Extra material for "Quantum Theory for Mathematicians"

Brian C. Hall

June 2013

This file contains a number of topics that were omitted from my book, "Quantum Theory for Mathematicians," in the interests of space, but which I wanted to make available to the interested reader.

1. Time-reversal invariance in classical mechanics

One intriguing aspect of classical mechanics is the phenomenon of *time-reversal invariance.* In terms of trajectories in the configuration space $\mathbb{R}^n$ (the position of the particle), time-reversal invariance means the following. If a curve $\mathbf{x}(t)$ in $\mathbb{R}^n$ is a solution to Newton's Law in the form

$$m\ddot{\mathbf{x}} = -\nabla V(\mathbf{x}), \tag{1}$$

then so is the curve $\mathbf{x}(-t)$, as is easily verified by direct computation. (Note that $d^2/dt^2(\mathbf{x}(-t)) = \ddot{\mathbf{x}}(-t)$, since the two derivatives pull out two minus signs.)

In terms of trajectories in phase space, we note that $d\mathbf{x}(-t)/dt = -\dot{\mathbf{x}}(-t)$. This means that the momentum of the particle along the time-reversed trajectory is the negative of the momentum along the original trajectory. We may therefore formulate time-reversal in phase space as follows.

Proposition 1. *Suppose the Hamiltonian of a particle is of the form in Eq. (2.24) in the text and that the curve $(\mathbf{x}(t), \mathbf{p}(t))$ in $\mathbb{R}^{2n}$ is a solution of Hamilton's Equations. Then the curve $(\mathbf{x}(-t), -\mathbf{p}(-t))$ is also a solution of Hamilton's Equations.*

This form of time-reversal invariance is again easily verified simply by plugging into Hamilton's Equations.

If the force law is of a more general form than $\mathbf{F} = -\nabla V$, then time-reversal invariance may fail to hold. If, for example, the force law depends on the velocity of the particle (as in the case of a particle in a magnetic field), then we typically will not get time-reversal invariance, simply because reversing the time reverses the sign of the velocity. In the case of a one-dimensional particle with damping, for example, the energy always decreases with time: $dE/dt \leq 0$ with equality only when $\dot{x} = 0$. Since the energy of the system is unchanged when we replace p by $-p$, we see that $E(x(-t), -p(-t)) = E(x(-t), p(-t))$ and so the energy along the curve $(x(-t), -p(-t))$ is *increasing* with time, and so this curve cannot be a solution to Newton's Law. (The only exception is if $x(t)$ and $p(t)$ are constant.) In the general Hamiltonian setting, it is not hard to see that time-reversal invariance, in the sense of Proposition 1, holds precisely if $H(\mathbf{x}, -\mathbf{p}) = H(\mathbf{x}, \mathbf{p})$ for all $\mathbf{x}$ and $\mathbf{p}$.

Time-reversal invariance is a surprising phenomenon, simply because real-world physical phenomena seem to behave very differently forward in time than they do backward in time. The occurrence of "practical" irreversibility in reversible systems is a fascinating possibility that has to do with statistical properties of systems with many degrees of freedom, as well as with the choice of initial condition. Roughly speaking, if we start off the system in a very atypical state (say one that is highly "ordered") it is likely to run toward a more typical state (one that is "disordered"). This is true whether we run the system forward in time or backward in time, and is therefore a time-reversal invariant statement. Thus it is possible to have a sort of irreversibility (the tendency to move from atypical states to typical states) even in time-reversal invariant systems. Although we cannot delve further into this topic here, the reader can find an excellent conceptual discussion of the issue in the following article by Joel L. Lebowitz: Microscopic origins of irreversible macroscopic behavior, STATPHYS 20 (Paris, 1998), *Phys. A* **263** (1999), 516–527.

There is an analog of Proposition 1 in quantum mechanics, as we will now see.

2. Time-reversal invariance in quantum mechanics

In the previous section, we examined time-reversal invariance in classical mechanics. For Hamiltonians of the usual form ($H(\mathbf{x},\mathbf{p}) = |\mathbf{p}|^2/(2m) + V(\mathbf{x})$), Proposition 1 tells us that if we time-reverse a trajectory in phase space (by changing t to $-t$) and also apply the map $(\mathbf{x},\mathbf{p}) \longmapsto (\mathbf{x},-\mathbf{p})$, then we get again a trajectory. In quantum mechanics, the analog of the map $(\mathbf{x},\mathbf{p}) \longmapsto (\mathbf{x},-\mathbf{p})$ is taking the complex conjugate of wave function. After all, if $\phi = \bar{\psi}$, then the position probability density $|\phi|^2$ is the same as $|\psi|^2$. On the other hand, the reader may easily verify that $|\tilde{\phi}(p)|^2 = |\tilde{\psi}(-p)|$, where $\tilde{\psi}$ is the momentum wave function given by Definition 6.1 in the text.

With this analogy in mind, we can easily formulate a version of time-reversal invariance in quantum mechanics.

Proposition 2. *Suppose the Hamiltonian of a particle in* $\mathbb{R}^n$ *is of the form*

$$\hat{H} = -\frac{1}{2m}\sum_{j=1}^{n} P_j^2 + V(\mathbf{X}),$$

where V *is a real-valued function. Then if* $\psi(\mathbf{x},t)$ *is a solution of the Schrödinger equation* $\partial\psi/\partial t = -(i/\hbar)\hat{H}\psi$*, so is*

$$\overline{\psi(\mathbf{x},-t)}.$$

This result is verified simply by plugging into the Schrödinger equation, using the simple observation that $\hat{H}$ commutes with complex conjugation. As in the classical case, there are quantum systems that do not exhibit time-reversal invariance, such as the Schrödinger equation for a charged particle in a magnetic field.

3. Asymptotics of the number of eigenvalues for a particle in a square well

In this section, we examine the approximate number of eigenvalues for a particle in a square well, in the case of a very deep well, where the number of eigenvalues is large. We show that the asymptotics of the number of eigenvalues is governed by the *volume of phase space* for the relevant range of energies, namely, $-C < E < 0$.

As seen, the structure of solutions to the matching conditions (5.9) and (5.10) (for even and odd solutions, respectively) is determined by the quantity

$$\sqrt{c}A = \frac{A\sqrt{2mC}}{\hbar}. \tag{2}$$

If $\sqrt{c}A$ is at least $3\pi/2$, then as ε decreases from c, the tangent function in (5.9) will go from 0 to $+\infty$ and then from $-\infty$ to $+\infty$, in which case we will have at least two different values of ε and hence of E. More generally, if

$$\left(n-\frac{1}{2}\right)\pi \leq \sqrt{c}A \leq (n+\frac{1}{2})\pi, \tag{3}$$

then the number of solutions to (5.9)—and hence the number of even eigenvectors—will be either n or $n+1$.

Similarly, if

$$n\pi \leq \sqrt{c}A \leq (n+1)\pi,$$

then the number of solutions to (5.10) will be either n or $n+1$. What we see, then, is that the total number of solutions to (5.9) and (5.10) will behave asymptotically like $2\sqrt{c}A/\pi$ for large values of $\sqrt{c}A$.

Proposition 3. *Let "Area" denote the area of a region in the classical phase space $\mathbb{R}^2$, and let $H(x,p) = p^2/(2m) + V(x)$, where $V(x)$ is given by Eq. (5.1) in the text. Then the number of independent square-integrable eigenvectors for $\hat{H}$ with eigenvalue less than zero behaves asymptotically like*

$$\frac{\text{Area}\{(x,p)\,|H(x,p)<0\}}{2\pi\hbar} \tag{4}$$

as the quantity $\sqrt{c}A$ tends to infinity.

Proof. For $H(x,p)$ to be negative, x has to be in the range $-A < x < A$. Assuming x is in this range, $H(x,p)$ will be negative provided that $p^2/(2m) < C$, or, equivalently, $|p| < \sqrt{2mC}$. Thus, the region $\{H(x,p) < 0\}$ is a rectangle with area $(2A)(2\sqrt{2mC})$. Thus,

$$\frac{\text{Area}\{(x,p)\,|H(x,p)<0\}}{2\pi\hbar} = \frac{2A\sqrt{2mC}}{\pi\hbar} = \frac{2\sqrt{c}A}{\pi},$$

which is what we said that the number of eigenvectors is asymptotic to. ■

Note that the numerical value of Area$\{H(x,p) < 0\}$ depends entirely on the system of units being used (meters versus centimeters, etc.). Similarly, the numerical value of $\hbar$ depends on the system of units. But both Area$\{H(x,p) < 0\}$ and $\hbar$ have units of position times momentum. Thus, the ratio in (4) *does not* depend on the system of units; it is a "dimensionless" quantity in the usual physic terminology.

With a little bit of additional effort, one can obtain a stronger result: for any two values E_0 and E_1 with $-C < E_0 < E_1 < 0$, the number of eigenvalues E in the range $E_0 < E < E_1$ behaves asymptotically like

$$\frac{\text{Area}\{(x,p)\,|E_0 < H(x,p) < E_1\}}{2\pi\hbar}$$

in the limit as the above quantity tends to infinity.

4. IDENTITIES INVOLVING WICK, ANTI-WICK, AND WEYL QUANTIZATIONS

In this section, we derive a beautiful formula (Theorem 5) that relates the three most important quantization schemes for Euclidean space, the Weyl, Wick-ordered, and anti-Wick-ordered quantizations.

Proposition 4. *The Wick-ordered and anti-Wick-ordered quantizations—viewed as linear maps of the space of polynomials on* $\mathbb{R}^2$ *into operators on* $C_c^\infty(\mathbb{R})$*—are uniquely characterized by the conditions that* $Q_{\text{Wick}}(1) = Q_{\text{anti-Wick}}(1) = I$ *and the following identities:*

$$Q_{\text{Wick}}((x - i\alpha p)f(x,p)) = (X - i\alpha P)Q_{\text{Wick}}(p) \tag{5}$$
$$Q_{\text{Wick}}((x + i\alpha p)f(x,p)) = Q_{\text{Wick}}(p)(X + i\alpha P) \tag{6}$$

and

$$Q_{\text{anti-Wick}}((x - i\alpha p)f(x,p)) = Q_{\text{Wick}}(p)(X - i\alpha P) \tag{7}$$
$$Q_{\text{anti-Wick}}((x + i\alpha p)f(x,p)) = (X + i\alpha P)Q_{\text{Wick}}(p) \tag{8}$$

for all polynomials f.

Proof. It is easily seen that (5) and (6) actually hold, and that they imply the defining condition of the Wick-ordered quantization, and similarly for (7) and (8). ■

We now turn to a fascinating result relating the Wick, anti-Wick, and Weyl quantizations. Let D be the operator on $\mathbb{R}^2$ given by

$$D = \alpha\frac{\partial^2}{\partial x^2} + \frac{1}{\alpha}\frac{\partial^2}{\partial p^2}, \tag{9}$$

which is just a rescaled version of the Laplacian. We may then define the forward and backward heat operators e^{tD} and e^{-tD} on polynomials by

$$e^{tD}(f) = \sum_{m=0}^{\infty}\frac{t^m}{m!}D^m f \tag{10}$$
$$e^{-tD}(f) = \sum_{m=0}^{\infty}\frac{(-t)^m}{m!}D^m f \tag{11}$$

where if f is a polynomial on $\mathbb{R}^2$ of degree j, we will have $D^m f = 0$ once $2m > j$.

Theorem 5. *The Wick, anti-Wick, and Weyl quantization schemes are related as follows. For every polynomial* $f(x,p)$ *on* $\mathbb{R}^2$*, we have*

$$Q_{\text{Wick}}(f) = Q_{\text{Weyl}}(e^{-\hbar D/4}f) \tag{12}$$
$$Q_{\text{anti-Wick}}(f) = Q_{\text{Weyl}}(e^{\hbar D/4}f). \tag{13}$$

It is not hard to show that each of the quantization schemes in Definition 13.1 in the text is invertible as a map from polynomials on $\mathbb{R}^2$ into differential operators on $\mathbb{R}$ with polynomial coefficients. It then follows from Theorem 5 that

$$Q_{\text{Wick}}^{-1}(Q_{\text{antiWick}}(f)) = e^{\hbar D/2}f. \tag{14}$$

The map $Q_{\text{Wick}}^{-1} \circ Q_{\text{antiWick}}$ is generally called the *Berezin transform*.

Lemma 6. *Let* $z = x - i\alpha p$ *and* $\bar{z} = x + i\alpha p$*, and consider the operators*

$$\frac{\partial}{\partial z} := \frac{1}{2}\left(\frac{\partial}{\partial x} + \frac{i}{\alpha}\frac{\partial}{\partial p}\right); \quad \frac{\partial}{\partial \bar{z}} := \frac{1}{2}\left(\frac{\partial}{\partial x} - \frac{i}{\alpha}\frac{\partial}{\partial p}\right). \tag{15}$$

Then for every polynomial f on $\mathbb{R}^2$, we have

$$(e^{-\hbar D/4}M_z e^{\hbar D/4})f = \left(M_z - \alpha\hbar\frac{\partial}{\partial\bar{z}}\right)f, \tag{16}$$

where M_z is the operation of multiplication by z.

Proof. Following the argument in Exercise 2.19 in my Lie groups book, let us expand the exponentials on the left-hand side of (16) in power series. If we then group together all the terms in which the total power of D (on the left or right of M_z) is k, we obtain, after a little simplification

$$(e^{-\hbar D/4}M_z e^{\hbar D/4})f = \sum_{k=0}^{\infty}\left(\frac{\hbar}{4}\right)^k \frac{1}{k!}\sum_{l=0}^{k}\frac{k!}{l!(k-l)!}(-1)^l D^l M_z D^{k-l} f. \tag{17}$$

Note that for any polynomial f, only finitely many terms on either side of (17) are nonzero, so that we may rearrange the terms freely.

If A and B are operators, let $\mathrm{ad}_A(B) = [A,B]$, so that, say, $(\mathrm{ad}_A)^3(B) = [A,[A,[A,B]]]$. It is straightforward to verify by induction on k that the sum over l on the right-hand side of (17) is equal to $(-\mathrm{ad}_D)^k(M_z)$. Thus,

$$\begin{aligned}&(e^{-\hbar D/4}M_z e^{\hbar D/4})f \\ &= \left(\exp\left\{-\frac{\hbar}{4}\mathrm{ad}_D\right\}(M_z)\right)f \\ &= \left(M_z - \frac{\hbar}{4}[D,M_z] + \frac{\hbar^2}{2\cdot 4^2}[D,[D,M_z]] + \cdots\right)f. \end{aligned}\tag{18}$$

Observe, now, that

$$4\alpha\frac{\partial^2}{\partial z\partial\bar{z}} = \alpha\left(\frac{\partial^2}{\partial x^2} + \frac{1}{\alpha^2}\frac{\partial^2}{\partial p^2}\right) = D. \tag{19}$$

After verifying that $\partial/\partial z$ and $\partial/\partial\bar{z}$ behave in the expected way on powers of z and $\bar{z}$, we may easily calculate that

$$\begin{aligned}[D,M_z] &= 4\alpha\frac{\partial}{\partial\bar{z}} \\ [D,[D,M_z]] &= 0.\end{aligned}$$

Thus, (18) reduces to the desired expression for $(e^{-\hbar D/4}M_z e^{\hbar D/4})f$. ∎

Proof. [Proof of Theorem 5] The proof uses a similarity between a product formula for the heat operator and an product rule for the Weyl quantization, with respect to multiplying an arbitrary polynomial by a linear function of x and p. See (21) in the heat equation case and Proposition 13.4 in the text in the Weyl quantization case.

We begin by proving (12), which we do by induction on the degree of f. If the degree of f is 0 or 1, then the Weyl and Wick-ordered quantizations of f are equal, and, correspondingly, $Df = 0$ so that $e^{-\hbar D/4}f = f$. Assume, now, that the desired result holds for all polynomials of degree at most j. We will show that, on polynomials of degree $j+1$, the linear map $f \mapsto Q_{\mathrm{Weyl}}(e^{-\hbar D/4}f)$ satisfies the identity (5), which uniquely characterizes the Wick-ordered quantization.

Now, if f is a polynomial of degree at most j, then

$$e^{-\hbar D/4}(zf) = e^{-\hbar D/4}M_z f = (e^{-\hbar D/4}M_z e^{\hbar D/4})e^{-\hbar D/4}f. \tag{20}$$

In light of Lemma 6, this becomes

$$e^{-\hbar D/4}(zf) = ze^{-\hbar D/4}f - \hbar\alpha\frac{\partial}{\partial\bar{z}}\left(e^{-\hbar D/4}f\right). \tag{21}$$

Meanwhile, if we apply Eqs. (13.7) and (13.9) in Proposition 13.4 in the text with $g = e^{-\hbar D/4}$, we get

$$\begin{aligned} Q_{\mathrm{Weyl}}(ze^{-\hbar D/4}f) &= Q_{\mathrm{Weyl}}(z)Q_{\mathrm{Weyl}}(e^{-\hbar D/4}) \\ &\quad - \frac{i\hbar}{2}Q_{\mathrm{Weyl}}\left(\frac{\partial f}{\partial p}\right) + \frac{\hbar}{2}\alpha Q_{\mathrm{Weyl}}\left(\frac{\partial f}{\partial x}\right). \\ &= Q_{\mathrm{Weyl}}(z)Q_{\mathrm{Weyl}}(e^{-\hbar D/4}) + \hbar\alpha Q_{\mathrm{Weyl}}\left(\frac{\partial f}{\partial\bar{z}}\right). \end{aligned} \tag{22}$$

Therefore, if we apply the Weyl quantization to (21), a fortunate cancellation occurs:

$$\begin{aligned} Q_{\mathrm{Weyl}}(e^{-\hbar D/4}(zf)) &= Q_{\mathrm{Weyl}}(ze^{-\hbar D/4}f) - \hbar\alpha Q_{\mathrm{Weyl}}\left(\frac{\partial}{\partial\bar{z}}(e^{-\hbar D/4}f)\right) \\ &= Q_{\mathrm{Weyl}}(z)Q_{\mathrm{Weyl}}(e^{-\hbar D/4}f). \end{aligned}$$

We may now apply our induction hypothesis, since f—and therefore also $e^{-\hbar D/4}f$—has degree at most j. Thus,

$$\begin{aligned} Q_{\mathrm{Weyl}}(e^{-\hbar D/4}(zf)) &= Q_{\mathrm{Weyl}}(z)Q_{\mathrm{Wick}}(f) \\ &= Q_{\mathrm{Wick}}(zf), \end{aligned}$$

by (5). We have now established (12) for polynomials of degree $j+1$ of the form zf, with $\deg(f) = j$.

If use similar arguments for polynomials of the form $\bar{z}f$, with $\deg(f) = j$, we find that the analog of (21) is

$$e^{-\hbar D/4}(zf) = ze^{-\hbar D/4}f - \hbar\alpha\frac{\partial}{\partial\bar{z}}\left(e^{-\hbar D/4}f\right)$$

and the analog of (22) is, using Eqs. (13.8) and (13.10) in the text,

$$Q_{\mathrm{Weyl}}(\bar{z}e^{-\hbar D/4}f) = Q_{\mathrm{Weyl}}(e^{-\hbar D/4})Q_{\mathrm{Weyl}}(\bar{z}) + \hbar\alpha Q_{\mathrm{Weyl}}\left(\frac{\partial f}{\partial\bar{z}}\right).$$

From this and (6), we can verify that $Q_{\mathrm{Weyl}}(e^{-\hbar D/4}(\bar{z}f)) = Q_{\mathrm{Wick}}(\bar{z}f)$. We leave the proof of (13), which is entirely parallel to the proof of (12), as an exercise for the reader. ■

5. The Weyl quantization on tempered distributions

5.1. Tempered distributions. Let us consider the Schwartz space $\mathcal{S}(\mathbb{R}^k)$ and its topological dual space $\mathcal{S}'(\mathbb{R}^k)$, consisting of linear maps $T : \mathcal{S}(\mathbb{R}^k) \to \mathbb{C}$ that are continuous with respect to the natural topology on $\mathcal{S}(\mathbb{R}^{2n})$. (This topology is the one in which convergence of Schwartz functions χ_m to χ amounts to uniform convergence of $\mathbf{x}^{\mathbf{a}}\partial^{\mathbf{b}}\chi_m$ to $\mathbf{x}^{\mathbf{a}}\partial^{\mathbf{b}}\chi$, for every multi-index $\mathbf{a}$ and $\mathbf{b}$.) We refer to elements of $S'(\mathbb{R}^k)$ as **tempered distributions**. It is not hard to show that $C_c^\infty(\mathbb{R}^k)$ is dense in $\mathcal{S}(\mathbb{R}^k)$. Thus, an ordinary distribution $T \in (C_c^\infty(\mathbb{R}^k))'$ can have at most one extension to an continuous linear functional on $\mathcal{S}(\mathbb{R}^k)$. We say that an ordinary distribution T is **tempered** if such a continuous extension exists. Any locally integrable function with at most polynomial growth at infinity determines a tempered distribution.

Now, it is not hard to show that if f and χ are Schwartz functions, then

$$\int_{\mathbb{R}^k} \mathcal{F}(f)(\mathbf{k})\chi(\mathbf{k})\ d\mathbf{k} = \int_{\mathbb{R}^k} f(\mathbf{x})\tilde{\mathcal{F}}(\chi)(\mathbf{x})\ d\mathbf{x}, \tag{23}$$

where

$$\tilde{\mathcal{F}}(\chi)(\mathbf{x}) = (2\pi)^{-n/2}\int_{\mathbb{R}^n} e^{-i\mathbf{k}\cdot\mathbf{x}}\chi(\mathbf{k})\ d\mathbf{k}.$$

(Apply the Plancherel theorem, to $\bar{f}$ and χ and keep track of the conjugates.) Although $\tilde{\mathcal{F}}$ is technically the same as $\mathcal{F}$, it is better to use a different symbol, because the roles of the variables are reversed from what one would expect for the forward Fourier transform $\mathcal{F}$. One should, rather, think of $(\tilde{\mathcal{F}}\chi)(\mathbf{x})$ as being the *inverse* Fourier transform of χ evaluated at $-\mathbf{x}$. It is then easy to see that

$$\mathcal{F}\tilde{\mathcal{F}}(\chi)(\mathbf{k}) = \chi(-\mathbf{k}). \tag{24}$$

With (23) as motivation, we may make the following definition of the Fourier transform of a tempered distribution.

Definition 7. *If T is a tempered distribution on $\mathbb{R}^k$, we define the* ***Fourier transform*** *$\mathcal{F}(T)$ of T to be the tempered distribution given by*

$$\mathcal{F}(T)(\chi) = T(\tilde{\mathcal{F}}(\chi))$$

for all $\chi \in \mathcal{S}(\mathbb{R}^k)$. Similarly, if T is a tempered distribution on $\mathbb{R}^{k+l}$, we may define the ***partial Fourier transform*** *$\mathcal{F}_{\mathbf{y}}(T)$ of T by*

$$\mathcal{F}_{\mathbf{y}}(T)(\chi) = T(\tilde{\mathcal{F}}_{\mathbf{y}}(\chi)),$$

where $\mathcal{F}_{\mathbf{y}}$ denotes the map in (23) applied to $\chi(\mathbf{x},\mathbf{y})$ with respect to $\mathbf{y} \in \mathbb{R}^l$ with $\mathbf{x} \in \mathbb{R}^k$ fixed.

Note that we cannot replace tempered distributions by ordinary distributions in Definition 7, because the Fourier transform of a nonzero function $\chi \in C_c^\infty(\mathbb{R}^k)$ does not belong to $C_c^\infty(\mathbb{R}^k)$.

Example 8. *For any $\mathbf{a} \in \mathbb{R}^n$, we have*

$$\mathcal{F}(e^{i\mathbf{a}\cdot\mathbf{x}})(\mathbf{k}) = (2\pi)^{n/2}\delta_n(\mathbf{k}-\mathbf{a}), \tag{25}$$

where δ_n is an n-dimensional delta-function.

Equation (25) means, more explicitly, that the Fourier transform of $e^{i\mathbf{a}\cdot\mathbf{x}}$ is the distribution T such that $T(\chi) = (2\pi)^{n/2}\chi(\mathbf{a})$. We can easily obtain this result at a heuristic level by formally applying the inverse Fourier transform to the right-hand side of (25).

Proof. Applying the definition, we have

$$\begin{aligned}\mathcal{F}(e^{i\mathbf{a}\cdot\mathbf{x}})(\chi) &= \int_{\mathbb{R}^n} e^{i\mathbf{a}\cdot\mathbf{x}}\tilde{\mathcal{F}}(\chi)\ d\mathbf{x} \\ &= (2\pi)^{n/2}(2\pi)^{-n/2}\int_{\mathbb{R}^n} e^{-i(-\mathbf{a})\cdot\mathbf{x}}\tilde{\mathcal{F}}(\chi)\ d\mathbf{x} \\ &= (2\pi)^{n/2}\left(\mathcal{F}\tilde{\mathcal{F}}\chi\right)(-\mathbf{a}) \\ &= (2\pi)^{n/2}\chi(\mathbf{a}),\end{aligned}$$

where in the last line we have used (24). ∎

If T and χ are Schwartz functions and $A : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is an invertible linear map, then the change-of-variables formula tells us that

$$\int_{\mathbb{R}^{2n}} T(A\mathbf{x})\chi(\mathbf{x})\ d\mathbf{x} = \frac{1}{\det A}\int_{\mathbb{R}^{2n}} T(\mathbf{u})\chi(A^{-1}\mathbf{u})\ d\mathbf{u}.$$

This observation motivates the following definition.

Definition 9. *If T is a tempered distribution and A is an invertible linear transformation, we define $T \circ A$ to be the tempered distribution given by*

$$(T \circ A)(\chi) = \frac{1}{\det A}T(\chi \circ A^{-1}).$$

5.2. The Weyl quantization on tempered distributions. In this section, we observe that the Weyl quantization, which at the moment is defined for symbols that are either polynomials or square-integrable functions on $\mathbb{R}^{2n}$, can be extended to symbols that are tempered distributions. The price we pay for this extension is that the Weyl quantization of a general tempered distribution will map a very small space ($\mathcal{S}(\mathbb{R}^n)$) to a very large space ($\mathcal{S}'(\mathbb{R}^n)$). For an arbitrary tempered distribution f, the space of those $\psi \in \mathcal{S}(\mathbb{R}^n)$ for which $Q_{\mathrm{Weyl}}(f)$ belongs to $L^2(\mathbb{R}^n)$ may not be a dense subspace of $L^2(\mathbb{R}^n)$. On the other hand, if f is a polynomial in $\mathbf{x}$ and $\mathbf{p}$, then $Q_{\mathrm{Weyl}}(f)$ is a differential operator with polynomial coefficients, in which case, $Q_{\mathrm{Weyl}}(f)$ maps $\mathcal{S}(\mathbb{R}^n)$ into $L^2(\mathbb{R}^n)$, meaning that $Q_{\mathrm{Weyl}}(f)$ is at least densely defined as an operator on $L^2(\mathbb{R}^n)$.

Definitions 7 and 9 show that the operations going into the map $f \mapsto \kappa_f$, namely a partial Fourier transform and composition with an invertible linear transformation, make sense when f is a tempered distribution, with the result that κ_f is also a tempered distribution. This observation allows us to define $Q_{\mathrm{Weyl}}(f)$ for $f \in \mathcal{S}'(\mathbb{R}^n)$.

Definition 10. *Given $f \in \mathcal{S}'(\mathbb{R}^{2n})$ we define*

$$Q_{\mathrm{Weyl}}(f) : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$$

as follows. Given $\psi \in \mathcal{S}(\mathbb{R}^n)$, we define $Q_{\mathrm{Weyl}}(f)\psi$ to be the tempered distribution given by

$$[Q_{\mathrm{Weyl}}(f)\psi](\chi) = \kappa_f(\chi\psi),$$

where κ_f is given by Definition 13.7 in the text.

Now that the Weyl quantization is defined for symbols in $\mathcal{S}'(\mathbb{R}^{2n})$, which certainly contains both polynomials and functions of the form $\exp(i(\mathbf{a}\cdot\mathbf{X}+\mathbf{b}\cdot\mathbf{P}))$, the formal calculations that we did in Section 13.3.1 in the text can now be made rigorous, and we can verify that Q_{Weyl} behaves as expected on such functions.

Theorem 11. *For all* $\mathbf{a}$ *and* $\mathbf{b}$ *in* $\mathbb{R}^n$*, we have*

$$Q_{\mathrm{Weyl}}(e^{i(\mathbf{a}\cdot\mathbf{x}+\mathbf{b}\cdot\mathbf{p})}) = e^{i(\mathbf{a}\cdot\mathbf{X}+\mathbf{b}\cdot\mathbf{P})}, \tag{26}$$

where the operator $e^{i(\mathbf{a}\cdot\mathbf{X}+\mathbf{b}\cdot\mathbf{P})}$ *is as in Proposition 13.5 in the text. For all* $\mathbf{a}$ *and* $\mathbf{b}$ *in* $\mathbb{R}^n$*, we have*

$$Q_{\mathrm{Weyl}}((\mathbf{a}\cdot\mathbf{x}+\mathbf{b}\cdot\mathbf{p})^j) = (\mathbf{a}\cdot\mathbf{X}+\mathbf{b}\cdot\mathbf{P})^j, \tag{27}$$

and this relation uniquely determines the Weyl quantization on symbols that are polynomials in $\mathbf{x}$ *and* $\mathbf{p}$*. In particular, when* $n=1$*, the Weyl quantization of a polynomial* f *in* x *and* p *as defined by Definition 10 agrees with the expression for* $Q_{\mathrm{Weyl}}(f)$ *in Definition 13.1 in the text.*

Proof. For any $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^n$, we have, by Example 8 with $\mathbf{x}$ replaced by $\mathbf{p}$,

$$\mathcal{F}_{\mathbf{p}}\left(e^{i(\mathbf{a}\cdot\mathbf{x}+\mathbf{b}\cdot\mathbf{p})}\right) = (2\pi)^{n/2}e^{i\mathbf{a}\cdot\mathbf{x}}\delta_n(\boldsymbol{\xi}-\mathbf{b}),$$

where δ_n is an n-dimensional delta-function. Thus, the kernel associated to $f(\mathbf{x},\mathbf{y}) = e^{i(\mathbf{a}\cdot\mathbf{x}+\mathbf{b}\cdot\mathbf{p})}$ is

$$\begin{aligned}\kappa_f(\mathbf{x},\mathbf{y}) &= \hbar^{-n}e^{i\mathbf{a}\cdot(\mathbf{x}+\mathbf{y})/2}\delta_n((\mathbf{y}-\mathbf{x}-\hbar\mathbf{b})/\hbar)\\ &= e^{i\mathbf{a}\cdot\mathbf{x}/2}\delta_n(\mathbf{y}-(\mathbf{x}+\hbar\mathbf{b}))e^{i\mathbf{a}\cdot\mathbf{y}/2}.\end{aligned}$$

It follows that

$$\begin{aligned}\left[Q_{\mathrm{Weyl}}\left(e^{i(\mathbf{a}\cdot\mathbf{x}+\mathbf{b}\cdot\mathbf{p})}\right)\psi\right](\mathbf{x}) &= e^{i\mathbf{a}\cdot\mathbf{x}/2}\psi(\mathbf{x}+\hbar\mathbf{b})e^{i\mathbf{a}\cdot(\mathbf{x}+\hbar\mathbf{b})/2}\\ &= e^{i\mathbf{a}\cdot\mathbf{x}}\psi(\mathbf{x}+\hbar\mathbf{b})e^{i\hbar\mathbf{a}\cdot\mathbf{b}/2}.\end{aligned}$$

In light of Proposition 13.5 in the text, (26) holds, as expected.

Suppose now that f_m is a sequence of elements of $\mathcal{S}'(\mathbb{R}^{2n})$ with the property that for all $\chi\in\mathcal{S}(\mathbb{R}^{2n})$ we have

$$f_m(\chi)\to f(\chi),$$

for some fixed $f\in\mathcal{S}'(\mathbb{R}^{2n})$. It is not hard to verify that for all ψ and χ in $\mathcal{S}(\mathbb{R}^n)$, we have

$$[Q_{\mathrm{Weyl}}(f_m)\psi](\chi)\to[Q_{\mathrm{Weyl}}(f)\psi](\chi).$$

We may apply this result to the case where

$$f_m(\mathbf{x},\mathbf{p}) = \frac{e^{it_m(\mathbf{a}\cdot\mathbf{x}+\mathbf{b}\cdot\mathbf{p})}-e^{it(\mathbf{a}\cdot\mathbf{x}+\mathbf{b}\cdot\mathbf{p})}}{t_m-t},$$

where t_m is a sequence of real numbers converging to t. The upshot of this is that we can interchange the Weyl quantization (as applied to two fixed test functions ψ and χ) with differentiation with respect to the parameter t in $e^{it(\mathbf{a}\cdot\mathbf{x}+\mathbf{b}\cdot\mathbf{p})}$. In particular, since

$$(\mathbf{a}\cdot\mathbf{x}+\mathbf{b}\cdot\mathbf{p})^j = (-i)^j\left(\frac{d}{dt}\right)^j e^{it(\mathbf{a}\cdot\mathbf{x}+\mathbf{b}\cdot\mathbf{p})}\bigg|_{t=0},$$

we conclude that for each $\psi, \chi \in \mathcal{S}(\mathbb{R}^n)$, we have

$$\left[Q_{\mathrm{Weyl}}((\mathbf{a}\cdot\mathbf{x}+\mathbf{b}\cdot\mathbf{p})^j)\psi\right](\chi) = (-i)^h \left(\frac{d}{dt}\right)^j \left[e^{it(\mathbf{a}\cdot\mathbf{X}+\mathbf{b}\cdot\mathbf{P})}\psi\right](\chi)\Big|_{t=0}. \tag{28}$$

Now, the domain of the infinitesimal generator of the one-parameter unitary group $e^{it(\mathbf{a}\cdot\mathbf{X}+\mathbf{b}\cdot\mathbf{P})}$ contains $\mathcal{S}(\mathbb{R}^n)$ and $\mathbf{a}\cdot\mathbf{X}+\mathbf{b}\cdot\mathbf{P}$ preserves $\mathcal{S}(\mathbb{R}^n)$. Thus, by Lemma 10.17 in the text, the derivative $e^{it(\mathbf{a}\cdot\mathbf{X}+\mathbf{b}\cdot\mathbf{P})}\psi$ with respect to t exists, with the limit in the L^2 norm, and may be computed by pulling down $i(\mathbf{a}\cdot\mathbf{X}+\mathbf{b}\cdot\mathbf{P})$ on the right-hand side of the exponential. Since $\chi \in L^2(\mathbb{R}^n)$ and the inner product on L^2 is continuous, the obvious formal way of computing the right-hand side of (28) is correct, and we obtain that

$$[Q_{\mathrm{Weyl}}((\mathbf{a}\cdot\mathbf{x}+\mathbf{b}\cdot\mathbf{p})^j)\psi](\chi) = [(\mathbf{a}\cdot\mathbf{X}+\mathbf{b}\cdot\mathbf{P})^j\psi](\chi), \tag{29}$$

as desired.

Now, by the n-dimensional generalization of Exercise 1 in Chapter 13, every homogeneous polynomial of degree j on $\mathbb{R}^{2n}$ can be expressed as a linear combination of polynomials of the form $(\mathbf{a}\cdot\mathbf{x}+\mathbf{b}\cdot\mathbf{p})^j$, so that (29) uniquely determines the Weyl quantization on polynomials. In particular, in the $n = 1$ case, we have agreement with Definition 13.1 in the text, by Proposition 13.3. ■

5.3. The Wick and anti-Wick quantizations. One can define Wick- and anti-Wick-ordered quantizations for $\mathbb{R}^{2n}$ by analogy to the $\mathbb{R}^2$ case. On polynomial symbols, these quantizations are related to the Weyl quantization by

$$Q_{\mathrm{Wick}}(f) = Q_{\mathrm{Weyl}}(e^{-\hbar D/4}f) \tag{30}$$

$$Q_{\mathrm{anti-Wick}}(f) = Q_{\mathrm{Weyl}}(e^{\hbar D/4}f), \tag{31}$$

where D is the obvious generalization to $\mathbb{R}^{2n}$ of the operator in Theorem 5. Now that we have defined the Weyl quantization on a very large class of symbols, it is natural to take (30) and (31) as the definitions of the Wick and anti-Wick quantizations on some larger class of symbols.

The operator $e^{\hbar D/4}$ is a forward heat operator and it makes sense on tempered distributions, where it may be defined by means of the Fourier transform. (If f is a polynomial, the Fourier transform of $e^{\hbar D/4}f$ is equal to the Fourier transform of f multiplied by a Gaussian. We may take this result as a definition of $e^{\hbar D/4}f$ for general $f \in \mathcal{S}'(\mathbb{R}^{2n})$.) Thus, the anti-Wick quantization naturally extends from polynomial symbols to arbitrary symbols in $\mathcal{S}'(\mathbb{R}^{2n})$. Indeed, because of the smoothing nature of the heat equation, $Q_{\mathrm{anti-Wick}}(f)$ is likely to be a "nicer" operator than $Q_{\mathrm{Weyl}}(f)$.

The operator $e^{-\hbar D/4}$, by contrast, is the *backward* heat operator. Although this operator can be defined on polynomials as a terminating power series, it is in general very singular. If f is Schwartz function, for example, there usually does not exist F such that $e^{\hbar D/4}F = f$, even if we allow F to be a tempered distribution. Thus, the Wick quantization is much more singular than the anti-Wick quantization, and the Wick quantization can only be defined on very special classes of symbols, such as the class of $f \in L^2(\mathbb{R}^{2n})$ for which the Fourier transform of f has compact support.

The idea that the Wick-ordered quantization is more singular than the anti-Wick-ordered quantization may be surprising to those who have encountered Wick-ordering in quantum field theory. In quantum field theory, after all, "Wick ordering" (that is, applying Wick-ordered quantization instead of Weyl quantization) frequently has a desingularizing effect, allowing the Wick-ordered version of certain operators to be well defined when the Weyl-quantized operator is singular. The resolution of this paradox is that in quantum field theory, divergences typically arise because there are infinitely many degrees of freedom. The functions one wishes to quantize in field theory are often "polynomial" in nature, with a typical example being $\phi(x)^4$, where ϕ is a classical field. Because $\phi(x)^4$ is simply the fourth power of a linear functional in ϕ, there is no difficulty in applying a backward heat operator. Although the Weyl quantization of $\phi(x)^4$ is definitely singular, because it gives a divergent result when applied to the "ground state" of the system, the Wick-ordered quantization of $\phi(x)^4$ puts all the annihilation operators acting first and is, therefore, zero on the ground state. Although there may still be singularities in the operator on excited states, Wick ordering definitely improves the situation.

6. Construction of basic line bundle over S^2

In this section, we construct the "basic" Hermitian line bundle L with connection over the unit sphere $S^2 \subset \mathbb{R}^3$, that is, the one where the integer in Eq. (23.8) has the value 1 when S is taken to be S^2 itself (with the standard orientation). All other line bundles on S^2 can be constructed either by similar methods—as a sub-bundle of the trivial vector bundle $S^2 \times V$, where V is an irreducible representation of $\mathsf{so}(3)$—or by taking tensor powers of L and the dual bundle to L.

Consider the Pauli spin matrices σ_1, σ_2, and σ_3, which are the self-adjoint matrices given by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(The matrices $i\sigma_j$, $j = 1, 2, 3$, form a basis for the Lie algebra $\mathsf{su}(2)$ of the group $\mathsf{SU}(2)$.) These satisfy

$$[\sigma_1, \sigma_2] = 2i\sigma_3$$

and identities obtained from this by cyclic permutation of the indices, as well as

$$\begin{aligned} \sigma_j^2 &= I \\ \sigma_j\sigma_k + \sigma_j\sigma_k &= 0, \quad j \neq k \\ \operatorname{trace}(\sigma_j) &= 0. \end{aligned}$$

Given any vector $\mathbf{x} \in \mathbb{R}^3$, we can form the matrix $\boldsymbol{\sigma} \cdot \mathbf{x} := x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3$. If $\mathbf{x}$ is a unit vector, then

$$(\boldsymbol{\sigma} \cdot \mathbf{x})^2 = |\mathbf{x}|^2 I = I.$$

Since also $\operatorname{trace}(\boldsymbol{\sigma} \cdot \mathbf{x}) = 0$, we conclude that the eigenvalues of $\boldsymbol{\sigma} \cdot \mathbf{x}$ are 1 and -1. Thus, the operator

$$\rho_{\mathbf{x}} := (\boldsymbol{\sigma} \cdot \mathbf{x} + I)/2 \tag{32}$$

has eigenvalues 1 and 0 and is self-adjoint, meaning that $\rho_{\mathbf{x}}$ is an orthogonal projection operator onto a 1-dimensional subspace of $\mathbb{C}^2$, for each $\mathbf{x} \in S^2$.

Definition 12. *Let L be the complex line bundle over S^2 where the fiber of L over the point $\mathbf{x} \in S^2$ is the range of the projection $\rho_{\mathbf{x}}$. We define a Hermitian structure on L by defining $(s_1(x), s_2(x))$ to be the ordinary inner product of $s_1(x)$ and $s_2(x)$ in $\mathbb{C}^2$. We then define a Hermitian connection ∇ on L by defining*

$$(\nabla_X s)(\mathbf{x}) = \rho_{\mathbf{x}}(Xs)$$

for every smooth vector field X on S^2 and every smooth section s of L. Here Xs denotes the ordinary derivative of s as a $\mathbb{C}^2$-valued function on S^2.

It is straightforward to check that L is actually a line bundle over S^2 and that ∇ is a Hermitian connection on L.

Proposition 13. *The curvature 2-form ω of the connection ∇ is given by*

$$\omega(X,Y)_{\mathbf{x}} = \frac{1}{2}((X \times Y)\cdot \mathbf{x}),$$

for all $\mathbf{x} \in S^2$ and all $X, Y \in T_{\mathbf{x}}S^2$. It follows that ω is one half of the standard area form on S^2 and that

$$\int_{S^2} \omega = 2\pi,$$

where we use the standard orientation on S^2 in computing the integral.

Proof. Applying the definition of the curvature gives

$$\begin{aligned}\omega(X,Y)s &= -i\left(\nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]}\right)s \\ &= -i\left(\rho_{\mathbf{x}}X\rho_{\mathbf{x}}Y - \rho_{\mathbf{x}}Y\rho_{\mathbf{x}}X - \rho_{\mathbf{x}}(XY - YX)\right)s \\ &= -i(\rho_{\mathbf{x}}X(\rho_{\mathbf{x}})Y - \rho_{\mathbf{x}}Y(\rho_{\mathbf{x}})X)s.\end{aligned}$$

Furthermore, from the definition (32) of $\rho_{\mathbf{x}}$, we compute that

$$X(\rho_{\mathbf{x}}) = \frac{1}{2}\boldsymbol{\sigma}\cdot X.$$

Meanwhile, given any fixed vector $v \in \mathbb{C}^2$, we can define a section s of L by defining $s(\mathbf{x}) = \rho_{\mathbf{x}}v$, so that $Xs = (X\rho_{\mathbf{x}})v$. Given any point $\mathbf{x}_0 \in S^2$, we can choose v to lie in $L_{\mathbf{x}_0}$, in which case, s will be nonzero in a neighborhood of $\mathbf{x}_0$. Applying ω to gives

$$\omega(X,Y)s = -\frac{i}{4}\rho_{\mathbf{x}}\left[(\boldsymbol{\sigma}\cdot X)(\boldsymbol{\sigma}\cdot Y) - (\boldsymbol{\sigma}\cdot Y)(\boldsymbol{\sigma}\cdot X)\right]v.$$

Now, calculation with the commutation relations for the σ_j's shows that $[(\boldsymbol{\sigma}\cdot X), (\boldsymbol{\sigma}\cdot Y)] = 2i(X \times Y)\cdot\boldsymbol{\sigma}$. Furthermore, since X and Y are tangent to S^2, their values at $\mathbf{x}$ are orthogonal to $\mathbf{x}$. Hence, $X \times Y$ is a multiple of $\mathbf{x}$, specifically, $X \times Y = ((X \times Y)\cdot\mathbf{x})\mathbf{x}$, and so

$$(\boldsymbol{\sigma}\cdot X)(\boldsymbol{\sigma}\cdot Y) - (\boldsymbol{\sigma}\cdot Y)(\boldsymbol{\sigma}\cdot X) = 2i((X \times Y)\cdot\mathbf{x})(\boldsymbol{\sigma}\cdot\mathbf{x}).$$

Since $\boldsymbol{\sigma}\cdot\mathbf{x}$ commutes with $\rho_{\mathbf{x}}$, we obtain

$$\begin{aligned}\omega(X,Y)s &= \frac{1}{2}((X \times Y)\cdot\mathbf{x})(\mathbf{x}\cdot\boldsymbol{\sigma})\rho_{\mathbf{x}}v \\ &= \frac{1}{2}((X \times Y)\cdot\mathbf{x})s.\end{aligned}$$

In the last step we have used that $\rho_{\mathbf{x}}v$ is an eigenvector for $\rho_{\mathbf{x}}$ with eigenvalue 1, and thus also an eigenvector for $\boldsymbol{\sigma}\cdot\mathbf{x}$ with eigenvalue 1. ■

7. Momentum maps and symplectic group actions

7.1. Momentum maps. In this section, we introduce a general notion of "momentum" that generalizes both ordinary (linear) momentum and angular momentum. The notion of a momentum map is an important tool in understanding the role of symmetry in symplectic geometry. In particular, we will give a general reason that the components of the angular momentum satisfy the same commutation relations (under the Poisson bracket) as the Lie algebra $\mathsf{so}(3)$. (Compare Exercise 19 in Chapter 2 to the commutation relations (16.3) for the matrices F_j in (16.2).) Momentum maps will also give a method of generating constants of motion in the case that the Hamiltonian on N is invariant under a group action that preserves the symplectic structure of M. (Compare the second part of Conclusion 2.31 in the case of the rotation group $\mathsf{SO}(2)$.)

Once the classical theory of momentum maps has been understood, one can use the methods of geometric quantization (Chapter 23 in the text) to construct quantum counterparts to these functions. The operators obtained by quantizing the components of the momentum map serve as analogs of the linear or angular momentum operators in quantum mechanics on Euclidean space.

Definition 14. *If G is a Lie group and M is a smooth manifold, then a smooth* ***action*** *of G on M is a smooth map of $G\times M$ into M, denoted $(g,x)\mapsto g\cdot x$, such that for all $g,h\in G$ and $x\in M$ we have*

$$g\cdot(h\cdot x)=(gh)\cdot x$$

and

$$e\cdot x=x,$$

where e is the identity element of G.

If N is a symplectic manifold, a smooth action of a Lie group G on N is ***symplectic*** *if for each fixed $g\in G$, the map*

$$x\mapsto g\cdot x$$

is a symplectomorphism of N to itself.

Since the letter X is currently being used to represent a vector field on N, we will, in this section, use Greek letters such as ξ and η to represent elements of the Lie algebra $\mathfrak{g}$ of G. Given any $\xi\in\mathfrak{g}$, the map Φ_ξ given by

$$\Phi_{\xi,t}(x)=e^{t\xi}\cdot x \tag{33}$$

is a flow on M. We will let $-X^\xi$ denote the vector field generating this flow.

Proposition 15. *Given a smooth action of a group G on a manifold M the vector fields X^ξ generating the flows Φ_ξ in (33) satisfy the relations*

$$X^{c\xi+\eta}=cX^\xi+X^\eta \tag{34}$$

and

$$X^{[\xi,\eta]}=[X^\xi,X^\eta] \tag{35}$$

for all $\xi,\eta\in\mathfrak{g}$ and $c\in\mathbb{R}$.

Proof. As with any flow, the vector field X^ξ generating Φ_ξ may be computed by differentiation:

$$(X^\xi)_x = \left.\frac{d}{dt}e^{t\xi}\cdot x\right|_{t=0}.$$

Thus, for any fixed x, the quantity $(X^\xi)_x$ is the differential at the origin of the map $\eta \mapsto e^\eta \cdot x$, applied to the vector ξ. Since the differential of a smooth map is linear, we have (34).

In light of Point 3 of Theorem 16.15 and the definition of a group action, we have

$$\Phi_{e^{s\xi}\eta e^{-s\xi},t} = \Phi_{\xi,s}\circ\Phi_{\eta,t}\circ\Phi_{\xi,-s}.$$

Now,

$$[\xi,\eta] = \left.\frac{d}{ds}e^{s\xi}\eta e^{-s\xi}\right|_{s=0}.$$

For a smooth function f and a flow Φ, let $\Phi_t \cdot f = f\circ\Phi_t$. Then the generator X of Φ may be computed as

$$Xf = \left.\frac{d}{dt}\Phi_t f\right|_{t=0}.$$

Note that $(\Phi_t\circ\Psi_t)\cdot f = \Psi_t\cdot\Phi_t\cdot f$. Thus, recalling that Φ_ξ is generated by the vector field $-X^\xi$, we have

$$\begin{aligned} X^{[\xi,\eta]}f &= \left.\frac{d}{ds}\left(\left.\frac{d}{dt}\Phi_{e^{s\xi}\eta e^{-s\xi},t}f\right|_{t=0}\right)\right|_{s=0} \\ &= \left.\frac{d^2}{ds\,dt}(\Phi_{\xi,s}\Phi_{\eta,t}\Phi_{\xi,-s})f\right|_{s=t=0} \\ &= \left.\frac{d^2}{ds\,dt}\Phi_{\xi,-s}\cdot\Phi_{\eta,t}\cdot\Phi_{\xi,s}\cdot f\right|_{s=t=0} \\ &= (X^\xi(-X^\eta)+(-X^\eta)(-X^\xi))f, \end{aligned}$$

which is what we are trying to prove. ■

Definition 16. *A smooth, symplectic action of a Lie group G on a symplectic manifold N is said to* ***admit a momentum map*** *if for each $\xi\in\mathfrak{g}$, the vector field X^ξ is globally Hamiltonian. A* ***momentum map*** *for such an action is a choice ϕ_ξ of Hamiltonian generator for each X^ξ with the property that the map $\xi\mapsto\phi_\xi$ is linear.*

If each X^ξ is globally Hamiltonian, it is possible to construct a momentum map, as follows. Choose a basis $\xi_1,\ldots,\xi_N$ for $\mathfrak{g}$ and choose a Hamiltonian generator ϕ_{ξ_j} for each X^{ξ_j}. Then extend the map $\xi\mapsto\phi_\xi$ to all of $\mathfrak{g}$ by linearity. In light of the linearity in (34), ϕ_ξ will be the Hamiltonian generator for X^ξ, for all $\xi\in\mathfrak{g}$. Note that if $\{\phi_\xi\}$ is a momentum map for a symplectic action of G, then so is $\phi_\xi + c_\xi$ for any family of constants c_ξ depending linearly on ξ.

In symplectic geometry, momentum maps are often referred to as "moment maps." Nevertheless, the terminology "momentum maps" is conceptually preferable, since the prototypical examples of momentum maps are, as we shall see, provided by the ordinary and angular momentum of a particle moving in $\mathbb{R}^n$.

Note that for each $z \in N$, the quantity $\phi_\xi(z)$ is linear in ξ. If we let $\phi_\cdot(z)$ denote the linear functional $\xi \mapsto \phi_\xi(z)$, then $\phi_\cdot(z)$ is an element of the dual space $\mathfrak{g}^*$. Thus, we may think of ϕ as defining a map of N into $\mathfrak{g}^*$, i.e., the map $z \mapsto \phi_\cdot(z)$. This is the way that momentum maps are often defined in symplectic geometry.

Proposition 17. *Assume N is connected. Then if ϕ is a momentum map for a symplectic action of G on N, we have, for each $\xi, \eta \in \mathfrak{g}$,*

$$\{\phi_\xi, \phi_\eta\} = \phi_{[\xi,\eta]} + d_{\xi,\eta}$$

for some constant $d_{\xi,\eta}$.

Proof. By definition, the Hamiltonian vector fields associated to ϕ_ξ and ϕ_η are X^ξ and X^η, respectively. Thus, general properties Hamiltonian vector fields, we have

$$X_{\{\phi_\xi,\phi_\eta\}} = [X^\xi, X^\eta] = X^{[\xi,\eta]} = X_{\phi_{[\xi,\eta]}}.$$

Thus, $\{\phi_\xi, \phi_\eta\}$ and $\phi_{[\xi,\eta]}$ generate the same Hamiltonian vector field. It follows that $d(\{\phi_\xi, \phi_\eta\}) = d\phi_{[\xi,\eta]}$ which, since N is connected, means that $\{\phi_\xi, \phi_\eta\}$ and $\phi_{[\xi,\eta]}$ differ by a constant. ■

As we have already remarked, a momentum map is unique only up to adding a family of constants c_ξ depending linearly in ξ. It is reasonable to attempt to choose the constants c_ξ in such a way that the constants $d_{\xi,\eta}$ in Proposition 17 are zero.

Definition 18. *A momentum map ϕ for a symplectic action of G on N is said to be **equivariant** if*

$$\{\phi_\xi, \phi_\eta\} = \phi_{[\xi,\eta]}$$

for all $\xi, \eta \in \mathfrak{g}$.

It is not always possible to choose the constants in the construction of a momentum map so as to make the momentum map equivariant. It is possible for there to exist multiple different equivariant momentum maps (differing by constant, for each ξ) for the same symplectic action. If we view the momentum map as a map of N into $\mathfrak{g}^*$, then it can be shown that ϕ is equivariant if and only if ϕ intertwines the action of G on N with the coadjoint action of G on $\mathfrak{g}^*$. (The coadjoint action of G on $\mathfrak{g}$ is given by $g \cdot \psi = (\mathrm{Ad}_{g^{-1}})^{\mathrm{tr}}\psi$ for all $\psi \in \mathfrak{g}^*$.)

Example 19. *Let $N = \mathbb{R}^{2n}$ with the standard symplectic form $\omega = dp_j \wedge dx_j$. Let $G = \mathbb{R}^n$, acting on $\mathbb{R}^{2n}$ by*

$$\mathbf{a} \cdot (\mathbf{x}, \mathbf{p}) = (\mathbf{x} + \mathbf{a}, \mathbf{p}).$$

This action admits an equivariant momentum map given by

$$\phi_{\boldsymbol{\xi}}(\mathbf{x}, \mathbf{p}) = \boldsymbol{\xi} \cdot \mathbf{p}$$

for each $\boldsymbol{\xi}$ in the (commutative) Lie algebra of $\mathbb{R}^n$.

Proof. We have already shown, in Proposition 2.29 in the text that the Hamiltonian flow Φ generated by $\boldsymbol{\xi} \cdot \mathbf{p}$ is given by $\Phi_t(\mathbf{x}, \mathbf{p}) = (\mathbf{x} + t\boldsymbol{\xi}, \mathbf{p})$. This is precisely the property required of the momentum map for the given action of $\mathbb{R}^n$ on $\mathbb{R}^{2n}$. Since G is commutative, ϕ will be equivariant if $\{\phi_\xi, \phi_\eta\} = 0$ for all ξ and η, which is true. ■

Note that in Example 19, we could have replaced $\boldsymbol{\xi} \cdot \mathbf{p}$ by $\boldsymbol{\xi} \cdot \mathbf{p} + c_\xi$, for any family c_ξ of constants depending linearly on ξ, and obtained another *equivariant* momentum map, since the Poisson bracket of any two components of the momentum map would still be zero.

Example 20. *Let $N = \mathbb{R}^6$ with the standard symplectic form $\omega = dp_j \wedge dx_j$. Let $G = \mathsf{SO}(3)$, acting on $\mathbb{R}^6$ by*

$$R \cdot (\mathbf{x}, \mathbf{p}) = (R\mathbf{x}, R\mathbf{p}).$$

If we identify the Lie algebra $\mathsf{so}(3)$ of $\mathsf{SO}(3)$ with $\mathbb{R}^3$ by means of the basis $\{F_j\}$ in Eq (16.2) in the text, this action admits an equivariant momentum map given by

$$\phi_{\boldsymbol{\xi}}(\mathbf{x}, \mathbf{p}) = \boldsymbol{\xi} \cdot \mathbf{J},$$

where $\mathbf{J} = \mathbf{x} \times \mathbf{p}$ is the usual angular momentum vector.

Proof. See Exercise 1. ■

Example 21. *Let $N = \mathbb{R}^{2n}$ with the standard symplectic form $\omega = dp_j \wedge dx_j$. Let $G = \mathbb{R}^{2n}$, acting on $\mathbb{R}^{2n}$ by*

$$(\mathbf{a}, \mathbf{b}) \cdot (\mathbf{x}, \mathbf{p}) = (\mathbf{x} + \mathbf{a}, \mathbf{p} + \mathbf{b}).$$

This action admits a momentum map given by

$$\phi_{(\boldsymbol{\xi}, \boldsymbol{\eta})}(\mathbf{x}, \mathbf{p}) = \boldsymbol{\xi} \cdot \mathbf{p} - \boldsymbol{\eta} \cdot \mathbf{x}.$$

This momentum map is, however, not *equivariant; indeed, the action does not admit an equivariant momentum map.*

Proof. The claimed form of the momentum map follows easily from Proposition 2.29. To see that there is no *equivariant* momentum map, note that G is commutative, so all of the brackets in the Lie algebra are zero. Now, any momentum map for the action of $\mathbb{R}^{2n}$ must differ from the one in the proposition by a family of constants. But

$$\{x_j + c, p_j + d\} = 1 \neq 0.$$

Thus, no matter how the constants are chosen, the different components of the momentum map will not all Poisson commute. ■

Example 22. *Let $N \subset \mathbb{R}^3$ be the sphere of radius r, with symplectic form ω given by*

$$\omega = x_1 \; dx_2 \wedge dx_3 + x_2 \; dx_3 \wedge dx_1 + x_3 \; dx_1 \wedge dx_2.$$

Let $\mathsf{SO}(3)$ act on N by rotations. There exists an equivariant momentum map ϕ given by

$$\phi_\xi(\mathbf{x}) = -r^2 \boldsymbol{\xi} \cdot \mathbf{x}, \quad \mathbf{x} \in N.$$

Here we identify $\mathsf{so}(3)$ with $\mathbb{R}^3$ using the basis $\{F_j\}$ in Eq. (16.2) in the text.

The form ω is a constant multiple of the area form on the sphere, and is invariant under rotations. The proof of this result is left as an exercise to the reader.

A Lie algebra $\mathfrak{g}$ is said to be **simple** if $\mathfrak{g}$ has no nontrivial ideals *and* $\dim \mathfrak{g} \geq 2$. (That is to say, a one-dimensional Lie algebra is, by decree, not "simple," even though it has no notrivial ideals.) A Lie algebra is said to be **semisimple** if it is a direct sum of simple algebras, and a Lie group G is said to be **semisimple** if the Lie algebra $\mathfrak{g}$ of G is semisimple. Many of the familiar examples of (matrix) Lie groups are semisimple, including $\mathsf{SO}(n)$ for $n \geq 3$, $\mathsf{SU}(n)$ for $n \geq 2$, and $\mathsf{SL}(n;\mathbb{R})$ for $n \geq 2$.

Theorem 23. *Suppose a semisimple Lie group G acts symplectically on N. Then the action admits a momentum map. Furthermore, there is a unique equivariant momentum map.*

Proof. If $\mathfrak{g}$ is the (semisimple) Lie algebra of G, it is not hard to show (Exercise 2) that every element ξ of $\mathfrak{g}$ can be expressed as

$$\xi = \sum_j [\eta_j, \zeta_j]$$

for some elements $\{\eta_j, \zeta_j\}$ in $\mathfrak{g}$. Now, the vector fields X^{η_j} and X^{ζ_j} have *local* Hamiltonian generators f_j and g_j, and these are unique up to a constant. But the value of $\{f_j, g_j\}$ is independent of the choice of constants in the definition of f_j and g_j. Thus, $\{f_j, g_j\}$ is actually a globally defined function. We may define, then,

$$\phi_\xi = \sum_j \{f_j, g_j\}. \tag{36}$$

Arguing as in the proof of Proposition 17, we will see that ϕ_ξ is a global Hamiltonian generator of X^ξ. This shows that the action of G admits a momentum map.

Let $\mathfrak{h} = \{\xi \in \mathfrak{g} | X^\xi = 0\}$, which is an ideal in $\mathfrak{g}$, by Proposition 15. Let $\mathfrak{g}' = \mathfrak{g}/\mathfrak{h}$. It is not hard to verify that $\mathfrak{g}'$ is also semisimple. Pick a basis ξ_j for $\mathfrak{g}'$ and choose generators ϕ_j for each X^{ξ_j}. Now let $V \subset C^\infty(N)$ be defined by

$$V = \operatorname{span}(\phi_1, \ldots, \phi_d, \mathbf{1}), \tag{37}$$

where $d = \dim \mathfrak{g}'$ and $\mathbf{1}$ is the constant function on N. By Proposition 17, V will form a Lie algebra under the Poisson bracket. Since $\{f, \mathbf{1}\} = 0$ for any f, the span $\mathbb{R}\mathbf{1}$ of $\mathbf{1}$ is an ideal in V, and the quotient algebra $V/\mathbb{R}\mathbf{1}$ will be isomorphic to $\mathfrak{g}'$.

Thus, the Lie algebra V is a "one-dimensional central extension" of $\mathfrak{g}'$. It follows, however, from the Levi decomposition of V (see, for example, the book of Varadarjan) that every one-dimensional central extension of a semisimple Lie algebra splits. This means that V decomposes as a Lie algebra (not just as a vector space) as

$$V \cong V' \oplus \mathbb{R}\mathbf{1}.$$

We may now obtain the desired equivariant momentum map by choosing the generator of each X^ξ, $\xi \in \mathfrak{g}$, to be in V'.

Finally, the argument in the first paragraph of the proof shows that there can be only one equivariant momentum map. If ϕ is going to be equivariant, then we must have (36). Since the value of $\{f_j, g_j\}$ is unchanged under addition of a constant, there is only one equivariant possibility for ϕ_ξ. ■

7.2. Quantization of symplectic group actions. When quantizing a symplectic manifold N, if we have a connected Lie group G acting symplectically on N, then we may reasonably hope that there is a projective unitary representation Π of G acting on the associated quantum Hilbert space. We may use the technique of momentum maps together the methods of geometric quantization to construct such an action, at least in favorable cases.

Suppose that we can choose a polarization on N that is invariant under the action of G. Suppose at first that G admits an equivariant momentum map ϕ. Then the components ϕ_ξ of ϕ are quantizable functions. It follows that the quantizations $Q(\phi_\xi)$ of these functions are self-adjoint (or at least symmetric) operators on the quantum Hilbert space, whether we use half-forms or not in our quantization. Furthermore, these operators satisfy the commutation relations of the Lie algebra $\mathfrak{g}$ of G. It is then reasonable to hope that we can exponentiate the operators $Q(\phi_\xi)$ to obtain a unitary representation Π of the *universal cover* $\tilde{G}$ of G, acting on the quantum Hilbert space. Since the quantum Hilbert space is infinite-dimensional, there are some technical issues associate to carrying out this exponentiation process, which we will not delve into here. If it should happen that $\Pi(A)$ is a multiple of the identity for each A in the kernel of the covering map $\tilde{G} \to G$, then we can think of Π as a projective unitary representation of G.

If G acts transitively on N, then experience with examples suggests that the representation Π of $\tilde{G}$ will be irreducible. (Irreducibility is to be thought of as the quantum counterpart of transitivity.) In that case, Schur's lemma tells us that $\Pi(A)$ must be multiple of the identity for each A in the center of $\tilde{G}$, and thus, in particular for each A in the kernel of the covering map. The process of quantizing symplectic manifolds with transitive group actions is an important source of irreducible unitary representations of groups. (This approach is the "orbit method" of Kirillov and Kostant.)

Suppose now that the action of G on N admits a momentum map ϕ, but that the action does not admit an *equivariant* momentum map. (We may think, for example, of the action of $\mathbb{R}^{2n}$ on itself by translation.) For notational simplicity, let us assume that ϕ_ξ is nonzero for each nonzero $\xi \in \mathfrak{g}$. Then the space V of functions in (37) forms a Lie algebra under the Poisson bracket, and this Lie algebra is a one-dimensional central extension of $\mathfrak{g}$. If G^V denotes the unique simply connected Lie group with Lie algebra V, then we may hope to form a representation Π of G^V by exponentiating the operators $Q(\phi_\xi)$ and the operators $Q(\mathbf{1}) = I$. If H denotes the one-dimensional subgroup of G^V whose Lie algebra corresponds to the multiples of the identity in $\mathfrak{g}'$, then, by construction, $\Pi(h)$ will be a multiple of the identity for all $h \in H$. Thus, at the projective level, Π will factor through the quotient group G^V/H (assuming H is a closed subgroup of G^V). Now, the Lie algebra of G^V/H is the quotient algebra $V/\mathfrak{h} \cong \mathfrak{g}$. Thus, G^V/H is covered by the universal cover $\tilde{G}$ of G. Thus, we obtain in the end a projective unitary representation of $\tilde{G}$, just as in the case where an equivariant momentum map exists.

In the case of the group $\mathbb{R}^{2n}$ acting on itself by translations, there is no equivariant momentum map, but there is a momentum map, whose components are the functions x_j and p_j, $j = 1, \ldots, n$. Thus, the Lie algebra V is, in this case, the space of functions spanned by the x_j's, the p_j's, and the constant function $\mathbf{1}$. This algebra is isomorphic to the Lie algebra of the Heisenberg group. (See Exercise 6 in Chapter 14 of the text.) The associated representation Π of the group G' is then given in terms of the exponentiated position and momentum operators described in Chapter 14 of the text. These operators do, in fact, form a projective representation of the group $\mathbb{R}^{2n}$, as described in Example 16.56 in the text.

1. Verify the claim in Example 20.

   Hint: To show that ϕ is a momentum map, imitate the proof of Proposition 2.30. To verify equivariance, use Proposition 17 and evaluate at the origin.

2. (a) Show that in any Lie algebra $\mathfrak{g}$, the space of linear combinations of commutators forms an ideal in $\mathfrak{g}$, called the **commutator ideal**.

   (b) Show that if $\mathfrak{g}$ is semisimple, then the commutator ideal of $\mathfrak{g}$ is equal to $\mathfrak{g}$.

3. Assume N is compact and suppose a symplectic action of a group G on N admits a momentum map. Show that the action of G admits an *equivariant* momentum map.

4. (a) Verify the function ϕ in Example 22 is an momentum map for the action of $\mathsf{SO}(3)$ on the sphere N of radius r.

   Hint: Consider, for example, action of e^{tF_3} on the sphere. The vector field generating this flow is easily computed as

$$x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1} = -X^{F_3}.$$

   Show that on N we have

$$\omega(X^{F_3}, \cdot) = -r^2 dx_3 - \frac{1}{2} x_3 \; d(|\mathbf{x}|^2).$$

   (b) Using Exercise 4 in Chapter 21 in the text, show that ϕ is equivariant.

8. Conservative forces

In this section, we briefly sketch the argument for (the harder direction of) Proposition 2.7 in the text. Let U be a simply connected domain in $\mathbb{R}^n$ and let $\mathbf{F} : U \to \mathbb{R}^n$ be a smooth, vector valued function on U. Assume $\mathbf{F}$ satisfies

$$\frac{\partial F_j}{\partial x_k} - \frac{\partial F_k}{\partial x_j} = 0 \tag{38}$$

for all j, k at every point in U. Then there exists a smooth function $V : U \to \mathbb{R}$ such that

$$\mathbf{F} = -\nabla V \tag{39}$$

on U.

The condition (38) can be understood in the language of differential forms. Let us identify the vector-valued function $\mathbf{F}$ with the 1-form α given by

$$\alpha = \sum_{j=1}^{n} F_j(\mathbf{x}) \; dx_j. \tag{40}$$

Then the *differential* of α is the 2-form $d\alpha$ given by

$$d\alpha = \sum_{j,k=1}^{n} \frac{\partial F_j}{\partial x_k} \; dx_k \wedge dx_j = \sum_{k<j} \left(\frac{\partial F_j}{\partial x_k} - \frac{\partial F_k}{\partial x_j} \right) \; dx_k \wedge dx_j,$$

since $dx_j \wedge dx_k = -dx_j \wedge dx_k$ and $dx_j \wedge dx_j = 0$. We see then that (38) holds if and only if $d\alpha = 0$.

We are looking for a function V satisfying (39). If such a function exists, we will have

$$\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{F} \cdot d\mathbf{x} = -\int_{\mathbf{a}}^{\mathbf{b}} \nabla V \cdot d\mathbf{x} = -\left(V(\mathbf{b}) - V(\mathbf{a})\right), \tag{41}$$

for all $\mathbf{a}$ and $\mathbf{b}$ in U, where the integral is over an smooth path connecting $\mathbf{a}$ to $\mathbf{b}$. Note that $\int_{\mathbf{a}}^{\mathbf{b}} \nabla V \cdot d\mathbf{x}$ is just the integral of the derivative of V along the chosen path, which accounts for the second equality in (41). We would like to try to *define* V by means of (41); that is, we fix some basepoint $\mathbf{a} \in U$ and then define $V(\mathbf{b})$, for any $\mathbf{b} \in U$, by

$$V(\mathbf{b}) = -\mathbf{F}(\mathbf{a}) - \int_{\mathbf{a}}^{\mathbf{b}} \mathbf{F} \cdot d\mathbf{x}. \tag{42}$$

We now argue that the value of V is independent of the choice of path, provided that (38) holds. Given any two paths connecting $\mathbf{a}$ to $\mathbf{b}$, we can form a loop by going out along the first path and then back along the second path with the orientation of the second path reversed. The integral of $\mathbf{F} \cdot d\mathbf{x}$ along the two paths will be equal provided that the integral of $\mathbf{F} \cdot d\mathbf{x}$ over the loop is zero. Since U is simply connected, then every loop is the boundary of some oriented surface S. (As we shrink the loop to a point, it traces out the surface S.) The integral of $\mathbf{F} \cdot d\mathbf{x}$ over a curve is the same as the integral of the 1-form α constructed from $\mathbf{F}$ in (40). By the general form of Stoke's Theorem,

$$\int_{\partial S} \mathbf{F} \cdot d\mathbf{x} = \int_S d\alpha = 0. \tag{43}$$

Thus, the integral around the loop is zero and so the value of $V(\mathbf{b})$ is independent of the choice of path. Using independence of the path, it is not hard to check that indeed $-\nabla V = \mathbf{F}$.