from them.)

4.2 The general idea of one-particle structure

We consider a symplectic vector space (V, ω) (aka: (M, ω)), and we fix a dynamics by determining a Hamiltonian vector field X_h generated by $h : V \to \mathbb{R}$. We also write D_t ('d' for dynamics) for this vector field; and we write Φ_t for the one-parameter group of motions (i.e. symplectomorphisms) generated by $X_h \equiv D_t$, i.e. the motions along the integral curves of X_h. This set of integral curves is the 'solution space' $\mathcal{S}$ for this system. Given Conditions 1-2 of the previous section, $\mathcal{S}$ is itself a vector space, and has an associated symplectic form Ω determined by ω. Thus we begin with the triple, $(\mathcal{S}, \Omega, \Phi_t)$. We also write elements of V and of $\mathcal{S}$ as z, to signal the symplectic structure.

We now find the corresponding quantum system by "Hilbertizing" S and "unitarizing" D_t. This will give a single-particle Hilbert space, called a *one-particle structure*, $(\mathcal{H}, \langle \cdot, \cdot \rangle, U(t))$: i.e. a Hilbert space equipped with an inner product and a one-parameter family of unitaries, representing the quantum dynamical evolution. Note that the quantization assumes a classical dynamics D_t, and yields a quantum (unitarized) "cousin" of it. We recall that this was prefigured in Section 2.4's discussion of the complex structure being non-unique and as encoding aspects of dynamics. After all, $\mathcal{S}$ is a vector space over the reals; and so our quantization, in order to make sense of complex multiplication of the elements of $\mathcal{S}$, needs a *complex structure* $J : \mathcal{S} \to \mathcal{S}$. Recall Section 2.2!

In addition to a complex structure J, the key ingredient in the quantization—in the relationship between the one-particle structure and our original symplectic vector space—is a map $K : \mathcal{S} \to \mathcal{H}$ subject to the conditions that:

(i) $\mathrm{ran}(K)$ is dense in $\mathcal{H}$;

(ii) $2\mathrm{Im}\, m\langle Kz_1, Kz_2 \rangle = \Omega(z_1, z_2)$;

(iii) $U(t)K = K\Phi_t$, where $U(t) = e^{-itA}$ and A is a positive operator. (We write A not H to avoid suggesting that A must be the 'true' quantum Hamiltonian.) So K intertwines the classical and quantum dynamics.

This definition goes back to Segal (1959-1963) and Weinless (1969). We said '*a* map $K : \mathcal{S} \to \mathcal{H}$ such that ...'. But in fact, if a map K exists, it is unique up to unitary equivalence; see Kay (1979) (which is short and almost readable!).

We now discuss how a suitable complex structure J, and a K satisfying conditions (i) to (iii), yield $\mathcal{H}$; in four steps. (A): First, we assume that the complex structure J (so that $J^2 = -\mathbb{1}$; cf. Section 2.2.A) satisfies two conditions:

(a) J is a symplectomorphism; i.e. $\Omega(Jz_1, Jz_2) = \Omega(z_1, z_2)$; (compatibility with Ω: cf. Section 2.3.A); it follows that $[J, \Phi_t] = 0$, i.e. J is equivariant under the classical dynamics.

(b) $\Omega(z, Jz) > 0$, for all $z \neq 0$ ('taming': cf. Section 2.3.A).

(B): Then we can define complex scalar multiplication in the range of the map K by saying: for any $a, b \in \mathbb{R}$ and any $z \in \mathcal{S}$:

$$(a + ib)K(z) := K(az) + K(bJz). \tag{4.1}$$

We may then write (iii) in infinitesimal form to derive a "Schrödinger equation"

$$AK(z) = iKD_t(z) = KJD_t(z). \tag{4.2}$$

(In some cases, K will be the embedding map, i.e. $K(z) = z$: which simplifies this equation to $Az = JD_t z$.)

(C): Given this J, we may also follow the definitions in Section 2.3. First, we define a complex inner product on (S, Ω, Φ_t, J):

$$\langle z_1, z_2 \rangle_{\mathcal{S}} = \frac{1}{2}\Omega(z_1, Jz_2) + \frac{1}{2}i\Omega(z_1, z_2), \tag{4.3}$$

Here we recall: from Section 2.3.B: the definition of the *complex-linear* but *real-valued* symmetric bilinear form g_J on the complex vector space V_J by eq. 2.26. It was (now reverting to lower-case ω):

$$g_J(u, v) := \omega(u, Jv) \ . \tag{4.4}$$

And we recall from Section 2.3.C: the definition of the *sesquilinear*, *complex-valued* function on $V \times V$, i.e. complex inner product, in terms of g_J and ω, by eq. 2.28. It was

$$\langle u, v \rangle \equiv \langle u, v \rangle_{\omega, J} := g_J(u, v) + i\omega(u, v) \ . \tag{4.5}$$

(D): We now define an inner product on the range $K[\mathcal{S}]$ of K, by demanding that $\langle Kz_1, Kz_2 \rangle := \langle z_1, z_2 \rangle_{\mathcal{S}}$ on $K[\mathcal{S}]$. Then the definition eq. 4.3 guarantees condition (ii) on K. By completing $K[\mathcal{S}]$ in the norm induced by this inner product (if required)—and so guaranteeing condition (i) on K—we obtain $\mathcal{H}$.

4.3 Example: the simple harmonic oscillator

Perhaps surprisingly, the above strategy works—and in a very simple way—for $(\mathcal{S}, \Omega, \Phi_t)$ taken as the theory of the classical simple harmonic oscillator. K will map points $(q, p) \in \mathbb{R}^2$ to points in $\mathbb{C}$—with a slight "squeeze and-or stretch" We will see that this yields a single-particle Hilbert space with unitary dynamics.

Recall that for the simple harmonic oscillator, $S = \mathbb{R}^2 \ni (q, p)$, Ω is defined as usual, i.e.

$$\Omega((q_1, p_1), (q_2, p_2)) = q_1 p_2 - q_2 p_1, \tag{4.6}$$

and Φ_t is generated by the Hamiltonian

$$H = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2 q^2. \tag{4.7}$$

Hamilton's equations yield

$$\dot{q} = \frac{\partial H}{\partial p} = \frac{p}{m}; \qquad \dot{p} = -\frac{\partial H}{\partial q} = -m\omega^2 q; \tag{4.8}$$

and so $\ddot{q} + \omega^2 q = 0$. Solutions take the form

$$q = \frac{1}{\sqrt{2m\omega}}\left(ae^{-i\omega t} + a^* e^{i\omega t}\right), \qquad p = -i\sqrt{\frac{m\omega}{2}}\left(ae^{-i\omega t} - a^* e^{i\omega t}\right), \tag{4.9}$$

which defines trajectories in the phase space which are ellipses centred at $(0, 0)$. (The reason for our choice of constants will be clear soon.) These trajectories already look *almost* like unitary evolution in $\mathbb{C}$; i.e. circular motion centred at $(0, 0)$. Heuristically, we need to rescale q and p to send these ellipses to circles: suggesting that we define K as mapping points $(q, p) \in \mathbb{R}^2$ to points in $\mathbb{C}$—with an appropriate"squeeze and-or stretch".

One way of doing so is to define the map $K : \mathbb{R}^2 \to \mathbb{C}$ by

$$K(q, p) = \sqrt{\frac{m\omega}{2}}q + \frac{i}{\sqrt{2m\omega}}p \ =: \ \xi_{(q,p)}. \tag{4.10}$$

Solutions in $\mathbb{C}$ are then defined by $K(q(t),p(t)) = ae^{-i\omega t}$. This of course suggests that the quantum Hamiltonian be taken as $A := \omega$; and that we define the complex structure J as "the usual flip q and p and add a minus sign, but with rescaling":

$$J(q,p) := \left(-\frac{p}{m\omega}, m\omega q\right); \tag{4.11}$$

for then we have

$$JD_t(q,p) = J(\dot{q},\dot{p}) = J\left(\frac{p}{m}, -m\omega^2 q\right) = (\omega q, \omega p) = \omega(q,p), \tag{4.12}$$

which is the classical counterpart of our "Schrödinger equation" in $\mathcal{S}$.

It is interesting to note that if we define the "positive frequency" component of $(q(t),p(t))$:

$$q^{(+)}(t) := \frac{1}{\sqrt{2}}\left(q(t) + \frac{i}{m\omega}p(t)\right) = \frac{1}{\sqrt{m\omega}}ae^{-i\omega t}; \tag{4.13}$$

$$p^{(+)}(t) := \frac{1}{\sqrt{2}}\left(p(t) - im\omega q(t)\right) = -i\sqrt{m\omega}ae^{-i\omega t} = -im\omega q^{(+)}(t); \tag{4.14}$$

then

$$J\left(q^{(+)}(t), p^{(+)}(t)\right) = J\left(1, -im\omega\right) q^{(+)}(t) = (i, m\omega)\, q^{(+)}(t) = i\left(q^{(+)}(t), p^{(+)}(t)\right); \tag{4.15}$$

i.e., J acts as multiplication by i; while for the "negative frequency" component:

$$q^{(-)}(t) := \frac{1}{\sqrt{2}}\left(q(t) - \frac{i}{m\omega}p(t)\right) = \frac{1}{\sqrt{m\omega}}a^*e^{i\omega t}; \tag{4.16}$$

$$p^{(-)}(t) := \frac{1}{\sqrt{2}}\left(p(t) + im\omega q(t)\right) = i\sqrt{m\omega}a^*e^{i\omega t} = im\omega q^{(-)}(t); \tag{4.17}$$

we have

$$J\left(q^{(-)}(t), p^{(-)}(t)\right) = J\left(1, im\omega\right) q^{(-)}(t) = (-i, m\omega)\, q^{(-)}(t) = -i\left(q^{(-)}(t), p^{(-)}(t)\right); \tag{4.18}$$

i.e., J acts as multiplication by $-i$. This a toy analogue of the solution to the "negative energy problem" in quantum field theory: for negative-frequency solutions, the complex structure has the opposite sign, allowing for positive-energy solutions in all cases.

Our inner product in S is given by

$$\begin{aligned}
\langle (q_1,p_1),(q_2,p_2)\rangle_{\mathcal{S}} &= \frac{1}{2}\Omega\left((q_1,p_1), J(q_2,p_2)\right) + \frac{1}{2}i\Omega\left((q_1,p_1),(q_2,p_2)\right) && (4.19)\\
&= \frac{1}{2}\Omega\left((q_1,p_1), \left(-\frac{p_2}{m\omega}, m\omega q_2\right)\right) + \frac{1}{2}i\Omega\left((q_1,p_1),(q_2,p_2)\right) && (4.20)\\
&= \frac{p_1p_2}{2m\omega} + \frac{1}{2}m\omega q_1q_2 + \frac{i}{2}\left(q_1p_2 - q_2p_1\right) && (4.21)\\
&= \left(\sqrt{\frac{m\omega}{2}}q_1 - \frac{i}{\sqrt{2m\omega}}p_1\right)\left(\sqrt{\frac{m\omega}{2}}q_2 + \frac{i}{\sqrt{2m\omega}}p_2\right) && (4.22)\\
&= K(q_1,p_1)^*K(q_2,p_2), && (4.23)
\end{aligned}$$

so our inner product in $\mathbb{C}$ is just

$$\langle \xi_1, \xi_2\rangle = \xi_1^*\xi_2. \tag{4.24}$$

This elegant form for the inner product is a consequence of our particular choice of the map K.

4.4 The free bosonic field on any one-particle structure: the idea

In this Section, we present the idea of Fock space in a mathematical way. Similar presentations are in e.g. Baez et al.(1992, Sections 1.8, 1.9), Araki (1993, Section 3.5), Folland (2008, Section 4.5), and De Faria and De Melo. More physics-oriented presentations include e.g.: Schweber, *Introduction to Relativistic Quantum Field Theory* (1961, Chapters 6 and 7), Loudon, *The Quantum Theory of Light* (1973: Chapters 6 and 7), Itzykson and Zuber, *Quantum Field Theory* (1987, Section 3.1), Coleman *Lectures on Quantum Field Theory* (2019, Chapters 2,3,4).

Once we have our one-particle system $(\mathcal{H}, \langle\cdot,\cdot,\rangle, U(t))$, we may define the *free boson field over* it. This quantum theory will provide a representation of our Weyl algebra. (The following prescription is unique, up to unitary equivalence; see Baez *et al* 1992, pp. 49-56, Theorem 1.10.) The free boson field over $\mathcal{H}$ is the system $(\mathfrak{F}_+(\mathcal{H}), W, \Gamma, \nu)$ where

$$\mathfrak{F}_+(\mathcal{H}) := \bigoplus_{n=0}^{\infty} \mathcal{S}_n \left(\otimes^n \mathcal{H}\right) \tag{4.25}$$

is the Hilbert space of all symmetric tensors on $\mathcal{H}$, and for any linear operator $Q \in \mathcal{B}(\mathcal{H})$,

$$\Gamma(Q) := 1 \oplus Q \oplus (Q \otimes Q) \oplus (Q \otimes Q \otimes Q) \oplus \ldots|_{\mathfrak{F}_+(\mathcal{H})}\,. \tag{4.26}$$

We assume a strongly continuous one-parameter family $U(t)$ of unitaries on $\mathcal{H}$, which is generated by some self-adjoint operator A. The corresponding family $\Gamma(U(t))$, is generated by a self-adjoint operator which we call $\mathrm{d}\Gamma(A)$. It satisfies

$$\Gamma(U(t)) = \Gamma(e^{itA}) = e^{it\mathrm{d}\Gamma(A)} \tag{4.27}$$

and

$$\mathrm{d}\Gamma(A) := 0 \oplus A \oplus (A \otimes \mathbb{1} + \mathbb{1} \otimes A) \oplus \ldots|_{\mathfrak{F}_+(\mathcal{H})}\,. \tag{4.28}$$

Finally, the *vacuum state* ν is defined by

$$\nu = 1 \oplus \mathbf{0} \oplus \mathbf{0} \oplus \ldots \tag{4.29}$$

Note that it then satisfies

$$\Gamma(U(t))\nu = \nu. \tag{4.30}$$

The free bosonic field $(\mathfrak{F}_+(\mathcal{H}), W, \Gamma, \nu)$ provides a representation for the Weyl algebra as follows. We need to define, for every $\xi \in \mathcal{H}$, creation and annihilation operators $a^\dagger(\xi), a(\xi)$; $\mathfrak{F}_+(\mathcal{H})$ is the closed linear span of arbitrary combinations of these acting on ν. To this end we define the operators $a^\dagger_{(N)}(\xi) : \otimes^{N-1}\mathcal{H} \to \otimes^N \mathcal{H}$ and $a_{(N)}(\xi) : \otimes^N \mathcal{H} \to \otimes^{N-1}\mathcal{H}$ for all $N \in \mathbb{N}$:

$$\begin{array}{rcl} a^\dagger_{(N)}(\xi)\left(\psi_1 \otimes \ldots \otimes \psi_{N-1}\right) & := & \xi \otimes \psi_1 \otimes \ldots \otimes \psi_{N-1} \\ a_{(N)}(\xi)\left(\psi_1 \otimes \psi_2 \otimes \ldots \otimes \psi_N\right) & := & \langle \xi, \psi_1 \rangle\ \psi_2 \otimes \ldots \otimes \psi_N \end{array} \tag{4.31}$$

where $\chi_k(j) = \delta_{jk}$. Now we may define $a^\dagger(\xi), a(\xi) : \mathfrak{F}_+(\mathcal{H}) \to \mathfrak{F}_+(\mathcal{H})$ by

$$\begin{array}{rcccccccc} a^\dagger(\xi) & := & a^\dagger_{(1)}(\xi) & \oplus & \sqrt{2}\mathcal{S}_2 a^\dagger_{(2)}(\xi) & \oplus & \sqrt{3}\mathcal{S}_3 a^\dagger_{(3)}(\xi) & \oplus & \ldots \\ a(\xi) & := & 0 & \oplus & a_{(1)}(\xi) & \oplus & \sqrt{2}a_{(2)}(\xi) & \oplus & \sqrt{3}a_{(3)}(\xi) \ \oplus \ \ldots \end{array} \tag{4.32}$$

It may be checked that

$$[a(\xi_1), a(\xi_2)] = [a^\dagger(\xi_1), a^\dagger(\xi_2)] = 0; \qquad [a(\xi_1), a^\dagger(\xi_2)] = \langle \xi_1, \xi_2 \rangle; \tag{4.33}$$

this will be crucial for representing the Weyl algebra. We also have, for any projector P on $\mathcal{H}$,

$$\mathrm{d}\Gamma(P) = \sum_i \mathrm{d}\Gamma(\Pi(\xi_i)) = \sum_i a^\dagger(\xi_i)a(\xi_i), \tag{4.34}$$

where the ξ_i are an orthonormal basis for $\mathrm{ran}(P)$ and $\Pi(\xi_i)$ projects onto the ray spanned by ξ_i.

We now define the (unbounded) field operators for all $z \in S$:

$$\Phi(z) := a(K(z)) + a^\dagger(K(z)), \tag{4.35}$$

where $K : S \to \mathcal{H}$ is our map from the classical phase space to the single-particle Hilbert space. It follows from (4.33) that, for all $z_1, z_2 \in S$ in a dense domain,

$$\begin{aligned}[\Phi(z_1), \Phi(z_2)] &= [a(K(z_1)), a^\dagger(K(z_2))] + [a^\dagger(K(z_1)), a(K(z_2))] \\ &= -2i\mathrm{Im}\, m\langle K(z_1), K(z_2)\rangle = -i\Omega(z_1, z_2),\end{aligned} \tag{4.36}$$

Equation (4.36) is none other than our Weyl relations in infinitesimal form. The representation $W : S \to \mathfrak{B}\left[\mathfrak{F}_+(\mathcal{H})\right]$ of the Weyl algebra is then provided by

$$W(z) := e^{i\Phi(Jz)}. \tag{4.37}$$

The "particle picture"
For any projector P on $\mathcal{H}$, the operator $\mathrm{d}\Gamma(P)$ is the particle number operator associated with P. The total particle number operator is $N := \mathrm{d}\Gamma(\mathbb{1})$. Eigenstates of N are states of the field with definite particle number.
The "real wave picture"
For each $z \in S$, the field operator $\Phi(Jz)$ is the unique self-adjoint operator which generates the strongly continuous one-parameter family of unitaries $W(tz)$, where $t \in \mathbb{R}$. Eigenstates of $\Phi(Jz)$ do not, strictly speaking, exist, but $\Phi(Jz)$ admits of a spectral decomposition, in analogy with $\mathbf{Q}$ and $\mathbf{P}$ in elementary nonrelativistic quantum mechanics.
The "complex wave picture"
Here the relevant operators are the creation and annihilation operators, for any $z \in S$:

$$a^\dagger(K(z)) = \frac{1}{2}\left(\Phi(z) - i\Phi(Jz)\right); \qquad a(K(z)) = \frac{1}{2}\left(\Phi(z) + i\Phi(Jz)\right) \tag{4.38}$$

The relevant "eigenstates" are of $a(K(z))$ (a misleading term, since $a(K(z))$ is not a normal operator). These are coherent states.

Note that there is a natural sense in which the field operator is a function over the *classical* phase space S, while the creation and annihilation operators are functions over the *quantum* one-particle Hilbert space $\mathcal{H}$.

4.5 Example: the simple harmonic oscillator again

Here we simply apply the above general prescription to the case where $\mathcal{H} = \mathbb{C}$, $\langle \xi_1, \xi_2\rangle = \xi_1^* \xi_2$ and the unitary evolution is generated by the Hamiltonian $A = \omega$. Our Fock space is

$$\mathfrak{F}_+(\mathbb{C}) = \bigoplus_{n=0}^{\infty} \mathcal{S}_n(\otimes^N \mathbb{C}) = \mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C} \oplus \ldots = l^2(\mathbb{N}) \tag{4.39}$$

Unitary evolution is governed in this Fock space by

$$\Gamma(e^{-i\omega t}) = 1 \oplus e^{-i\omega t} \oplus e^{-2i\omega t} \oplus \ldots, \tag{4.40}$$

which is generated by the Hamiltonian

$$\mathrm{d}\Gamma(\omega) = 0 \oplus \omega \oplus 2\omega \oplus \dots. \tag{4.41}$$

For each $\xi \in \mathbb{C}$, the creation and annihilation operators $a^\dagger(\xi), a(\xi) : l^2(\mathbb{N}) \to l^2(\mathbb{N})$ satisfy

$$[a(\xi_1), a(\xi_2)] = [a^\dagger(\xi_1), a^\dagger(\xi_2)] = 0; \qquad [a(\xi_1), a^\dagger(\xi_2)] = \xi_1^* \xi_2; \tag{4.42}$$

Since $a^\dagger(\xi)$ is complex-linear and $a(\xi)$ is complex-antilinear, we may define $a^\dagger := a^\dagger(1), a := a(1)$, and then $a^\dagger(\xi) = \xi a^\dagger$ and $a(\xi) = \xi^* a$, and $[a, a^\dagger] = 1$. The operator $a^\dagger a$ is a number operator—in fact, the *only* number operator, up to a complex constant—and it may be checked that $\mathrm{d}\Gamma(\omega) = \omega a^\dagger a$. (*Note:* no zero-point energy!)

Self-adjoint field operators over $S = \mathbb{R}^2 \ni (\alpha, \beta)$ are then defined by

$$\begin{aligned} \Phi(J(\alpha,\beta)) &= a(KJ(\alpha,\beta)) + a^\dagger(KJ(\alpha,\beta)) \\ &= iK(\alpha,\beta)a^\dagger - iK(\alpha,\beta)^* a \\ &= -\frac{\beta}{\sqrt{2m\omega}}(a + a^\dagger) - i\sqrt{\frac{m\omega}{2}}\alpha(a - a^\dagger). \end{aligned} \tag{4.43}$$

We now recover the familiar self-adjoint operators

$$Q := \Phi(J(0,-1)) = \frac{1}{\sqrt{2m\omega}}(a + a^\dagger); \tag{4.44}$$

$$P := \Phi(J(1,0)) = -i\sqrt{\frac{m\omega}{2}}(a - a^\dagger); \tag{4.45}$$

from which we recover the familiar Heisenberg relation $[Q, P] = i$.

Pause for a moment to consider the identity $P = \Phi(J(1,0))$. Recall that $\Phi(J\cdot)$ is a function from the classical phase space S to (unbounded) operators on the "field" Hilbert space $\mathfrak{F}_+(\mathcal{H})$. But remember that elements of S are "really" proxies for vectors which determine vector fields over S. (We are lucky enough that S is a symplectic vector space, so this use of proxies is possible.) The vector $(1,0)$ determines the vector field $\frac{\partial}{\partial q}$, i.e. translations in the position q. We know that these translations are generated by momentum, so it is fitting that $P = \Phi(J(1,0))$. Similarly, the vector $(0,-1)$ determines the vector field $-\frac{\partial}{\partial p}$, i.e. negative translations in the momentum p, which we know are generated by position; so it is fitting that $Q = \Phi(J(0,-1))$. Quite generally, the self-adjoint operator $\Phi(Jz)$ is the quantum observable corresponding to the classical generator of phase space translations in the direction z. This identification will be important in identifying the local field operators for the quantum field, in the successor document ...

We may also express the $a, a^\dagger$ in terms of Q and P:

$$a = \sqrt{\frac{m\omega}{2}}\left(Q + \frac{i}{m\omega}P\right); \quad a^\dagger = \sqrt{\frac{m\omega}{2}}\left(Q - \frac{i}{m\omega}P\right); \tag{4.46}$$

allowing us to similarly re-express the field operators, for any $(\alpha, \beta) \in \mathbb{R}^2$:

$$\Phi(J(\alpha,\beta)) = \Phi\left(-\frac{\beta}{m\omega}, m\omega\alpha\right) = \alpha P - \beta Q. \tag{4.47}$$

In terms of Q and P, the (normal-ordered!) Hamiltonian for the bosonic field is

$$\mathrm{d}\Gamma(\omega) = \omega a^\dagger a = \frac{m\omega^2}{2}\left(Q - \frac{i}{m\omega}P\right)\left(Q + \frac{i}{m\omega}P\right) = \frac{1}{2m}P^2 + \frac{1}{2}m\omega^2 Q^2 - \frac{1}{2}\omega. \tag{4.48}$$

The "particle picture"
There is only one non-zero projector on $\mathbb{C}$, and the operator $\mathrm{d}\Gamma(1)$ is the total particle number operator $N = a^\dagger a$. Eigenstates of N are states of the field with definite particle number.
The "real wave picture"
Eigenstates of $\Phi(J(\alpha,\beta)) = \alpha P - \beta Q$ do not, strictly speaking exist, but it is familiar that we may define spectral resolutions for Q and P.
The "complex wave picture"
The annihilation operator, for any a has the "eigenstates"

$$|\xi\rangle := e^{-\frac{1}{2}|\xi|^2} e^{a^\dagger(\xi)}\nu = e^{-\frac{1}{2}|\xi|^2} e^{\xi a^\dagger}\nu, \tag{4.49}$$

for any $\xi \in \mathbb{C}$, using the complex-linearity of $a^\dagger$. We have that $a|\xi\rangle = e^{-\frac{1}{2}|\xi|^2} a e^{\xi a^\dagger}\nu = e^{-\frac{1}{2}|\xi|^2}\left(e^{\xi a^\dagger}a + \xi e^{\xi a^\dagger}\right)\nu = \xi e^{-\frac{1}{2}|\xi|^2} e^{\xi a^\dagger}\nu = \xi|\xi\rangle$. The most familiar coherent state is $|0\rangle \equiv \nu$ (i.e. when $\xi = 0$), the "Fock space vacuum", which yields probability distributions in both Q and P that are gaussians centred at zero. And in general it may be checked that

$$W(z)|0\rangle := e^{i\Phi(Jz)}|0\rangle = |K(z)\rangle. \tag{4.50}$$

The state $|K(z)\rangle$ yields probability distributions in Q and P that are both gaussians, centred at α and β respectively, where $z = (\alpha, \beta)$. These states are crucial to defining the classical limit of the theory: specifically, as $\hbar \to 0$, the behaviour of $|K(z)\rangle$ approaches that of the classical state z.

4.6 (Apparently) rival quantizations

The story just given for the simple harmonic oscillator may be run again, this time starting with a classical system with a different Hamiltonian:

$$H_2 = \frac{1}{2m}p^2 + \frac{1}{2}m\omega_2^2 q^2. \tag{4.51}$$

(Set $\omega_1 = \omega$, etc. in the above discussion.) The new classical dynamics induced by this new Hamiltonian results in a different map $K_2 : \mathbb{R}^2 \to \mathbb{C}$, different complex structure J_2 and different quantum Hamiltonian $A_2 = \omega_2$. A shortcut to the new "bosonic field" is to transform

$$Q \mapsto \sqrt{\frac{\omega_2}{\omega_1}}Q; \qquad P \mapsto \sqrt{\frac{\omega_1}{\omega_2}}P. \tag{4.52}$$

This gives rise to new creation and annihilation operators $a_2^\dagger, a_2$, related to the previous ones $a_1^\dagger (= a^\dagger), a_1 (= a)$ by

$$a_2 = \frac{1}{2}\left(\sqrt{\frac{\omega_2}{\omega_1}} + \sqrt{\frac{\omega_1}{\omega_2}}\right)a_1 + \frac{1}{2}\left(\sqrt{\frac{\omega_2}{\omega_1}} - \sqrt{\frac{\omega_1}{\omega_2}}\right)a_1^\dagger; \tag{4.53}$$

$$a_2^\dagger = \frac{1}{2}\left(\sqrt{\frac{\omega_2}{\omega_1}} - \sqrt{\frac{\omega_1}{\omega_2}}\right)a_1 + \frac{1}{2}\left(\sqrt{\frac{\omega_2}{\omega_1}} + \sqrt{\frac{\omega_1}{\omega_2}}\right)a_1^\dagger. \tag{4.54}$$

It follows from these relations that the vacuum for $A = \omega_1$ is not the vacuum for $A = \omega_2$; specifically,

$$\langle \nu_1, N_2 \nu_1\rangle = \langle \nu_1, a_2^\dagger a_2 \nu_1\rangle = \frac{(\omega_1 - \omega_2)^2}{4\omega_1\omega_2}. \tag{4.55}$$

We know from the Stone-von Neumann theorem that, since the ω_1 representation and the ω_2 representation both provide a representation of the Weyl algebra over $\mathbb{R}^2$, they must be

unitarily equivalent. In fact, the equations (4.53) & (4.54) specify the unitary which intertwines them.

In the position representation, the unitary transformation between the two representations is implemented by

$$\psi(x) \mapsto \left(\frac{\omega_2}{\omega_1}\right)^{\frac{1}{4}} \psi\left(\sqrt{\frac{\omega_2}{\omega_1}}x\right). \tag{4.56}$$

5 "Unitarization": complex structures, metrics, and the 2-out-of-3 property

"Unitarization" of linear Hamiltonian systems appeals to the "2-out-of-3" property of the group $U(n)$ of unitary transformations in n complex dimensions:

$$U(n) = Sp(2n,\mathbb{R}) \cap GL(n,\mathbb{C}) \cap O(2n) \tag{5.1}$$

The identity above continues to hold for any two of the three groups on the righthand side: i.e. $Sp(2n,\mathbb{R}) \cap GL(n,\mathbb{C}) = GL(n,\mathbb{C}) \cap O(2n) = O(2n) \cap Sp(2n,\mathbb{R}) = U(n)$.

- $Sp(2n,\mathbb{R})$ is the *symplectic group in* $2n$ *dimensions*, the group of automorphisms of the symplectic vector space $\langle \mathbb{R}^{2n}, \Omega \rangle$, the *linear symplectic transformations*, and so is the most general group of symmetries pertaining to a linear Hamiltonian system. It is the group of linear transformations $A : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ which preserve some symplectic product Ω, according to $\Omega(Au, Av) = \Omega(u, v)$ for all $u, v \in \mathbb{R}^{2n}$. The symplectic product Ω is any bilinear, alternating and non-degenerate map $\Omega : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \to \mathbb{R}$. If Ω is represented as a matrix $\mathbf{\Omega}$, so that $\Omega(u, v) \equiv u^T \mathbf{\Omega} v$, then this condition is equivalent to $A^T \mathbf{\Omega} A = \mathbf{\Omega}$. (The bilinearity of $\mathbf{\Omega}$ is baked in, since it is a matrix; the non-degeneracy of Ω means that $\mathbf{\Omega}^{-1}$ exists; Ω's being alternating means that $u^T \mathbf{\Omega} v = -v^T \mathbf{\Omega} u$ for all $u, v \in \mathbb{R}^n$, which is equivalent to $\mathbf{\Omega}^T = -\mathbf{\Omega}$.) By means of some similarity transformation the matrix $\mathbf{\Omega}$ may be brought to the form

  $$\mathbf{\Omega} = \begin{pmatrix} & \mathbb{1} \\ -\mathbb{1} & \end{pmatrix}. \tag{5.2}$$

  All symplectic matrices $A \in Sp(2n,\mathbb{R})$ have unit determinant, which means that they preserve area in the symplectic vector space—the specialisation of preserving the symplectic form to vector spaces. In the case $n = 1$ the elements of $Sp(2,\mathbb{R})$ just are all and only the area-preserving 2×2 matrices.

- $GL(n,\mathbb{C})$ is the group of general linear transformations on $\mathbb{C}^n$, i.e. the group of all invertible $n \times n$ matrices with complex entries. This group is isomorphic to the group of invertible $2n \times 2n$ matrices A with *real* entries which commute with some *complex structure* $J : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$, so that $AJ = JA$. The complex structure J here is some $2n \times 2n$ matrix such that $J^2 = -\mathbb{1}$, so that $J^{-1} = -J$. J allows us to pass from the real vector space $\mathbb{R}^{2n}$ to the complex vector space $\mathbb{C}^n$, where J takes the role of multiplication by i.

- $O(2n)$ is the group of orthogonal transformations on $\mathbb{R}^{2n}$, which can be thought of as the group of real $2n \times 2n$ matrices A such that $A^{-1} = A^T$. More generally, given some positive-definite, symmetric bilinear form $M : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \to \mathbb{R}$, $O(2n)$ is the group of linear transformations $A : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ such that $M(Au, Av) = M(u, v)$ for all $u, v \in \mathbb{R}^{2n}$. If M is represented as a matrix $\mathbf{M}$, so that $M(u, v) \equiv u^T \mathbf{M} v$, then this condition is equivalent to $A^T \mathbf{M} A = \mathbf{M}$. (The bilinearity of $\mathbf{M}$ is baked in, since it is a matrix; the positive-definiteness of M means that $\mathbf{M}$ has only positive eigenvalues; M's being symmetric means that $u^T \mathbf{M} v = v^T \mathbf{M} u$ for all $u, v \in \mathbb{R}^n$, which is equivalent to $\mathbf{M}^T = \mathbf{M}$.) By

means of some similarity transformation the matrix $\mathbf{M}$ may be brought to the form $\mathbb{1}$; then we recover the condition $AA^T = A^TA = \mathbb{1}$, i.e. $A^{-1} = A^T$. The crucial idea is that the elements of $O(2n)$ preserve the lengths of all vectors in $\mathbb{R}^{2n}$ according to *some* metric, given by M.

The intersection of any two of these groups generates a subgroup of the third by means of the identification $M(\cdot,\cdot) = \Omega(\cdot, J\cdot)$, or $\mathbf{M} = \mathbf{\Omega}J$, with the crucial compatibility requirement that $\Omega(J\cdot, J\cdot) = \Omega(\cdot,\cdot)$, or $J^T\mathbf{\Omega}J = \mathbf{\Omega}$; equivalent to $J^T\mathbf{M}J = \mathbf{M}$ and $\mathbf{M}^{-1}\mathbf{\Omega} = -\mathbf{\Omega}^{-1}\mathbf{M}$. It may then be checked that any two of: (i) $\mathbf{\Omega}^T = -\mathbf{\Omega}$; (ii) $J^{-1} = -J$; and (iii) $\mathbf{M}^T = \mathbf{M}$ implies the third. For example: $\mathbf{M}^T = (\mathbf{\Omega}J)^T = J^T\mathbf{\Omega}^T = -J^T\mathbf{\Omega} = J^T\mathbf{\Omega}J^2 = \mathbf{\Omega}J = \mathbf{M}$, where we use the identification $\mathbf{M} = \mathbf{\Omega}J$, the property $\mathbf{\Omega}^T = -\mathbf{\Omega}$, the property $J^2 = -\mathbb{1}$ and the compatibility condition $J^T\mathbf{\Omega}J = \mathbf{\Omega}$, respectively.

Starting with any linear Hamiltonian system, we are given $Sp(2n,\mathbb{R})$ and $\mathbf{\Omega}$ at the outset. Our goal is $U(n)$. The "2-out-of-3" identity above entails that we can get there so long as we can find some metric $\mathbf{M}$ or some complex structure J obeying the compatibility conditions (and, if we can find one of these, then we automatically have the other). In other words, if we can find some natural way to give the symplectic vector space $\langle\mathbb{R}^{2n}, \Omega\rangle$ the structure of a real inner-product space $\langle\mathbb{R}^{2n}, M\rangle$ or the structure of the complex vector space $\mathbb{C}^n \cong \langle\mathbb{R}^{2n}, J\rangle$, then we will end up with the Hilbert space $\langle\mathbb{C}^n, \langle\cdot,\cdot\rangle\rangle$ and its associated group of automorphisms $U(n)$.

It turns out that the metric $\mathbf{M}$ is determined by the classical Hamiltonian *so long as* the associated Hamiltonian matrix $\mathbf{H}$ is positive-definite (it is already symmetric and bilinear). This corresponds to the existence of a unique global vacuum at $\mathbf{0}$ in the symplectic vector space. The compatibility requirement relies on the dynamics decomposing into normal modes. For then we can find some invertible symmetric matrix A, diagonalised in the normal-mode basis, such that $[A, \mathbf{H}] = [A, \mathbf{\Omega}] = 0$ and such that we can set $\mathbf{M} = \mathbf{H}A^{-1}$. It follows that $J = \mathbf{\Omega}^{-1}\mathbf{H}A^{-1}$ and $[A, J] = 0$. It turns out that A is proportional to the *quantum* Hamiltonian of the resulting quantum theory.

These conditions on A give us a recipe for finding it. For in coordinates such that $\mathbf{\Omega}^{-1} = \mathbf{\Omega}^T = -\mathbf{\Omega}$ (which we can always find) we have $(\mathbf{\Omega H})^2 = \mathbf{\Omega H\Omega H} = \mathbf{\Omega}A\mathbf{M\Omega}A\mathbf{M} = A\mathbf{\Omega M\Omega M}A = A\mathbf{\Omega}^2A = -A^2$. We may therefore set $A := (-(\mathbf{\Omega H})^2)^{\frac{1}{2}}$. Crucial here is that $\mathbf{H}$ is at least positive-semi-definite; otherwise A is not a real matrix. We now rely on the fact that we can find coordinates in which not only $\mathbf{\Omega}$ takes the standard form in which $\mathbf{\Omega}^{-1} = \mathbf{\Omega}^T = -\mathbf{\Omega}$, but also $\mathbf{H}$ takes the form

$$\mathbf{H} = \begin{pmatrix} B & \\ & C \end{pmatrix} \tag{5.3}$$

(so that there are no q-p cross terms), and where $[B, C] = 0$. (Usually we expect $C = \mathbb{1}$.) In this case we find that

$$A = \begin{pmatrix} (BC)^{\frac{1}{2}} & \\ & (BC)^{\frac{1}{2}} \end{pmatrix} . \tag{5.4}$$

It may now be verified that $[A, \mathbf{H}] = [A, \mathbf{\Omega}] = 0$. The metric $\mathbf{M}$ then takes the form

$$\mathbf{M} = \begin{pmatrix} B^{\frac{1}{2}}C^{-\frac{1}{2}} & \\ & B^{-\frac{1}{2}}C^{\frac{1}{2}} \end{pmatrix} , \tag{5.5}$$

and it may be checked that the compatibility requirement $\mathbf{M}^{-1}\mathbf{\Omega} = -\mathbf{\Omega}^{-1}\mathbf{M} = \mathbf{\Omega M}$ is satisfied. Crucial here is that $\mathbf{H}$ is not only positive-semi-definite but positive-definite: for, both $B^{\frac{1}{2}}$ and $C^{\frac{1}{2}}$ must exist and be invertible; otherwise $\mathbf{M}$ is not well defined. The complex structure J then takes the form

$$J = \begin{pmatrix} & -B^{-\frac{1}{2}}C^{\frac{1}{2}} \\ B^{\frac{1}{2}}C^{-\frac{1}{2}} & \end{pmatrix} . \tag{5.6}$$

In the case $B = \omega^2, C = \mathbb{1}$, we have

$$\mathbf{H} = \begin{pmatrix} \omega^2 & \\ & \mathbb{1} \end{pmatrix} \; ; \quad A = \begin{pmatrix} \omega & \\ & \omega \end{pmatrix} \; ; \quad \mathbf{M} = \begin{pmatrix} \omega & \\ & \omega^{-1} \end{pmatrix} \; ; \quad J = \begin{pmatrix} & -\omega^{-1} \\ \omega & \end{pmatrix} . \tag{5.7}$$

We may then define the inner product

$$\langle \cdot, \cdot \rangle = \lambda \left(M(\cdot, \cdot) + i\Omega(\cdot, \cdot) \right) \; , \tag{5.8}$$

where λ is some constant. Since $\Omega(\mathfrak{z}_1, \mathfrak{z}_2)$ has units of action, it seems sensible to set $\lambda \sim \frac{1}{\hbar}$. The standard choice is $\lambda = \frac{1}{2\hbar}$ (this will end up yielding the usual ladder operators in the case of of the s.h.o.).

6 The Stone-von Neumann and Jordan-Wigner uniqueness theorems

In this Section, we will: (i) state these two theorems and discuss Mackey's generalization of the Stone-von Neumann theorem (Section 6.1); then (ii) discuss proofs, and sketch one proof (Section 6.2); then (iii) list ways that they can break down (i.e. physically relevant ways their assumptions can fail; Section 6.3)). These will especially include the system being infinite: so that we look ahead to the successor document . . .

6.1 The two theorems

The Stone-von Neumann uniqueness theorem for CCRs guarantees that, in certain cases, the representation of the Weyl algebra is effectively unique (i.e., unique up to unitary equivalence). Therefore there is a reasonable sense in which, in these cases, there is only one quantization of a classical system. The theorem applies to *weakly continuous* representations of the Weyl algebra. A representation on some Hilbert space $\mathcal{H}$, with inner product $\langle \cdot, \cdot \rangle$, is weakly continuous iff it is continuous in the weak topology provided by $\langle \cdot, \cdot \rangle$. I.e. for all $\psi \in \mathcal{H}$, $z_1, z_2 \in S$ and $\epsilon \in \mathbb{R}^+$,

$$\langle \psi, W(z_1 + \epsilon z_2)\psi \rangle \to \langle \psi, W(z_1)\psi \rangle \quad \text{as} \quad \epsilon \to 0. \tag{6.1}$$

Theorem 6.1 (Stone-von Neumann Uniqueness Theorem)**.** *Let (S, Ω) be a symplectic vector space, with $S = \mathbb{R}^{2n}$. Every weakly continuous irreducible representation of the Weyl algebra over (S, Ω) is unitarily equivalent to the Schrödinger representation, in which, for all $\psi(\mathbf{x}) \in L^2(\mathbb{R}^n)$,*

$$(W(\mathbf{a}, \mathbf{b})\psi)(\mathbf{x}) := e^{-i\mathbf{a}.(\mathbf{x} - \frac{1}{2}\mathbf{b})}\psi(\mathbf{x} - \mathbf{b}). \tag{6.2}$$

Note as special cases that $(W(\mathbf{a}, \mathbf{0})\psi)(\mathbf{x}) \equiv (U(\mathbf{a})\psi)(\mathbf{x}) = e^{-i\mathbf{a}.\mathbf{x}}\psi(\mathbf{x})$ and $(W(\mathbf{0}, \mathbf{b})\psi)(\mathbf{x}) \equiv (V(\mathbf{b})\psi)(\mathbf{x}) = \psi(\mathbf{x} - \mathbf{b})$. In fact, the Schrödinger representation is strongly continuous, so by Stone's Theorem there are $2n$ self-adjoint operators, Q^i and P_i, such that $U(\mathbf{a}) = e^{-i\mathbf{a}.\mathbf{Q}}$, $V(\mathbf{b}) = e^{-i\mathbf{b}.\mathbf{P}}$ and for all $\psi(\mathbf{x}) \in L^2(\mathbb{R}^n)$ in suitable domains,

$$(\mathbf{Q}\psi)(\mathbf{x}) = \mathbf{x}\psi(\mathbf{x}); \qquad (\mathbf{P}\psi)(\mathbf{x}) = -i\nabla\psi(\mathbf{x}). \tag{6.3}$$

For the Jordan-Wigner theorem for the CARs, we consider first a sequence of quantum theories, each corresponding to a chain of spin-$\frac{1}{2}$ systems. The first theory describes a single spin-$\frac{1}{2}$ system, with observables $\{\sigma(x), \sigma(y), \sigma(z)\}$, which satisfy the Pauli relations

$$[\sigma(x), \sigma(y)] = 2i\sigma(z) \quad \text{and cyclic perms}; \qquad \sigma^2 := \sigma(x)^2 + \sigma(y)^2 + \sigma(z)^2 = 3\mathbb{1}. \tag{6.4}$$

This is equivalent to satisfying the canonical *anti*-commutation relations (CARs; see eg Ruetsche (2011, p.60-61))

$$d^2 = \left(d^\dagger\right)^2 = 0; \qquad [d, d^\dagger]_+ = 1; \tag{6.5}$$

where

$$\sigma(x) = d + d^\dagger; \quad \sigma(y) = -i\left(d - d^\dagger\right); \quad \sigma(z) = dd^\dagger - d^\dagger d. \tag{6.6}$$

We now consider a theory describing a linear chains of n spin-$\frac{1}{2}$ systems, with observables $\{\sigma_k(x), \sigma_k(y), \sigma_k(z) \mid k \in \{1, 2, \ldots n\}\}$, satisfying

$$[\sigma_j(x), \sigma_k(y)] = 2i\delta_{jk}\sigma_k(z) \quad \text{and cyclic perms}; \qquad \sigma_k^2 := \sigma_k(x)^2 + \sigma_k(y)^2 + \sigma_k(z)^2 = 3\mathbb{1}. \tag{6.7}$$

Of course, our theory falls outside the scope of the Stone-von Neumann theorem, because it is characterized by CARs, rather than CCRs. However, there is an analogous uniqueness theorem:

Theorem 6.2 (Jordan-Wigner Uniqueness Theorem)**.** *For each finite n, every irreducible representation of the CARs (equivalently, the Pauli relations) is unitarily equivalent to the Pauli representation, in which*

$$\begin{aligned}
\sigma_k^P(x) &= \underbrace{\mathbb{1} \otimes \ldots \otimes \mathbb{1}}_{k-1} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \underbrace{\mathbb{1} \otimes \ldots \otimes \mathbb{1}}_{n-k}; \\
\sigma_k^P(y) &= \underbrace{\mathbb{1} \otimes \ldots \otimes \mathbb{1}}_{k-1} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \otimes \underbrace{\mathbb{1} \otimes \ldots \otimes \mathbb{1}}_{n-k}; \\
\sigma_k^P(z) &= \underbrace{\mathbb{1} \otimes \ldots \otimes \mathbb{1}}_{k-1} \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \underbrace{\mathbb{1} \otimes \ldots \otimes \mathbb{1}}_{n-k}.
\end{aligned} \tag{6.8}$$

We earlier saw two unitarily equivalent quantizations of the simple harmonic oscillator. Similarly, here for the CARs: an alternative—though by the Jordan-Wigner theorem: unitarily equivalent—representation S (for 'switch') defines the spin matrices according to

$$\sigma_k^S(x) = \sigma_k^P(y); \qquad \sigma_k^S(y) = \sigma_k^P(z); \qquad \sigma_k^S(z) = \sigma_k^P(x); \tag{6.9}$$

i.e. the switch representation of $\sigma_k(x)$ in $\mathcal{H}_S$ has the same matrix elements as the Pauli representation of $\sigma_k(x)$ in $\mathcal{H}_P$, etc. Now let $U : \mathbb{C}_P^2 \to \mathbb{C}_S^2$ be the unitary such that $U\sigma_k^P(x)U^\dagger = \sigma_k^S(x)$, etc. Then the unitary $\otimes^n U : \mathcal{H}_P \to \mathcal{H}_S$ establishes the unitary equivalence between the switch and Pauli representations.

This equivalence extends to all operators in $\mathcal{B}(\mathcal{H}_S)$ and $\mathcal{B}(\mathcal{H}_P)$. In particular, let $\{f_i(\{\sigma_k^P(i)\})\}$ be a sequence of linear functions of the $\{\sigma_k^P(i)\}$ which converges in $\mathcal{H}_P$'s weak topology to the operator F_P. Each $f_i(\{\sigma_k^P(i)\}) \in \mathcal{B}(\mathcal{H}_P)$ and $\mathcal{B}(\mathcal{H}_P)$ is closed under weak convergence; so $F_P \in \mathcal{B}(\mathcal{H}_P)$. Similarly, let $\{f_i(\{\sigma_k^S(i)\})\}$ be a sequence of linear functions of the $\{\sigma_k^S(i)\}$, where

$$f_i(\{\sigma_k^S(i)\}) = U f_i(\{\sigma_k^P(i)\})U^\dagger. \tag{6.10}$$

Weak convergence is preserved under unitary transformations, so the $\{f_i(\{\sigma_k^S(i)\})\}$ converge in $\mathcal{H}_S$'s weak topology to some operator $F_S \in \mathcal{B}(\mathcal{H}_S)$, and $F_S = UF_PU^\dagger$.

In the Pauli representation $\mathcal{H}_P \cong \mathbb{C}^{2n}$, we may define the *polarization* observable $\hat{\mathbf{m}}^P := (m_x^P, m_y^P, m_z^P)$, where

$$m_x^P := \frac{1}{n}\sum_{k=1}^{n} \sigma_k^P(x), \quad \text{etc.} \tag{6.11}$$

Clearly, $\hat{\mathbf{m}}^P \in \mathcal{B}(\mathcal{H}_P)$, and the spectrum of $\hat{\mathbf{m}}^P$ is parametrized by points on the unit sphere. From the above considerations, we know that the similarly defined polarization observable $\hat{\mathbf{m}}^S := (m_x^S, m_y^S, m_z^S)$ in the switch representation satisfies

$$\hat{\mathbf{m}}^S = U\hat{\mathbf{m}}^P U^\dagger, \tag{6.12}$$

and so expectation values in S are identical to corresponding (given U) expectation values in P.

The work of Mackey:—
We will summarize the contributions of Mackey to generalizing the Stone-von Neumann theorem. There are two themes, which we treat in order. (A) First: to generalize from $\mathbb{R}$, or $\mathbb{R}^n$, as the underlying classical configuration space, to a locally compact group G; (we will focus on the abelian case). (B) Second: to see the theorem as a special case of a bijective correspondence between *systems of imprimitivity* (in a sense that generalizes what we discussed in (A) of Section 1.1) and unitary representations.

We follow, respectively: Section 3 of S. Summers, 'On the Stone-von Neumann Uniqueness theorem and its ramifications' (Research Gate: 1998, publ. 2001); and Section 4.1 of N. Landsman, 'Between Classical and Quantum', in J Butterfield and J Earman eds., *Handbook of the Philosophy of Physics*, Elsevier 2006. We recommend both articles, especially the latter at: https://arxiv.org/abs/quant-ph/0506082.

(A): Given a locally compact abelian group G, define $L^2(G, d\mu)$, where μ is an appropriate measure (Haar measure). Define the topological character group G^*, i.e. the set of continuous homomorphisms from G to the multiplicate group $U(1)$, and endow it with a natural topology. Then the Weyl form of the Schroedinger representation is described by the following analogue of eq. 1.9, (now with the subscript S for 'Schroedinger'):

$$(U_S(g)\psi)(x) = \psi(g^{-1} \cdot x) \;\; ; \;\; (V_S(\tau)\psi)(x) = \tau\psi(\mathbf{x}) \tag{6.13}$$

So here we do not have an explicit definition of U_S and V_S as exponentiations of operators P, Q respectively, as we had in eq. 1.8. But we recall that eq. 1.9 read as follows:

$$(U(\mathbf{a})\psi)(\mathbf{x}) = \psi(\mathbf{x} - \mathbf{a}) \;\; ; \;\; (V(\mathbf{b})\psi)(\mathbf{x}) = e^{-i\mathbf{b}\cdot\mathbf{x}/\hbar}\psi(\mathbf{x})\,; \tag{6.14}$$

and the Weyl commutation relations, eq. 1.14, read as follows:

$$U(\mathbf{a})V(\mathbf{b}) = e^{i\mathbf{a}.\mathbf{b}/\hbar}V(\mathbf{b})U(\mathbf{a}). \tag{6.15}$$

Thus the expected analogous form of the Weyl commutation relations is

$$U(g)V(\tau) = \tau(g)V(\tau)U(g)\;. \tag{6.16}$$

And indeed Mackey proved (1949)

Theorem 6.3 (Mackey's Theorem)**.** *Let G be a locally compact abelian group, and G^* its topological character group. Let U be a weakly continuous representation of G in a separable Hilbert space $\mathcal{H}$, and V be a weakly continuous representation of G^* in $\mathcal{H}$, such that the following analogue of the Weyl relations, i.e. of eq. 1.14, holds:*

$$U(g)V(\tau) = \tau(g)V(\tau)U(g)\;. \tag{6.17}$$

Then $\mathcal{H}$ is a direct sum of at most countably many closed subspaces $\mathcal{H}_n$, each closed under $\{U(g) : g \in G\} \cup \{V(\tau) : \tau \in G^*\}$. If we define U_n, V_n as the restrictions of U, V to $\mathcal{H}_n$, then there is a Hilbert space isomorphism $W_n : \mathcal{H}_n \to L^2(G, d\mu)$ such that for all $g \in G, \tau \in G^*$:

$$W_n U(g) W_n^{-1} = U_S(g) \;\; ; \;\; W_n V(\tau) W_n^{-1} = V_S(\tau)\,. \tag{6.18}$$

(B): In (A) in Section 1.1, we discussed how the familiar Heisenberg form of the canonical commutation relations can be viewed as arising from a *transitive system of imprimitivity* for $\mathbb{R}^n$. Recall that this was defined as given by:

(1): a PVM on Borel sets $\Delta \subset \mathbb{R}^n$; i.e. a countably additive map $\Delta \mapsto E_\Delta$, where we write Q^i for the associated self-adjoint operators, defined by the spectral theorem: $Q^i := \int_{-\infty}^{\infty} \lambda \, dE^i_\lambda$, for each $i = 1, \ldots, n$; and we write $\mathbf{Q}$ for the vector-operator $\langle Q^1, Q^2, ..., Q^n \rangle$: where 'vector-operator' means as usual that the components transform as expected under rotations; together with:

(2): a strongly continuous unitary group $U(\mathbf{a})$ indexed by $\mathbf{a} \in \mathbb{R}^n$ such that: (i) $U(\mathbf{0}) = I$, $U(\mathbf{a} + \mathbf{a}') = U(\mathbf{a})U(\mathbf{a}')$ for $\mathbf{a}, \mathbf{a}' \in \mathbb{R}^n$; with the additivity and strong continuity being of course understood component-wise, $i = 1, 2, ...n$; and also (ii):

$$U(\mathbf{a}) E_\Delta U(\mathbf{a})^* = E_{\Delta - \mathbf{a}} \, . \tag{6.19}$$

The pair $(\Delta \mapsto E_\Delta, \mathbf{a} \mapsto U(\mathbf{a}))$ satisfying (1) and (2) is called a *transitive system of imprimitivity* following Mackey (1976, Section 3.7).

It then followed that this unitary $U(\mathbf{a})$ 'translates' the 'position' vector operator $\mathbf{Q}$ as expected. That is: $U(\mathbf{a})\mathbf{Q}U(\mathbf{a})^* = \mathbf{Q} + \mathbf{a}I$; again, of course understood component-wise, $i = 1, 2, ...n$. Stone's theorem then implied that there is a unique self-adjoint vector-operator $P \equiv \langle P_1, P_2, ..., P_n \rangle$ such that $U(\mathbf{a}) = e^{-i\mathbf{a}\cdot\mathbf{P}}$ for all $\mathbf{a} \in \mathbb{R}^n$. Besides, there is a dense domain of vectors $D_{QP} \subseteq \mathcal{H}$ on which both $\mathbf{Q}$ and $\mathbf{P}$ are defined, and such that,

$$[Q^j, P_k]\psi = i\delta^j_k \psi \text{ for all } \psi \in D_{QP} \tag{6.20}$$

With this background, we can now state Mackey's general definition of a *system of imprimitivity*. We envisage that a group $G \ni g$ on a space $Q \ni x$: $x \mapsto g \cdot x$; that there is a a unitary representation U of G on a Hilbert space $\mathcal{H}$, and a projection-valued measure $\Delta \mapsto E_{|Delta}$ on Q with values in $\mathcal{H}$. So the expected analogue of eq. 6.19 is

$$U(g) E_\Delta U(g)^* = E_{g\cdot\Delta} \, . \tag{6.21}$$

Thus Mackey defines:

> A *system of imprimitivity* $(\mathcal{H}, U, P)$ for a given action of a group G on a space Q consists of a Hilbert space $\mathcal{H}$, a unitary representation U of G on $\mathcal{H}$, and a projection-valued measure $\Delta \mapsto E_{|Delta}$ on Q with values in $\mathcal{H}$, such that eq. 6.21 holds for all $g \in G$ and all Borel sets $\Delta \subset Q$. (We assume all groups and spaces are locally compact, and actions and representations are assumed continuous.)

In physics such a system describes the quantum mechanics of a particle moving on a configuration space Q on which G acts by symmetry transformations. When everything is smooth (I.e. G is a Lie group, Q is a manifold, and the G-action is smooth), each element X of the Lie algebra $\mathfrak{g}$ of G defines a generalized momentum operator, as follows. We now write the 'quantization map' of Section 1.1 as $\mathcal{Q}_\hbar$

$$\mathcal{Q}_\hbar(X) = i\hbar dU(X) \tag{6.22}$$

on $\mathcal{H}$.[13] These operators satisfy the generalized canonical commutation relations

$$[\mathcal{Q}_\hbar(X), \mathcal{Q}_\hbar(Y)] = i\hbar \mathcal{Q}_\hbar([X, Y]). \tag{6.23}$$

[13]This operator is defined and self-adjoint on the domain of vectors $\psi \in \mathcal{H}$ for which $dU(X)\psi := \lim_{t\to 0} t^{-1}(U(\exp(tX)) - 1)\psi$ exists.

Furthermore, in terms of the operators

$$\mathcal{Q}_\hbar(f) = \int_Q dE(x)\, f(x), \tag{6.24}$$

where f is a smooth function on Q and $X \in \mathfrak{g}$, one in addition has

$$[\mathcal{Q}_\hbar(X), \mathcal{Q}_\hbar(f)] = i\hbar \mathcal{Q}_\hbar(\xi_X^Q f), \tag{6.25}$$

where ξ_X^Q is the canonical vector field on Q defined by the G-action (i.e. $\xi_X^Q f(y) = d/dt|_{t=0}[f(\exp(-tX)y)]$), and

$$[\mathcal{Q}_\hbar(f_1), \mathcal{Q}_\hbar(f_2)] = 0. \tag{6.26}$$

Elementary quantum mechanics on $\mathbb{R}^n$ corresponds to the special case $Q = \mathbb{R}^n$ and $G = \mathbb{R}^n$ with the usual additive group structure. To see this, we denote the standard basis of $\mathbb{R}^3$ (in its guise as the Lie algebra of $\mathbb{R}^3$) by the name (p_j), and furthermore take $f_1(q) = q^j$, $f_2(q) = f(q) = q^k$. Eq. (6.23) for $X = p_j$ and $Y = p_k$ then reads $[\mathcal{Q}_\hbar(p_j), \mathcal{Q}_\hbar(p_k)] = 0$, eq. (6.25) yields the Heisenberg canonical commutation relations, and (6.26) states the commutativity of the position operators, i.e. $[\mathcal{Q}_\hbar(q^j), \mathcal{Q}_\hbar(q^k)] = 0$.

In order to incorporate spin, one picks $G = E(3) = SO(3) \ltimes \mathbb{R}^3$ (i.e. the Euclidean motion group), acting on $Q = \mathbb{R}^3$ in the obvious (defining) way. The Lie algebra of $E(3)$ is $\mathbb{R}^6 = \mathbb{R}^3 \times \mathbb{R}^3$ as a vector space; we extend the basis (p_j) of the second copy of $\mathbb{R}^3$ (i.e. the Lie algebra of $\mathbb{R}^3$) by a basis (J_i) of the first copy of $\mathbb{R}^3$ (in its guise as the Lie algebra of $SO(3)$) , and find that the $\mathcal{Q}_\hbar(J_i)$ are just the usual angular momentum operators.[14]

Mackey's generalization of von Neumann's (1931) uniqueness theorem for the irreducible representations of the Heisenberg canonical commutation relations is his *imprimitivity theorem*. This theorem applies to the special case where $Q = G/H$ for some (closed) subgroup $H \subset G$, and states that (up to unitary equivalence) there is a bijective correspondence between:

1. Systems of imprimitivity $(\mathcal{H}, U, P)$ for the left-translation of G on G/H;

2. Unitary representations U_χ of H.

This correspondence preserves irreducibility. Specifically, given U_χ the triple $(\mathcal{H}^\chi, U^\chi, P^\chi)$ is a system of imprimitivity, where $\mathcal{H}^\chi = L^2(G/H) \otimes \mathcal{H}_\chi$ carries the representation $U^\chi(G)$ induced by $U_\chi(H)$, and the P^χ act like multiplication operators. Conversely, if $(\mathcal{H}, U, P)$ is a system of imprimitivity, then there exists a unitary representation $U_\chi(H)$ such that the triple $(\mathcal{H}, U, P)$ is unitarily equivalent to the triple $(\mathcal{H}^\chi, U^\chi, P^\chi)$ just described. For example, for $G = E(3)$ and $H = SO(3)$ one has $\chi = j = 0, 1, 2, \ldots$ and $\mathcal{H}^j = L^2(\mathbb{R}^3) \otimes \mathcal{H}_j$ (where $\mathcal{H}_j = \mathbb{C}^{2j+1}$ carries the given representation $U_j(SO(3))$).

For example, von Neumann's theorem is recovered as a special case of Mackey's by making the choice $G = \mathbb{R}^3$ and $H = \{e\}$ (so that $Q = \mathbb{R}^3$, as above): the uniqueness of the (regular) irreducible representation of the canonical commutation relations here follows from the uniqueness of the irreducible representation of the trivial group. A more illustrative example is $G = E(3)$ and $H = SO(3)$ (so that $Q = \mathbb{R}^3$), in which case the irreducible representations of the associated system of imprimitivity are classified by spin $j = 0, 1, \ldots$. (And by the usual arguments, one may replace $SO(3)$ by $SU(2)$, so as to obtain $j = 0, 1/2, \ldots$.)

[14]The commutation relations in the previous paragraph are now extended by the familiar relations $[\mathcal{Q}_\hbar(J_i), \mathcal{Q}_\hbar(J_j)] = i\hbar\epsilon_{ijk}\mathcal{Q}_\hbar(J_k)$, $[\mathcal{Q}_\hbar(J_i), \mathcal{Q}_\hbar(p_j)] = i\hbar\epsilon_{ijk}\mathcal{Q}_\hbar(p_k)$, and $[\mathcal{Q}_\hbar(J_i), \mathcal{Q}_\hbar(q^j)] = i\hbar\epsilon_{ijk}\mathcal{Q}_\hbar(q^k)$.

6.2 Proving the theorems ...

There are various proofs, at various levels of abstraction. Here is a partial guide.

(1): S. Summers, 'On the Stone-von Neumann Uniqueness theorem and its ramifications' (Research Gate: 1998, publ. 2001) gives an insightful overview of the history since 1931, emphasising (i) the breakdown for infinite systems, i.e. quantum field theory, and (ii) the generalization by Mackey (1949)

(2): Proofs by elementary methods, i.e. without use of algebraic states etc. (apparently the same as each other), are in:
(i) E Prugovecki *Quantum Mechanics in Hilbert Space*, Section 6.5, p. 342-348.
(ii) L. Takhtajan, *Quantum Mechanics for Mathematicians*, American Mathematical Society 2008, Section 3.1, p. 119-125
(iii) We recommend the approach of B.C. Hall, *Quantum Theory for Mathematicians*, Springer 2013, Chapter 14. He gives a heuristic argument, and then a formal proof (his Sections 14.1 and 14.2 respectively). This approach has the merit of exploiting our previous acquaintance with ladder operators.

(3): Advanced proofs, with use of algebraic states etc:
(i) O. Bratteli and D. Robinson, *Operator Algebras and Quantum Statistical Mechanics* volume 2: (first Chapter = Chapter 5): pp. 30-36, Theorem 5.2.14 and its corollary 5.2.15 (To orient you:— Earlier on, they discuss the uniqueness up to unitary equivalence, (in C*-algebra sense) of the CAR algebra and of the Weyl ie CCR algebra: treating CARs in Theorem 5.2.5 p. 15 f. ; then CCRs in Theorem 5.2.8 p. 19 f; with comparison of CARs and CCRs on p. 22.)
(ii) B. Simon, 'Topics in functional analysis', in R. Streater ed., *Mathematics of Contemporary Physics* 1972; (uses J. Slawny approach: cf. his 'Factor representations and CCRS', *Comm. Math. Phys.* 1972).

Most of the items in (1) to (3) discuss CCRs only. For CARs:

(i) we note that Jordan and Wigner (1928) proved this (in the same paper that introduced anti-commutation relations!). (Cf. also E Wigner, *Group Theory and its application to the quantum mechanics of atomic spectra* Chapter 20, Sections 4 to 9; pp. 223-232 of the 1959 English translation of 1931 German original: not clear to JNB whether this amounts to a proof of uniqueness!)

(ii) Cf. e.g. B C Hall, *Lie Groups Lie Algebras and Representations*, Section 4.4, pp. 101-107 Springer 2003.

(4): Helpful discussions of the theorem include, in addition to Summers above: G. Folland 2008, *Quantum Field Theory: a tourist guide for mathematicians*, pp. 44-46, pp. 57-58.

6.3 Going beyond these theorems ...

The Stone-von Neumann theorem fails to apply if either of its antecedent conditions fail; i.e. if either the classical phase space is not $\mathbb{R}^{2n}$, or else the representation of the Weyl algebra is not weakly continuous. Following Ruestche (2011, Ch. 3), it is helpful to break the various possible failures into three cases:

(i) weak continuity fails;

(ii) classical phase space is finite-dimensional, but not $\mathbb{R}^{2n}$;

(iii) classical phase space is infinite-dimensional: corresponding to a *field theory*.

In each of these cases, we have no guarantee that the quantization of our classical system is unique. In fact, for each of these cases we know that the quantization is not unique. This is considered in Part II.

For the break-down of the Jordan-Wigner theorem, we will in Part II consider the theory of the *infinite spin-chain*, in which we have a spin-$\frac{1}{2}$ system for every integer in $\mathbb{Z}$. This theory has observables satisfying the Pauli relations (6.7). Representations of the Pauli relations in such a theory will be carried by a separable Hilbert space only if we make some hard choices about which of the uncountably many *prima facie* possible states are to be excluded.

We will see that with the obvious ferromagnetic ("neighbours like to be aligned") Hamiltonian, we get a ground state for each spatial direction (unit vector) $\mathbf{u}$ in $\mathbb{R}^3$. Namely: the state in which all the spins are aligned in the direction $\mathbf{u}$. On each such ground state, we build a representation. But they are *unitarily inequivalent*. Thus we return to general philosophico-mathematical theme mentioned at the end of Section 2.4 (end of 2.4.C): *singular limits*. That is: for every finite n (the length of the spin-chain), we have unitary equivalence; but for $n = \infty$, there is unitary inequivalence.

7 Further Reading

In addition to the references below, the following resources are rich with related material.

- H. Araki. *Mathematical Theory of Quantum Fields*: up to Chapter 4. Oxford University Press, 1999.
- Baez, J. C., Segal, I. E. and Zhou, Z. (1992), *Introduction to Algebraic and Constructive Quantum Field Theory.* Princeton: Princeton University Press. (Available for free as a pdf or djvu file on John Baez's website: `http://math.ucr.edu/home/baez/bsz.html`.)
- Clifton. R. *Quantum Entanglements*, edited by J. Butterfield and H. Halvorson: Chapters 6 to 9. Oxford University Press 2004.
- Duncan, A. (2012), *The Conceptual Framework of Quantum Field Theory.* Oxford: OUP.
- Geroch, R. (2005), *Special Topics in Particle Physics.* Lecture notes. (Available online at: `http://strangebeautiful.com/other-texts/geroch-qft-lectures.pdf`.)
- Greiner, W. & Reinhardt, J. (1996), *Field Quantization.* Berlin: Springer.
- R. Haag. *Local Quantum Physics: fields, particles, algebras*: Chapters I, II, III and V.4: Springer 1992.
- K. Hannabuss. *An Introduction to Quantum Theory*: up to Chapter 11. Oxford University Press, 1997.
- T. Jordan. *Linear Operators for Quantum Mechanics*: especially Chapters 3 to 5. John Wiley 1969; Dover 2006.
- N. Landsman. *Foundations of Quantum Theory.* Springer 2017: especially Chapters 5, 6,7,9,10. Open access: downloadable at: https://link.springer.com/book/10.1007/978-3-319-51777-3 .
- Morandi, G. (1992), *The Role of Topology in Classical and Quantum Physics.* Berlin: Springer.
- Wallace, D. (2009), 'QFT, Antimatter, and Symmetry', *Studies in the History and Philosophy of Modern Physics* **40**, pp. 209-222.
  (Available online at: *http://arxiv.org/pdf/0903.3018v2.pdf.*)

References

Ali, S. T. and Engliš, M. (2005). Quantization methods: a guide for physicists and analysts, *Reviews in Mathematical Physics* **17**(04): 391–490. https://arxiv.org/abs/math-ph/0405065.

Butterfield, J. (2006). On Symmetry and Conserved Quantities in Classical Mechanics, *in* W. Demopoulos and I. Pitowsky (eds), *Physical Theory and its Interpretation: Essays in Honor of Jeffrey Bub*, Springer Netherlands, pp. 43–100. http://philsci-archive.pitt.edu/2362/.

Butterfield, J. (2007). On Symplectic Reduction in Classical Mechanics, *in* J. Butterfield and J. Earman (eds), *Philosophy of Physics Part A*, Handbook of the Philosophy of Science, Elsevier, pp. 1–132.
URL: *http://philsci-archive.pitt.edu/2373/*

Da Silva, A. C. (2001). *Lectures on symplectic geometry*, number 1764 in *Lecture Notes in Mathematics*, Berlin and Heidelberg: Springer Verlag.

De Faria, E. and De Melo, W. (2010). *Mathematical aspects of quantum field theory*, Vol. 127 of *Cambridge studies in advanced mathematics*, Cambridge: Cambridge university press.

Dirac, P. A. M. (1927). The quantum theory of the emission and absorption of radiation, *Proceedings of the Royal Society of London, Series A: Containing Papers of a Mathematical and Physical Character* **114**(767): 243–265.

Dirac, P. A. M. (1930). *The Principles of Quantum Mechanics*, 1st edn, Oxford: The Clarendon Press. https://archive.org/details/in.ernet.dli.2015.177580/.

Dirac, P. A. M. (1947). *The Principles of Quantum Mechanics*, 3rd edn, Oxford: The Clarendon Press.

Dixmier, J. (1958). Sur la relation $i(pq - qp) = 1$, *Compositio Mathematica* **13**: 263–269. http://www.numdam.org/item/CM_1956-1958__13__263_0/.

Dyson, F. (1996). *Selected Papers of Freeman Dyson with Commentary*, Providence, RH: American Mathematical Society.

Emch, G. G. (1972). *Algebraic methods in statistical mechanics and quantum field theory*, Interscience Monographs and Texts in Physics and Astronomy, New York, London, Sydney, Toronto: John Wiley & Sons.

Folland, G. B. (2008). *Quantum Field Theory: A tourist guide for mathematicians*, Vol. 149 of *Mathematical Surveys and Monographs*, Providence, Rhode Island: The American Mathematical Society.

Gotay, M. J. (1999). On the groenewold–van hove problem for $\mathbb{R}^{2n}$, *Journal of Mathematical Physics* **40**(4): 2107–2116.

Gotay, M. J., Grundling, H. B. and Tuynman, G. M. (1996). Obstruction results in quantization theory, *Journal of Nonlinear Science* **6**(5): 469–498. https://arxiv.org/abs/dg-ga/9605001.

Halvorson, H. (2001). Reeh-Schlieder Defeats Newton-Wigner: On Alternative Localization Schemes in Relativistic Quantum Field Theory, *Philosophy of Science* **68**(1): 111–133.

Jauch, J. M. (1968). *Foundations of Quantum Mechanics*, Addison-Wesley Series in Advanced Physics, Reading, MA, Menlo Park, CA, London, Don Mills, ON: Addison-Wesley Publishing Company, Inc.

Kay, B. S. (1979). A uniqueness result in the Segal–Weinless approach to linear Bose fields, *Journal of Mathematical Physics* **20**(8): 1712–1713.

Landsman, N. (2007). Between classical and quantum, *in* J. Butterfield and J. Earman (eds), *The Handbook of Philosophy of Physics*, Amsterdam: Elsevier B.V., pp. 417–554. `http://philsci-archive.pitt.edu/2328/`.

Mackey, G. W. (1976). *The theory of unitary group representations*, Chicago and London: The University of Chicago Press.

Marsden, J. and Ratiu, T. (2010). *Introduction to mechanics and symmetry: a basic exposition of classical mechanical systems*, 2nd edn, Springer-Verlag New York, Inc.

Prugovečki, E. (1971). *Quantum mechanics in Hilbert space*, Volume 41 of Pure and Applied Mathematics, New York and London: Academic Press, Inc.

Ruetsche, L. (2011). *Interpreting Quantum Theories*, New York: Oxford University Press.

Segal, I. (1967). Representations of the canonical commutation relations, *in* F. Lurçat (ed.), *Applications of mathematics to problems in theoretical physics: Cargèse Lectures in Theoretical Physics*, New York: Gordon and Breach. `http://philsci-archive.pitt.edu/2328/`.

Segal, I. and Mackey, G. (1963). *Mathematical problems of relativistic physics*, Volume II of Lectures in Applied Mathematics: Proceedings of the Summer Seminar, Boulder, Colorado, 1960, Providence, RI: American Mathematical Society.

Vogan, D. A. (1987). *Unitary representations of reductive Lie groups*, Princeton, NJ: Princeton University Press.

Wald, R. M. (1994). *Quantum field theory in curved spacetime and black hole thermodynamics*, University of Chicago Press.

Weinless, M. (1969). Existence and uniqueness of the vacuum for linear quantized fields, *Journal of Functional Analysis* **4**(3): 350–379.