The Quantization of Linear Dynamical Systems I: Finite Systems

Butterfield, Caulton and Roberts

Philosophical Aspects of QFT, Michaelmas 2020: version for Tues 10 November 2020

This document, and its succcessor 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 means we postpone to the second half of term, coverage of algebraic quantum theory; which will include e.g. inequivalent representations, 'getting out of Fock space', Haag's theorem etc.; for which (cf. 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 idea will be a *one particle structure*, both in general and for the harmonic oscillator as an example (Section 4). The key to successful quantization, which see at work in the harmonic oscillator example, turns out to be the *two out of three property* of the unitary group: which concerns its relation to certain orthogonal and symplectic groups (Section 5). Then we treat the case of finitely many harmonic oscillators, and so the occupation number representation: which can be described in a "Fock-space way" (Section 6). 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 7).

> "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

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

$$\{q^i, q^j\} = \{p_i, p_j\} = 0; \qquad \{q^i, p_j\} = \delta^i_j, \tag{1}$$

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

$$[Q^i, Q^j] = [P_i, P_j] = 0; \qquad [Q^i, P_j] = i\hbar\delta^i_j \mathbb{1}; \tag{2}$$

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

$$Q^i\psi := q_i\psi, \qquad P_j\psi := -\frac{ih}{2\pi}\frac{\partial\psi}{\partial q_j} \qquad \text{for } \psi \in L^2(\mathbb{R}^n, d\mathbf{q}) \text{ and } i = 1, \dots, n. \tag{3}$$

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 7. 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:

$$[Q(f), Q(g)] = i\hbar\, Q(\{f, g\}) \tag{4}$$

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:

$$E_\Delta \mapsto E_{\Delta'} = E_{\Delta - \mathbf{a}} \, ; \tag{5}$$

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 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):

$$U(\mathbf{a})E_\Delta U(\mathbf{a})^* = E_{\Delta - \mathbf{a}}\,. \tag{6}$$

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,

$$[Q^j, P_k]\psi = i\delta^j_k\psi \text{ for all } \psi \in D_{QP} \tag{7}$$

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. 3. Now let us define for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$,

$$U(\mathbf{a}) := e^{-i\mathbf{a}\cdot\mathbf{P}/\hbar} \quad ; \qquad V(\mathbf{b}) := e^{-i\mathbf{b}\cdot\mathbf{Q}/\hbar}; \tag{8}$$

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

$$(U(\mathbf{a})\psi)(\mathbf{x}) = \psi(\mathbf{x} - \mathbf{a}) \;\; ; \;\; (V(\mathbf{b})\psi)(\mathbf{x}) = e^{-i\mathbf{b}\cdot\mathbf{x}/\hbar}\psi(\mathbf{x}) \tag{9}$$

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:

$$U(\mathbf{a})U(\mathbf{a}') = U(\mathbf{a}')U(\mathbf{a}) = U(\mathbf{a} + \mathbf{a}') \qquad V(\mathbf{b})V(\mathbf{b}') = V(\mathbf{b}')V(\mathbf{b}) = V(\mathbf{b} + \mathbf{b}') \tag{10}$$

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:

$$e^A e^B = e^{A+B+\frac{1}{2}[A,B]} \equiv e^{A+B} e^{\frac{1}{2}[A,B]} \qquad (11)$$

To apply (11) to (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

$$U(\mathbf{a})V(\mathbf{b}) = \exp(\frac{1}{2}i(\mathbf{a}\cdot\mathbf{b})/\hbar).\exp(-i(\mathbf{a}.\mathbf{P}/\hbar + \mathbf{b}.\mathbf{Q}/\hbar)) \; ; \qquad (12)$$

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

$$V(\mathbf{b})U(\mathbf{a}) = \exp(-\frac{1}{2}i(\mathbf{a}\cdot\mathbf{b})/\hbar).\exp(-i(\mathbf{a}.\mathbf{P}/\hbar + \mathbf{b}.\mathbf{Q}/\hbar)). \qquad (13)$$

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

$$U(\mathbf{a})V(\mathbf{b}) = e^{i\mathbf{a}.\mathbf{b}/\hbar}V(\mathbf{b})U(\mathbf{a}). \qquad (14)$$

1.2 The Weyl algebra

So from now on, we take as our CCRs, not the Heisenberg form (2), but (14) together with the trivial commutations of Us and Vs alone i.e. (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 (14) and (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 (14) wrong. (See later for discussion of different choices of sign in the two definitions of (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 (14), and in (105) below.

(3): *A more abstract formulation*:— Equation (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

$$W(z) := e^{\frac{1}{2}i\mathbf{a}\cdot\mathbf{b}}U(\mathbf{a})V(\mathbf{b}). \tag{15}$$

Then the Weyl form of the CCRs, i.e. (14) and (10), are equivalent to the following, which is thus also called the *Weyl algebra*: for all $z, z_1, z_2 \in \mathbb{R}^{2n}$,

$$\begin{array}{rcl} W(z_1)W(z_2) & = & e^{\frac{1}{2}i\Omega(z_1,z_2)}W(z_1+z_2); \\ W^{\dagger}(z) & = & W(-z); \end{array} \tag{16}$$

where Ω is the *symplectic product*:

$$\Omega(z_1, z_2) := \mathbf{a}_2 \cdot \mathbf{b}_1 - \mathbf{a}_1 \cdot \mathbf{b}_2, \tag{17}$$

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 $\mathbb{C}$

$$V^{\mathbb{C}} := V \otimes \mathbb{C} \ . \tag{18}$$

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

$$v = v_1 \otimes 1 + v_2 \otimes i \tag{19}$$

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

$$\alpha(v \otimes \beta) = v \otimes (\alpha\beta) \ \text{ for all } v \in V \text{ and } \alpha, \beta \in \mathbb{C} \ ; \tag{20}$$

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

$$\alpha(v \otimes \beta + u \otimes \gamma) := \alpha(v \otimes \beta) + \alpha(u \otimes \gamma) \equiv v \otimes (\alpha\beta) + u \otimes (\alpha\gamma) \ . \tag{21}$$

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

$$v = v_1 + iv_2. \tag{22}$$

One then checks that the definition eq. 18, equivalently eq. 19, implies that the complex scalar multiplication defined by eq. 20, can be written in the usual-looking form. Namely: for a complex number $\alpha = a + ib$ with $a, b \in \mathbb{R}$

$$(a + ib)(v_1 + iv_2) = (av_1 - bv_2) + i(bv_1 + av_2). \tag{23}$$

So we regard $V^{\mathbb{C}}$ as the direct sum of two copies of V, equipped with a complex scalar multiplication defined by eq. 23 .

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

$$v \mapsto v \otimes 1 \ . \tag{24}$$

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:

$$\dim_{\mathbb{C}} V^{\mathbb{C}} = \dim_{\mathbb{R}} V. \tag{25}$$

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

$$V^{\mathbb{C}} := V \oplus V \tag{26}$$

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

$$J(v, w) := (-w, v) \ . \tag{27}$$

Here J encodes multiplication by i in the sense that setting $a = 0, b = 1$ in eq. 23 yields

$$i(v_1 + iv_2) = -v_2 + iv_1 = -v_2 \otimes 1 + v_1 \otimes i \tag{28}$$

where the last expression on the right is in the notation of eq. 19.

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

$$J = \begin{pmatrix} \mathbf{0} & -\mathbb{1}_{\mathbf{V}} \\ \mathbb{1}_{\mathbf{V}} & \mathbf{0} \end{pmatrix} \ . \tag{29}$$

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. 26 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. 25): the *complex* dimension of $V^{\mathbb{C}}$ is equal to the *real* dimension of V, which is half the *real* dimension of $V \oplus V$:

$$\dim_{\mathbb{C}} V^{\mathbb{C}} = \dim_{\mathbb{R}} V = \frac{1}{2} \dim_{\mathbb{R}} (V \oplus V) \ . \tag{30}$$

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. 19 would become

$$v = v_1 \otimes i + v_2 \otimes 1 \tag{31}$$

Then the definition of complex scalar multiplication, eq. 20 and 21, remain as they are. But the notation that drops the tensor product, i.e. eq. 22, becomes

$$v = iv_1 + v_2 \ ; \tag{32}$$

and the usual-looking form of the complex scalar multiplication that we now deduce is the following analogue of eq. 23: for a complex number $\alpha = a + ib$ with $a, b \in \mathbb{R}$

$$(a + ib)(iv_1 + v_2) = (av_2 - bv_1) + i(av_1 + bv_2). \tag{33}$$

Similarly, for the alternative direct sum approach of 2.1.B. Instead of eq. 27, we define the complex structure J on the direct sum $V \oplus V$ by

$$J(v, w) := (w, -v) \ . \tag{34}$$

Then, setting $a = 0, b = 1$ in eq. 33 yields

$$i(iv_1 + v_2) = -v_1 + iv_2 = iv_2 - v_1 = v_2 \otimes i - v_1 \otimes 1 \tag{35}$$

where the last expression on the right is in the notation of eq. 31. This J as defined by eq. 34 is of course just minus the J defined by eq. 27. The matrix form of J as defined by eq. 34 is thus the negative of eq. 29. That is:

$$J = \begin{pmatrix} \mathbf{0} & \mathbb{1}_{\mathbf{V}} \\ -\mathbb{1}_{\mathbf{V}} & \mathbf{0} \end{pmatrix} \ . \tag{36}$$

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:

$$(x + iy)v := xv + yJ(v) \text{ for all } v \in V \text{ and } x, y \in \mathbb{R} \; ; \tag{37}$$

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. 27, ; and the matrix form of J is as in eq. 29. 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 $\{e_1, e_2, \ldots, e_{2n}\}$ of V can be divided in to n pairs, say $\{e_1, e_2\}, \ldots, \{e_{2n-1}, e_{2n}\}$, and then one can define J as the 'swap with a minus' on each such pair, i.e. $J(e_1) := e_2, J(e_2) := -e_1, \ldots, J(e_{2n-1}) := e_{2n}, J(e_{2n}) := -e_{2n-1}$, 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. 37).

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. 39 and 40 in (1) and (2) below.

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

(cf. eq. 30):

$$\frac{1}{2}\dim_{\mathbb{R}}(V \oplus V)_J = \dim_{\mathbb{C}}(V \oplus V)_J = \dim_{\mathbb{R}} V = \frac{1}{2}\dim_{\mathbb{R}}(V \oplus V) \ . \tag{38}$$

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):

$$J_{2n} := \begin{pmatrix} J_2 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & J_2 & \ldots & \mathbf{0} \\ & & \ddots & \\ \mathbf{0} & \mathbf{0} & \ldots & J_2 \end{pmatrix} \ . \tag{39}$$

(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:

$$J_{2n} := \begin{pmatrix} \mathbf{0} & -\mathbb{1}_{\mathbf{n}} \\ \mathbb{1}_{\mathbf{n}} & \mathbf{0} \end{pmatrix} \ : \tag{40}$$

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. 40 is the same as eq. 29.

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 $u, v \in V$

$$B(Ju, Jv) = B(u, v) \ . \tag{41}$$

Recall that since J is an automorphism with $J^2 = -1$, we have $J^{-1} = -J$. This implies that eq. 41 is equivalent to J being skew-adjoint with respect to B. That is:

$$B(Ju, v) = -B(u, Jv) \ . \tag{42}$$

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:

$$g_J(u, v) := \omega(u, Jv) \; . \tag{43}$$

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. 42, implies that the rhs of eq. 43, 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. 37, we write:

$$\begin{aligned} g_J((x + iy)u, v) := \omega((x + iy)u, J(v)) \equiv \omega((xu + yJ(u)), J(v)) \\ \equiv \omega(xu, J(v)) + \omega(yJ(u), J(v)) \equiv x\omega((u, J(v)) + y\omega(J(u), J(v)) \\ \equiv xg_J(u, v) + yg_J(Ju, v) \; . \end{aligned} \tag{44}$$

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. 43.

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

$$\langle u, v \rangle \equiv \langle u, v \rangle_{\omega, J} := g_J(u, v) + i\omega(u, v) \tag{45}$$

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 $x, y \in \mathbb{R}$:

$$\langle (x + iy)u, v \rangle = x\langle u, v \rangle - iy\langle u, v \rangle \;\text{ and }\; \langle u, (x + iy)v \rangle = x\langle u, v \rangle + iy\langle u, v \rangle \; . \tag{46}$$

The check of eq. 46 uses most of the properties we have postulated. Namely: the definition eq. 43 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. 29 (cf. Section 2.1, and Section **??** 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 **??**). So for each $i = 1, ..., n$, and each q_i in the basis yielding the matrix form in eq. 29, 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. 29. 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. 29 (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 6), 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:

$$\alpha * w := \overline{\alpha} \cdot w \ , \quad \text{for all } \alpha \in \mathbb{C}, w \in \overline{W} \tag{47}$$

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,

$$id(\alpha \cdot w) = \alpha \cdot w \equiv \overline{\alpha} * w = \overline{\alpha} * id(w) \tag{48}$$

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:

$$f(\alpha * w) = f(\overline{\alpha} \cdot w) = \overline{\overline{\alpha}} \cdot f(w) = \alpha \cdot f(w) \ ; \tag{49}$$

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:

$$g(\alpha \cdot w) \equiv g(\overline{\alpha} * w) = \overline{\alpha} \cdot \cdot g(w) \ , \tag{50}$$

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

$$\overline{f}(\alpha * w) := f(\overline{\alpha} \cdot w) = \overline{\alpha} \cdot f(w) = \alpha * \overline{f}(w) \ . \tag{51}$$

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*

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q^i} \quad ; \quad \frac{dq^i}{dt} = \frac{\partial H}{\partial p_i} \ , \tag{52}$$

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

$$\xi^\alpha = q^\alpha, \ \ \alpha = 1, ..., n \quad ; \quad \xi^\alpha = p_{\alpha-n}, \ \ \alpha = n+1, ..., 2n \tag{53}$$

Hamilton's equations become

$$\dot{\xi}^\alpha = \frac{\partial H}{\partial \xi^{\alpha+n}}, \ \ \alpha = 1, ..., n \quad ; \quad \dot{\xi}^\alpha = -\frac{\partial H}{\partial \xi^{\alpha-n}}, \ \ \alpha = n+1, ..., 2n \ . \tag{54}$$

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

$$\omega := \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix} \ . \tag{55}$$

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

$$\tilde{\omega} = -\omega = \omega^{-1} \quad \text{so that} \quad \omega^2 = -\mathbf{1} \quad ; \quad \text{also} \quad \det \omega = 1. \tag{56}$$

Using ω, Hamilton's equations eq. 54 get the more symmetric form, in matrix notation

$$\dot{\xi} = \omega \frac{\partial H}{\partial \xi} \ . \tag{57}$$

In terms of components, writing $\omega^{\alpha\beta}$ for the matrix elements of ω, and $\partial_\alpha := \partial \ /\partial \xi^\alpha$, eq. 54 become

$$\dot{\xi}^\alpha = \omega^{\alpha\beta} \partial_\beta H. \tag{58}$$

Eq. 57 and 58 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. 57 can be written

$$X_H(z) = \omega \nabla H(z). \tag{59}$$

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

$$\omega := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{60}$$

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

$$A : ((q^1, p_1), (q^2, p_2)) \in \mathbb{R}^2 \times \mathbb{R}^2 \mapsto q^1 p_2 - q^2 p_1 \in \mathbb{R} \tag{61}$$

since

$$q^1 p_2 - q^2 p_1 = \begin{pmatrix} q^1 & p_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} q^2 \\ p_2 \end{pmatrix} = \det \begin{pmatrix} q^1 & q^2 \\ p_1 & p_2 \end{pmatrix} \quad . \tag{62}$$

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. 55 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:

$$\Sigma_{i=1}^n \; q^i p'_i - q'^i p_i \quad . \tag{63}$$

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. 55.)

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. 59; and we will see it again, for example in terms of Poisson brackets (eq. 100), and in geometric terms (Section 3.4). But throughout, the main idea will be as suggested by eq. 59: 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

$$e_i \mapsto [e_i] \in V^{**} : e^j \in V^* \mapsto e^j(e_i) = \delta^j_i \tag{64}$$

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

$$\forall \alpha \in W^*, \forall v \in V : \quad \tilde{A}(\alpha)(v) \equiv < \tilde{A}(\alpha) ; v > := \alpha(A(v)) \equiv (\alpha \circ A)(v) . \tag{65}$$

[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 $A : V \rightarrow W$ 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:

$$A(e_i) = A^j_i f_j \;\; ; \quad \text{i.e. with } \; v = v^i e_i, \;\; (A(v))^j = A^j_i v^i \;\; . \tag{66}$$

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

$$A_{ij} := A(e_i, f_j) \tag{67}$$

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 $A : V \times W \rightarrow \mathbb{R}$, we define the *associated linear map* $A^\flat : V \rightarrow W^*$ by

$$A^\flat(v)(w) \;:=\; A(v, w) \;\; . \tag{68}$$

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}:

$$[A^\flat]_{ij} = A_{ij}. \tag{69}$$

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

$$B^\sharp(v, w) \;\;=\; < B(v) \; ; \; w > \;\; . \tag{70}$$

If we put $A^\flat$ for B in eq. 70, its associated bilinear form, acting on vectors $v = v^i e_i, w = w^j f_j$, yields, by eq. 68:

$$(A^\flat)^\sharp(v, w) \;\;=\; < A^\flat(v) \; ; \; w > \;=\;\; A(v, w) \; . \tag{71}$$

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

$$(B^\sharp)^\flat(v)(w) \equiv < (B^\sharp)^\flat(v) \; ; \; w > = \; B(v)(w) \equiv < B(v) \; ; \; w > \;\; \text{so that} \;\; (B^\sharp)^\flat = B \; . \tag{72}$$

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 $\alpha \in V^*, \beta \in W^*$:

$$(v \otimes w)(\alpha, \beta) \;\;:=\;\; v(\alpha) w(\beta) \;\equiv < v \; ; \; \alpha >< w \; ; \; \beta > \;\; . \tag{73}$$

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 $V \times W$:

$$(\alpha \otimes \beta)(v, w) \;\;:=\;\; \alpha(v) \beta(w) \;\equiv < v \; ; \; \alpha >< w \; ; \; \beta > \;\; . \tag{74}$$

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 $V \times W^*$:

$$(\alpha \otimes w)(v, \beta) \;\;:=\;\; \alpha(v) w(\beta) \;\equiv < v \; ; \; \alpha >< w \; ; \; \beta > \;\; . \tag{75}$$

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

$$A(e_i) = A^j_i f_j \quad \text{iff} \quad < A(e_i); f^j > \;=\; A^j_i \tag{76}$$

eq. 75 implies that

$$A \;=\; A_i^j \, e^i \otimes f_j \;\; . \tag{77}$$

Similarly, a bilinear form $A : V \times W \rightarrow \mathbb{R}$ with matrix $A_{ij} := A(e_i, f_j)$ (cf. eq. 67) is:

$$A \;=\; A_{ij} \, e^i \otimes f^j \tag{78}$$

The definitions of tensor product eq. 73, 74 and 75 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 $\omega : V \times V \rightarrow \mathbb{R}$ 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 $\omega^\flat : V \rightarrow V^*$ now have a square matrix ω_{ij} (cf. eq. 69). 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. 74 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

$$\alpha \wedge \beta : (v, w) \in V \times V \mapsto (\alpha(v))(\beta(w)) - (\alpha(w))(\beta(v)) \in \mathbb{R} \, . \tag{79}$$

(The connection with the interpretation of the symplectic matrix in terms of areas, especially eq. 63, 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 $\omega : V \times V \rightarrow \mathbb{R}$: ω is non-degenerate iff the matrix ω_{ij} is non-singular iff $\omega^\flat : V \rightarrow V^*$ 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 $(\omega^\flat)^{-1} : V^* \rightarrow V$. The isomorphism $\omega^\sharp : V^* \rightarrow V$ 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. 70, 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. 71;

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

$$\omega^\sharp \circ \omega^\flat = id_V \;\text{ and }\; \omega^\flat \circ \omega^\sharp = id_{V^*} \; . \tag{80}$$

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. 80, is equivalent to a natural expression:

$$\forall \alpha, \beta \in V^* : \; < \omega^\sharp(\alpha), \beta > := \; \omega((\omega^\flat)^{-1}(\alpha), (\omega^\flat)^{-1}(\beta)) \; . \tag{81}$$

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

$$\omega \ = \ \Sigma_{i=1}^{s} \, e^i \wedge e^{i+s} \ \ ; \tag{82}$$

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

$$\omega \ = \ \begin{pmatrix} \mathbf{0} & \mathbf{1} & \mathbf{0} \\ -\mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \ . \tag{83}$$

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. 83 loses its bottom row and right column consisting of zero matrices, and reduces to the form of the naive symplectic matrix, eq. 55, in **(1)** of Section 3.1. Equivalently: eq. 82 reduces to eq. 63.

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 **??**); 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

$$\omega((v_1, \alpha_1), (v_2, \alpha_2)) \ \ := \ \ \alpha_2(v_1) - \alpha_1(v_2) \ \ . \tag{84}$$

So defined, ω is by construction a symplectic form, and so has the normal form given by eq. 55.

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. 55. 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. 65) and using the second definition eq. 80 of $\sharp$, the following maps (both from Z^* to Z) are equal:

$$A \circ \omega^\sharp \circ \tilde{A} = \omega^\sharp \ \ ; \tag{85}$$

or in matrix notation, with the *matrix* ω given by eq. 55, and again writing $\tilde{}$ for the transpose of a matrix

$$A \omega \tilde{A} = \omega \ \ . \tag{86}$$

(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. 55;

(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. 80).

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

$$\frac{df}{dt} = \dot{q}^i \frac{\partial f}{\partial q^i} + \dot{p}_i \frac{\partial f}{\partial p_i} \; . \tag{87}$$

(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

$$\frac{df}{dt} = \frac{\partial H}{\partial p_i}\frac{\partial f}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial f}{\partial p_i} \; . \tag{88}$$

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

$$\{f, g\} := \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} \; ; \tag{89}$$

so that the rate of change of f is given by

$$\frac{df}{dt} = \{f, H\} \; . \tag{90}$$

In terms of the $2n$ coordinates ξ^α (eq. 53) and the matrix elements $\omega^{\alpha\beta}$ of ω (eq. 58), we can write eq. 88 as

$$\frac{df}{dt} = (\partial_\alpha f)\dot{\xi}^\alpha = (\partial_\alpha f)\omega^{\alpha\beta}(\partial_\beta H) \;\; ; \tag{91}$$

and so we can define the Poisson bracket by

$$\{f, g\} := (\partial_\alpha f)\omega^{\alpha\beta}(\partial_\beta g) \equiv \frac{\partial f}{\partial \xi^\alpha}\omega^{\alpha\beta}\frac{\partial g}{\partial \xi^\beta} \;\; . \tag{92}$$

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 $z = (q, p) \in \Gamma$:

$$\{f, g\}(z) = \tilde{\nabla f}(z).\omega.\nabla g(z). \tag{93}$$

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:

$$\frac{dH}{dt} = \{H, H\} \equiv 0 \ . \tag{94}$$

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

$$\{f, h \cdot g\} = \{f, h\} \cdot g + h \cdot \{f, g\} \ . \tag{95}$$

(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:

$$\frac{d}{dt}\{f, g\} = \{\frac{df}{dt}, g\} + \{f, \frac{dg}{dt}\} \ . \tag{96}$$

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

$$\{\{f, h\}, g\} + \{\{g, f\}, h\} + \{\{h, g\}, f\} = 0 \ \ . \tag{97}$$

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

$$\{\xi^\alpha, \xi^\beta\} = \omega^{\alpha\beta} \ \ ; \ \text{i.e.} \tag{98}$$

$$\{q^i, p_j\} = \delta^i_j \ \ , \quad \{q^i, q^j\} = \{p_i, p_j\} = 0 \ . \tag{99}$$

Eq. 99 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 $(q, p) \rightarrow (Q(q, p), P(q, p))$, 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:

$$D := X_H = \frac{d}{dt} = \dot{q}^i \frac{\partial}{\partial q^i} + \dot{p}_i \frac{\partial}{\partial p_i} = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i} = \{\cdot, H\} \; . \tag{100}$$

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

$$X_f = \frac{d}{ds} = \{\cdot, f\} \; . \tag{101}$$

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. 53

$$\dot{\xi}^\alpha = X^\alpha(\xi) \;\; , \tag{102}$$

there need be no scalar $H : \Gamma \rightarrow \mathbb{R}$ such that, as required by eq. 58,

$$X^\alpha = \omega^{\alpha\beta}\partial_\beta H \; . \tag{103}$$

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 (17): which we repeat here:

$$\Omega(z_1, z_2) := \mathbf{a}_2.\mathbf{b}_1 - \mathbf{a}_1.\mathbf{b}_2, \tag{104}$$

Then—now looking briefly at the topic of quantization—the Weyl form of the CCRs, i.e. (14) and (10), are equivalent, as we saw in Section 1.2, to the following: for all $z, z_1, z_2 \in \mathbb{R}^{2n}$,

$$\begin{array}{rcl} W(z_1)W(z_2) & = & e^{\frac{1}{2}i\Omega(z_1,z_2)}W(z_1 + z_2); \\ W^\dagger(z) & = & W(-z) \ . \end{array} \tag{105}$$

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 (1: repeated as **??**) may be written

$$\{\Omega(z_1, \cdot), \Omega(z_2, \cdot)\} = -\Omega(z_1, z_2) \ . \tag{106}$$

So the corresponding Heisenberg form of the CCRs are

$$[\hat{\Omega}(z_1, \cdot), \hat{\Omega}(z_2, \cdot)] = -i\Omega(z_1, z_2)\mathbb{1} \ . \tag{107}$$

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

$$W(z) := e^{i\hat{\Omega}(z,\cdot)}. \tag{108}$$

obey the Weyl algebra, eq. 105.

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)

$$X_z = \sum_{i=1}^{n} b_i \frac{\partial}{\partial q^i} - a^i \frac{\partial}{\partial p_i}, \tag{109}$$

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

$$\omega(X_{z_1}, X_{z_2}) = \mathbf{a}_2.\mathbf{b}_1 - \mathbf{a}_1.\mathbf{b}_2. \tag{110}$$

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 (105), 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 (105) 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

$$\alpha \wedge \beta : (v, w) \in \mathbb{R}^2 \times \mathbb{R}^2 \mapsto (\alpha(v))(\beta(w)) - (\alpha(w))(\beta(v)) \in \mathbb{R} \ . \tag{111}$$

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. 61, 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. 63. 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 $\mathbb{R}^n \times \mathbb{R}^n$:

$$(q^i \frac{\partial}{\partial q^i} + p_i \frac{\partial}{\partial p_i}, q'^i \frac{\partial}{\partial q^i} + p'_i \frac{\partial}{\partial p_i}) \mapsto \Sigma_{i=1}^n \, q^i p'_i - q'^i p_i \ . \tag{112}$$

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:

$$\mathbf{i}_v\alpha(v_2, ..., v_k) := \alpha(v, v_2, ..., v_k) \ . \tag{113}$$

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

$$\mathbf{i}_v(\alpha \wedge \beta) = (\mathbf{i}_v \alpha) \wedge \beta + (-1)^k \alpha \wedge (\mathbf{i}_v \beta) \ . \tag{114}$$

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

$$df(X) := X(f) \tag{115}$$

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 Section **??**, 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 Section **??**, can be extended to $\Omega(M)$. On analogy with eq. 113, 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:

$$\mathbf{i}_X \alpha(x) : (v_2, ..., v_k) \mapsto \alpha(x)(X(x), v_2, ..., v_k) \in \mathbb{R} \ . \tag{116}$$

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. 114.)

(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

$$\mathbf{d}(\alpha \mid_U) = (\mathbf{d}\alpha) \mid_U \quad . \tag{117}$$

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

$$\alpha = \alpha_{i_1 \dots i_k} \, dx^{i_1} \wedge \cdots \wedge dx^{i_k} \quad \text{(sum on } i_1 < i_2 < \cdots < i_k\text{)}, \tag{118}$$

then one proves that the exterior derivative is

$$\mathbf{d}\alpha \; = \; \frac{\partial \alpha_{i_1 \dots i_k}}{\partial x^j} \, dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \quad \text{(sum on all } j \text{ and } i_1, \cdots < i_k\text{)}, \tag{119}$$

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

$$\mathcal{L}_X \alpha = \mathbf{d}\mathbf{i}_X \alpha + \mathbf{i}_X \mathbf{d}\alpha \; . \tag{120}$$

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. 55, 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, $\sigma, \tau \in T_x$:

$$\mathbf{d}\omega = 0 \quad \text{and} \quad \forall \, \tau \neq 0, \;\; \exists \sigma : \quad \omega(\tau, \sigma) \neq 0 \;\; . \tag{121}$$

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. 83).

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. 68. 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

$$\omega^\flat(\tau)(\sigma) := \omega(\sigma, \tau) \quad \forall \sigma \in T_x \ . \tag{122}$$

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. 59.

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

$$dp \wedge dq := dp_i \wedge dq^i := \Sigma_{i=1}^n dp_i \wedge dq^i \ . \tag{123}$$

To show that eq. 123 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

$$\pi : (q, p) \in T^*Q \mapsto q \in Q. \tag{124}$$

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:

$$D\pi : \tau \in T_{(q,p)}(T^*Q) \mapsto (D\pi(\tau)) \in T_q \ . \tag{125}$$

Now define a 1-form θ_H on T^*Q by

$$\theta_H : \tau \in T_{(q,p)}(T^*Q) \mapsto p(D\pi(\tau)) \in \mathbb{R} \ ; \tag{126}$$

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. 126, 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

$$\theta_H = p_i dq^i. \tag{127}$$

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

$$\mathbf{d}(\theta_H) := \mathbf{d}(p_i dq^i) \equiv dp_i \wedge dq^i \ . \tag{128}$$

where the last equation follows immediately from eq. 119. 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. 63 or eq. 84, both in Section 3.1, or eq. 112 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. 55.

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. 57 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. 59 of **(1)** of Section 3.1, viz.

$$X_H(z) = \omega \nabla H(z) \ ; \tag{129}$$

and then via the Poisson bracket, in eq. 100 of Section 3.2, viz.

$$D := X_H = \frac{d}{dt} = \dot{q}^i \frac{\partial}{\partial q^i} + \dot{p}_i \frac{\partial}{\partial p_i} = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i} = \{\cdot, H\} \ . \tag{130}$$

The symplectic structure and Poisson bracket were related by eq. 93, viz.

$$\{f, g\}(z) = \tilde{\nabla} f(z).\omega.\nabla g(z). \tag{131}$$

And to this earlier discussion, **(3)** above has added the identification of the canonical symplectic form of a cotangent bundle, eq. 128.

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. 80 and 122). Then Hamilton's equations, eq. 59 or 129, may be written as:

$$\dot{z} = X_H(z) = \omega^\sharp(\mathbf{d}H(z)) = \omega^\sharp(dH(z)) \ . \tag{132}$$

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

$$\omega^\flat X_H(z) = dH(z) \ . \tag{133}$$

In terms of the symplectic form ω at z, this is (cf. eq. 68): for all vectors $\tau \in T_z$

$$\omega(X_H(z), \tau) = dH(z) \cdot \tau \ ; \tag{134}$$

or in terms of the contraction defined by eq. 113, with $\cdot$ marking the argument place of $\tau \in T_z$:

$$\mathbf{i}_{X_H}\omega := \omega(X_H(z), \cdot) = dH(z)(\cdot) \ . \tag{135}$$

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

$$\mathbf{i}_{X_f}\omega = df \ . \tag{136}$$

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. 101 and 130: viz.

$$\mathcal{L}_{X_f} g = dg(X_f) = X_f(g) = \{g, f\} \ . \tag{137}$$

Combining this with eq. 136, we can reformulate the relation between the symplectic form and Poisson bracket, eq. 131, in the form:

$$\{g, f\} = dg(X_f) = \mathbf{i}_{X_f} dg = \mathbf{i}_{X_f}(\mathbf{i}_{X_g}\omega) = \omega(X_g, X_f) \ . \tag{138}$$

(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. 120, applied to ω becomes

$$\mathcal{L}_X\omega \equiv \mathbf{d}\mathbf{i}_X\omega + \mathbf{i}_X\mathbf{d}\omega = \mathbf{d}\mathbf{i}_X\omega \ . \tag{139}$$

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 $f : M \rightarrow \mathbb{R}$ such that

$$\mathbf{i}_X\omega = df \ \text{ i.e. } \ X = X_f \ . \tag{140}$$

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:*

$$\frac{d}{dt}q_i(t) = \frac{\partial H}{\partial p_i} \qquad\qquad \frac{d}{dt}p_i(t) = -\frac{\partial H}{\partial q_i} \tag{141}$$

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:*

$$\iota_X\omega = dH \tag{142}$$

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 (142) 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 (142) 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,

$$\omega := \begin{pmatrix} & -\mathbb{1} \\ \mathbb{1} & \end{pmatrix} \tag{143}$$

(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 (142) is just the familiar form of Hamilton's equations — it is just Equation (141) 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,

$$R_\theta := \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{144}$$

Comparing this to the matrix expression of ω for $n = 2$ in Equation (143) 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, $\sum_{i=1}^n u_i v_i' - v_i' u_i$.

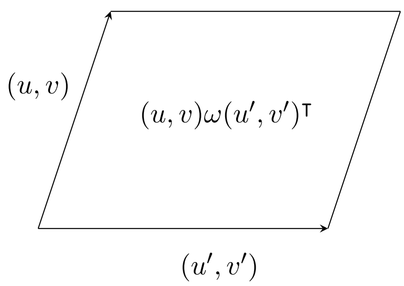

Figure 1: The symplectic form returns the signed area of a parallelogram.