GEOMETRIC QUANTIZATION AND REPRESENTATION THEORY

LECTURES BY AKSHAY VENKATESH,
NOTES BY TONY FENG AND NICCOLO RONCHETTI

CONTENTS

1. A FIRST EXAMPLE

1.1. **Representations of** SO_3 **as spherical harmonics.** Today we'll talk about the characters and the representations of $\mathrm{SO}_3(\mathbf{R})$, the Lie group of rotations in $\mathbf{R}^3$. We start by constructing its irreducible representations.

Definition 1.1.1. Let P_n be the $\mathbf{R}$-vector space of homogeneous polynomials of degree n in $\mathbf{R}^3$, viewed as a representation of $\mathrm{SO}_3(\mathbf{R})$ via its natural action on $\mathbf{R}^3$.

Example 1.1.2. $P_0 \cong \mathbf{R}$ consists of the constant functions. One can easily see that $\dim_{\mathbf{R}} P_1 = 3$, $\dim_{\mathbf{R}} P_2 = 6$ and so on.

Remark 1.1.3. The representations P_n are not irreducible in general! For example, P_2 contains the element $x^2+y^2+z^2$ which is the square radius r^2 and hence is preserved by any rotation in $\mathrm{SO}_3(\mathbf{R})$.

More generally, there are maps

$$\times r^2 \colon P_{n-2} \to P_n \qquad \text{multiply by } r^2 = x^2+y^2+z^2$$

which are $\mathrm{SO}_3(\mathbf{R})$-equivariant. We also have maps in the opposite directions

$$\Delta \colon P_n \to P_{n-2} \quad \text{apply Laplacian } \Delta = \partial_x^2+\partial_y^2+\partial_z^2$$

Consider the direct sum $P := \bigoplus_{n\geq 0} P_n$. Then the operators $\times r^2$ and Δ define a Lie algebra action of $\mathfrak{sl}_2$ on P.

This is not just a curiosity, but a specialization of a more general fact that we'll discuss later in the course.

Fact 1.1.4. The representations

$$V_n := \ker(\Delta : P_n \to P_{n-2}) \cong P_n / r^2(P_{n-2})$$

are a complete collection of irreducible $\mathrm{SO}_3(\mathbf{R})$-representations. The V_n's are sometimes called *spherical harmonics*.

This will be shown below.

Example 1.1.5. Explicitly, $V_2 = \mathrm{Span}_{\mathbf{R}}\{xy, xz, yz, x^2-y^2, y^2-z^2\}$.

One can easily compute that $\dim_{\mathbf{R}} V_n = 2n+1$.

Remark 1.1.6. The restriction of the functions on V_n to the unit sphere in $\mathbf{R}^3$ yields an eigenspace for the Riemannian Laplacian. This is a well-understandable model of the representation V_n, and in fact it is how they arise in physics.

1.2. **Characters.** Our next goal is to find a character formula for the V_n's. Recall that the *character* of a finite-dimensional representation

$$\rho : G \to \mathrm{GL}\, V$$

is the class function

$$\mathrm{Ch} : g \mapsto \mathrm{Tr}(\rho(g)).$$

When later we will talk about infinite-dimensional representations, we'll need to be more careful. Since the trace is additive, we clearly have

$$\mathrm{Ch}(V_n) = \mathrm{Ch}(P_n) - \mathrm{Ch}(P_{n-2}).$$

Any $g \in \mathrm{SO}_3(\mathbf{R})$ is conjugate to a rotation around the z-axis of angle θ - let's call this element g_θ, so it's enough to determine $\mathrm{Ch}(V_n)$ on the elements g_θ.

Fact 1.2.1. We have

$$\mathrm{Ch}(V_n)(g_\theta) = e^{in\theta} + e^{i(n-1)\theta} + \ldots + e^{-in\theta}.$$

More precisely, the eigenvalues of g_θ on V_n are $e^{in\theta}, e^{i(n-1)\theta}, \ldots, e^{-in\theta}$, each with multiplicity one.

Proof. This is an exercise worth doing. It's just a routine computation by writing down basis of P_n and P_{n-2} and explicitly computing how g_θ acts on it. □

Some algebraic manipulations yield

$$\chi_n(g_\theta) := \mathrm{Ch}(V_n)(g_\theta) = \frac{e^{i(n+1)\theta} - e^{-in\theta}}{e^{i\theta} - 1} = \frac{e^{i(n+\frac{1}{2})\theta} - e^{-i(n+\frac{1}{2})\theta}}{e^{i\theta/2} - e^{-i\theta/2}}. \tag{1.2.1}$$

If we normalize a Haar measure on $\mathrm{SO}_3(\mathbf{R})$ so that the total volume is 1, then the χ_n's are orthogonal with respect to it. In coordinates g_θ, this measure is proportional to $|e^{i\theta/2} - e^{-i\theta/2}|^2$, where the constant of proportionality is simply to make the Haar measure be a probability measure.

Proof of Fact 1.1.4. The irreducibility of the representations V_n follows by orthonormality of the characters χ_n. To show that the V_n's are a complete collection, one checks that the span of the χ_n's is dense in the space of class functions on $SO_3(\mathbf{R})$. □

That orthogonality of characters alone is enough to find and classify irreducibile representations for compact, connected Lie groups is a great insight by Weyl.

1.3. **Kirillov's formula.** Kirillov found (in another contest) a reformulation of Fact 1.2.1 which holds in a greater generality - for example allowing non-compact groups and many infinite-dimensional representations. The slogan is

$$\boxed{\chi = \left(\begin{array}{c} \text{Fourier transform of surface} \\ \text{measure on a sphere in } \mathbf{R}^3 \end{array} \right).} \tag{1.3.1}$$

Clearly this slogan "as is" does not make sense, so we now modify the two sides to get a reasonable statement.

Our first step is to make the left hand side take values on a linear space, since this is certainly true for the Fourier transform on the right hand side. In other words, we need to pull χ back to the Lie algebra.

Recall that in coordinates we have

$$SO_3(\mathbf{R}) = \left\{ A \in GL_3(\mathbf{R}) \mid AA^t = \mathrm{Id}_3 \right\}.$$

By thinking of A as "infinitesimally close" to the identity matrix Id_3, we get by "differentiating" a description of the Lie algebra

$$\mathrm{Lie}(SO_3(\mathbf{R})) = \left\{ X \in \mathrm{Mat}_3(\mathbf{R}) | X + X^t = 0. \right\}.$$

Under this description, the exponential map

$$\exp : \mathrm{Lie}(SO_3(\mathbf{R})) \to SO_3(\mathbf{R})$$

is simply given by exponentiating the matrix X - so in fact in the left hand side of our slogan we mean $\chi(e^X)$ for $X \in \mathrm{Lie}(SO_3(\mathbf{R}))$.

Let's coordinatize $\mathrm{Lie}(SO_3(\mathbf{R}))$. Let J_a be the rotation of 90 degrees around the axis a. More explicitly, we have

$$J_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}, \quad J_y = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad J_z = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

so that $[J_x, J_y] = J_z$, $[J_y, J_z] = J_x$ and $[J_z, J_x] = J_y$.

Any element of the Lie algebra can then be written in coordinates as $X = x J_x + y J_y + z J_z$. In fact, this choice of coordinates determines a map $\mathrm{Lie}(SO_3(\mathbf{R})) \to \mathbf{R}^3$ which is an isomorphism of $SO_3(\mathbf{R})$-representations: on the left we have the adjoint action by conjugation and on the right the tautological action by rotation.

In particular, since every point on the sphere $x^2 + y^2 + z^2 = r^2$ can be transformed to $(0, 0, r)$ by a rotation in $SO_3(\mathbf{R})$, for the purpose of computing the class function $\chi(e^X)$ we can assume that $X = r J_z$. In this case, we can use the formula 1.2.1 to obtain

$$\chi_n(e^X) = \frac{e^{i(n+\frac{1}{2})r} - e^{-i(n+\frac{1}{2})r}}{e^{ir/2} - e^{-ir/2}} \tag{1.3.2}$$

where $r = \sqrt{x^2 + y^2 + z^2} = |X|$ is the length of X with respect to the standard coordinates.

The right hand side of (1.3.2) can be thus thought of as a radial function on $\mathbf{R}^3$. Its Fourier transform is a mess. Kirillov's insight was then by putting a little twist to it, one can turn the Fourier transform into a much nicer function.

To explain the genesis of this twist, notice that if $f, g : \mathrm{SO}_3(\mathbf{R}) \to \mathbf{R}^3$ are nice functions, we can pull them back to $\mathrm{Lie}(\mathrm{SO}_3(\mathbf{R}))$ and integrate:

$$\int_{\mathrm{Lie}(\mathrm{SO}_3(\mathbf{R}))} f(e^X) g(e^X)\, \mathrm{d}X$$

with respect to the Lebesgue measure $\mathrm{d}X$. On the other hand, we can also integrate them against the Haar measure μ mentioned before:

$$\int_{\mathrm{SO}_3(\mathbf{R})} f(y) g(y)\, \mathrm{d}\mu(y).$$

These two integrals are not equal to each other. The issue is not simply that the exponential map is finite-to-one (rather than one-to-one), in fact the two integrals do not coincide even locally. The problem is that the map $X \mapsto e^X$ has nontrivial Jacobian.

In order to retain orthogonality of characters when pulling back to the Lie algebra, we need then to twist by a quantity related to the Jacobian of the exponential map. More precisely, let $j(X)$ be the Jacobian of the exponential map at $X \in \mathrm{Lie}(\mathrm{SO}_3(\mathbf{R}))$, and instead of $\chi(e^X)$ we look at $\chi(e^X)\sqrt{j(X)}$. Then the twist factors will cancel out with the Jacobian when comparing the two integrals above, and they will make the two integrals match up locally.

Fact 1.3.1. In the system of coordinates chosen above, one has

$$j(X) = \left(\frac{e^{ir/2} - e^{-ir/2}}{ir} \right)^2.$$

Proof. We'll prove this next time, using a more general formula worth seeing. □

Let's see what this buys us. By substituting the expression from (1.3.1) into (1.3.2), one gets that

$$\chi_n(e^X)\sqrt{j(X)} = \frac{e^{i(n+\frac{1}{2})r} - e^{-i(n+\frac{1}{2})r}}{ir},$$

so to prove our slogan it remains to show that the right hand side is the Fourier transform of some surface measure for a sphere in $\mathbf{R}^3$.

Definition 1.3.2. Consider the sphere of radius R in $\mathbf{R}^3$, centered at the origin, and denote it S^2_R. Given a surface measure μ on it, its *Fourier transform* is

$$\widehat{\mu}(k) := \int_{S^2_R} e^{i\langle k, x\rangle}\, \mathrm{d}\mu(x)$$

where $k, x \in \mathbf{R}^3$ and $\langle k, x\rangle$ is the usual dot product.

Lemma 1.3.3. *For μ the standard measure on the sphere of radius R,*

$$\widehat{\mu}(k) = (2\pi R)\frac{e^{iR\lambda} - e^{-iR\lambda}}{i\lambda}$$

where $\lambda = |k|$ is the length of $k \in \mathbf{R}^3$.

This completes the explanation of our slogan. We also record that we obtain that the χ_n's are Fourier transforms of surface measures of sphere with half-integral radius. One explanation is that the area of the sphere matches up the dimension of the representation - we discuss this more later in the course.

Proof of Lemma 1.3.3. We compute the Fourier transform of the surface measure. Since the surface measure is radial, so is its Fourier transform. Therefore it suffices to compute $\widehat{\mu}(k)$ at $k = (0,0,\lambda)$. The projection onto the z-axis of the sphere surjects onto the segment $[-R, R]$, and the pushforward of the surface measure on the sphere is $(2\pi R)dz$: we obtain then that the Fourier transform is

$$\int_{-R}^{R} e^{i\lambda z}(2\pi R\, dx) = 2\pi R \frac{e^{iR\lambda} - e^{-iR\lambda}}{i\lambda}.$$

□

1.4. **Summary.** Here's a summary of what we've done today:

- We listed irreducible representations of $\mathrm{SO}_3(\mathbf{R})$.
- We pulled back their characters to $\mathrm{Lie}(\mathrm{SO}_3(\mathbf{R}))$.
- We twisted them by a quantity related to the Jacobian of the exponential map.
- Up to pinning down a few constants, we obtain that the character of a representation of $\mathrm{SO}_3(\mathbf{R})$ is indeed the Fourier transform of a surface measure on a sphere in $\mathbf{R}^3$.

2. The general picture

2.1. **Review of the example.** Last time we described a model V_N for the irreducible $2N+1$-dimensional representation of $\mathrm{SO}_3(\mathbf{R})$, and a formula for its character χ_N. Denoting by g_θ the rotation of angle θ around the z-axis, we had

$$\chi_N(g_\theta) = \frac{e^{i(N+\frac{1}{2})\theta} - e^{-i(N+\frac{1}{2})\theta}}{e^{i\theta/2} - e^{-i\theta/2}}.$$

After massaging this formula, we reformulated it as

$$\left(\chi_N \circ \sqrt{j}\right)(e^X) = \frac{e^{i(N+\frac{1}{2})s} - e^{-i(N+\frac{1}{2})s}}{is} \tag{2.1.1}$$

where

- $X \in \mathrm{Lie}(\mathrm{SO}_3(\mathbf{R}))$, and $s = |X|$ is the length of X in the coordinate system described last time in §1.3.
- j is the Jacobian at X of the exponential map $\exp : \mathrm{Lie}(\mathrm{SO}_3(\mathbf{R})) \to \mathrm{SO}_3(\mathbf{R})$

Last time we compared this formula (2.1.1), in particular the right hand side, to the Fourier transform of the surface measure of a sphere. In fact, we calculated that the Fourier transform of the area measure of a sphere of radius R in $\mathbf{R}^2$ is

$$\widehat{\mu}(k) = \frac{e^{iR\lambda} - e^{-iR\lambda}}{i\lambda} 2\pi R$$

where $\lambda = |k|$ is the length of $k \in \mathbf{R}^3$.

Remark 2.1.1. The twist by the Jacobian $\sqrt{j}$ is important algebraically, but today we focus on a heuristic description of what's behind the scenes of the equality above, and for this purpose we can disregard this twist.

In other words, we obtain that $\chi_N \cdot \sqrt{j}$ is the Fourier transform of the area measure of a sphere of radius $N + \frac{1}{2}$, with the measure normalized so that the $(2\pi R)$ factor disappears and the total area is $2N + 1 = \dim V_N$. We can then rewrite (2.1.1) as

$$(\chi_N \sqrt{j})(e^X) = \int_{\xi \in S_{N+\frac{1}{2}}} e^{i\langle \xi, X\rangle} \, \mathrm{d}\,\xi. \tag{2.1.2}$$

which specializes for $X = 0$ to

$$2N + 1 = \dim V_N = \chi_N \sqrt{j}(e^0) = \int_{\xi \in S_{N+\frac{1}{2}}} \mathrm{d}\,\xi = \mathrm{Area}\left(S^2_{N+\frac{1}{2}}\right). \tag{2.1.3}$$

2.2. **Some speculations.** Why do these formulas hold? One could speculate that (2.1.3) is true because there is a basis of V_N that is in bijection with bits of S_{2N+1} each having area 1.

Speculation 2.2.1. There exists a basis $\{v_1, \dots, v_{2N+1}\}$ of V_{2N+1} in bijection with a partition

$$S_{2N+1} = \coprod_{k=1}^{2N+1} \mathscr{O}_k$$

such that $\mu(\mathscr{O}_k) = 1$ for each k.

To explain the more general formula (2.1.2), we need to go beyond this. We could ask how G or $\mathrm{Lie}(G)$ act on the basis $\{v_k\}$ and we may hope that they are approximate eigenvectors.

Speculation 2.2.2. For some $\xi_k \in \mathscr{O}_k$, we have that each $X \in \mathrm{Lie}(\mathrm{SO}_3(\mathbf{R}))$ yields

$$X \cdot v_k \approx i\langle \xi_k, X\rangle v_k.$$

Notice that this latter speculation would explain formula (2.1.2), as it implies that[1]

$$\mathrm{Tr}(e^X) \approx \sum_k e^{i\langle \xi_k, X\rangle} \approx \int_{S_{N+\frac{1}{2}}} e^{i\langle \xi, X\rangle} \, \mathrm{d}\,\xi.$$

Together, Speculations 2.2.2 and 2.2.1 say that the decomposition $S_{N+\frac{1}{2}} = \coprod_k \mathscr{O}_k$ should correspond to an approximate diagonalization of the $\mathrm{SO}_3(\mathbf{R})$-action.

[1] Ignoring $\sqrt{j}$ as usual

Remark 2.2.3. The above picture is really an approximation: it cannot be true literally as it imply that V_N is a sum of 1-dimensional representation, which is not.

In particular, Speculation 2.2.2 cannot be true as stated, because it would imply that for two non-commuting elements $X, Y \in \mathrm{Lie}(\mathrm{SO}_3(\mathbf{R}))$ we have on one hand

$$[X, Y] \cdot v_k = i\langle \xi_k, [X, Y]\rangle v_k \neq 0$$

and on the other hand

$$[X, Y] \cdot v_k = X \cdot Y v_k - Y \cdot X v_k = 0.$$

One of the goal of the course is to flesh out a picture like this for a general Lie group G, where the role of these spheres is taken by more general spaces (we'll describe the one for $\mathrm{SO}_4(\mathbf{R})$ later). The Kirillov formula makes the above speculations precise, but it is even more important because of the picture behind it - which is one of the most useful aspects from the viewpoint of representation theory.

Remark 2.2.4. Notice that we have been very imprecise about how to choose $\xi_k \in \mathcal{O}_k$. This adds some fuzziness to the picture.

The best interpretation of Speculation 2.2.1 is saying that if we decompose v_k as a sum of X-eigenvectors, the eigenvalues appearing are of the form $i\langle \xi, X\rangle$ for some $\xi \in \mathcal{O}_k$. This yields a map

$$\begin{aligned} \mathcal{O}_k &\to \mathbf{R} \\ \xi &\mapsto \langle \xi, X\rangle \end{aligned}$$

whose image is a closed small interval in $\mathbf{R}$, let's say of length l_X: this can be thought of "length of $\mathcal{O}_k$ in the X-direction".

With this description, Speculation 2.2.2 is saying that X acts on v_k via some range of eigenvalues around $i\langle \xi_k, X\rangle$ of length approximately l_X. The argument with the commutators given before implies that to have no contradiction we must get

$$l_X \cdot l_Y > \langle \xi_k, [X, Y]\rangle.$$

The upshot is that Speculation 2.2.2 is a sensible statement as long as $\mathcal{O}_k$ is "not too small" - in the sense given by the inequality above.

Remark 2.2.5. Notice that this "non-smallness" condition is defined purely in terms of the Lie brackets; we did not have to introduce any non-canonical measure.

2.3. **The general formula.** Let's come back to reality, after our speculations. Making formal the reasoning explained above, one gets

Fact 2.3.1. Let G be a Lie group, and $\xi \in \mathrm{Lie}(G)^*$ be an element of the dual Lie algebra. Let $\mathcal{O}$ be the G-orbit of ξ.

Then there is a natural G-invariant non-degenerate symplectic[2] form ω on $\mathcal{O}$ defined by the following rule: for each $t_1, t_2 \in \mathrm{Tan}_\xi(\mathcal{O})$ choose $X, Y \in \mathrm{Lie}(G)$ such that $t_1 = X \cdot \xi$, $t_2 = Y \cdot \xi$ and set

$$\omega(t_1, t_2) = \xi([X, Y]).$$

[2]That is to say, an alternating non-degenerate bilinear form on the tangent bundle of $\mathcal{O}$, varying G-equivariantly.

The condition above obviously determines ω, one needs to show that it is well-defined, for example independent of the choices of X, Y.

Remark 2.3.2.
(1) This canonical form is completely canonical. In particular, one can prove that it does not depend on the choice of $\xi \in \mathcal{O}$.
(2) It is hard to grasp what this symplectic form "is". We'll come back to this point later.

Example 2.3.3. As seen last time, in the case of $G = \mathrm{SO}_3(\mathbf{R})$, the orbits $\mathcal{O}$ turn out to be spheres in $\mathbf{R}^3$.

Corollary 2.3.4. *The orbits $\mathcal{O}$ are even-dimensional.*

This is immediate, as any symplectic manifold with a non-degenerate symplectic form is even-dimensional.

Theorem 2.3.5 (Kirillov, others)**.** *Let G be a connected Lie group which is either nilpotent or semisimple (the latter includes the compact case). Let π be a tempered representation of G. Then there is an orbit $\mathcal{O}$ of G on $\mathrm{Lie}(G)^*$ such that the twist $\left(\chi_\pi\sqrt{j}\right)(e^X)$ is the Fourier transform of the measure $\left(\frac{\omega}{2\pi}\right)^d$ on $\mathcal{O}$, where ω is the symplectic form associated to $\mathcal{O}$ as above, and $2d$ is the real dimension of $\mathcal{O}$.*

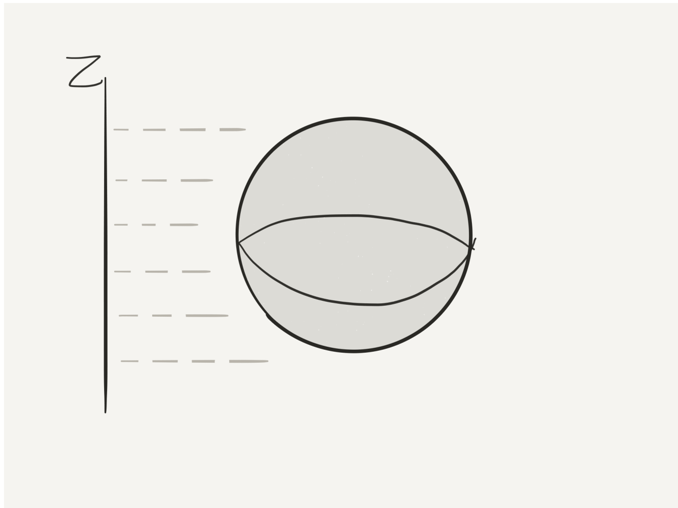
The upshot is that the canonical choice of the symplectic form ω fixes the area discrepancy that we have seen in the example of $\mathrm{SO}_3(\mathbf{R})$.

Remark 2.3.6. The dimension of an orbit $\mathcal{O}$ is the dimension of G minus the rank of G, where the latter should be thought of the dimension of the centralizer of a typical element of G.

Exercise 2.3.7. In the case $G = \mathrm{SO}_3(\mathbf{R})$, check that the measure $\frac{\omega}{2\pi}$ gives area $2N+1$ to the sphere of radius $N+\frac{1}{2}$. You could even get rid of the 2π in the expression of the measure if you normalize the Fourier transform accordingly.

Exercise 2.3.8. What does orthogonality of the characters χ's of G say about the Fourier transforms of the twists $\chi\sqrt{j}$'s? You should expect to recover properties of "areas of spheres".

Exercise 2.3.9. Here's the decomposition in orbits of the sphere $S_{N+\frac{1}{2}}$ inside $\mathbf{R}^3 \cong \mathrm{Lie}(\mathrm{SO}_3(\mathbf{R}))$. Centering the sphere at the origin, the projection onto the z-axis surjects onto the interval $\left[-\left(N+\frac{1}{2}\right), \left(N+\frac{1}{2}\right)\right]$, which we partition in $2N+1$ subintervals of length 1 in the obvious way.



Then the orbit decomposition $\coprod \mathcal{O}_k = S_{N+\frac{1}{2}}$ is given by the preimages of these intervals. In particular, you can check that each of them has area 1. Which basis of V_N does this orbit decomposition correspond to?

Later in the course we'll prove the promised general formula for the Jacobian $j(X)$.

2.4. **Vista.** Where do we go from here? The Kirillov formula is interesting from several points of view:

(1) In representation theory, it is related to the *Duflo isomorphism*, a completely algebraic way of describing the center of the universal enveloping algebra of $\mathrm{Lie}(G)$ that involves factors of $\sqrt{j}$.
(2) From the point of view of symplectic analysis, one could ask why the Fourier transform of the measure on $\mathcal{O}$ is so nice (having properties resembling those of a character of a representation). An explanation has been given by Diustermaat and Heckmann, with a later explanation using equivariant cohomology by Atiyah and Bott.
(3) In general, the formula for the Jacobian is

$$j(X) = \det\left(\frac{\operatorname{ad} X}{e^{\operatorname{ad} X} - 1}\right).$$

This expression $\frac{u}{e^u - 1}$ appears in many areas of mathematics. Why does it show up here?

In the rest of the course we will try to explain why the correspondence between representations V of G and orbits $\mathcal{O}$ of G on $\mathrm{Lie}(G)^*$ is a special case of quantization: a correspondence between Hilbert spaces H and symplectic manifolds (M, ω). Our first goal starting next time is to study the case of $M = \mathbf{R}^2$ via the theory of pseudodifferential operators.

3. Pseudo-differential operators

3.1. **Motivation.** Recall the picture from last time: we had the $SO_3(\mathbf{R})$-representation V_N on the space of harmonic polynomials of degree N in $\mathbf{R}^3$. This was related with a sphere $\mathscr{O} \subset \mathrm{Lie}(SO_3(\mathbf{R}))^*$ whose normalized area is $2N+1 = \dim V_N$, via the Kirillov character formula:

$$\left(\chi_{V_N}\sqrt{j}\right)(e^X) = \int_{\mathscr{O}} e^{i\xi\cdot x}\,\mathrm{d}\,\xi.$$

The heuristic picture for this (which we will make more precise today
setting) is the following: we can dissect the sphere in pieces $\mathscr{O}_i$ of area 1, in bijection with a basis $\{v_i\}$ of V_N of approximate eigenvectors - in the sense that

$$X.v_i \approx \langle \xi_i, X\rangle v_i \text{ for some } \xi_i \in \mathscr{O}_i.$$

This formula is not quite well-defined because we have not specified which ξ we pick in $\mathscr{O}_i$, so there is some fuzziness going on.

The above description should hold for a large class of groups G and representations V. Today we will instead established a related picture, in a different context: V_N is replace by $L^2(\mathbf{R})$ and $\mathscr{O}$ by $\mathbf{R}^2$. There will be no groups involved (for now, but in a few lectures we'll see that this picture is related to the Heisenberg group), so the connection between $L^2(\mathbf{R})$ and $\mathbf{R}^2$ will come from pseudo-differential operators - in fact, the Kirillov character formula will be "replaced" by the formula for the trace of a pseudo-differential operator.

3.2. **Fourier transform.** The discussion today will belong to the realm of microlocal analysis, but we will be a bit informal, not worrying about convergence issues or the exact regularity of our functions - this will be made more precise next time. Our next mini-goal is to write every $f \in L^2(\mathbf{R})$ as a sum of functions $f_i \in L^2(\mathbf{R})$ which are localized (that is to say, they are supported in a small interval) and whose Fourier transforms $\widehat{f_i}$ is also localized.

To fix normalizations, recall that the *Fourier transform* of $f(x) \in L^2(\mathbf{R})$ is

$$\widehat{f}(\xi) = \int f(x)e^{ix\xi}\,\mathrm{d}\,x$$

and its inverse is

$$f(y) = \int \widehat{f}(\xi)e^{-iy\xi}\,\mathrm{d}\,\xi.^{3}$$

In particular, the above normalizations yield that

$$\widehat{\left(\frac{\mathrm{d}}{\mathrm{d}\,x}f\right)}(\xi) = (-i\xi)\widehat{f}(\xi).$$

[3] For ease of notation, we are sweeping under the rug all the factors of (2π) that usually appears when dealing with Fourier transforms. To make the statements literally true, one needs to normalize the measure $\mathrm{d}\,\xi$ to be the Lebesgue measure divided by 2π.

Remark 3.2.1. Why are we shooting for the goal of decomposing f into localized functions? The reason is that such functions f_i's will be approximate eigenfunctions of a differential operator.

More generally, a constant-coefficient differential operator acts on the Fourier transforms as multiplication by a polynomial in ξ. This is not true anymore for differential operators with non-constant coefficients, but the situation is very similar when considering localized functions.

Lemma 3.2.2. *We cannot have both g and $\widehat{g}$ be compactly supported.*

Exercise 3.2.3. Describe a quantitative version of the failure above. More precisely, if g is mainly supported in an interval of length N, how localized is $\widehat{g}$?

Since the lemma says that our mini-goal is impossible when stated as above, we will content ourselves with coming up with f_i's that are "mainly" localized, and similarly for the $\widehat{f_i}$'s.

Let's be a bit more precise. Fix an interval I of $\mathbf{R}$ of length L and an interval J of length M, and let χ_I, χ_J be the two characteristic functions. Then $\chi_I f$ is supported on I, and $\left(\chi_J \widehat{f}\right)^\vee$ has Fourier transform[4] supported on J. The properties of the Fourier transform also yields that

$$\left(\chi_J \widehat{f}\right)^\vee = \chi_J^\vee * f$$

is the convolution of f and the Fourier transform of the characteristic function of J. The upshot is that we have two possible 'operations' to do on f:

(1) $f \mapsto \chi_I f$ localizes f, while
(2) $f \mapsto \chi_J^\vee * f$ localizes $\widehat{f}$.

We'd like to do both (to enforce both f and $\widehat{f}$ to be localized), but these two operations do not commute! So we just ignore this failure of commutative, do it anyway, and hope for the best. We'll get a pseudo-differential operator which will have good properties.

3.3. **Example of pseudo-differential operators.** Let's get down to the computations. First, pick smooth approximations $a, b \in C^\infty(\mathbf{R})$ for the characteristic functions χ_I and χ_J, so that convolution and Fourier transform will have nice properties. Applying $b(\xi)$ to f we get

$$\left(b(\xi)\widehat{f}\right)^\vee(x) = \int b(\xi)\widehat{f}(\xi)e^{-ix\xi}\,\mathrm{d}\xi$$

and multiplying by $a(x)$ yields

$$\int a(x)b(\xi)\widehat{f}(\xi)e^{-ix\xi}\,\mathrm{d}\xi.$$

Definition 3.3.1. More generally, let $a(x,\xi)$ be a function in two variables x and ξ. We can write down a similar formula, which will have nice properties if $C_c^\infty(\mathbf{R}^2)$. We define $\mathrm{Op}(a) : L^2(\mathbf{R}) \to L^2(\mathbf{R})$ by the rule

$$\left(\mathrm{Op}(a)f\right)(x) = \int a(x,\xi)\widehat{f}(\xi)e^{-ix\xi}\,\mathrm{d}\xi. \tag{3.3.1}$$

[4]We denote by $g^\vee$ the inverse of the Fourier transform.

This is a *pseudo-differential operator with symbol a*.

Morally, we want to take $a(x,\xi)$ to be the characteristic function of the rectangle $I\times J$, and we hope that $\mathrm{Ob}(a)$ localizes f at I and $\widehat{f}$ at J. But as the two operations do not commute, something delicate needs to happen for our hope to be (approximately) fulfilled. Since clearly localizing a function $f\mapsto \chi_I f$ is a projection operator on $L^2(\mathbf{R})$, we expect $\mathrm{Op}\big(\chi_{I\times J}\big)$ to be a projection and hence $\mathrm{Op}(a)$ to be an approximate projection:

$$\mathrm{Op}(a)^2 \approx \mathrm{Op}(a)$$

if a is constructed from the two smooth approximations of χ_I, χ_J as in the discussion above.

Fact 3.3.2. The larger the area of $I\times J$ is (in particular, larger than 1), the better the approximation $\mathrm{Op}(a)^2\approx \mathrm{Op}(a)$ will be. In particular, $\mathrm{Op}(a)$ will almost localize f at $I\times J$.

Fix $a(x,\xi)$ and $b(y,\nu)$ two "regular" functions: we want to compute $\mathrm{Op}(a)\mathrm{Op}(b)$. We have

$$\begin{aligned}
\big(\mathrm{Op}(a)\mathrm{Op}(b)f\big)(x) &= \int a(x,\xi)\big(\widehat{\mathrm{Op}(b)f}\big)(\xi)e^{-ix\xi}\,\mathrm{d}\xi\\
&= \int a(x,\xi)\bigg(\int \mathrm{Op}(b)f(y)e^{iy\xi}\,\mathrm{d}y\bigg)e^{-ix\xi}\,\mathrm{d}\xi\\
&= \int a(x,\xi)\bigg(\int\bigg(\int b(y,\nu)\widehat{f}(\nu)e^{-iy\nu}\,\mathrm{d}\nu\bigg)e^{iy\xi}\,\mathrm{d}y\bigg)e^{-ix\xi}\,\mathrm{d}\xi.
\end{aligned}$$

Notice that multiplying all the exponentials in the last integral yields $e^{i(y-x)(\xi-\nu)}e^{ix\nu}$. If we define

$$c(x,\nu) := \int\int a(x,\nu)e^{i(y-x)(\xi-\nu)}b(y,\nu)\,\mathrm{d}\xi\,\mathrm{d}y$$

we finally obtain that

$$\big(\mathrm{Op}(a)\mathrm{Op}(b)f\big)(x) = \int c(x,\nu)\widehat{f}(\nu)e^{-ix\nu}\,\mathrm{d}\nu = \big(\mathrm{Op}(c)f\big)(x).$$

Shifting variables, we obtain equivalently that

$$c(x,\nu) = \int\int a(x,\nu+s)e^{ist}b(x+t,\nu)\,\mathrm{d}s\,\mathrm{d}t.$$

Denote now by ∂_2 the derivative in the second variable, we can expand a as a Taylor series in the second variable to get

$$a(x,\nu+s) = \sum_{l\geq 0}\frac{\partial_2^l a(x,\nu)}{l!}s^l$$

and plugging this into the formula for $c(x,\nu)$ one gets

$$c(x,\nu) = \sum_{l\geq 0}\frac{\partial_2^l a(x,\nu)}{l!}\int s^l b(x+t,\nu)e^{ist}\,\mathrm{d}s\,\mathrm{d}t.$$

The last integral is easy:

Fact 3.3.3. Let g be a Schwartz function. Then

$$\int s^l e^{ist} g(t) \mathrm{d}s\,\mathrm{d}t = \int s^l \widehat{g}(s) \mathrm{d}s = \left(i \frac{\mathrm{d}}{\mathrm{d}t}\right)^l g(0).$$

Therefore, we obtain

$$c(x,\nu) = \sum_{l \geq 0} \frac{i^l}{l!} \partial_2^l a(x,\nu) \partial_1^l b(x,\nu). \tag{3.3.2}$$

Remark 3.3.4. Clearly the above passages will need to be justified for a, b, f in some class of functions - we'll do this next time.

Explicitly, we have

$$c(x,\nu) = a(x,\nu) b(x,\nu) + \underbrace{i \partial_\nu a(x,\nu) \partial_x b(x,\nu) + \frac{1}{2} \partial_\nu^2 a \partial_x^2 b + \ldots}_{=: a \star b}$$

and we denote the first-order and higher terms by $a \star b$.

Remark 3.3.5. This is quite a nice expression, and yet it is not literally true - it needs to be interpreted as an asymptotic series.

Notice that $a + b \neq b * a$, and we can measure the failure of commutativity:

$$a * b - b * a = i(\partial_\nu a \partial_x b - \partial_x a \partial_\nu b) + \text{ (higher order terms)}$$

where the second-degree term is called the *Poisson bracket*, which plays an important role in symplectic geometry.

Let's go back to our question: to what extent is $\mathrm{Op}(\xi_{I \times J})$ a projection operator? To be precise, let's assume $I = [0, L]$ and $J = [0, M]$. Fix a smooth approximation ϕ of $\xi_{[0,1]}$, so that

$$a(x,\nu) = \phi\left(\frac{x}{I}\right) \phi\left(\frac{\nu}{J}\right)$$

is an approximation of $\chi_{I \times J}$. Then the explicit formula (3.3.2) yields that

$$\mathrm{Op}(a)^2 = \mathrm{Op}(a^2) + \frac{1}{LM} \cdot (\text{derivatives of } \phi) + \left(\frac{1}{LM}\right)^2 \cdot (\text{second derivatives of } \phi) + \ldots$$

The upshot is that as long as $LM >> 1$, we have

$$\mathrm{Op}(a)^2 \approx \mathrm{Op}(a^2) \approx \mathrm{Op}(a)$$

where the last approximation is due to the fact that $a^2 \approx a$, since a is a smooth approximation of a characteristic function.

Remark 3.3.6. Again, we are postponing to next time the formal, rigorous version of the above reasoning, that takes care of all the convergence and well-posedness issues.

3.4. **Trace of pseudo-differential operators.** Going back to our heuristic picture, here's what we found. Given a decomposition

$$\mathbf{R}^2 = \coprod_{i\in\mathbf{N}} \mathcal{O}_i$$

in blocks of very large area, we obtain

$$\mathrm{Id} = \mathrm{Op}(1) = \sum_{i\in\mathbf{N}} \mathrm{Op}(a_i)$$

where each $\mathrm{Op}(a_i)$ is approximately a projection operator. In particular, this yields morally a decomposition of $L^2(\mathbf{R})$ which is completely geometric in nature!

$$L^2(\mathbf{R}) = \bigoplus_{i\in\mathbf{N}} V_i$$

and moreover the dimension of the V_i is related to the trace of the pseudo-differential operator. More precisely:

Lemma 3.4.1. *We have*

$$\mathrm{Tr}(\mathrm{Op}(a_i)) = \int_{\mathcal{O}_i} a(x,\xi)\,\mathrm{d}x\,\mathrm{d}\xi \approx \mathrm{Area}(\mathcal{O}_i).$$

Proof. Assuming f is Schwartz, we have

$$\begin{aligned}\big(\mathrm{Op}(a)f\big)(x) &= \int a(x,\xi)\widehat{f}(\xi)e^{-ix\xi}\,\mathrm{d}\xi\\ &= \int a(x,\xi)\left(\int e^{iy\xi}f(y)\,\mathrm{d}y\right)e^{-ix\xi}\,\mathrm{d}\xi\\ &= \int a(x,\xi)e^{i(y-x)\xi}f(y)\,\mathrm{d}y\,\mathrm{d}\xi\\ &= \int K(x,y)f(y)\,\mathrm{d}y,\end{aligned}$$

where we defined the kernel

$$K(x,y) := \int a(x,\xi)e^{i(y-x)\xi}\,\mathrm{d}\xi.$$

Assuming the standard conditions for $\mathrm{Op}(a)$ to be trace class, we have that its trace is the integral of its kernel on the diagonal:

$$\mathrm{Tr}(\mathrm{Op}(a)) = \int K(x,x)\,\mathrm{d}x = \int a(x,\xi)\,\mathrm{d}x\,\mathrm{d}\xi.$$

The last approximation $\int a_i(x,\xi)\,\mathrm{d}x\,\mathrm{d}\xi \approx \mathrm{Area}(\mathcal{O}_i)$ is clear, since $a_i(x,\xi)$ is an approximation of the characteristic function of $\mathcal{O}_i$. □

We'll see more examples, but it's worthwhile (and nice!) to see that in this example we were able to compute everything.

3.5. **Making this rigorous.** We now want to sketch how to make the preceding discussion more rigorous. We wrote down a formal infinite sum for the composition of two pseudo-differential operators:

$$\mathrm{Op}(a)\mathrm{Op}(b) = \mathrm{Op}(ab + i\partial_\xi a \partial_x b + \dots) \tag{3.5.1}$$

To make this rigorous, we need to impose analytic conditions on the functions involved. So we specify a class of "allowable symbols" $a(x,\xi)$. A convenient class S_m of such is "differentiable operators of order $\leq m$ with bounded coefficients". The simplest version of the definition is: $a(x,\xi) \in S^m$ if

$$\sup_{x,\xi} \frac{|\partial_x^i \partial_\xi^j a(x,\xi)|}{(1+|\xi|)^{m-j}} \leq \mathrm{const}(i,j) \text{ for all } i,j \geq 0.$$

Note that we don't force i,j to be less than m. Note also that there is an asymmetry between the ξ and x variables. This is because we are thinking of the symbol as a differential operator, with ξ being the "derivative symbol", and we want to model an "almost constant coefficient" differential operator, meaning that its coefficients and their derivatives are bounded.

Fact 3.5.1. Here are the important facts about this definition.

- If $a \in S_m$ and $b \in S_{m'}$, then there exists $c \in S_{m+m'}$ with $\mathrm{Op}(a)\mathrm{Op}(b) = \mathrm{Op}(c)$.
- Moreover,

$$c - \sum_{N=0}^{w} \frac{i^N}{N!} \partial_\xi^N a \partial_x^N b \in S^{m+m'-w-1}.$$

Informally, this says that if w is large then $c - \sum_{N=0}^{w} \frac{i^N}{N!} \partial_\xi^N a \partial_x^N b$ decays very quickly in the second variable. Thus the right interpretation of the formal infinite sum (3.5.1) is to truncate the sum at some large w.

The formula (3.3.1) works well a priori if $a \in C_c^\infty$. But we can extend it to $a \in S^0$, and for such a we have that $\mathrm{Op}(a)$ is L^2-bounded.

Remark 3.5.2. The theory is not very symmetric. It is adapted to our specific system of coordinates in $\mathbf{R}^2$. We have a decomposition corresponding to the dissection of $\mathbf{R}^2$ into rectangles parallel to the axes, and wouldn't work as well for rotated rectangles.

4. The Kirillov paradigm for the Heisenberg group

4.1. **Review.** Last time we defined a pseudo-differential operator $\mathrm{Op}(a)$ for a function a on $\mathbf{R}^2$:

$$\mathrm{Op}(a)f(x) = \int a(x,\xi)\widehat{f}(\xi)e^{-ix\xi}\, d\xi.$$

Recall that if a is (approximately) the characteristic function of $I \times J$, then $\mathrm{Op}(a)$ is an approximate projector onto the space of functions f with $\mathrm{supp}(f) \subset J$ and $\mathrm{supp}(\widehat{f}) \subset I$, as long as the area is large, i.e. $\ell(I)\ell(J) \gg 1$.

In particular, dividing up $\mathbf{R}^2$ into rectangles of pretty large area, we get an approximate decomposition

$$L^2(\mathbf{R}) \approx \bigoplus V_i.$$

This decomposition is meant to be the same relationship as that between representations of a group and the decomposition into orbits $\mathscr{O}$ of its action on $\mathrm{Lie}(G)^*$.

4.2. **Where we are headed.** The next goal is to show that this discussion is a special case of something that fits into the Kirillov paradigm: we'll exhibit $(L^2(\mathbf{R}), \mathbf{R}^2)$ as a pair

$$(\text{representation of some Lie group H, orbit of } H \text{ on } \mathrm{Lie}(H)^*).$$

In this situation there will be an exact Kirillov character formula, corresponding to theh formula (3.4.1)

$$\mathrm{Tr}\,\mathrm{Op}(a) = \int_{\mathbf{R}^2} a(x,\xi)\,dx\,d\xi$$

which we discussed last time.

It's not very hard to guess what H is, using that the approximate decomposition of $L^2(\mathbf{R})$ into localized functions should correspond to approximate eigenvectors for the action of H. For a function to be localized in the $\mathbf{R}^2$ picture means that both the function and its Fourier transform are localized. In the Kirillov picture, being localized means being approximately eigenvector. (Roughly speaking, dividing an orbit of H into little pieces should correspond to a basis of the representation.)

The condition that $\widehat{f}$ is localized is equivalent to $f(x) \approx e^{i\xi x}$, which is equivalent to being an *eigenvector for translation*. This suggests that the H should contain translations. Let τ_y be the translation operator

$$(\tau_y \cdot f)(t) = f(t+y).$$

We can take the Fourier transform of this condition to get the story for f. Let m_y be the multiplication operator

$$(m_y f)(x) = e^{ixy} f(x).$$

Essentially H will be the group generated by the τ_y's and m_y's. We'll now set this up abstractly.

4.3. **The Heisenberg group.**

Definition 4.3.1. The *Heisenberg group* is

$$\mathrm{Heis} := \left\{ \begin{pmatrix} 1 & x & z \\ & 1 & y \\ & & 1 \end{pmatrix} \mid x, y, z \in \mathbf{R} \right\}$$

with the group operation being matrix multiplication:

$$\begin{pmatrix} 1 & x & z \\ & 1 & y \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & x' & z' \\ & 1 & y' \\ & & 1 \end{pmatrix} = \begin{pmatrix} 1 & x+x' & z+z'+xy' \\ & 1 & y+y' \\ & & 1 \end{pmatrix}$$

Let's define some subgroups of Heis. We define

$$U_x = \begin{pmatrix} 1 & x & 0 \\ & 1 & 0 \\ & & 1 \end{pmatrix}$$

$$V_y = \begin{pmatrix} 1 & 0 & 0 \\ & 1 & y \\ & & 1 \end{pmatrix}$$

$$W_z = \begin{pmatrix} 1 & 0 & z \\ & 1 & 0 \\ & & 1 \end{pmatrix}$$

Then $\{U_x\}, \{V_y\}$, and $\{W_z\}$ all form subgroups. Furthermore, W_z is *central* in H (in fact, it is the center). These operators satisfy the commutation relation

$$U_x V_y = V_y U_x W_{xy}.$$

The Heisenberg group is the basic example of a "2-step nilpotent group": it has a filtration by a central subgroup, and the commutators land in it.

Remark 4.3.2. There is a better system of coordinates on H. We have

$$\text{Lie(Heis)} = \left\{ \begin{pmatrix} 0 & x & z \\ & 0 & y \\ & & 0 \end{pmatrix} \mid x, y, z \in \mathbf{R} \right\}$$

The matrix exponential sends this to

$$\exp \begin{pmatrix} 0 & x & z \\ & 0 & y \\ & & 0 \end{pmatrix} = \begin{pmatrix} 1 & x & z + \frac{xy}{2} \\ & 1 & y \\ & & 1 \end{pmatrix}$$

This coordinate system is more symmetric.

Fact 4.3.3. The rule

$$\begin{aligned} U_x &\mapsto \tau_x \quad (f(t) \mapsto f(x+t)) \\ V_y &\mapsto m_y \quad (f(t) \mapsto e^{iyt} f(t)) \\ W_z &\mapsto \text{scalar multiplication by } e^{2\pi i z} \end{aligned}$$

extends to a representation

$$H \to \mathrm{U}(L^2(\mathbf{R})).$$

The remark essentially gives a presentation of the Heisenberg group, so you just check that it satisfies the commutation relations. We'll see later that this is very connected with the theory of pseudo-differential operators. Furthermore, it is characterized in a very intrinsic way: the Stone-von Neumann theorem says that it is the *unique* irreducible representation where W_z acts by the scalar $e^{2\pi i z}$.

We're heading towards establishing the Kirillov character formula for Heis.

4.4. **Infinite-dimensional unitary representations.** When dealing with compact groups, there are enough finite-dimensional representations for all intents and purposes (Peter-Weyl theorem). But *noncompact* groups (e.g. the Heisenberg group) typically don't have "enough" finite-dimensional representations.

It is a little subtle to say what an infinite-dimensional representation should be. Once the representation is not finite-dimensional, you need to pick a topology, since otherwise it would be equivalent to talking about representations of the underlying *discrete group*, which is obviously not what we want to do. So we want to impose some kind of continuity on the map

$$G \to \mathrm{End}(V).$$

There is no strict need for V to be a Hilbert space, but it gives a convenient theory. So in the rest of the course, infinite-dimensional representations will always be on Hilbert spaces.

What should we mean by continuity? The model case is $G = \mathbf{R}$, acting by translation on $L^2(\mathbf{R})$. This should be a representation under whatever definition we decide on. This is not automatic: the strongest topology on the unitary group of H is the operator norm, but this is way too strong because a small translation is very far from the identity operator, so the preceding model case would not be continuous in this topology.

Definition 4.4.1. A *unitary representation of G on a (separable) Hilbert space H* is a homomorphism

$$G \to \mathrm{Unitary}(H)$$

such that the action map

$$G \times H \to H$$

is continuous.

> **Policy.** Henceforth, if we ever refer to infinite-dimensional representations then we will mean a unitary representation on a Hilbert space, as above.

Definition 4.4.2. A unitary representation H of G is *irreducible* if H has no *closed* G-invariant subspaces.

Example 4.4.3. The model example $L^2(\mathbf{R})$, with $G = \mathbf{R}$ acting by translation is not irreducible. This is a little tricky to see (you might think that constant functions form an invariant subspace, but they aren't in $L^2(\mathbf{R})$!). However, it's much clearer on the Fourier side, where the action is multiplication. The subspace of functions whose Fourier transform is supported on, for instance, $[0, \infty)$ is invariant.

Definition 4.4.4. Let $\widehat{G}$ be the set of irreducible unitary representations of G, up to isomorphism.

We will be able to understand nicely the set of irreducible representations of the Heisenberg group. We want to give you an example that shows that $\widehat{G}$ can be a really pathological object. Consider the "discrete Heisenberg group"

$$\mathrm{Heis}_{\mathbf{Z}} := \left\{ \begin{pmatrix} 1 & x & z \\ & 1 & y \\ & & 1 \end{pmatrix} \mid x, y, z \in \mathbf{Z} \right\}$$

It has "pathological" representations, so there is no nice structure on its space of representations. To see this, for $x, y, z \in \mathbf{Z}$ denote by U_x, W_y, V_z the same operators as before.

Pick $\alpha, \beta \in \mathbf{R}$. We make an action of $\Gamma_{\mathbf{Z}}$ on $L^2(\mathbf{Z})$ as follows:

$$(U_x f)(t) = f(t+x)$$
$$(V_y f)(t) = (m(t) = e^{2\pi i(\alpha t+\beta)y}) \cdot f(t)$$
$$(W_z f)(t) = e^{2\pi i \alpha z} f(t)$$

Exercise 4.4.5. Check that this defines a representation, which we call $V(\alpha,\beta)$.

Fact 4.4.6. If α is irrational, then $V(\alpha,\beta)$ is irreducible.

Fixing α, we get a family of irreducible representations parametrized by $\beta \in \mathbf{R}/\langle \mathbf{Z} + Z\alpha\rangle$. Note that $\mathbf{Z}+\mathbf{Z}\alpha$ is dense in $\mathbf{R}$, so this index set is something horrifying, and certainly doesn't have a reasonable topology.

How do you show that $V(\alpha,\beta)$ is irreducible? We have a version of Schur's Lemma for infinite-dimensional representations (of a group with countable basis):

Lemma 4.4.7 (Schur's Lemma)**.** *A unitary representation H of G is irreducible if and only if any bounded operator $A \in \operatorname{End}(H)$ commuting with the action of G is scalar.*

Proof. If H isn't irreducible, it has a closed subspace H_1 which is G-stable. Since H is a *unitary* representation, we get an orthogonal decomposition

$$H = H_1 \oplus H_1^{\perp}.$$

So we can take A to be projection to H_1 to get a non-scalar endomorphism.

What about the converse? Following the proof of the finite-dimensional version of Schur's Lemma, we would like to take an "eigenspace" of A. However, this requires some more care in the infinite-dimensional theory.

First we reduce to the case where A is self-adjoint. If A commutes with the G-action then so does its adjoint A^*:

$$gA = Ag \implies A^*g^* = g^*A^*$$

and $g^* = g$, since the G-action is unitary. So we can express A in terms of self-adjoint operators:

$$2A = (A+A^*) + i\left(\frac{A-A^*}{i}\right)$$

thus reducing to the case where A is self-adjoint. An informal expression of the spectral theorem for self-adjoint operators says:

> Let A be a self-adjoint operator on a Hilbert space H. For each Borel set $T \subset \mathbf{R}$, it makes sense to talk about H_T ="sum of eigenspaces with eigenvalues in T". (To do this, you approximate the characteristic function of T by polynomials, and apply them to A.) In particular, the spectrum of A on H_T is contained in $\overline{T}$.

Assume that H is an irreducible representation. Then $H_T = 0$ or H for each Borel set T, because it is a closed G-invariant subspace (since it is intrinsically defined). These behave as you would expect: if $T = T' \sqcup T''$ then $H_T = H_{T'} \oplus H_{T''}$.

Now we can play the following game. If we take T to be a sufficiently large interval, then $H_T = H$. Then we can divide T up and narrow in on the spectrum: we get a decreasing sequence of intervals $T_1 \supset T_2 \supset T_3 \supset \ldots$ with $H_{=}H$. The spectrum of A is then contained in the intersection of these intervals T_i, which is a single point. This implies that $A = \lambda_0 \,\mathrm{Id}$.

$\square$

Exercise 4.4.8. Check that $V(\alpha,\beta)$ is irreducible. [Hint: The idea is that the matrix of V_y is a big diagonal matrix with different entries. If you commute with it, then you have to be diagonal. If you commute with a shift as well, then your entries have to be all the same.]

Corollary 4.4.9. *If G is abelian, then every irreducible representation of G is 1-dimensional.*

Example 4.4.10. If $G = \mathbf{R}$, then the irreducible representations of all of the form $x \mapsto e^{itx}$ for some $t \in \mathbf{R}$. Thus, $\widehat{G} = \mathbf{R}$.

Corollary 4.4.11. *In any irreducible representation of G, the center $Z(G)$ acts by scalars.*

This is going to be important for us. After we understand the Kirillov formula for Heis acting on $L^2(\mathbf{R})$, we'll check it for *every* representation of Heis.

Recall that we defined subgroups

$$U_x = \begin{pmatrix} 1 & x & 0 \\ & 1 & 0 \\ & & 1 \end{pmatrix}$$

$$V_y = \begin{pmatrix} 1 & 0 & 0 \\ & 1 & y \\ & & 1 \end{pmatrix}$$

$$W_z = \begin{pmatrix} 1 & 0 & z \\ & 1 & 0 \\ & & 1 \end{pmatrix}$$

The action on $L^2(\mathbf{R})$ is given by

$$\begin{aligned} U_x &\mapsto \tau_x \quad (f(t) \mapsto f(x+t)) \\ V_y &\mapsto m_y \quad (f(t) \mapsto e^{iyt} f(t)) \\ W_z &\mapsto \text{scalar multiplication by } e^{iz} \end{aligned}$$

Remark 4.4.12. According to our preceding discussion, in any representation of Heis the operators W_z must act by scalars of this form.

In other words,

$$(V_y U_x W_z f)(t) = \begin{pmatrix} 1 & x & z \\ & 1 & y \\ & & 1 \end{pmatrix} f(t) = f(t+x) e^{it+z}.$$

This is a unitary representation of Heis on $L^2(\mathbf{R})$.

Exercise 4.4.13. Prove that this representation is irreducible. [Hint: check that an operator commuting with all V_y is a multiplication operator. Then check that commuting with U_x forces it to be constant.]

Now we want to consider the character. Recall that we are aiming for a formula

$$\mathrm{char}(e^X) = \int e^{i\langle \xi, X\rangle}\, d\xi.$$

But we have a problem. The element W_z acts as a scalar, so can't have any meaningful notion of trace. In fact, no $g \in \mathrm{Heis}$ gives a trace class operator.

However, elements of the "group algebra" will have a trace. In the end, we will get the character as a distribution.

4.5. **The group algebra.** Let G be a (locally compact) group, and an irreducible representation $\pi\colon G \to \mathrm{U}(H)$. For a finite group, the group algebra can be regarded as functions on the group. Fix a Haar measure on G. We can extend the action of G to $L^1(G)$as follows: for $A \in L^1(G)$, we define

$$\pi(A)v = \int_G A(g)(g\cdot v)\, dg.$$

Then $\pi(A)$ gives a bounded operator $H \to H$, and in fact $||\pi(A)v|| \le ||A||_{L^1} \cdot ||v||$. (The algebra structure is by convolution.)

Example 4.5.1. Let $G = \mathbf{R}$, acting on $L^2(\mathbf{R})$ as translation. Then

$$\pi(A)f(t) = \int A(x)f(t+x)\, dx.$$

If we denote $\check{A}(x) = A(-x)$, then we can rewrite this as

$$\pi(A)f(t) = (f * \check{A})(t).$$

Thus, the $L^1(\mathbf{R})$ action on $L^2(\mathbf{R})$ is basically by convolution.

Remark 4.5.2. In usual functional analysis, if you want to approximate a function by a smooth function, you can convolve with a smooth approximation to the δ function. The same works here.

From here on G is a Lie group.

Definition 4.5.3. A vector $v \in H$ is *smooth* if the map $g \mapsto gv$ defines a smooth function $G \to H$. We define the subspace of smooth vectors to be H^∞.

Then $\mathrm{Lie}(G)$ acts on H^∞. Of course, you don't see the distinction come up for finite-dimensional representations.

Fact 4.5.4. H^∞ is dense in H.

Proof. Take $A \in C_c^\infty(G)$ be an approximation to the delta function at e. Then $\pi(A)v \approx v$ is smooth. $\square$

This shows that $H^\infty = H$ in the finite-dimensional case.

Example 4.5.5. For $\mathbf{R}$ acting on $L^2(\mathbf{R})$ by translation,

$$H^\infty = \{f \in C^\infty \mid \partial^i f \in L^2(\mathbf{R}) \forall i\}.$$

You can sort of guess this because the derivative of the translation action on f looks just like the derivative of f. However, it is not completely formal because the limit is being taken in $L^2(\mathbf{R})$, not pointwise.

Example 4.5.6. For Heis acting on L^2 as above, H^∞ is the Schwartz space

$$\{f \in C^\infty(\mathbf{R}) \colon x^j \partial^i f \text{ are bounded for all } i, j\}.$$

For $f \in H^\infty$,

$$\begin{aligned}\frac{d}{dx}(U_x f) &= f' \\ \frac{d}{dy}(V_y f) &= \frac{d}{dy}(e^{ity} f) = (it)f \\ 2\frac{d}{dz}(W_z f) &= if.\end{aligned}$$

i.e. the action of

$$\begin{pmatrix} & x & z \\ & & y \\ & & \end{pmatrix} \cdot f = (xf' + ityf + izf).$$

Remark 4.5.7. If G acts unitarily, then $\operatorname{Lie} G$ acts skew-adjointly. This lets you figure out where the factors of i should be.

4.6. **Computing the character.** The hope to define the "trace" of an operator in $L^1(G)$ on H as a distribution.

Start with a function $A \in C_c^\infty(\text{Heis})$. We'll write a formula for $\pi(A) \cdot f$ for $f \in L^2(\mathbf{R})$:

$$\pi(A) \cdot f = \int A \begin{pmatrix} 1 & x & z \\ & 1 & y \\ & & 1 \end{pmatrix} f(t+x) e^{i(ty+z)}\, dx\, dy\, dz$$

To simplify this, we can first integrate out the z. Letting

$$B(x,y) = A \begin{pmatrix} 1 & x & z \\ & 1 & y \\ & & 1 \end{pmatrix} e^{iz} \in C_c^\infty(\mathbf{R}^2).$$

we have

$$\pi(A) f(t) = \int B(x,y) f(t+x) e^{ity}\, dx\, dy. \tag{4.6.1}$$

We claim that this is essentially $\operatorname{Op}(\widehat{B})f$, up to a few signs, where $\widehat{B}$ is the Fourier transform in both variables.

The equation (4.6.1) obviously involves a Fourier transform of B in the y variable. To rewrite it in terms of the Fourier transform of B in both variables, we need to put in an inverse Fourier transform in the x-variable. Denoting

$$\widehat{B}(u,t) = \int B(x,y) d^{i(xu+ty)}\, dx\, dy.$$

Equation (4.6.1) then becomes

$$\pi(A) f(t) = \int \widehat{B}(u,t) e^{-iux} f(t+x)\, dx\, du. \tag{4.6.2}$$

Integrating $e^{-iux}f(t+x)$ in x yields $e^{iut}\widehat{f}(-u)$, so (4.6.2) can be rewritten as

$$\pi(A)f(t)=\int \widehat{B}(u,t)e^{iut}\hat{f}(-u)\,du$$

which, up to signs, coincides with $\mathrm{Op}(\widehat{B})(f)$.

Now we compute the trace. In fact we saw earlier that

$$\mathrm{Tr}\,\mathrm{Op}(a)=\int a.$$

(The picture was that if a approximates the characteristic function then $\mathrm{Op}(a)$ approximates a projection.) So

$$\mathrm{Tr}\,\pi(A)\sim\int \widehat{B}(u,v)\,du\,dv=B(0,0)=\int A(0,0,z)e^{iz}\,dz.$$

In fact we were a little careless with the constants of normalization. Taking care with the factors of 2π, the result is

$$\mathrm{Tr}\,\pi(A)=2\pi\int A(0,0,z)e^{iz}\,dz. \tag{4.6.3}$$

What does this tells us about the trace of π as a distribution? If π had a character $\chi(g)$, then we would have

$$\mathrm{Tr}\,\pi(A)=\int \chi(g)A(g)\,dg.$$

So we should interpret the formula (4.6.3) as saying that

$$\chi=(2\pi)e^{iz}\delta_{z\text{-axis}}.$$

Remark 4.6.1. This in particular implies τ_x (resp. m_y) should have trace 0 if $x\neq 0$ (resp. $y\neq 0$). If you think about it, you'll see that this makes sense for τ_x: you can choose a basis in which τ_x has entries supported strictly above the diagonal. It's a little harder to see why the trace should be 0 if y is non-zero.

4.7. **The character on the Lie algebra.** We now pull the character back to the Lie algebra using the exponential map:

$$\exp\begin{pmatrix}0 & x & z\\ & 0 & y\\ & & 0\end{pmatrix}=\begin{pmatrix}1 & x & z+\frac{xy}{2}\\ & 1 & y\\ & & 1\end{pmatrix}$$

For the Heisenberg group, the exponential map enjoys a few special properties. It defines a (global) *diffeomorphism*

$$\mathrm{Lie}(\mathrm{Heis})\xrightarrow{\sim}\mathrm{Heis}.$$

Moreover, the Haar measures are also preserved:

$$dx\,dy\,dz\mapsto dx\,dy\,dz$$

i.e. the Jacobian factor j is equal to 1 in this case.

If we pull back the character by the exponential map, we get a distribution on Lie(Heis): for

$$X = \begin{pmatrix} 0 & x & z \\ & 0 & y \\ & & 0 \end{pmatrix} \implies \chi(e^X) = (2\pi) e^{iz} \delta_{z\text{-axis}}.$$

(One might object that the exponential factor should be $e^{i(z+xy/2)}$, but on the z-axis we have $x = y = 0$.)

This means that for $\psi \in C_c^\infty(\operatorname{Lie}\operatorname{Heis})$,

$$\operatorname{Tr}\left(\int \psi(X) e^X \, dX\right) = 2\pi \int \psi(0,0,z) e^{iz} \, dz.$$

Taking Fourier transform, we have

$$\chi(e^X) = \frac{1}{2\pi} \int_{\alpha,\beta \in \mathbf{R}} e^{(i\alpha x + \beta y + z)} \, d\alpha \, d\beta. \tag{4.7.1}$$

Here we are using that

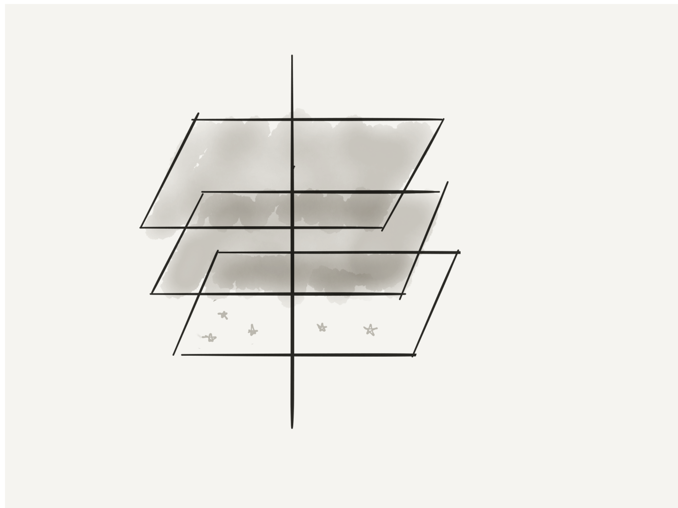
$$\begin{aligned} \int \psi(0,0,z) e^{iz} \, dz \sim \int \widehat{\psi}(\alpha, \beta, 1) \, d\alpha \, d\beta \\ &= \int_z \left(\int_{\alpha,\beta \in \mathbf{R}} e^{(i\alpha x + \beta y + z)} \psi(x,y,z) \, dx \, dy \, dz \right) d\alpha \, d\beta \\ &= \int_z \left(\int_{\alpha,\beta \in \mathbf{R}} e^{(i\alpha x + \beta y + z)} \, d\alpha \, d\beta \right) \psi(x,y,z) \, dx \, dy \, dz. \end{aligned}$$

4.8. **Orbits of** Heis **on** $(\operatorname{Lie} H)^*$**.** In terms of the x, y, z coordinates on $\operatorname{Lie} H$, put coordinates α, β, γ on $(\operatorname{Lie} H)^*$ by

$$X = \begin{pmatrix} 0 & x & z \\ & 0 & y \\ & & 0 \end{pmatrix} \mapsto \alpha x + \beta y + \gamma z. \tag{4.8.1}$$

Then H acts on $(\operatorname{Lie} H)^*$, with the following types of orbits:

- For each $\gamma_0 \in \mathbf{R} - \{0\}$, we have an orbit
$$\mathscr{O}_{\gamma_0} = \{(\alpha, \beta, \gamma) \mid \gamma = \gamma_0\}.$$
- If $\gamma_0 = 0$, then the functional (4.8.1) is actually a character of the Lie algebra (i.e. preserves the Lie bracket). So you get for each $(\alpha_0, \beta_0) \in \mathbf{R}^2$ an orbit
$$\mathscr{O}_{\alpha_0, \beta_0} = \{(\alpha_0, \beta_0, 0)\}.$$



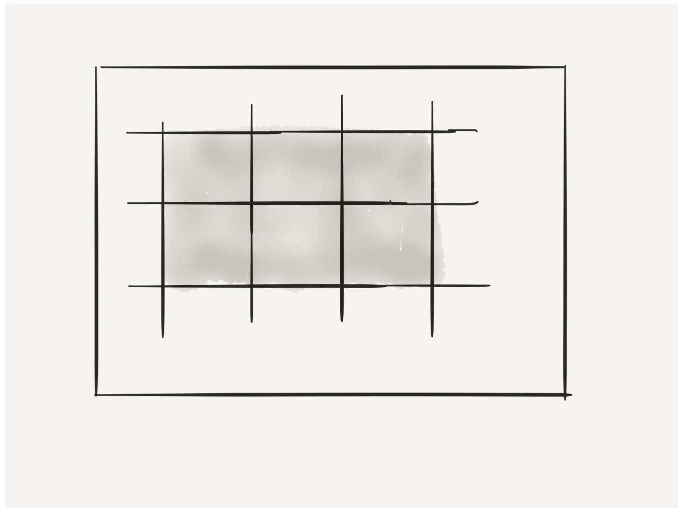
So we can rewrite (4.7.1) as

$$\chi(e^X) = \frac{1}{2\pi}\int_{\alpha,\beta\in\mathbf{R}} e^{(i\alpha x+\beta y+z)}\,d\alpha\,d\beta$$
$$= \int_{\xi\in\mathcal{O}_1} e^{i\langle\xi,X\rangle}\frac{d\alpha\,d\beta}{2\pi}.$$

4.9. **Summary.** So we've seen that the action of Heis on $L^2(\mathbf{R})$ corresponds to the orbit $\mathcal{O}_1 = \{(\alpha,\beta,1)\} \subset (\mathrm{Lie}\,H)^*$ in the sense that

$$\chi(e^X) = \int_{\mathcal{O}_1} e^{i\langle\xi,X\rangle}\,d\xi.$$

Also, from $B \in C_c^\infty(\mathcal{O}_1 = \mathbf{R}^2)$ we get an operator on $L^2(\mathbf{R})$, which realizes the heuristic picture from early on: if you decompose $\mathcal{O}_1$ into rectangles, then you get something like a basis for the representation.



If the dissection is into rectangles of area about 1, then the trace of the projection has rank 1, hence you get projection onto a rank 1 thing. (This is inaccurate, but gets more accurate as the areas get larger.) These rank 1 spaces are spanned by approximate eigenfunctions v_i in the sense that

$$X v_i \approx \langle \xi, X \rangle v_i.$$

The picture is of course not right, although it gets better as the area gets larger. The Kirillov formula is manifestation of this picture that miraculously happens to be *exact*, rather than asymptotic.

5. THE STONE-VON NEUMANN THEOREM

5.1. **Kirillov's theorem for** Heis**.**

Theorem 5.1.1. *There is a 1-1 correspondence between* $\widehat{\mathrm{Heis}}$ *(the space of irreducible unitary representations of* Heis*) and orbits $\mathcal{O}$ of* Heis *on* $\mathrm{Lie}(H)^*$*.*

The correspondence is determined by the following condition: for each $\mathcal{O}$ there exists a unique irreducible representation $\pi_{\mathcal{O}}$ satisfying the Kirillov character formula

$$\operatorname{Tr} \pi_{\mathcal{O}}(e^X) = \int_{\xi \in \mathcal{O}} e^{i\langle \xi, X \rangle}\, d\xi \tag{5.1.1}$$

Equation (5.1.1) is meant in the sense of distributions, which means concretely that for any $f \in C_c^\infty(\mathrm{Lie})$,

$$\operatorname{Tr}\left(\int f(X) e^X \, dX \right) = \int_{\mathcal{O}} \widehat{f}(\xi)\, d\xi.$$

Remark 5.1.2. We make the convention that when $\mathcal{O}$ is a point, the measure on $\mathcal{O}$ is a point mass of measure 1.

Concerning the proof: it is easy to make the representations. The hard part is the uniqueness aspect, saying that there is only one representation with a given central character. For Heis this is the "Stone-von Neumann Theorem"

Remark 5.1.3. This holds for *any* simply-connected, nilpotent Lie group, i.e. a connected Lie subgroup of an upper-triangular group

$$\begin{pmatrix} 1 & * & * & * \\ & 1 & * & * \\ & & \ddots & * \\ & & & 1 \end{pmatrix}.$$

Remark 5.1.4. In general we have a pairing between representations and conjugacy classes (trace). But you don't know how to attach representations to conjugacy classes. However, in this case the theorem is telling us that we can parametrize representations by conjugacy classes in Lie(Heis)*.

5.2. **Construction of $\pi_{\mathscr{O}}$.** For $\mathscr{O} = \mathscr{O}_{\alpha_0,\beta_0}$, the representation (necessarily 1-dimensional) is

$$\begin{pmatrix} 1 & x & z \\ & 1 & y \\ & & 1 \end{pmatrix} \mapsto e^{i(\alpha_0 x + \beta_0 y)}.$$

For $\mathscr{O} = \mathscr{O}_1$, we've seen that $L^2(\mathbf{R})$ works.

For $\mathscr{O} = \mathscr{O}_{\Gamma_0}$ for $\gamma_0 \neq 1$, we make a representation that looks very similar to the one we constructed on $L^2(\mathbf{R})$. Notice that $\mathbf{R}^*$ acts on Heis by conjugating by the matrix

$$\begin{pmatrix} u & & \\ & 1 & \\ & & 1 \end{pmatrix}.$$

This takes $\mathscr{O}_{\gamma_0} \to \mathscr{O}_{u\cdot\gamma_0}$. You can also do this explicitly by replacing "1" by γ_0 judiciously in the definition of the action on $L^2(\mathbf{R})$.

5.3. **Uniqueness.** Now we need to classify all $\pi \in \widehat{\text{Heis}}$. By Schur's Lemma 4.4.7, the center

$$\begin{pmatrix} 1 & 0 & z \\ & 1 & 0 \\ & & 1 \end{pmatrix}$$

acts by the scalar $e^{i\gamma z}$ for some γ. If $\gamma = 0$, then the representation factors through $\text{Heis}/Z(\text{Heis}) \cong (\mathbf{R}^2, +)$. Since this is abelian, all the representations are 1-dimensional.

If $\gamma \neq 0$, then we may as well assume that $\gamma = 1$, by composing with an automorphism of Heis. It is a theorem of Stone - von Neumann that the only irreducible representation with $\gamma = 1$ is, up to isomorphism, the one that we wrote down on $L^2(\mathbf{R})$. We are going to describe two proofs: von Neumann's original proof, and a later proof by George Mackey.

Remark 5.3.1. Why might Stone and von Neumann been thinking about this? The motivation could have come from quantum mechanics.

If we have an action of Heis on a Hilbert space H, then we get an action of Lie(Heis) on H^∞. Now, Lie(Heis) is generated by the operators

$$U = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$V = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$W = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

which satisfy the relation $[U,V] = W$.

We know that W acts by a scalar since it is central; in our normalization it is $i\,\mathrm{Id}$. So such a representation is the same as giving two operators U, V on a Hilbert space H such that $[U,V] = i\,\mathrm{Id}$. (Note that this is impossible to achieve in finite-dimensional representations, since the trace of a commutator is 0.) This is a situation that naturally arises in quantum mechanics, with the operators corresponding to position and momentum.

5.4. **von Neumann's proof.** We begin with the following observation.

> **Observation.** For suitable Gaussian $a(x,\xi)$, meaning that a is the exponential of a quadratic function, $\mathrm{Op}(a)$ is a rank 1 projection.

We had discussed that if a approximates the characteristic function of a region of area 1, then $\mathrm{Op}(a)$ is approximately a rank 1 projection. The claim is that in one miraculous case, $\mathrm{Op}(a)$ is an exact projection.

An indication of the subtlety of this situation is that it doesn't hold for all Gaussians. A quadratic form on $\mathbf{R}^2$ has one invariant with respect to the symplectic form, which is its area, and there is exactly one value for the area that works.

Earlier we noted that for a function $B(x,y)$ on $\mathbf{R}^2$,

$$\int B(x,y)\pi \begin{pmatrix} 1 & x & 0 \\ & 1 & y \\ & & 1 \end{pmatrix} dx\,dy = \frac{1}{2\pi}\mathrm{Op}(\widehat{B}).$$

For convenience, we would like to use a slight variant of this. Note that this formula essentially depends on choosing a section $\mathbf{R}^2 \to$ Heis. This isn't the best section, for instance because it isn't stable under inversion. A better section is given by the exponential map

$$\exp \begin{pmatrix} 0 & x & 0 \\ & 0 & y \\ & & 0 \end{pmatrix} = \begin{pmatrix} 1 & x & xy/2 \\ & 1 & y \\ & & 1 \end{pmatrix}$$

So we prefer to use a slightly twisted version of the pseudo-differential operator:

$$\int B(x,y)\pi\begin{pmatrix}1 & x & xy/2\\ & 1 & y\\ & & 1\end{pmatrix}dx\,dy =: \frac{1}{2\pi}\operatorname{Op}^W(\widehat{B}).$$

(There isn't really any significant difference, however.)

von Neumann showed that if you take $C = \frac{1}{\pi}e^{-x^2-y^2}$, or any $\frac{1}{\pi}e^{-Q(x,y)}$ with $\det Q = 1$, then

$$\operatorname{Op}^W(C) = \text{ rank one projection.}$$

(This is consistent with C having total integral 1.)

Exercise 5.4.1. Check that this property breaks if you alter Q a bit. For instance, what do you get if you scale Q down first, then scale up the exponential to have area 1?

von Neumann proves this by just computing

$$\operatorname{Op}^W(C)\operatorname{Op}^W(C) = \operatorname{Op}^W(C).$$

Now let's see how this is useful. The upshot is that there exists some function B such that

$$\int B(x,y)\pi\begin{pmatrix}1 & x & xy/2\\ & 1 & y\\ & & 1\end{pmatrix}dx\,dy = \text{ rank one projection.}$$

Now take π' to be another representation with the same central character; we want to show that $\pi' \cong \pi$. We can consider the same operator, with π' in place of π. Morally, the argument is that we get a vector from this rank one projection, and we can reconstruct the representation from that vector. So let

$$P = \int B(x,y)\pi'\begin{pmatrix}1 & x & xy/2\\ & 1 & y\\ & & 1\end{pmatrix}dx\,dy.$$

von Neumann shows that

(1) P is non-zero, and is a rank one projection. Let $\operatorname{Im} P = \mathbf{C}v_0$.
(2) Compute explicitly $\langle gv_0, v_0\rangle$ for $g \in \text{Heis}$. This can be done by a universal computation in the group algebra of G; the answer is a Gaussian in the coordinates of g.

We can then recover π' from the knowledge of $\langle gv_0, v_0\rangle$ for all $g \in \text{Heis}$. Why? If we know all the matrix coefficients $\langle gv_0, v_0\rangle$, then we also know

$$\langle (a_1g_1 + \ldots + a_ng_n)v_0, (b_1g_1 + \ldots + b_ng_n)v_0\rangle.$$

From this you can reconstruct the Hilbert space: as a vector space it should consist of formal linear combinations $(\sum a_ig_i)v_0$, and since we know the matrix coefficients we know the inner product. Technically, we also need to take the completion.

5.5. **A finite model.** We are now going to move on to discuss Mackey's proof. To motivate it, we first discuss a finite model, replacing $\mathbf{R}$ by $\mathbf{Z}/p\mathbf{Z}$. Define

$$\mathrm{Heis}_p = \left\{ \begin{pmatrix} 1 & x & z \\ & 1 & y \\ & & 1 \end{pmatrix} \mid x, y \in \mathbf{Z}/p\mathbf{Z} \right\}.$$

Theorem 5.5.1. *The irreducible representations of* Heis_p *are, up to isomorphism:*

- *characters* $e^{2\pi i(ax+by)/p}$ *for* $a, b \in \mathbf{Z}/p\mathbf{Z}$*, and*
- $(p-1)$ *representations of dimension* p*, one for each non-trivial character of the center.*

Proof. If the center acts trivially, then the representation factors through $\mathrm{Heis}_p / Z(\mathrm{Heis}_p) \cong (\mathbf{Z}/p)^2$, hence is one of the characters enumerated above.

We'll show that there is a unique irreducible representation where the center acts by $e^{2\pi i z/p}$. Take such a representation V. Since we can twist any other non-trivial central character to this one, that will prove the theorem.

A general strategy for understanding representations is to first restrict them to the largest abelian subgroup. So look at the abelian subgroup

$$A = \left\{ \begin{pmatrix} 1 & 0 & z \\ & 1 & y \\ & & 1 \end{pmatrix} \right\}.$$

Then

$$V|_A \cong \bigoplus \text{characters } \psi \text{ of } A.$$

Pick one character ψ_0. There is a map $V|_A \to \psi_0$, so by Frobenius reciprocity we have a non-zero map

$$V \to \mathrm{Ind}_A^{\mathrm{Heis}_p}(\psi_0). \tag{5.5.1}$$

Fact 5.5.2. For every character ψ of A extending the central character, $\mathrm{Ind}_A^{\mathrm{Heis}_p}$ is irreducible, and (up to isomorphism) independent of ψ.

The proof of this fact is left as an exercise. Since (5.5.1) is a non-zero map of irreducible representations it must be an isomorphism, so we have characterized V, concluding the proof.

□

In fact we can be a little more precise. Since A is a *normal* subgroup of Heis_p, the set of characters appearing is acted on by Heis_p (by conjugation). There are p characters extending a given central character form a single orbit under this action, so we may as well look at a single ψ_0:

$$\begin{pmatrix} 1 & & z \\ & 1 & y \\ & & 1 \end{pmatrix} \mapsto e^{2\pi i z/p}.$$

The definition of $\mathrm{Ind}_A^{\mathrm{Heis}_p}(\psi_0)$ is the set of functions

$$f\colon \mathrm{Heis}_p \to \mathbf{C}$$

such that $f(ag) = \psi_0(a) f(g)$ for $g \in \mathrm{Heis}_p$, $a \in A$. If we restrict such a function f to

$$\begin{pmatrix} 1 & x & 0 \\ & 1 & 0 \\ & & 1 \end{pmatrix} \cong \mathbf{Z}/p$$

then we get a function in $L^2(\mathbf{F}_p)$. This is the exactly analogous to our model of the "standard representation", replacing $\mathbf{R}$ by $\mathbf{F}_p$.

In summary: to construct the unique irreducible representation of Heis_p where the center acts by some fixed non-trivial character ω, we extended ω to a larger abelian subgroup and then induced to Heis_p. It will turn out that some analog of this procedure works for all nilpotent groups.

5.6. **Disintegration into irreducibles.** We now give Mackey's proof, which is an analogue of this argument. To transpose the argument to Heis, we need an analogue of the decomposition

$$V|_A = \bigoplus \psi.$$

In other words, we need to know how to decompose a unitary representation into irreducibles.

Example 5.6.1. Consider our model example, $\mathbf{R}$ acting on $L^2(\mathbf{R})$ by translation. This is certainly not a direct sum of irreducible representations. Indeed, since $\mathbf{R}$ is abelian an irreducible representation would be 1-dimensional, and then an eigenfunction would be $f(x) = e^{i\alpha x}$. But these aren't in L^2.

So we need to enlarge the notion of "direct sum" to "direct integral".

Let G be a locally compact topological group. Recall that the *unitary dual* $\widehat{G}$ is the set of irreducible unitary representations. There is a σ-algebra of "Borel" sets on $\widehat{G}$, which we will define shortly.

To get a sense of what this σ-algebra looks like, we first consider the space of *framed* irreducible n-dimensional representations FrIrr_n. It is the set of unitary irreducible G-representations on $\mathbf{C}^n$. The space Irr_n of irreducible representations of dimension n is a quotient of FrIrr_n by an equivalent relation.

The space FrIrr_n has a reasonable topology, which is cut out by the following functions: for each $1 \le i, j \le n$ and each $g \in G$ we get a function on Irr_n, taking

$$(\pi, \mathbf{C}^n) \mapsto \langle \pi(g) v_i, v_j \rangle$$

where $\{v_i\}$ is a basis for $\mathbf{C}^n$. If $n = \infty$, then by $\mathbf{C}^n$ we mean $\ell^2(\mathbf{N})$.

This defines a Borel σ-algebra on FrIrr_n, which we can then push down to get a σ-algebra on $\widehat{G}$. The resulting thing can be quite pathological, however. The problem is the quotient process: it can look like a quotient by a dense equivalence relation. Recall that we saw this issue arise for the representations of the discrete Heisenberg group, in §4.4.

We now spell out the result of this discussion:

Definition 5.6.2. Let Irr_n be the set of homomorphisms $G \to U(\mathbf{C}^n)$ which define irreducible representations. (For $n = \infty$, we mean $\ell^2(\mathbf{N})$ for $\mathbf{C}^\infty$.) We define a σ-algebra on $\mathrm{Irr}(n)$ by requiring all the coordinate functions to be measurable. (The coordinate functions are the matrix coefficients indexed by group elements and bases of $\mathbf{C}^n$). A set $S \subset \widehat{G}$ is *measurable* if for every n, its pre-image under $\mathrm{Irr}(n) \to \widehat{G}$ is measurable.

This σ-algebra is called the *Borel σ-algebra.*

Remark 5.6.3. In good cases one can put a topology on $\widehat{G}$ such that this is the Borel σ-algebra in the usual sense.

We now make a key assumption on G that rules out pathologies.

> **Assumption (*).** For every irreducible unitary representation π and every function $f \in L^1(G)$, the operator $\pi(f)$ is compact (i.e. a limit of finite-rank operators in the norm topology).

This assumption implies that there is a good theory of decomposition of representations. For example, it implies that the σ-algebra on $\widehat{G}$ is "nice". (A reference for all this is Dixmier's book on C^*-algebras [Dix96].)

Theorem 5.6.4. *Under the assumption (*),*

(1) The map
$$\bigsqcup_n \mathrm{FrIrr}(n) \to \widehat{G}$$
has a Borel section.

(2) Any unitary representation of G on a Hilbert space H is isomorphic to a direct integral
$$\bigoplus_{m \in \mathbf{N} \cup \{\infty\}} \left[\int_{\widehat{G}} \pi \, d\nu_m \right]^{\oplus m}$$
for some mutually singular measures ν_m on $\widehat{G}$.

(3) The measures ν_m are uniquely determined up to equivalence. (The equivalence relation is $\nu \sim \nu'$ if each is absolutely continuous with respect to the other, meaning that they share the same measure-0 sets.)

Remark 5.6.5. We make some remarks on what the statement of the theorem means.

(1) Concretely this means that we can make a model for each representation, which varies in a nice way.

(2) We'll explain what a direct integral of representations is shortly. When $G = \mathbf{Z}$, this recovers the spectral theorem for unitary operators. However, it looks slightly different because the usual statement of the spectral theorem doesn't have a uniqueness clause. Note that the index m represents the multiplicity of the representation.

Direct integrals. What is the meaning of a direct integral of representations? By (1) we can choose a Borel section of the space of irreducible representations, which means that there is a reasonable notion of a measurable assignment from an irreducible representation $\pi \in \widehat{G}$ to a vector v_π in the space of π.

Definition 5.6.6. We define the direct integral

$$\int \pi \, d\nu$$

to be the space of measurable assignments $\pi \rightsquigarrow v_\pi$, completed for the norm

$$\int |v_\pi|^2 \, d\nu.$$

Then G acts on an assignment $(\pi \rightsquigarrow v_\pi)$ by acting pointwise:

$$g \cdot (\pi \rightsquigarrow v_\pi) = (\pi \rightsquigarrow g \cdot v_\pi).$$

Remark 5.6.7. If you replace ν by an equivalent ν', the resulting directing integrals are isomorphic. Imagine the simplest case where $\nu = 2\nu'$; then we can just scale the norm to make the two spaces isomorphic. In general, if ν and ν' are equivalent then morally they differ by a reasonable function at all points of $\widehat{G}$, and we can use this to scale the norms to be equivalent.

Example 5.6.8. Let $G = \mathbf{R}$, so $\widehat{G} = \mathbf{R}$, with $\xi \in \mathbf{R}$ corresponding the representation π_ξ of G on $\mathbf{C}$ where x acts as multiplication by $e^{i\xi x}$. The Borel σ-algebra here turns out to coincide with the standard one. Take the standard norm on $\mathbf{C}$.

Consider the direct integral with respect to Lebesgue measure:

$$\int \pi_\xi \, d\xi.$$

What is this? As a set it consists of assignments: for each ξ an element of the space of π_ξ which we identify with $\mathbf{C}$, with finite norm. So as a vector space the direct integral is identified with the set complex-valued function $f(\xi)\colon \mathbf{R} \to \mathbf{C}$ such that

$$\int |f(\xi)|^2 \, d\xi < \infty.$$

How does the group act? By replacing $f(\xi)$ with $e^{ix\xi} f(\xi)$. This is isomorphic to the G-action on $L^2(\mathbf{R})$ by translation; the isomorphism is given by the Fourier transform.

In summary, the Fourier transform gives an explicit isomorphism

$$(L^2(\mathbf{R}), \text{translation}) \cong \int_{\xi \in \widehat{G}} \pi_\xi \, d\xi.$$

Exercise 5.6.9. What does the theorem say for $G = \mathbf{Z}$?

Example 5.6.10. We want to illustrate the kind of pathology that can occur with the "$\pi(f)$ compact" assumption. Once you drop this condition, the structure of $\widehat{G}$ becomes very bad and you also lose any hope of uniqueness.

Let's go back to the discrete Heisenberg group $\mathrm{Heis}_{\mathbf{Z}}$. One presentation for it is that it's generated by operators

$$U = \begin{pmatrix} 0 & 1 & 0 \\ & 1 & 0 \\ & & 1 \end{pmatrix}$$
$$V = \begin{pmatrix} 0 & 0 & 0 \\ & 1 & 1 \\ & & 1 \end{pmatrix}$$
$$W = \begin{pmatrix} 0 & 0 & 1 \\ & 1 & 0 \\ & & 1 \end{pmatrix}$$

with the relation $[U, V] = W$.

Consider the representation V_β of $\mathrm{Heis}_{\mathbf{Z}}$ on $L^2(\mathbf{R})$ given by

$$\begin{aligned} Uf(t) &= f(t+1) \\ Vf(t) &= e^{i\beta t} f(t) \\ Wf(t) &= e^{i\beta} f. \end{aligned}$$

Take β to be irrational. (We previously constructed a similar representaiton on $L^2(\mathbf{Z})$.)

We claim that this representation is not irreducible. Take $J \subset [0,1)$, and we can take the set of functions f whose support is contained in $J + \mathbf{N}$. Such a support condition is obviously preserved by the $\mathrm{Heis}_{\mathbf{Z}}$-action, so this is a closed invariant subspace.

We want to decompose V_β into irreducibles. We can decompose it by shrinking J. Morally we want to shrink J down to a point. If we do this, we'll get that V_J is a direct integral of the spaces where J is a point. But are really just like the representations $V(\alpha, \beta)$ that we discussed earlier in §4.4: they are irreducible but there are many isomorphisms between them. The space of irreducible representations looks like $\mathbf{R}$ quotiented out by a dense subset.

So we do get a decomposition into irreducibles, but it looks when we try to express it as a direct integral over $\widehat{\mathrm{Heis}_{\mathbf{Z}}}$ since the collapsing of isomorphism type is very violent. There's a second problem, which is that this decomposition is not unique. If we take the Fourier transform we get a similar decomposition, since the roles of U and V are reversed, but the two decompositions don't have anything to do with each other.

Remark 5.6.11. Another way to see that this space is not irreducible is to write down non-trivial operators that commute with it. Clearly any translation commutes with U. If we translate by the right amount, it will translate with V as well. Define $U' : f(t) \mapsto f(t + 2\pi/\beta)$; this commutes with U and V (hence also with W). Similarly, if we define $V' : f(t) \mapsto e^{it} f(t)$, then it commutes with the $\mathrm{Heis}_{\mathbf{Z}}$-action. Thus we find two representations of $\mathrm{Heis}_{\mathbf{Z}}$ that commute with each other!

5.7. **Mackey's proof.** Let π be a representation of Heis, where the center acts as

$$\begin{pmatrix} 1 & & z \\ & 1 & \\ & & 1 \end{pmatrix} \mapsto \text{multiplication by } e^{iz}.$$

Let σ_α be the representation of the (commutative) subgroup

$$A := \left\{ \begin{pmatrix} 1 & & z \\ & 1 & y \\ & & 1 \end{pmatrix} \right\}$$

on $\mathbf{C}$ given by the character $e^{i(\alpha y+z)}$. The Decomposition Theorem 5.6.4 implies that

$$\pi|_A \cong \bigoplus_m \left[\int \sigma_\alpha \, d\,\nu_m(\alpha) \right]$$

where ν_m is a measure on $\mathbf{R}$. (We've snuck a step here: $\widehat{A} \cong \mathbf{R} \times \mathbf{R}$, but the subset corresponding to a fixed central character is just $\mathbf{R}$.)

Only one summand is non-zero, because the splitting is intrinsic and the subgroup A is normal. (Conjugation by G preserves the multiplicities, so each multiplicity summand defines a G-invariant subspace of π.) Say m_0 is the non-zero contributing multiplicity.

Consider how the U operator acts. Conjugation by

$$\begin{pmatrix} 1 & x & \\ & 1 & \\ & & 1 \end{pmatrix}$$

sends σ_α to $\sigma_{\alpha+x}$. On the other hand this conjugation doesn't change the isomorphism type of π, so we conclude that ν_m is invariant under translation by x, for all x. That implies that ν_m must be equivalent to the Lebesgue measure. So at this point we have shown that

$$\pi|_A \cong \left(\int \sigma_\alpha \, d\alpha \right)^{\bigoplus m}.$$

Now we just need to show that $m = 1$.

What is the representation

$$\int \sigma_\alpha \, d\alpha?$$

To specify a vector in this representaiton, you have to specify a complex number for each $\mathbf{C}$, with the collection having bounded norm, so as a vector space this is just $L^2(\mathbf{R})$. You can check that the action is given by

$$\begin{pmatrix} 1 & & z \\ & 1 & y \\ & & 1 \end{pmatrix} f(t) = e^{i(ty+z)} f(t).$$

We want to show that $m = 1$, and that

$$\begin{pmatrix} 1 & x & \\ & 1 & \\ & & 1 \end{pmatrix} \text{ acts by translation } \tau_x f(t) = f(t+x).$$

Let

$$U_x = \text{ action of } \begin{pmatrix} 1 & x & \\ & 1 & \\ & & 1 \end{pmatrix}$$

$$V_y = \text{action of } \begin{pmatrix} 1 & & \\ & 1 & y \\ & & 1 \end{pmatrix}$$

Then

$$[U_x, V_y] = e^{ixy}$$

and the same commutation relation holds for τ_x:

$$[\tau_X, V_Y] = e^{ixy}.$$

So this tells us that $\tau_{-x} U_x$ commutes with A. As we've discussed, anything that commutes with A must be a multiplication operator, so $\tau_{-x} U_x$ is of the form

$$f(t) \mapsto m_x(t) f(t)$$

where $m_x(t) \in \mathrm{U}(\mathbf{C}^n)$. Let's re-index slightly: write $M(t, x+t) := m_x(t)$. The fact that $U_x U_y = U_{x+y}$ translates into a cocycle condition $M(a,b)M(b,c) = M(a,c)$. Taking $b = 0$, we get

$$M(a,0)M(0,c) = M(a,c).$$

Remark 5.7.1. The careful reader will notice that we're being a little sloppy here, because the functions are really only up to equivalence (away from a measure 0 set). This is problematic because we're restricting to a measure 0 set. However, it works for almost all values of "0".

This implies $M(a,c) = M(a,0)M(c,0)^{-1}$. So if we replace $f(t)$ by $M(t,0)f(t)$, then in these coordinates U_x acts by τ_x.

The conclusion is that $\pi = L^2(R, \mathbf{C}^m)$ with our favorite action

$$\begin{pmatrix} 1 & x & z \\ & 1 & y \\ & & 1 \end{pmatrix} f(t) = f(t+x) e^{i(ty+z)}.$$

This is $L^2(R)^{\oplus m}$, so irreducibility forces $m = 1$.

6. THE WEIL REPRESENTATION

6.1. **Functoriality of the Kirillov correspondence.** We've now completed the dictionary between representations of Heis and orbits in Lie(Heis)*:

$$\pi_{\mathscr{O}} \longleftrightarrow \mathscr{O}.$$

We next ask the question: how functorial is this? We'll see that this leads somewhere interesting.

More precisely, we'll shortly see that $\mathrm{SL}_2(\mathbf{R})$ acts on Heis, preserving all the orbits $\mathscr{O}$ of positive dimension. This raises the question: does $\mathrm{SL}_2(\mathbf{R})$ also act on $\pi_{\mathscr{O}}$?

We'll now shift to a more convenient set of coordinates for Heis. Write

$$(x, y; z) := \begin{pmatrix} 1 & x & \frac{xy}{2}+z \\ & 1 & y \\ & & 1 \end{pmatrix} = \exp\left(\begin{pmatrix} & x & z \\ & & y \\ & & \end{pmatrix}\right).$$

In these coordinates the multiplication is

$$(x, y; z)\cdot(x', y'; z') := (x+x', y+y'; z+z'+\frac{xy'-x'y}{2}).$$

The quantity $\frac{xy'-x'y}{2}$ is a symplectic form! It is the pairing given by

$$\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} & 1 \\ -1 & \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix}.$$

It follows that for $g \in \mathrm{SL}_2(\mathbf{R})$, the rule $[x, y] \mapsto g[x, y]$ gives an automorphism of Heis.

Remark 6.1.1. More generally, if V is a vector space over $\mathbf{R}$ with symplectic form $\langle x, y\rangle$ then we can define a group $\mathrm{Heis}_V = V \oplus \mathbf{R}$ with multiplication

$$(v; t)(v'; t) = (v+v', t+t'+\frac{\langle v, v'\rangle}{2}).$$

Then $\mathrm{Lie}(\mathrm{Heis}_V) = V \oplus \mathbf{R}$, with the Lie bracket being

$$[(v; t), (v', t')] = (0; \langle v, v'\rangle).$$

Then $\mathrm{Sp}(V) = \{g \in \mathrm{GL}(V) \mid \langle gv, gw\rangle = \langle v, w\rangle\}$ acts on Heis_V.

Everything we've said generalizes to these types of groups. In particular, there is a unique irreducible representation corresponding to a fixed non-trivial central character.

Now check that $\mathrm{SL}_2(\mathbf{R})$ preserves each positive-dimensional orbit. This implies that for every $g \in \mathrm{SL}_2(\mathbf{R})$, we must have

$$\pi_{\mathscr{O}}^g \cong \pi_{\mathscr{O}}$$

where

$$\pi_{\mathscr{O}}^g(h) = \pi_{\mathscr{O}}(g\cdot h).$$

This implies that there exists some $A_g \in$ Unitary such that

$$\pi_{\mathscr{O}}(g(h)) = A_g \pi_{\mathscr{O}}(h) A_g^{-1}.$$

This A_g is unique up to scalars, by Schur's Lemma. How do they multiply? It is easy to check that $A_g A_{g'}$ has the same formal property as $A_{gg'}$. So

$$A_g A_{g'} = (\text{scalar}) \cdot A_{gg'}.$$

Moreover, this scalar is in S^1, since the operators are unitary. (Obviously we couldn't have done any better, since A_g is only defined up to scalars in S^1.)

The upshot is that we get a *projective* representation

$$\mathrm{SL}_2 \to U(L^2(\mathbf{R}))/S^1$$

sending $g \mapsto A_g$.

Let $G = \{g \in \mathrm{SL}_2(\mathbf{R}),\ \text{choice of } A_g\}$. This is a Lie group under composition (although technically we haven't justified this yet), and it has a map

$$G \to \mathrm{SL}_2(\mathbf{R})$$

whose kernel is a central S^1. This G acts on $L^2(\mathbf{R})$.

So we have an extension

$$1 \to S^1 \to G \to \mathrm{SL}_2(\mathbf{R}) \to 1.$$

This gives an extension at the level of Lie algebras:

$$0 \to \mathbf{R} \to \mathrm{Lie}(G) \to \mathrm{Lie}(\mathrm{SL}_2\,\mathbf{R}) \to 0$$

which is (uniquely) *split*. This is a general fact about central extensions of semisimple (or even reductive) Lie algebras, but here is a way to see it concretely (which also generalizes). Let

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$
$$f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$
$$h = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$$

be the standard basis of $\mathfrak{sl}_2(\mathbf{R})$, which satisfies the commutation relations

$$[h,e] = 2e$$
$$[h,f] = -2f$$
$$[e,f] = h.$$

To lift $\mathfrak{sl}_2$ to Lie(g): note that for any $x, y \in \mathfrak{sl}_2$ and lifts $\widetilde{x}, \widetilde{y} \in \mathrm{Lie}(G)$, the Lie bracket $[\widetilde{x}, \widetilde{y}]$ is *independent* of lift, because the ambiguity in the lift is central. So we can lift commutators uniquely, but everything in $\mathfrak{sl}_2$ is a commutator.

Exercise 6.1.2. Check that this defines a lift of the Lie algebra.

At the group level, this implies that we can find a lift $\widetilde{\mathrm{SL}_2}(\mathbf{R}) \to G$ from the universal cover $\widetilde{\mathrm{SL}_2}(\mathbf{R})$ of $\mathrm{SL}_2(\mathbf{R})$. The conclusion is that we get

$$\rho : \widetilde{\mathrm{SL}_2}(\mathbf{R}) \to \mathrm{Unitary}(L^2(\mathbf{R}))$$

such that

$$\rho(g)\pi_{\mathscr{O}}(h)\rho(g^{-1}) = \pi_{\mathscr{O}}(\overline{g}\cdot h) \tag{6.1.1}$$

where $\overline{g}$ is the image in $\mathrm{SL}_2\,\mathbf{R}$. To see the nature of this universal cover, note that $\mathrm{SL}_2(\mathbf{R})$ has a map to the bottom row

$$\begin{array}{ccc} \mathrm{SL}_2(\mathbf{R}) & & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ \big\downarrow & & \big\downarrow \\ \mathbf{R}^2 - 0 & & (c,d) \end{array}$$

whose fibers are lines, which are contractible. So there is a map $(c,d)\colon \mathrm{SL}_2\,\mathbf{R} \to \mathbf{R}^2 - \{0\}$ showing $\pi_1(\mathrm{SL}_2\,\mathbf{R}) = \mathbf{Z}$.

Remark 6.1.3. The group $\widetilde{\mathrm{SL}_2}(\mathbf{R})$ doesn't have any faithful finite-dimensional representations.

So far everything is formal. Now a miracle happens: ρ is trivial on $2\mathbf{Z}$, so it descends to a double cover of $\mathrm{SL}_2\,\mathbf{R}$:

$$0 \to \mathbf{Z}/2 \to \mathrm{Mp}_2(\mathbf{R}) \to \mathrm{SL}_2(\mathbf{R}) \to 0.$$

Remark 6.1.4. Mp_2 is the "metaplectic" group. This is parallel to the Spin double covers of special orthogonal groups.

6.2. **Summary.** We have constructed a "projective representation"

$$\rho\colon \mathrm{SL}_2(\mathbf{R}) \to \mathrm{U}(L^2(\mathbf{R}))/S^1$$

or alternately a representation of the universal cover

$$\widetilde{\rho}\colon \widetilde{\mathrm{SL}}_2(\mathbf{R}) \to \mathrm{U}(L^2(\mathbf{R})).$$

This is *characterized* by the relation

$$\rho(g)\pi(h)\rho(g^{-1}) = \pi(g(h)).$$

We claim that this miraculously factors through the double cover of $\mathrm{SL}_2(\mathbf{R})$, which we will show later.

6.3. **A heuristic picture.** For $g \in \mathrm{SL}_2(R)$ and $h \in \mathrm{Heis}$, we had

$$\rho(g)\pi(h)\rho(g^{-1}) = \pi(g\cdot h).$$

This implies that for any $B \in L^1(\mathrm{Heis})$, we have

$$\rho(g)\pi(B)\rho(g^{-1}) = \pi(g\cdot B).$$

We saw that $\pi(B)$ had something to do with $\mathrm{Op}(B)$. So we conclude that for $a \in C_c^\infty(\mathbf{R}^2)$,

$$\rho(g)\,\mathrm{Op}(a)\rho(g^{-1}) \sim \mathrm{Op}(g\cdot a).$$

Remark 6.3.1. We only get equality if we replace Op with the Weyl-twisted verison Op^W:

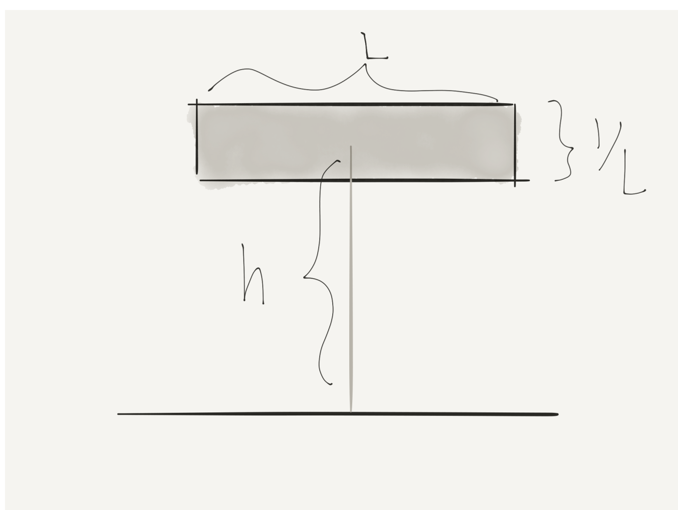
$$\rho(g)\mathrm{Op}^W(a)\rho(g^{-1}) = \mathrm{Op}^W(g\cdot a)$$

where

$$\mathrm{Op}^W(\hat{a}) = \int a(x,y)\pi(x,y;xy/2)\,dx\,dy$$

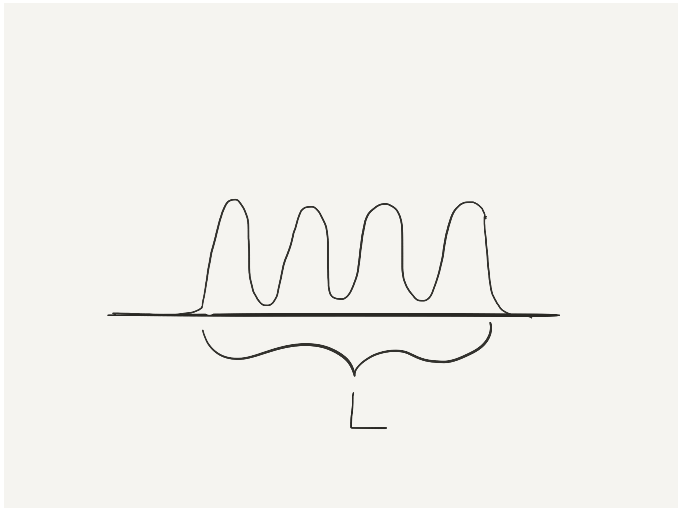
We will present some heuristic reasoning to figure out what, roughly speaking, $\rho(g)$ should look like for each g.

Let a be the characteristic function of a rectangle of length L and height $1/L$, centered at y-coordinate h.



Then $\mathrm{Op}(a)$ is approximately a projector onto functions which are localized on this rectangle, i.e. which have support in an interval I of length L and Fourier support in an interval of length $1/L$. For large L, this means that $\mathrm{Op}(a)$ should be an approximate projector onto the line spanned by the function: e^{ihx} truncated to the interval I.

You can pretty much guess the formulas we're about to write down by thinking in terms of these pictures.



- First consider $g \in \mathrm{SL}_2(\mathbf{R})$ of the form

$$g = \begin{pmatrix} s & \\ & s^{-1} \end{pmatrix}.$$

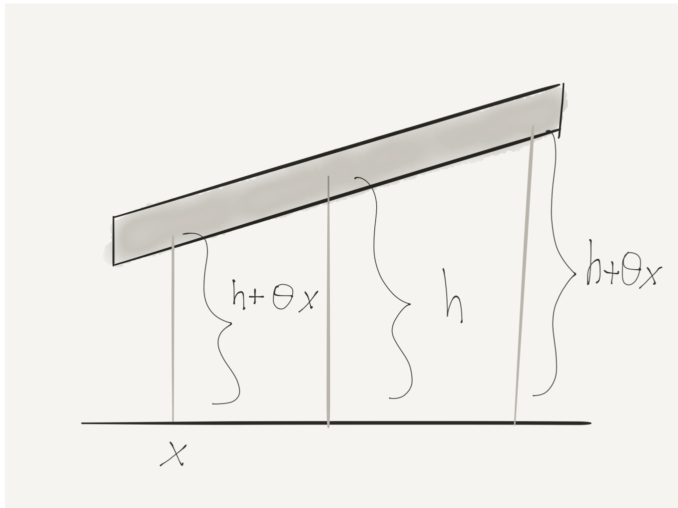
Then $g \cdot a = a \circ g$ is the characteristic function of a stretched rectangle of length as and height $1/sL$ centered at height sh. So $\rho(g)$ should be approximately

$$\rho \begin{pmatrix} s & \\ & s^{-1} \end{pmatrix} \sim f(x) \mapsto \frac{1}{\sqrt{s}} f(x/s)$$

- Next consider g of the form

$$g = \begin{pmatrix} 1 & \\ \theta & 1 \end{pmatrix}.$$

Then $\rho(g)$ should be an approximate projector onto the line spanned by $e^{ih'(x)x}$ where the frequency $h'(x)$ now varies with x.



Since the height is $h+\theta x$, we should have $h'(x) = h+\theta x$ the upshot is that $\rho(g)$ should be an approximate projector onto $e^{i(h+\theta x)x}$. Since $\mathrm{Op}(a)$ is itself a projector onto e^{ihx}, we conclude that

$$\rho\begin{pmatrix} 1 & \\ \theta & 1 \end{pmatrix} \sim \text{ multiplication by } e^{-i\theta x^2/2}.$$

- Finally we consider $\rho(g)$ for $g = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}$. Since this interchanges the support and Fourier support, we guess that $\rho(g)$ is something like the Fourier transform.

6.4. **The Weil representation.** We're now going to write down explicit formulas for $\rho(g)$, guessed from the preceding heuristics. For $g \in \mathrm{SL}_2(\mathbf{R})$, we need to write down $\rho(g)$ such that

$$\rho(g)\pi(h)\rho(g^{-1}) = \pi(g\cdot h) \tag{6.4.1}$$

or in other words the following diagram commutes up to scalars:

$$\begin{tikzcd} L^2(\mathbf{R}) \arrow[r, "\varphi\mapsto\varphi(t+x)e^{ity}"] \arrow[d, "\rho(g)"] & L^2(\mathbf{R}) \arrow[d, "\rho(g)"] \\ L^2(\mathbf{R}) \arrow[r, "\varphi\mapsto\varphi(t+x')e^{ity'}"] & L^2(\mathbf{R}) \end{tikzcd}$$

where $[x', y'] = g[x, y]$.

Example 6.4.1. Let's use this to compute $\rho(g)$ for

$$g = \begin{pmatrix} 1 & \\ \theta & 1 \end{pmatrix}$$

so $[x', y'] = [x, \theta x + y]$.

If $x=0$, then $[x',y']=[x,y]$ so the condition (6.4.1) is that $\rho(g)$ commutes with multiplication by e^{ity}, which implies - as we have discussed -

$$\rho\begin{pmatrix}1 & \\ \theta & 1\end{pmatrix} = \text{multiplication by some } m(t).$$

To figure out the function $m(t)$, we make the diagram commute:

$$m(t)=m(t+\theta x)e^{i\theta t x}.$$

So $m(t)=e^{-i\theta t^2/2-i\theta x^2/2}$. Since really $m(t)$ is only defined up to a scalar in S^1, this is just as good as $m(t)=e^{-i\theta t^2/2}$.

We claim that the following formulas work:

$$\rho\begin{pmatrix}s & \\ & s^{-1}\end{pmatrix}: \varphi(t)\mapsto \frac{1}{\sqrt{s}}\varphi(t/s)$$
$$\rho\begin{pmatrix}1 & \\ \theta & 1\end{pmatrix}: \varphi(t)\mapsto e^{-i\theta t^2/2}\varphi(t)$$
$$\rho\begin{pmatrix} & 1\\ -1 & \end{pmatrix}: \varphi(t)\mapsto \widehat{\varphi}(t)$$

We can check them by explicit computation. In fact these matrices generate $\mathrm{SL}_2(\mathbf{R})$, so they tells us the action of all of $\mathrm{SL}_2(\mathbf{R})$. However, it's easier to just go to the Lie algebra.

6.5. **Action of the Lie algebra.** When a Lie group G acts on a Hilbert space H, we get a Lie algebra action of $\mathrm{Lie}(G)$ on the (dense) subspace of smooth vectors $H^\infty\subset H$, by the rule: for $X\in \mathrm{Lie}(G)$,

$$X\cdot v=\frac{d}{du}|_{u=0}(e^{uX}\cdot v)$$

Let's determine the action of $\mathrm{Lie}(\mathrm{SL}_2)$ on $L^2(\mathbf{R})$. The Lie algebra is generated by

$$h=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$
$$f=\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}$$
$$e=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$$

The action of $h\in \mathrm{Lie}(\mathrm{SL}_2)$ is given by

$$\varphi\mapsto \frac{d}{du}|_{u=0}\left(e^{-u/2}\varphi(e^{-u}t)\right)=\left(-\frac{1}{2}-t\frac{d}{dt}\right)\varphi.$$

The action of f is given by

$$\varphi\mapsto \frac{d}{du}|_{u=0}\left(e^{-iut^2/2}\varphi(t)\right)=\frac{-it^2}{2}\varphi(t).$$

The action of e is essentially the action of f on the Fourier side. So the conclusion is that

$$h\cdot\varphi(t)=\left(-\frac{1}{2}-t\frac{d}{dt}\right)\varphi$$
$$f\cdot\varphi(t)=\frac{-it^2}{2}\varphi(t)$$
$$e\cdot\varphi(t)=-\frac{iD^2}{2}\varphi.$$

Actually what we've been doing is not quite well-defined, because we only had a representation of $\mathrm{SL}_2(\mathbf{R})$ up to scalars. However, we know that there is a way to promote this to an honest representation of $\mathrm{Lie}(\mathrm{SL}_2)$, because there is an honest representation of $\widetilde{\mathrm{SL}}_2(\mathbf{R})$ which has the same Lie algebra.

To pin down the representation on the nose, we just need to check that the correct commutation relations are satisfied. In fact it will turn out that we were fortunate enough to get the scalars right in the first place.

Recall that we constructed the representation $\widetilde{\mathrm{SL}}_2$ by lifting the commutator relations $[h,e]=2e$, $[h,f]=-2f$, and $[e,f]=h$. To confirm that our formulas are correct, we just need to verify that these relations hold. For instance,

$$[e,f]\varphi(t)=-\frac{1}{4}((t^2f)''-t^2\cdot f'')=-1/4(2f+2tf')$$

while

$$h\varphi(t)=-\frac{1}{2}(\varphi+t\varphi').$$

We can now check that ρ descends from $\widetilde{\mathrm{SL}_2}$ to the double cover. We have the defining extension

$$0\to\mathbf{Z}\to\widetilde{\mathrm{SL}}_2(\mathbf{R})\to\mathrm{SL}_2(\mathbf{R})\to 0. \tag{6.5.1}$$

We have the subgroup $\mathrm{SO}_2(\mathbf{R})\subset\mathrm{SL}_2\mathbf{R}$, and if we pull this back along (6.5.1) then we get (in suitable coordinates):

$$0\to\mathbf{Z}\to\mathbf{R}\to\mathrm{SO}_2\mathbf{R}\to 0.$$

So we know that $\theta=2\pi$ generates the kernel of $\widetilde{\mathrm{SL}}_2\to\mathrm{SL}_2$. Let Z be the action of $\theta=2\pi$. We know that it acts as a scalar in $L^2(\mathbf{R})$ since it projects to the identity in $\mathrm{SL}_2(\mathbf{R})$. Let's compute this scalar.

We have

$$\mathrm{Lie}(\mathrm{SO}_2(\mathbf{R}))=\mathbf{R}\begin{pmatrix}0&1\\-1&0\end{pmatrix}$$

which acts by

$$\varphi\mapsto-\frac{i}{2}(D^2\varphi-t^2\varphi).$$

The function $\varphi(t)=e^{-t^2/2}$ is an eigenfunction (with eigenvalue 1) of the operator, by the calculation:

$$D^2\varphi(t)=(-te^{-t^2/2})'=(t^2-1)e^{-t^2/2}.$$

So the conclusion is that
$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot \varphi(t) = i/2 \cdot \varphi(t).$$
Then
$$\exp\begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix} \varphi(t) = e^{i\theta/2}\varphi(t).$$
(Similar conclusion the other eigenvectors.) So the element
$$Z = \exp\begin{pmatrix} 0 & 2\pi \\ -2\pi & 0 \end{pmatrix}$$
acts as multiplication by -1.

Now let's try to understand Lie(Heis) as a representation of Lie(SL_2).
$$\mathrm{Lie(Heis)} = \begin{pmatrix} 0 & x & z \\ & 0 & y \\ & & 0 \end{pmatrix}$$
is generated by
$$U = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$V = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$
$$W = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
with the commutation relation $[U, V] = W$. Because U generates translation and V generates multiplication by e^{ity}. So in terms of the action on $L^2(\mathbf{R})$, we have:
$$f \longleftrightarrow \frac{i}{2}V^2$$
$$e \longleftrightarrow -\frac{i}{2}U^2$$
and (writing the h more symmetrically)
$$h \longleftrightarrow \frac{i}{2}(UV + VU).$$

Remark 6.5.1. More generally, we can identify
$$\mathrm{Lie}(\mathrm{Sp}_V) \cong \mathrm{Sym}^2(V)$$
Indeed, $\mathrm{Lie}(\mathrm{Sp}_V)$ consists of $A \in \mathrm{End}(V)$ that infinitesimally preserve the symplectic form on V:
$$\langle Ax, y\rangle + \langle x, Ay\rangle = 0.$$

From A we can make a quadratic form on V: $\langle Ax, x\rangle$. Using the symplectic form, we can identify this with an element of $\mathrm{Sym}^2(V)$.

6.6. **Action of a general element.** The formulas

$$\rho\begin{pmatrix} s & \\ & s^{-1}\end{pmatrix} : \varphi(t) \mapsto \frac{1}{\sqrt{s}}\varphi(t/s)$$

$$\rho\begin{pmatrix} 1 & \\ \theta & 1\end{pmatrix} : \varphi(t) \mapsto e^{-i\theta t^2/2}\varphi(t)$$

$$\rho\begin{pmatrix} & 1 \\ -1 & \end{pmatrix} : \varphi(t) \mapsto \widehat{\varphi}(t)$$

specify the action of all of $\mathrm{SL}_2(\mathbf{R})$, since the special matrices

$$\left\langle \begin{pmatrix} s & \\ & s^{-1}\end{pmatrix}, \begin{pmatrix} 1 & \\ \theta & 1\end{pmatrix}, \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}\right\rangle_{s,\theta}$$

generate all of $\mathrm{SL}_2(\mathbf{R})$.

Nonetheless, it will be useful to explain the what the action of a general element of $\mathrm{SL}_2(\mathbf{R})$ looks like. Consider the map

$$L^2(\mathbf{R}) \to L^2(\mathbf{R})$$

given by

$$f(x) \mapsto \int f(x) e^{iQ(x,y)}\, dx$$

where $Q(x,y)$ is a (real-value) quadratic form on $\mathbf{R}^2$. Strictly speaking, one might complain that the right side doesn't clearly lie in $L^2(\mathbf{R})$, but the following discussion will make it clear what we mean by this.

If

$$Q(x,y) = ax^2 + bxy + cy^2$$

(assume $b \neq 0$) then

$$f \mapsto f(x) e^{i(ax^2+bxy+cy^2)}\, dx$$

is the composite of three operations:

$$f(x) \mapsto \underbrace{e^{iax^2} f(x)}_{h(x)} \mapsto \underbrace{\int h(x) e^{ibxy}}_{g(y)} \mapsto e^{icy^2} g(y).$$

In fact we can even think of factoring up the middle step into Fourier transform followed by rescaling by b. Each of the steps is the action of some special matrix in $\mathrm{SL}_2(\mathbf{R})$.

In other words, the map

$$f \mapsto \int e^{iQ(x,y)} f(x)\, dx$$

gives (up to scalar) the action of

$$\begin{pmatrix} 1 & -2c \\ 0 & 1\end{pmatrix}\begin{pmatrix} b^{-1} & 0 \\ 0 & b\end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0\end{pmatrix}\begin{pmatrix} 1 & 0 \\ -2a & 1\end{pmatrix}$$

which matrix product is

$$g_Q := -b^{-1}\begin{pmatrix} 2a & 1 \\ 4ac-b^2 & 2c \end{pmatrix} \in \mathrm{SL}_2(\mathbf{R}).$$

The map $Q \mapsto g_Q$ is birational, with image

$$\begin{pmatrix} * & * \neq 0 \\ * & * \end{pmatrix} \in \mathrm{SL}_2(\mathbf{R}).$$

So for any quadratic form Q on $\mathbf{R}^2$ with $b \neq 0$, we get $g_Q \in \mathrm{SL}_2(\mathbf{R})$ such that

$$g_Q \cdot f = (\text{scalar}) \int f(x) e^{iQ(x,y)}\, dx.$$

Where does the relation $Q \longleftrightarrow g_Q$ come from? We'll try to give a high-level overview. Given manifolds X, Y of the same dimension, and a smooth ($\mathbf{R}$-valued) function $K(x,y)$ on $X \times Y$ (which we'll specialize to $X = \mathbf{R}$, $Y = \mathbf{R}$, and $K = Q$), its derivative $dK(x,y) \in T^*_x \times T^*_y$ gives a graph $\Gamma(dK) \subset T^*X \times T^*Y$. In our case,

$$\Gamma(dQ) \subset \mathbf{R}^2 \times \mathbf{R}^2$$

and it turns out that $\Gamma(dQ) = \Gamma(g_{\mathbf{Q}})$. In the general case, this suggests that the map

$$\begin{aligned} C_c^\infty(X) &\to C_c^\infty(Y) \\ f &\mapsto \int_{C^\infty(Y)} e^{iK(x,y)} f(x)\, dx \end{aligned}$$

should be related to the graph of a symplectomorphism $T^*X \to T^*Y$. This is the start of the theory of Fourier integral operators.

6.7. **How to work explicitly with $\widetilde{\mathrm{SL}}_2(\mathbf{R})$.** We want to explain how to compute explicit cocycles describing the extension

$$0 \to \mathbf{Z} \to \widetilde{\mathrm{SL}}_2(\mathbf{R}) \to \mathrm{SL}_2(\mathbf{R}) \to 0$$

Given a topological group G acting on a connected topological space X, make an extension of G. Let $\widetilde{X}$ be a the universal cover of X and let

$$\widetilde{G} = \{\text{lifts of some } g \in G \text{ to } \widetilde{X}\}.$$

i.e. for any $g \in G$, $\widetilde{g}$ is some map $\widetilde{X} : \to \widetilde{X}$ making the following diagram commute:

$$\begin{tikzcd} \widetilde{X} \arrow[r, "\widetilde{g}"] \arrow[d] & \widetilde{X} \arrow[d] \\ X \arrow[r, "g"] & X \end{tikzcd}$$

By the theory of covering spaces, such a lift exists but is not unique. There is a map $\widetilde{G} \to G$, whose kernel is the ambiguity of the lift, which is a torsor for $\pi_1(X)$.

$$0 \to \pi_1(X) \to \widetilde{G} \to G \to 0.$$

Apply this to $\mathrm{SL}_2(\mathbf{R})$ acting on $X = \mathbf{R}^2 - \{0\}/\mathbf{R}_{>0}$, which you can think of as the space of *oriented* lines in $\mathbf{R}^2$. We can identify $X \cong S^1 \subset \mathbf{R}^2$. Then $\widetilde{G}$ is exactly the universal cover of $\mathrm{SL}_2(\mathbf{R})$.

Exercise 6.7.1. Check this. [Hint: restrict to $\mathrm{SO}_2(\mathbf{R})$.]

From here you can get an explicit cocycle for $\widetilde{SL}_2(\mathbf{R})$. We have $\widetilde{X} = \mathbf{R}$, and the map $\widetilde{X} \to X$ sends $\theta \in \mathbf{R}$ to the line at angle θ. Choose some set-theoretic section

$$X \to \widetilde{X}$$

(e.g. measure the angle on the x-axis in $[-\pi, \pi)$). Fix a basepoint $x_0 \in X$, say the positive x-axis. Write $\widetilde{x}$ for the lift of x. For each $g \in G$, we get a lift $\widetilde{g} \in \widetilde{G}$ by asking that $\widetilde{g}\widetilde{x}_0 = \widetilde{g x_0}$. This defines a set-theoretic section $G \to \widetilde{G}$.

We want to find an explicit formula for

$$\widetilde{g}\widetilde{h} = \widetilde{gh} \cdot (\text{something in } \pi_1).$$

How does $\widetilde{g}$ act on general $\widetilde{x}$? Assume that π_1 is abelian for simplicity. For any path p in X from x to y, we can lift to a path $\widetilde{p}$ in $\widetilde{X}$ from $\widetilde{x}$ to to some lift of y, which is $\widetilde{y} + I(p)$ where $I(p) \in \pi_1(X)$. (In the example of $\mathrm{SL}_2(\mathbf{R})$ acting on X, $I(p)$ measures the number of times that p crosses the negative x-axis with suitable multiplicities.)

We'll just state the result and leave it to you to check it:

$$\widetilde{g}\widetilde{x} = \widetilde{gx} + I(gp) - I(p)$$

where p is a path from x_0 to x in X.

So for $g_1, g_2 \in G$ we get

$$\widetilde{g}_1\widetilde{g}_2\widetilde{x}_0 = \widetilde{g}_1(\widetilde{g_2 x_0}) = \widetilde{g_1 g_2 x} + I(g_1 p) - I(p)$$

where p is a path from x_0 to $g_2 x_0$. The cocycle is then

$$\widetilde{g}_1\widetilde{g}_2 = \widetilde{g_1 g_2} + \underbrace{(I(g_1 p) - I(p))}_{\in \pi_1}.$$

6.8. **The Weil representation in general.** Now we'll go to the general setting. We're going to try to write down an explicit realization of the Weil representation, in a form that makes it clear why it is only defined on a cover.

Let $(V, \langle , \rangle)$ be a symplectic vector space, meaning that V a finite-dimensional $\mathbf{R}$-vector space with a symplectic (i.e. nondegenerating and alternating) form $\langle x, y \rangle$. We can form the group

$$\mathrm{Heis}_V = V \oplus \mathbf{R}$$

with multiplication

$$(v, t)(v', t') = \left(v + v', t + t' + \frac{\langle v, v' \rangle}{2} \right).$$

As before, Heis_V has a unique irreducible representation where $(0, t)$ acts by e^{it}.

In the previous case $V = \mathbf{R}^2$ and the representation was realized on $L^2(\mathbf{R})$. In general, if $\dim V = 2n$ then the representation can be realized on $L^2(\mathbf{R}^n)$. However, we want to phrase this more intrinsically.

Any symplectic vector space $(V,\langle,\rangle)$ is isomorphic to a direct sum of $(\mathbf{R}^2, \mathrm{std})$. Said differently, we can choose bases $(e_1, f_1), (e_2, f_2), \ldots, (e_n, f_n)$ such that

$$\begin{aligned}\langle e_i, e_j\rangle &= 0\\ \langle f_i, f_j\rangle &= 0\\ \langle e_i, f_j\rangle &= \delta_{ij}.\end{aligned}$$

So $X := \mathrm{Span}(e_i)$ is an n-dimensional subspace of V where $\langle,\rangle$ vanishes, and $Y := \mathrm{Span}(f_j)$ has the same property ("Lagrangian").

In fact we can go backwards from this: suppose we're given $V = X \oplus Y$ where X, Y are Lagrangian, then $\langle,\rangle$ gives a perfect pairing $X \times Y \to \mathbf{R}$ and we can choose dual bases (e_i) for X and (f_j) for Y. Given such a splitting, Heis_V acts on $L^2(X)$ by

$$\begin{aligned}(x,0)f(t) &= f(t+x)\\ (y,0)f(t) &= f(t)e^{i\langle t,y\rangle}.\end{aligned}$$

This determines the action of $\mathbf{R} \subset \mathrm{Heis}_V$, and it is the unique representation where

$$(0,v)f(t) = e^{iv}f.$$

We'll try to turn the collection of models into a canonical one, by establishing compatible isomorphisms between them. We won't quite be able to accomplish this, because of the need to pass to a spin cover; we'll only achieve compatibility up to an 8th root of unity.

The first step is to realize that the model just constructed depends only on Y, although we formulated it in a way that seems to depend on X as well. Recall that for the Heisenberg over $\mathbf{F}_p$, we constructed the analogous representation by inducing the character

$$\begin{pmatrix}1 & & z\\ & 1 & y\\ & & 1\end{pmatrix} \mapsto e^{2\pi i z/p}$$

am then restricting. (In fact we could have extended the central character arbitrarily to this subgroup.)

We now want to discuss induction more generally. It requires some care to write down a good definition for possibly infinite-dimensional representations.

Definition 6.8.1. If $H \subset G$ are topological groups such that $H\backslash G$ admits a G-invariant measure, and $\chi\colon H \to S^1$ is a character, we define the *induced representation*

$$\mathrm{Ind}_H^G \chi = \{f\colon G \to \mathbf{C} \mid f(hg) = \chi(h)f(g)\}$$

with the norm

$$||f|| = \int_{H\backslash G} |f|^2.$$

The action of G is by right translation. Note that since χ is valued in S^1, $|f|^2$ is well-defined on the quotient space $H\backslash G$.

Let $H \leq \mathrm{Heis}$ be the subgroup $Y \oplus \mathbf{R}$. (Note that H is abelian, because $\langle,\rangle$ vanishes on Y.) Let $\chi(y,v) = e^{iv}$ be a character of H. Then

Claim 6.8.2. There is an isomorphism

$$\mathrm{Ind}_H^G \chi \cong L^2(X) \text{ (with prior action)}$$

given by $f \mapsto f|_X$.

Proof. You can just check this by hand. We'll do one example: matching up the action of $(y,0)$. For $f \in \mathrm{Ind}_H^G \chi$,

$$((y,0)f)(t,0) = f((t,0)\cdot(y,0)) = f((y,\langle t,y\rangle)(t,0)) = e^{i\langle t,y\rangle} f(t,0).$$

$\square$

For the upcoming discussion on the Weil representation, a reference is [LV80].

We have constructed an irreducible representation $I_Y = \mathrm{Ind}_{Y\oplus\mathbf{R}}^{\mathrm{Heis}_V} \chi$ for each Lagrangian $Y \subset V$, which are all abstractly isomorphic. We now define isomorphisms

$$\varphi_{YY'}: I(Y) \to I(Y').$$

It's not hard to guess what such a map should look like. It should send f to some function invariant with respect to Y', and the most straightforward to arrange this is to integrate over Y':

$$\varphi_{YY'}(f): g \mapsto \int_{Y'\in Y'} f(y'g)dy'$$

We should specify the measure, but at the end it will be uniquely specified by the condition that it preserves the norm. So we will be quite careless throughout about these constant factors.

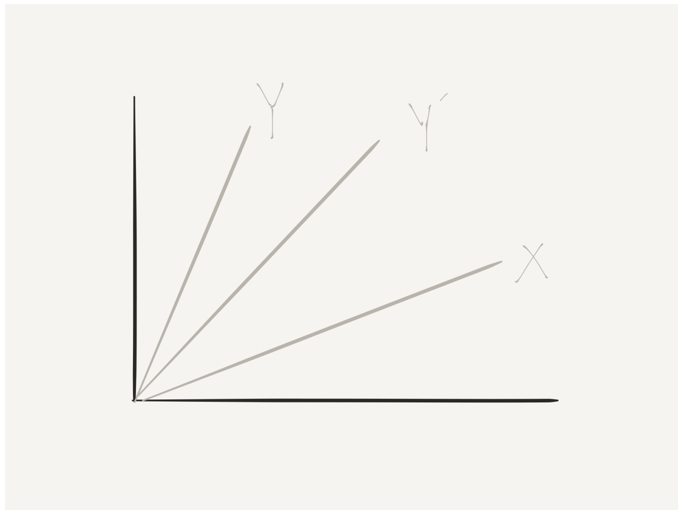
Actually this formula only works if $Y \cap Y'$ is transverse. In general, you replace Y' by $Y'/(Y\cap Y')$.

Remark 6.8.3. The definition suggests that $\varphi_{YY'}$ doesn't vary in a nice way, since the definition changes depending on the the nature of the intersection.

We want to compare

$$\varphi_{Y'Y''} \circ \varphi_{YY'} \quad \text{and} \quad \varphi_{Y''Y}.$$

Choose a third Lagrangian subspace X, with $X \cap Y = 0$ and $X \cap Y' = 0$.



Then we want to understand

$$\begin{array}{ccc} I(Y) & \xrightarrow{\varphi_{YY'}} & I(Y') \\ \downarrow & & \downarrow \\ L^2(X) & \xrightarrow{?} & L^2(X) \end{array}$$

It turns out that ? is convolution with a Gaussian. The important aspect is the phase which appears in the Gaussian.

By transversality, we can write Y' as the graph of some map

$$A\colon X \to Y.$$

In other words,

$$Y' = \{x + Ax : x \in X\}. \tag{6.8.1}$$

(There is a constraint on A coming from the fact that Y' is Lagrangian, which we'll flesh out later.) Let $f \in I(Y)$; we want to evaluate

$$\varphi_{YY'} f((t,0))$$

for $t \in X$. By definition, and using (6.8.1), we have

$$\varphi_{YY'} f((t,0)) = \int_{y' \in Y'} f((y',0)\cdot(t,0))dy' = \int_{x\in X} f((x+Ax,0)\cdot(t,0))\,dx.$$

Since f is invariant under Y, we want to pull everything involving Y to the left. Now

$$(x+Ax,0)\cdot(t,0) = (Ax, \langle Ax, x\rangle/2)(x+t,0)$$

so

$$f((x+Ax,0)\cdot(t,0)) = e^{i\langle x,Ax\rangle/2} f(x+t,0).$$

Therefore

$$\varphi_{YY'} f(t,0) = \int_X e^{i\langle x, Ax\rangle/2} f(x+t)\,dx$$

which is the convolution of f with the Gaussian $e^{i\langle x,Ax\rangle/2}$. Then

$$\text{FT(conv. of } f \text{ with } e^{i\langle x,Ax\rangle/2}) = \text{FT}(f)\cdot\text{FT}(e^{i\langle x,Ax\rangle/2}).$$

The Fourier transform of $e^{i\langle x,Ax\rangle/2}$ is another Gaussian. What are its parameters? On $\mathbf{R}^n$ if A is a nondegenerate symmetric $n\times n$ matrix,

$$\text{FT}(e^{ix^TAx/2}) = \underbrace{\gamma(A)}_{\text{phase}} e^{-ix^TA^{-1}x/2}$$

The Gaussian doesn't actually matter for us; the only important thing is the phase

$$\gamma(A) = e^{i\pi/4\,\text{sign}(A)}$$

where

$$\text{sign}(A) = \#\text{positive eigenvalues} - \#\text{negative eigenvalues}.$$

This shows that the map in question is quite simple in the Fourier-transformed picture. Let $g(A) = e^{-ix^TA^{-1}x/2}$ be the Gaussian associated with A. Then we have the commutative diagram

$$\begin{tikzcd}
I(Y) \arrow[r, "\varphi_{YY'}"] \arrow[d] & I(Y') \arrow[d] \\
L^2(X) \arrow[r, "\text{convolution}"] \arrow[d, "\text{FT}"] & L^2(X) \arrow[d, "\text{FT}"] \\
L^2(X) \arrow[r, "\cdot\gamma(A)g(A)"] & L^2(X)
\end{tikzcd}$$

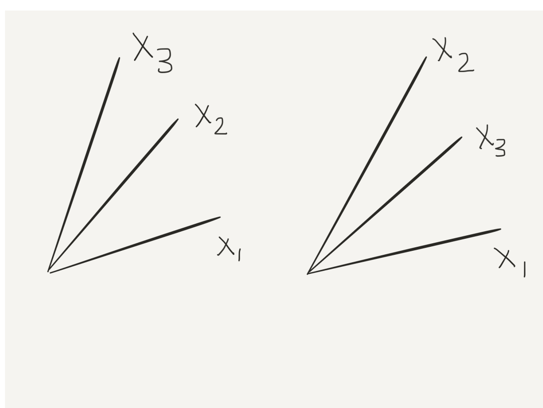
This is enough to see that the maps are compatible up to 8th root of unity. Indeed, suppose we choose another (transverse) Lagrangian Y'':

$$
\begin{tikzcd}
I(Y) \arrow[r, "\varphi_{YY'}"] \arrow[d] \arrow[rr, bend left, "\varphi_{YY''}"] & I(Y') \arrow[r, "\varphi_{Y'Y''}"] \arrow[d] & I(Y'') \arrow[d] \\
L^2(X) \arrow[r] \arrow[d, "\mathrm{FT}"] & L^2(X) \arrow[r] \arrow[d, "\mathrm{FT}"] & L^2(X) \arrow[d, "\mathrm{FT}"] \\
L^2(X) \arrow[r, "\cdot\gamma(A)g(A)"'] \arrow[rr, bend right, "\cdot\gamma(A'')g(A'')"'] & L^2(X) \arrow[r, "\cdot\gamma(A')g(A')"'] & L^2(X)
\end{tikzcd}
$$

If the Gaussians are g_1, g_2 then the composite Gaussian must be $g_1 g_2 = g_3$. Indeed, the difference between these sides commutes with the action of the Heisenberg group, hence must be a scalar. So only the phases can differ:

$$\varphi_{YY''} = \varphi_{Y'Y''}\varphi_{YY'} \cdot (\text{8th root of } 1).$$

6.9. **The Maslov index.** Let V be as before, and $X_1, X_2, X_3 \subset V$ be Lagrangian subspaces. The Maslov index is an integer $\tau(X_1, X_2, X_3) \in \mathbf{Z}$ which is invariant by the symplectic group. Before we give the general definition, let's say what it is in 2 dimensions.



Example 6.9.1. Suppose you have three pairwise-transverse lines $X_1, X_2, X_3 \subset V$. You need to attach an integer to the position. One way is to get the cyclic orientation of the three lines. It should be 1 in one case and -1 in another.

Now we give the general definition.

Definition 6.9.2 (Kashiwara)**.** Consider the quadratic form on $X_1 \oplus X_2 \oplus X_3$ given by

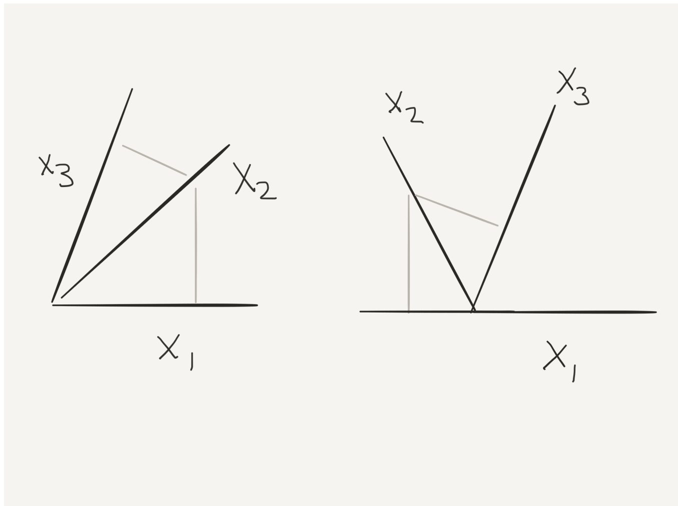
$$(x_1, x_2, x_3) \mapsto \langle x_1, x_2 \rangle + \langle x_2, x_3 \rangle + \langle x_3, x_1 \rangle$$

Then define the *Maslov index* $\tau(X_1, X_2, X_3)$ to be the signature of this form.

You can also do this algebraically. This is a quadratic form over a general field, and gives you an invariant in the Witt group.

Example 6.9.3. Let's see what form this takes in the tranverse case. Suppose X_i are pairwise transverse. We can write X_2 are the graph of a map $A: X_1 \to X_3$. Then

$$\tau(X_1, X_2, X_3) = \operatorname{sign}(\langle x, Ax\rangle).$$



Properties of the Maslov index:

(1) It is antisymmetric under S_3,
(2) If X_1, X_2, X_3, X_4 are four subspaces,

$$\tau(X_2, X_3, X_4) - \tau(X_1, X_3, X_4) + \tau(X_1, X_2, X_4) - \tau(X_1, X_2, X_3) = 0.$$

Exercise 6.9.4. Check these.

The phases for $I_{YY'}, I_{Y'Y''}$ and $I_{YY''}$ are

$$e^{i\frac{\pi}{4}\tau(X,Y',Y)}, e^{i\frac{\pi}{4}\tau(X,Y'',Y')} \quad \text{and} \quad e^{i\frac{\pi}{4}\tau(X,Y'',Y)}.$$

Therefore

$$\boxed{\varphi_{Y'Y''}\varphi_{YY'} = e^{-\frac{i\pi}{4}\tau(Y,Y',Y'')}\varphi_{YY''}.}$$

Remark 6.9.5. Strictly speaking we have only justified this when Y, Y', Y'' are transverse, but it is true in general.

Now we study the action of the symplectic group. Fix a Lagrangian Y_0. For any $g \in \mathrm{Sp}(V)$, we have a map

$$I_{Y_0} \xrightarrow{g} I_{gY_0}$$

given by

$$f \mapsto f(g^{-1}x).$$

We can use φ to go backwards, obtaining

$$I_{Y_0} \xrightarrow{g} I_{gY_0} \xrightarrow{\varphi} I_{Y_0}.$$

Call this composition $[g]$: it is a unitary operator on I_{Y_0}. Now we try to compute the relation of $[g_1][g_2]$ and $[g_1 g_2]$:

$$\begin{tikzcd}
 & & I_{g_1 g_2 Y_0} \arrow[dr] \arrow[ddrr, bend left, "\varphi_{g_1 g_2 Y_0, Y_0}"] & & \\
 & I_{g_2 Y_0} \arrow[ur, "g_1"] \arrow[dr] & & I_{g_1 Y_0} \arrow[dr, "\varphi_{g_1 Y_0, I_{Y_0}}"] & \\
I_{Y_0} \arrow[ur, "g_2"] \arrow[rr, "{[g_2]}"] \arrow[uurr, bend left, "g_1 g_2"] & & I_{Y_0} \arrow[ur, "g_2"] \arrow[rr, "{[g_1]}"] & & I_{Y_0}
\end{tikzcd}$$

It is easy to see that the middle square commutes, as does the left triangle. So the failure to commute is expressed by the righmost triangle, i.e.

$$[g_1 g_2] e^{-\frac{i\pi}{4}\tau(g_1 g_2 Y_0, g_1 Y_0, Y_0)} = [g_1][g_2].$$

From a slightly different perspective, this describes an action of an 8-fold cover of $\mathrm{Sp}(V)$ on I_{Y_0}. Explicitly, this 8-fold cover is

$$\{(g, \zeta) \in \mathrm{Sp}(V) \times \mu_8\}$$

with the action being $[g]\zeta$, and the multiplication law being

$$(g_1, \zeta_1)(g_2, \zeta_2) = (g_1 g_2, \zeta_1 \zeta_2 e^{-\frac{i\pi}{4}\tau(g_1 g_2 Y_0, g_1 Y_0, Y_0)}).$$

In fact $\widetilde{G}$ has 4 components, and the identity component is a double cover of $\mathrm{Sp}(V)$. The point is that the cocycle is actually valued in ± 1. We'll skip this computation for reasons of time, and leave you to look up the proof in [LV80].

Remark 6.9.6. The rule $g_1 g_2 \mapsto \tau(g_1 g_2 Y_0 k, g_1 Y_0, Y_0)$ is a cocycle

$$\mathrm{Sp} \times \mathrm{Sp} \to \mathbf{Z}.$$

The connected component of the associated central extension is the universal cover of $\mathrm{Sp}(V)$.

For the connection between this presentation and Maslov's, see [ES76].

7. Nilpotent groups

7.1. **Statement of Kirillov's Theorem.** Let G be a connected, simply-connected Lie group with nilpotent Lie algebra. The prototypical example is

$$\begin{pmatrix} 1 & * & * & * \\ & 1 & * & * \\ & & \ddots & * \\ & & & 1 \end{pmatrix}$$

Let's review what these terms mean. That $\mathfrak{g} := \mathrm{Lie}(G)$ be nilpotent means that the sequence

$$\mathfrak{g} \supset [\mathfrak{g}, \mathfrak{g}] \supset [\mathfrak{g}, [\mathfrak{g}, \mathfrak{g}]] \supset \dots$$

is eventually 0.

For such a G, the exponential map $\exp\colon \mathfrak{g} \to G$ is a *diffeomorphism*. Therefore we can think of the group structure as being on $\mathfrak{g}$, where it is given by the *Campbell-Baker-Hausdorff* formula:

$$\exp(X)\exp(Y) = \exp(X + Y + \frac{1}{2}[X,Y] + \ldots)$$

Since $\mathfrak{g}$ is nilpotent, the multiplication law on the Lie algebra is actually just a *polynomial*. This means that G is the $\mathbf{R}$-points of a unipotent algebraic group. In particular, this implies that G is a closed subgroup of a standard upper-triangular unipotent group

$$\begin{pmatrix} 1 & * & * & * \\ & 1 & * & * \\ & & \ddots & * \\ & & & 1 \end{pmatrix}.$$

Theorem 7.1.1 (Kirillov)**.** *The irreducible unitary representations of G are in bijection with G-orbits on $\mathfrak{g}^*$, and the correspondence is characterized by the Kirillov character formula.*

Remark 7.1.2. As with the Heisenberg group, most of these representations are ∞-dimensional.

Let's remind you what the correspondence is. An orbit $\mathcal{O} \subset \mathfrak{g}^*$ corresponds to a representation $\pi_{\mathcal{O}}$ satisfying: for $X \in \mathfrak{g}$,

$$\sqrt{j}(e^X)\,\mathrm{Tr}\,\pi(e^X) = \mathrm{FT}(\text{volume measure on orbit } \mathcal{O} \subset \mathfrak{g}^*). \tag{7.1.1}$$

We explicate some aspects of this formula.

- Here j is the Jacobian of the exponential map $\exp\colon \mathfrak{g} \to G$. (The Jacobian is defined by taking a volume form on $\mathfrak{g}$, transporting it to G by the diffeomorphism, and propagating it by invariance.)
- There's a symplectic form on $\mathcal{O}$ defined as follows: for $X, Y \in \mathcal{O}$,

  $$\omega(\overline{X}, \overline{Y}) = \lambda([X,Y])$$

  where $\overline{X}, \overline{Y}$ are the derivatives of $g \mapsto g\lambda$ at X, Y. The volume measure on $\mathcal{O}$ is then

  $$\left(\frac{\omega}{2\pi}\right)^d$$

  where $2d = \dim \mathcal{O}$.

Remark 7.1.3. The Jacobian simplifies to $j = 1$ in the nilpotent case.

The equality (7.1.1) is in the sense of distributions. Concretely, it means that for $f \in C_c^\infty(\mathfrak{g})$,

$$\mathrm{Tr}\left(\int_X f(X)\pi(e^X)\,dX\right) = \int_{\mathcal{O}} \mathrm{FT}(f).$$

Here

$$\mathrm{FT}(f)(\lambda) = \int f(X) e^{i\langle \lambda, X\rangle}\,dX.$$

The well-definedness (i.e. convergence) of either side is not obvious.

7.2. **Outline of a proof.** As for Heis_V, we consider restriction to a subgroup. Thanks to the nilpotence of $\mathfrak{g}$, the commutator subgroup is proper:

$$\mathfrak{g} \supsetneq [\mathfrak{g}, \mathfrak{g}]$$

so there exists a Lie algebra homomorphism $\mathfrak{g} \to \mathbf{R}$. Let $\mathfrak{g}_1$ be the kernel; exponentiating it gives a subgroup $G_1 = \exp(\mathfrak{g}_1)$. If you choose $X \in \mathfrak{g} - \mathfrak{g}_1$, then

$$G = G_1 \rtimes \underbrace{\exp(\mathbf{R} \cdot X)}_{\cong(\mathbf{R},+)}.$$

By induction, we can assume that we know the result for G_1. The main point is that you have a map $\mathfrak{g}^* \xrightarrow{\mathrm{pr}} \mathfrak{g}_1^*$.

The idea is to take a representation π of G, restrict it to G_1, and then decompose it into irreducibles by using the theory of disintegration. It turns out that the possibilities for this decomposition are limited: either the restriction remains irreducible or it splits up into a single orbit of an irreducible σ of G_1 under the action of $\mathbf{R}$. By the result for G_1, we match σ with an orbit $\overline{\mathscr{O}} \subset \mathfrak{g}_1^*$. Then we show that the union of the conjugates of $\mathrm{pr}^{-1}(\overline{\mathscr{O}})$ by $\exp(\mathbf{R} \cdot X_1)$ is a single G-orbit on $\mathfrak{g}^*$, and is matched with π.

I wrote down the argument, but the technical details are quite tricky. We won't go through it.

7.3. **Construction of the representation.** We now want to explain an explicit way to construct $\pi_{\mathscr{O}}$ from $\mathscr{O}$ via induction. Fix $\lambda \in \mathscr{O}$. The G-action on λ satisfies (by definition)

$$g\lambda(gX) = \lambda(X).$$

Differentiating this equation, we find that for $Y \in \mathfrak{g}$,

$$Y\lambda(X) + \lambda([Y, X]) = 0.$$

i.e.

$$Y\lambda(X) = -\lambda([Y, X]). \tag{7.3.1}$$

This tells us that if X, Y belong to the stabilizer $\mathfrak{g}_\lambda$ of $\lambda \in \mathfrak{g}$, then

$$\lambda([X, Y]) = 0$$

i.e. $\lambda([\mathfrak{g}_\lambda, \mathfrak{g}_\lambda]) = 0$. In other words, λ defines a Lie algebra homomorphism

$$\mathfrak{g}_\lambda \to \mathbf{R}.$$

For these simply-connected groups we can easily pass between the group and Lie algebra by exponentiation, so we get a corresponding Lie group homomorphism

$$e^{i\lambda} \colon G_\lambda \to S^1.$$

In other words, we've used the orbit to build a character of its stabilizer. Now you might like to induce it to a representation of G. However, that's not the right thing to do at this point. As a general philosophy, you need to extend the character as much as possible before you inducing it. For example, in the Heisenberg case $\mathfrak{g}_\lambda$ is the center, and we saw that we should first extend it to a 2-dimensional subalgebra.

Note that (7.3.1) implies that the pairing

$$X, Y \mapsto \lambda([X, Y])$$

descends to a pairing

$$\mathfrak{g}/\mathfrak{g}_\lambda \times \mathfrak{g}/\mathfrak{g}_\lambda \to \mathbf{R}$$

which is alternating and non-degenerate. (According to (7.3.1), if it vanishes for all X then by definition $Y\lambda = 0$.)

Under the identification $\mathfrak{g}/\mathfrak{g}_\lambda \cong T_\lambda \mathscr{O}$, this form $\lambda([X, Y])$ corresponds to the canonical symplectic structure of Fact 2.3.1.

We would like a sub-algebra $\mathfrak{q} \supset \mathfrak{g}_\lambda$ as large as possible so that $\lambda \colon \mathfrak{q} \to R$ is a Lie algeba homomorphism, meaning that $\lambda([\mathfrak{q}, \mathfrak{q}]) = 0$. In other words, $\mathfrak{q}$ should have *isotropic* image in $\mathfrak{g}/\mathfrak{g}_\lambda$. The largest we can hope for is half the dimension.

Definition 7.3.1. A *polarization* is a (Lie) sub*algebra* $\mathfrak{g}_\lambda \subset \mathfrak{q} \subset \mathfrak{g}$ such that $\mathfrak{q}/\mathfrak{g}_\lambda$ is Lagrangian (maximal isotropic) for the form

$$X, Y \mapsto \lambda[X, Y].$$

Fact 7.3.2 (Kirillov)**.** Polarizations exist.

Of course there is no problem in finding Lagrangian subspaces; the issue is to arrange them to be sub*algebras*.

If $\mathfrak{q}$ is a polarization, and $Q = \exp(\mathfrak{q})$, and let

$$\chi \colon Q \to S^1$$

be given by

$$\chi(e^q) = e^{i\lambda q}.$$

Then $\operatorname{Ind}_Q^G \chi$ is irreducible, and gives $\pi_\mathscr{O}$.

Remark 7.3.3. The polarizing subalgebras are not conjugate, yet somehow the representations are isomorphic. In the Heisenberg group, any two-dimensional subalgebra containing the center is a polarization.

We'll compute the character to verify that it is right. A polarization gives a picture of what the orbit looks like.

Exercise 7.3.4. Let $H \le G$ be Lie groups. If χ is a character of H, then the character of $\pi := \operatorname{Ind}_H^G \chi$ is

$$\theta_\pi(x) = \int_{g \in G/H} \chi(g^{-1} x g)\, dg.$$

This is to be interpreted in the sense of distributions: it means that for $f \in C_c^\infty(G)$,

$$\operatorname{Tr} \pi(f) = \int_{g \in G/H} \int_{x \in H} \chi(x) f(g x g^{-1})\, dx\, dg$$

You can think of this as taking χ on H and forcing it to be conjugacy-invariant on G. To prove this, you write down an integral kernel for $\pi(f)$. In other words, there is a finite-group proof involving summing over diagonal entries of a matrix, and here you do the same for an integral kernel.

We apply this discussion to $\mathrm{Ind}_Q^G \chi$ for $\mathfrak{q}$ a polarization. We want to work out the trace of the operator

$$\pi(f) = \int_{X \in \mathfrak{g}} f(X)\pi(e^X)\, dx.$$

By Exercise 7.3.4, applied with $H = \exp(\mathfrak{q})$ and in Lie algebra coordinates, says that the trace is

$$\mathrm{Tr}\,\pi(f) = \int_{g \in G/Q} \int_{Y \in \mathfrak{q}} e^{i\lambda(Y)} f(gYg^{-1})\, dy\, dg.$$

For $g = 1$, the contributing term is

$$\int_{Y \in \mathfrak{q}} e^{i\lambda(Y)} f(Y)\, dy$$

The Plancherel formula tells us that

$$\int_{\mathfrak{q}} f = \left(\frac{1}{2\pi}\right)^d \int_{\mathfrak{q}^\perp} \mathrm{FT}(f)$$

so, shifting this a little, we find that

$$\int_{Y \in \mathfrak{q}} e^{i\lambda(Y)} f(Y)\, dy = \left(\frac{1}{2\pi}\right)^d \int_{\mathfrak{q}^\perp + \lambda} \mathrm{FT}(f)$$

where $\mathfrak{q}^\perp = \{\mu \in \mathfrak{g}^* : \mu|_{\mathfrak{q}} = 0\}$ (note that $\dim \mathfrak{q} + \dim \mathfrak{q}^\perp = \dim \mathfrak{g}$) and $2d = \dim \mathcal{O} = \dim \mathfrak{g} - \dim \mathfrak{g}_\lambda$, so $d = \dim \mathfrak{g} - \dim \mathfrak{q}$.

Therefore,

$$\begin{aligned}
\mathrm{Tr}\,\pi(f) &= \int_{g \in G/Q} \int_{Y \in \mathfrak{q}} e^{i\lambda(Y)} f(gYg^{-1})\, dy\, dg \\
&= \int_{g \in G/Q} \int_{g(\lambda + \mathfrak{q}^\perp)g^{-1}} \mathrm{FT}(f) \\
&= \int \mathrm{FT}(f) d\mu
\end{aligned}$$

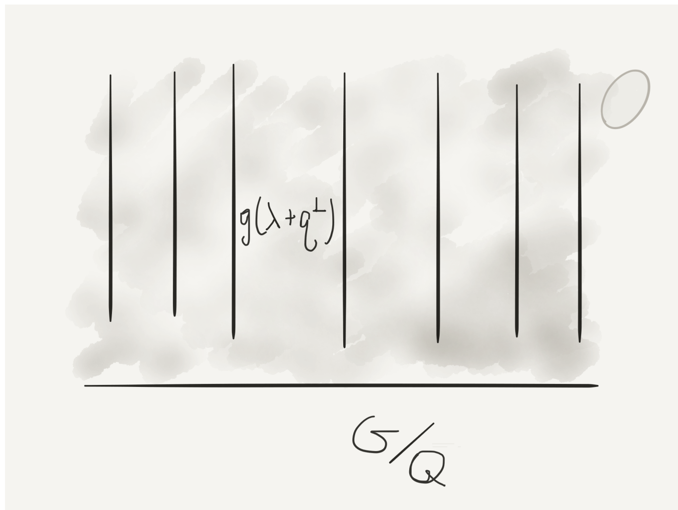
where μ is the measure on $\mathfrak{g}^*$ obtained by pushing forward

$$G/Q \times (\lambda + \mathfrak{q}^\perp) \to \mathfrak{g}^*$$

Exercise 7.3.5. Show that $\lambda + \mathfrak{q}^\perp$ is the Q-orbit of λ. Then show that $G/Q \times (\lambda + \mathfrak{q}^\perp)$ maps bijectively to the orbit $\mathcal{O}$, i.e. the union of the G-conjugates of $\lambda + \mathfrak{q}^\perp$ is $G\lambda = \mathcal{O}$ and

$$g_1(\lambda + \mathfrak{q}^\perp)g_1^{-1} = g_2(\lambda + \mathfrak{q}^\perp)g_2^{-1} \iff g_1 Q = g_2 Q.$$

This presents a picture of $\mathcal{O}$ as fibered over G/Q, with each fiber an affine space which is the corresponding translate of $\lambda + \mathfrak{q}^{\perp}$.



This picture interacts well with the symplectic form. It is a Lagrangian fibration, meaning that each fiber is Lagrangian (at every point of every fiber, the tangent space is a Lagrangian).

Note that

$$\pi_{\mathcal{O}} = \operatorname{Ind}_Q^G \chi.$$

Ignoring the character, this looks like the space of functions on G/Q, i.e. functions defined on the space of fibers, i.e. "functions on $\mathcal{O}$ constant along fibers". This is a useful picture to have. For example, it explains why for the Heisenberg group the induced representation has a model as functions on $\mathbf{R}^2$.

7.4. **A look ahead.** The next goal is to prove the Kirillov character formula for compact groups. We don't want to assume much familiarity with the general formalism of roots and weights, so we'll focus on the unitary group $U(N)$.

We first need some of the machinery of symplectic manifolds, including the Duistermaat-Heckman localization formula. To explain why we need some a formula, recall that in the Kirillov character formula one side is an integral of a Fourier transform. For nilpotent groups, the coadjoint orbit looks very linear. (In the Heisenberg case it was literally linear; in the general case it is foliated by linear spaces.) So the Fourier space is just a function on the dual space.In the compact case we have no linearity, so we need to identify the Fourier transform somehow. The Duistermaat-Heckman formula deals with this.

In summary, the unitary case looks very different from the nilpotent one. Afterwards, we'll turn to $\mathrm{SL}_2(\mathbf{R})$, which exhibits some features of both.

8. Background on symplectic geometry

8.1. **Symplectic manifolds.** A *symplectic manifold* (M,ω) is a manifold M equipped with a closed, non-degenerate 2-form $\omega \in \bigwedge^2 T_m^*$.

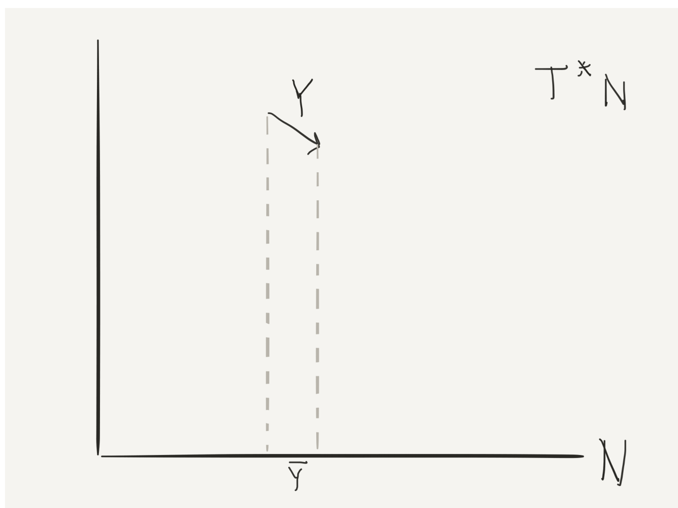
Example 8.1.1. $\mathbf{R}^{2n}$ with coordinates $(x_1,\dots,x_n,y_1,\dots,y_n)$; the form is $\sum_{i=1}^n dx_i \wedge dy_i$. Darboux's theorem says that any (M,ω) locally looks like this.

The following is the most important example for the course.

Example 8.1.2. For any manifold N, the cotangent bundle T^*N has a canonical symplectic structure. Note that T^*N carries a canonical 1-form η, specified by for a point $(n \in N, v \in T_n^*)$ the value on a tangent vector $Y \in T_{(n,v)}T^*N$ is

$$\eta(Y) = \langle v, \overline{Y} \rangle$$

where $\overline{Y}$ is the projection of Y to T_n^*N.



Then on T^*N we take the symplectic form $\omega = d\eta$.

Let $\mathcal{O}$ be the G-orbit of $\lambda \subset \mathfrak{g}^*$ with the form previously defined. We still need to show that ω is closed. For each $X \in \mathfrak{g}$, we get a vector field X^* on $\mathcal{O}$ given by

$$X_\lambda = \frac{d}{dt}|_{t=0}(e^{tX}\cdot\lambda)$$

We call this vector field X^*. The symplectic form ω is defined by: at λ

$$\omega(X^*, Y^*) = \lambda([X,Y]).$$

To show that $d\omega = 0$, we want to verify that

$$d\omega(X^*, Y^*, Z^*) \stackrel{?}{=} 0$$

(since the map $X \mapsto X^*_\lambda$ defines a surjection $\mathfrak{g} \to T_\lambda \mathcal{O}$). The formula for the derivative of a form gives

$$d\omega(X^*, Y^*, Z^*) = X^*\omega(Y^*, Z^*) - Y^*\omega(X^*, Z^*) + Z^*\omega(X^*, Y^*)$$
$$- \omega([X^*, Y^*], Z^*) - \omega([Y^*, Z^*], X^*) - \omega([Z^*, X^*], Y^*) \ldots$$

Here $[X^*, Y^*]$ is the bracket of vector fields on $\mathcal{O}$. We can also form $[X, Y]^*$. These are obviously going to be related, but in fact they are off by a sign:

$$[X^*, Y^*] = -[X, Y]^*.$$

Let's digest the formula. By the definition of ω, the first line is

$$X^*\lambda([Y, Z]) + \ldots = \lambda([X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]]) = 0$$

by the Jacobi identity. The second line is

$$\lambda([[X, Y], Z]) + \ldots$$

which again vanishes by the Jacobi identity.

8.2. **Hamiltonian flow.** We have just seen that G acts on the orbit $\mathcal{O}$, preserving the symplectic form.

More generally, suppose g_t is a 1-parameter group of diffeomorphisms of some symplectic manifold (M, ω) preserving ω: $g_t^*\omega = \omega$. Associated to g_t, one get a vector field X given by

$$\frac{d}{dt}|_{t=0}(g_t m).$$

What property of X corresponds to ω being fixed?

When we have a vector field, we have a Lie derivative $\mathscr{L}_X$ on differential forms, given by flowing the differential form through the vector field:

$$\mathscr{L}_X \nu = \frac{d}{dt}|_{t=0}(g_t)^* \nu.$$

In particular, since the symplectic form ω is invertible we have $\mathscr{L}_X \omega = 0$.

Remark 8.2.1. Don't confuse the Lie derivative with the differential. If ν is a k-form, then $\mathscr{L}_X \nu$ is a k-form.

An identity of Cartan says that

$$\mathscr{L}_X \nu = d(X \lrcorner \nu) + X \lrcorner d\nu.$$

Here the interior product is defined by: if Ω is a k-form, then $X \lrcorner \Omega$ is the $(k-1)$-form

$$X \lrcorner \Omega(Y_1, \ldots, Y_{k-1}) = \Omega(X, Y_1, \ldots, Y_{k-1}).$$

The interior product is sort of like an adjoint to wedging.

Returning to the situation with ω a symplectic form, by Cartan's formula

$$\mathscr{L}_X \omega = d(X \lrcorner \omega) + X \lrcorner d\omega$$

but $d\omega = 0$ because ω is already closed, so the condition that the flow of X preserves ω is that $X \lrcorner \omega$ is closed.

Now assume that $H^1(M,\mathbf{C})=0$. Then closed 1-forms are automatically exact, so

$$X \lrcorner \omega = dH \tag{8.2.1}$$

for some function $H: M\to R$. Concretely, this means (evaluating (8.2.1) at Y)

$$\omega(X,Y)=dH(Y). \tag{8.2.2}$$

Said differently, X assigns to every $m\in M$ an element of T_m. On the other hand, dH assigns to every $m\in M$ an element of T_m^*. The symplectic form ω can be thought of as an identification $T_m\cong T_m^*$ forevery m, and the condition is that the elements agree under this identification.

To summarize, under the assumption $H^1(M,\mathbf{R})=0$, if X preserves ω then X must arise from $H: M\to\mathbf{R}$ as $X\longleftrightarrow dH$ under ω. In this case we say that X is the "Hamiltonian flow" associated to H.

Example 8.2.2. Consider $(\mathbf{R}^2, dx\wedge dy)$ with $H=x^2+y^2$. This corresponds to a vector field X_H, which we will shortly determine. Note $dH=2x\,dx+2y\,dy$. The vector field X_H is characterized by

$$dH(Y)=\omega(X_H,Y).$$

If $Y=\partial_X$, then this says

$$2Y=\omega(X_H,\partial_Y).$$

Thus $X_H=2y\partial_x+?\partial_y$. Repeating this for the other direction, we find that

$$X_H=(2y)\partial_x-(2x)\partial_y.$$

Observe that H is constant along the orbits of the flow. This is not a coincidence! It is true in general that H is constant along the flow X_H.

8.3. **Poisson brackets and quantization.** Let (M,ω) be a symplectic manifold. For $f,g\in C^\infty(M)$, we can make the flows X_f,X_g. The rate of change of g along the X_f-flow is, by definition, $dg(X_f)$, which is by definition of X_g (8.2.2) is the same as $\omega(X_g,X_f)=-\omega(X_f,X_g)$.

Definition 8.3.1. For $f,g\in C^\infty(M)$, we define the *Poisson bracket*

$$\{f,g\}:=\omega(X_f,X_g).$$

Example 8.3.2. On $(\mathbf{R}^{2n},\sum dx_i\wedge dy_i)$ the Poisson bracket is

$$\{f,g\}=\sum_i\left(\frac{\partial f}{\partial x_i}\frac{\partial g}{\partial y_i}-\frac{\partial f}{\partial y_i}\frac{\partial g}{\partial x_i}\right)$$

So

$$\{x_i,x_j\}=0$$
$$\{y_i,y_j\}=0$$
$$\{x_i,y_j\}=\delta_{ij}$$

Quantum mechanics suggests that many naturally-occurring symplectic manifolds (M,ω) should have a *quantization*, which should be a Hilbert space $\mathscr{H}$. The example to keep in mind is that for $(M,\omega)=(\mathbf{R}^2, dx\wedge dy)$ the Hilbert space is $H=L^2(\mathbf{R})$. In general, $\dim\mathscr{H}$ is the symplectic volume (which could be infinite), and a function $a\in C^\infty(M)$ picks out an operator $\mathrm{Op}(a)$ on $\mathscr{H}$ with

$$\mathrm{Tr}\,\mathrm{Op}(a)=\int a\,d\mathrm{vol},$$

and

$$[\mathrm{Op}(a),\mathrm{Op}(b)]=i\,\mathrm{Op}(\{a,b\}).$$

We've already seen examples of this. The theme of the course is to quantize the orbits $\mathscr{O}$ of G on $\mathfrak{g}^*$. The story of the Weil representation showed the limits of this quantization: there is a topological obstruction to making it $\mathrm{SL}_2(\mathbf{R})$-equivariant.

Remark 8.3.3. The Poisson bracket $\{,\}$ is a Lie bracket on $C^\infty(M)$. We would like to view Lie brackets as coming from Lie groups. In fact, the association

$$f\mapsto X_f$$

defines a map

$$C^\infty(M)\to \mathrm{Lie}(\mathrm{Symplectomorphisms}(M,\omega))$$

(whose kernel consists of the constant functions) and we've seen that if $H^1(M,\mathbf{R})=0$ then everything comes from this construction. This map is a Lie algebra anti-homomorphism:

$$[X_f,X_g]=-X_{\{f,g\}}.$$

Now we want to understand the Poisson structure on coadjoint orbit $\mathscr{O}\subset\mathfrak{g}^*$. Here we have a very nice class of functions: every $Y\in\mathfrak{g}$ gives a function $f_Y\in C^\infty(\mathscr{O})$, defined by $f_Y(\lambda)=\lambda(Y)$. So we can ask: What is the flow of f_Y, and what are the Poisson brackets? I think the answer is that $X_{f_Y}=Y^*$, and

$$\{f_Y,f_Z\}=f_{[Y,Z]}.$$

The symplectic structure on $\mathscr{O}$ was set up to make this true.

8.4. **Summary.** We've seen that to a symplectic manifold (M,ω) and $H\in C^\infty(M)$ we can attach a flow (vector field) X_H preserving ω, which is characterized by $dH\in T^*$ corresponds via ω to $X_H\in T$:

$$dH(Y)=\omega(X_H,y).$$

In particular, on $\mathscr{O}$ one has functions f_Y for each $Y\in\mathfrak{g}$, such that

(1) the flow of f_Y is the vector field corresponding to Y, and
(2) $\{f_Y,f_Z\}=f_{[Y,Z]}$,
(3) the map $Y\mapsto f_Y$ is G-equivariant.

A G-action on a symplectic manifold is *Hamiltonian* if there exists such a system: a linear map

$$\mathfrak{g}\ni Y\to f_Y\in C^\infty(M)$$

satisfying (1)-(3).

Example 8.4.1. We explain a simple example, but one which will be important later. The complex projective space $\mathbf{CP}^n = \mathbf{C}^{n+1} - \{0\}/\mathbf{C}^*$ has a unique symplectic form invariant by $U(n+1)$. This is called the *Fubini-Study metric*. It looks a little messy when expressed in terms of the usual coordinates, so we'll explain a cleaner point of view.

Let $z_0, \ldots, z_n$ be the coordinates on $\mathbf{C}^{n+1}$. Start with the form

$$\sum_{k=0}^{n} i(dz_k \wedge d\overline{z}_k).$$

The i looks a little funny, but in coordinates $z_k = x_k + iy_k$ it becomes

$$i(dx_k + i\,dy_k) \wedge (dx_k - i\,dy_k) = 2dx_k \wedge dy_k$$

which is essentially the usual volume. This is not invariant by $\mathbf{C}^*$, but it is invariant under the unitary part:

$$(z_0, \ldots, z_n) \mapsto (e^{i\alpha} z_0, \ldots, e^{i\alpha} z_n).$$

We can think of $\mathbf{CP}^n = S^{2n+1}/S^1$ where $S^{2n+1} = \{(z_i) \mid \sum |z_k|^2 = 1\} \subset \mathbf{C}^{n+1}$. Obviously $\omega|_{S^{2n+1}}$ is not symplectic, since it's on an odd-dimensional manifold, but it turns out that the degenerate direction is the S^1, i.e. $\omega|_{S^{2n+1}}$ is the pullback of a symplectic form ω on $\mathbf{CP}^n$.

We'll now write out a formula for this form in coodinates. The usual affine coordinates are not well-adapted to this form, so we'll introduce some slightly different coordinates.

Let $H_k = |z_k|^2$, viewed as a function on S^{2n+1} invariant by S^1. Let $\theta_k = \arg(z_k)$, which is a function on S^{2n+1}, not invariant by rotation of course. Think about polar coordinates in the plane:

$$2dx\,dy = 2r\,dr\,d\theta = d(r^2)d\theta.$$

So we have (on S^{2n+1})

$$\sum dz_k \wedge d\overline{z}_k = \sum dH_k \wedge d\theta_k. \tag{8.4.1}$$

We want to descend this down to $\mathbf{CP}^n$. Since $H_0 + \ldots + H_n = 1$ on S^{2n+1}, we have

$$dH_0 = -dH_1 - dH_2 - \ldots - dH_n$$

which we can use to rewrite (8.4.1) as

$$\sum dz_k \wedge d\overline{z}_k = \sum_{k=1}^{n} dH_k \wedge d(\theta_k - \theta_0).$$

Now $\theta_k - \theta_0$ *is* invariant by rotation, so it descends to a function $\overline{\theta}_k := \theta_k - \theta_0$ on $\mathbf{CP}^n$. They give local coordinates where the $z_i \neq 0$. This implies that

$$\omega_{\mathbf{CP}^n} = \sum dH_k \wedge d\overline{\theta}_k.$$

Let's have an example of using this form. Suppose we want to compute the volume of $\mathbf{CP}^n$:

$$\int \omega_{\mathbf{CP}^n}^m = m! \int dH_1 \wedge \ldots \wedge dH_n \wedge d\overline{\theta}_1 \wedge \ldots \wedge d\overline{\theta}_n$$
$$= n!(2\pi)^n \int dH_1 \wedge \ldots \wedge dH_n.$$

The functions H_i map $\mathbf{CP}^n$ to a simplex.

$$\mathbf{CP}^n \xrightarrow{[H_0,\ldots,H_n]} \{(x_0,\ldots,x_n) \in \mathbf{R}^{n+1}_{\geq 0} \mid \sum x_i = 1\}$$

where we normalize $H_i = \frac{|z_i|^2}{\sum |z_i|^2}$. The way we've set things up, the volume form on the simplex is $dx_1 \wedge \ldots \wedge dx_n$ so the n-simplex has volume $\frac{1}{n!}$. So the volume of $\mathbf{CP}^n$ is $(2\pi)^n$.

The flow associated to H_k is $\theta_k \mapsto \theta_k + t$. In other words, in these coordinates the vector field X_{H_k} is ∂_{θ_k}.

In summary, on $\mathbf{CP}^n$, the flow associated to $H_k = \frac{|z_i|^2}{\sum |z_i|^2}$ rotates z_k:

$$(z_0,\ldots,z_n) \mapsto (z_0,\ldots,z_{k-1}, e^{it} z_k, z_{k+1},\ldots,z_n)$$

which is even true for $k = 0$.

9. Duistermaat-Heckman localization

This is a theorem about symplectic manifolds, but the original motivation was from representation theory. It will be used in the proof of Kirillov's theorem for compact groups.

9.1. **Stationary phase.**

Suppose M is a manifold, $f \in C_c^\infty(M)$ (which is just a technical device to regularize integrals), and $\varphi \in C^\infty(M)$. Consider the integral

$$I(\lambda) := \int_M f(x) e^{i\lambda\varphi(x)}\, dx.$$

The point of stationary phase is to study the behavior of this integral as $\lambda \to \infty$.

Lemma 9.1.1. *If $d\varphi \neq 0$ on* $\operatorname{supp}(f)$*, then $I(\lambda)$ decays faster than any power of λ as $\lambda \to \infty$.*

Proof. Choose a vector field X on M such that $X\varphi = 1$ on $\operatorname{supp}(f)$. Now

$$X(e^{i\lambda\varphi}) = (i\lambda) e^{i\lambda\varphi}.$$

Iterating this a thousand times (say), we find that

$$(i\lambda)^{1000} I(\lambda) = \int_m f(X^{1000} e^{i\lambda\varphi}) dx = \pm \int X^{1000} f \cdot e^{i\lambda\varphi}\, dx \ll 1.$$

□

Principle of stationary phase. By this argument, the asymptotic behavior of $I(\lambda)$ is controlled by critical points of φ.

Suppose that the critical points of φ are $P_1,\ldots,P_n$. We assume that they are all non-degenerate, i.e. that the Hessian of φ at the P_i is non-degenerate. Then near the critical points the function looks like a Gaussian:

$$\int_{\mathbf{R}^n} e^{i(\frac{1}{2}x^T H x)}\,dx = \frac{(2\pi)^{n/2} e^{\frac{i\pi}{4}\operatorname{sign}(H)}}{\sqrt{|\det H|}}.$$

So in this situation

$$I(\lambda) \sim \sum_{P_i} f(P_i) e^{i\lambda\varphi(P_i)} \frac{(2\pi)^{n/2} e^{\frac{i\pi}{4}\operatorname{sign}(\operatorname{Hess}_{P_i}(\varphi))}}{\lambda^{n/2}\sqrt{|\det \operatorname{Hess}_{P_i}(\varphi)|}} \tag{9.1.1}$$

where the Hessian is taken with respect to a system of coordinates $(x_1,\ldots,x_n)$ such that $dx = dx_1\ldots dx_n$. The meaning of $\sim$ in (9.3.1) is that

$$|\text{LHS-RHS of (9.3.1)}| = O(\lambda^{-n/2-1}).$$

Example 9.1.2. If $M = S^2 \subset \mathbf{R}^3$, with $f = 1$ and $\varphi(x,y,z) = z$. We want to compute

$$\int_{S^2} e^{i\lambda z}.$$

The critical points occur at the north and south poles. Since

$$z = \sqrt{1-x^2-y^2} \approx 1 - \frac{1}{2}(x^2+y^2)$$

the Hessian is

$$\begin{pmatrix} -1 & \\ & -1 \end{pmatrix}.$$

Plugging this into (9.3.1) gives

$$\int_{S^2} e^{i\lambda z} \sim 2\pi\left(\frac{e^{i\lambda}}{i\lambda} - \frac{e^{-i\lambda}}{i\lambda}\right) \tag{9.1.2}$$

so

$$\int_{S^2} e^{i\lambda z} \sim 4\pi\frac{\sin\lambda}{\lambda}.$$

But in fact this is even an *equality*, as we saw way back in Lemma 1.3.3.

9.2. **The Duistermaat-Heckman formula.** Duistermaat-Heckman understood the general context for this phenomenon.

Theorem 9.2.1 (Duistermaat, Heckman)**.** *Let (M,ω) be a compact symplectic manifold of dimension $2d$. Let $H \in C^\infty(M)$, with isolated and non-degenerate critical points, and consider*

$$\int e^{i\lambda H}\omega^d.$$

Assume that the flow X_H generates an action of S^1, i.e. all orbits of the flow return to themselves after exactly the same length of time. (In other words, all orbits are periodic with the same length.*) Then the stationary phase asymptotic* (9.3.1) *is exact:*

$$\int e^{i\lambda H}\omega^d = \text{(stationary phase asymptotic)} = \frac{(2\pi)^d}{t^d}\sum_i \frac{e^{itH(P_i)}}{\sqrt{|\mathrm{Hess}_{P_i}(H)|}} e^{\frac{i\pi}{4}\operatorname{sign}(\mathrm{Hess}_{P_i}(H))}.$$

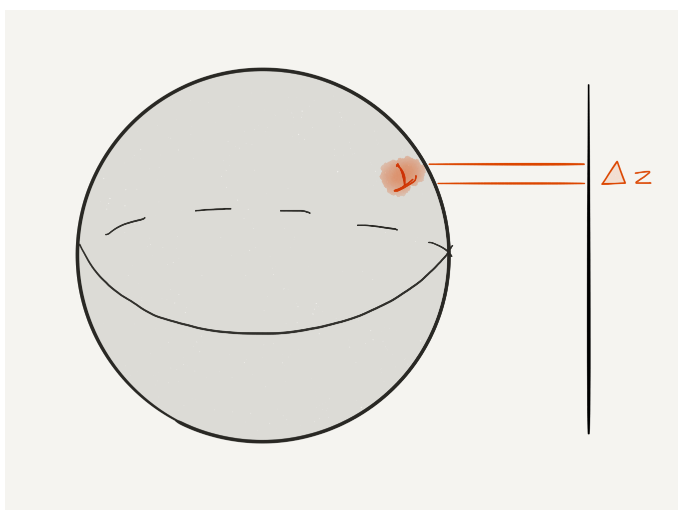
Remark 9.2.2. In general it is not even clear that the stationary phase asymptotic extends to all λ. For instance, (9.1.2) has an issue at 0 a priori.

Remark 9.2.3. This theorem will be applied to compute one side in the Kirillov character formula for compact groups.

Example 9.2.4. We claim that in Example 9.1.2, the flow attached to φ is the vector field R corresponding to z-axis rotation. We must check that

$$d\varphi(Y) = \omega(R, Y).$$

It suffices to check this as Y runs over a basis. If $Y = R$, then both sides are 0 because the z-coordinate is unchanging under rotation along the z-axis. It remains to check the claim for another Y. We choose Y to generate the longitudinal flow. Then $\omega(R, Y)$ is the area of the patch.



This area is $d\theta \cdot \Delta z$ so $\omega(R, Y) = dz(Y)$. The content of the equality is the same as that the pushforward of the area measure on S^2 to dz is uniform.

Example 9.2.5. Take S^1 acting on $\mathbf{CP}^n$ by

$$\theta \cdot (z_0, z_1, \ldots, c_n) = (e^{im_0\theta} z_0, e^{im_1\theta} z_1, \ldots)$$

with $m_0, m_1, \ldots, \in \mathbf{Z}$ all non-zero and distinct (this latter condition corresponds to the fixed points being non-degenerate). Consider the flow corresponding to $H = \sum m_i H_i$, with H_i as in Example 8.4.1. The Duistermaat-Heckman localization theorem computes

$$\int_{\mathbf{CP}^n} e^{i\lambda H}\omega^n$$

as the sum of $(n+1)$ terms. Naïvely one can write

$$\int_{\mathbf{CP}^n} e^{i\lambda H}\omega^n = (2\pi)^n \int_{\sum x_i=1} e^{i\lambda \sum m_i x_i}\, dx_1 \ldots dx_n$$

and the n-fold iterated integral has 2^n terms (two endpoints for each integral), so the Duistermaat-Heckman localization gives a significant amount of collapsing. In Example 8.4.1 we saw that under the map

$$\begin{aligned} \mathbf{CP}^n &\to \{(x_0,\ldots,x_n)\in \mathbf{R}^{n+1}_{\geq 0} : \sum x_i = 1\} \\ (z_i) &\mapsto \frac{(|z_0|^2,\ldots,|z_n|^2)}{\sum |z_i|^2} \end{aligned}$$

the volume form pushes forward as

$$\omega^n \mapsto (2\pi)^n n!\, dx_1 dx_n.$$

So we have

$$\int e^{itH}\omega^n = (2\pi)^n n! \int_{\sum x_i=1} e^{it(\sum m_i x_i)}\, dx_1\wedge \ldots \wedge dx_n. \tag{9.2.1}$$

You can compute this explicitly. For $n=2$, it comes out to

$$(2\pi)^2\cdot 2!\cdot e^{itm_0}\cdot \int_{x_1,x_2,0\leq x_1+x_2\leq 1} e^{it(m_1-m_0)x_1+it(m_2-m_0)x_2}\, dx_1 dx_2.$$

Evaluating the integral produces four terms, which collapse as

$$(2\pi)^2 2!\left[\frac{e^{itm_0}}{(it(m_0-m_1))(it(m_0-m_2))} + \text{symmetric terms}\right].$$

(Note that it looks singular at $t=0$. In fact, when $t\to 0$, this approaches a limit which is the area of the simplex.)

Let's compare this to the stationary phase asymptotic. What are the critical points? They are the points $Q\in\mathbf{CP}^n$ such that $dH(Q)=0$, but dH is identified under the symplectic form with X_H, so this is equivalent to $X_H(Q)=0$, i.e. Q is fixed by the S^1-flow. (As mentioned above, the m_i should be distinct to get non-degenerate critical points.)

So the critical points are $P_i = (0,0,\ldots,\underbrace{1}_{i},0,\ldots,0)$ with $0\leq i\leq n$. Now we have to compute the Hessian at these points. Put the following local coordinates near P_k

$$(z_0,\ldots,z_{k-1},\underbrace{1}_{k},z_{k+1},\ldots,z_n).$$

Then

$$H(z_0,\dots,z_{k-1},\underbrace{1}_{k},z_{k+1},\dots,z_n)=\frac{\sum_{j\neq k} m_j|z_j|^2+m_k}{\sum_{j\neq k}|z_j|^2+1}=1+\sum_{j\neq k}(m_j-m_k)|z_j|^2+\dots$$

Since each complex coordinate gives two real coordinates, one has in terms of the symplectic basis

$$\det \mathrm{Hess}_{P_k}(H)=\prod_j (m_j-m_k)^2$$

(because the symplectic volume form is 2^n times the usual one). You can check that this agrees with (9.2.1).

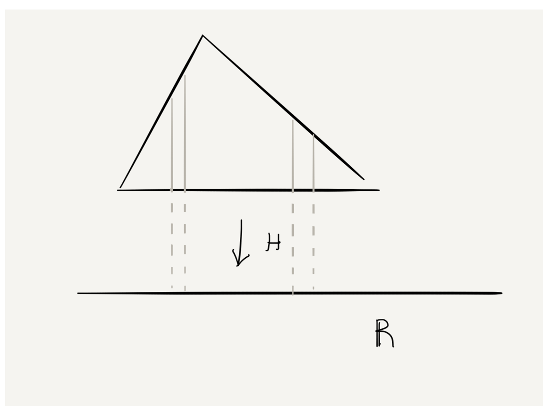
9.3. **Sketch of Duistermaat-Heckman's proof.** We now discuss what goes into the proof of Theorem 9.2.1. We want to show that

$$\int e^{itH}\omega^n = \text{stationary phase expansion} = \sum c_i \frac{e^{it\alpha_i}}{t^d}. \tag{9.3.1}$$

We can push forward the measure ω^n by the map $H: M\to \mathbf{R}$. The result will be $\nu(x)dx$ for some function $\nu(x)$ (which we can think of as the measure of $H^{-1}(x,x+dx)$) which is compactly supported. In these terms, the left side of (9.3.1) can be rewritten as

$$\int e^{itH}\omega^n=\int e^{itx}\nu(x)\,dx=\mathrm{FT}(\nu)(t). \tag{9.3.2}$$

In the example of $\mathbf{CP}^n$, we push forward the simplex Δ^n to $\mathbf{R}$ by the map $\sum m_i x_i$.



For $n=2$ it is clear that $\nu(x)$ is piecewise-linear; in general for $n=d$ it will be piecewise-polynomial of degree $d-1$.

In fact we claim that the formula (9.3.1) forces this sort of behavior. That is, *assuming* the Duistermaat-Heckman localization theorem we will show that $\nu(x)$ is piecewise-linear of degree at most $d-1$. Multiplying (9.3.1) by t^d and using (9.3.2) implies

$$t^d\,\mathrm{FT}(\nu)=\mathrm{FT}(\sum c_i\delta_{\alpha_i})$$

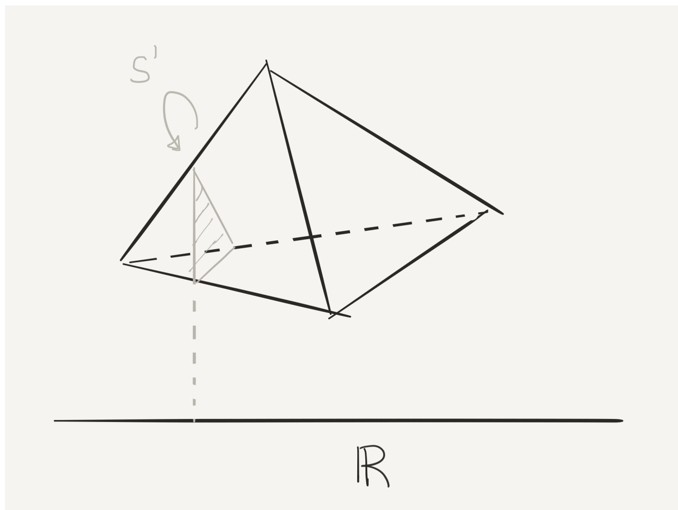
so

$$\mathrm{FT}\left((\frac{d}{dt})^d\nu\right)=\mathrm{FT}(\text{sum of }\delta\text{ measures})$$

so $\left(\frac{d}{dt}\right)^d \nu$ is supported at finitely many points, hence ν is piecewise polynomial.

The argument of Duistermaat-Heckman goes in the converse direction: they begin by showing that ν is piecewise-polynomial, and deduce the theorem from there. We'll outline their argument, and then switch to a different one.

Outline of why ν is piecewise-polynomial.



The first step is to relate $\nu(x)$ to a volume. (What we'll say is valid away from the vertices of the tetrahedron.) For each $x \in \mathbf{R}$, $H^{-1}(x)$ is a manifold of dimension $2d-1$. The restriction $\omega|_{H^{-1}(x)}$ is obviously not symplectic, but it is pulled back from a form on $M_x := H^{-1}(x)/S^1$, which *is* a $2d-2$ dimensional symplectic manifold. First one checks that

$$\nu(x) = \text{volume}(M_x, \omega^{d-1}).$$

Remark 9.3.1. We saw an example of this already (admittedly not in the compact case):

$$(\mathbf{C}^{n+1}, \sum dz_k \wedge d\overline{z}_k).$$

In this case H should be the function

$$H(z) = \sum |z_i|^2$$

This generates the S^1-action of rotating the coordinates (at the same rate). For example, $H^{-1}(S^1) = S^{2n+1}$ and $H|_{S^{2n+1}}$ is the pullback of the Fubini-Study metric on $\mathbf{CP}^n$. (The M_x is the *symplectic reduction* of (M, ω).

In this example

$$M_x = H^{-1}(x)/S^1$$

we can scale all the coordinates to transform M_x to M_1. But this doesn't preserve the symplectic form, i.e. doesn't take ω_x to ω_1. Instead, one finds that $\omega_x \mapsto x \cdot \omega_1$. This is indicative of the general behavior.

In the general case, near $x = x_0$ all the M_x are homeomorphic. So we can think of the cohomological class of the symplectic form $\omega_x \in H^2(M_x, \mathbf{R}) \cong H^2(M_{x_0}, \mathbf{R})$ as a varying class ω_x in the fixed group $H^2(M_{x_0}, \mathbf{R})$, with

$$\mathrm{vol}(M_x) = \int_{M_{x_0}} \omega_x^{d-1}.$$

The key observation of Duistermaat-Heckman was that the function

$$x \mapsto [\omega_x] \in H^2(M_{x_0}, \mathbf{R})$$

is linear. (So the core of the proof is to interpret things in a cohomological way.) So the derivative is a canonical class in

$$\frac{d}{dx}[\omega_x] \in H^2(M_{x_0}, \mathbf{R}).$$

This derivative is normalized by demanding that X_H is $d/d\theta$. What canonical cohomology class is this? It is a constant multiple of the Chern class of the S^1-fibration $H^{-1}(x_0) \to M_{x_0}$.

Once you know that you want to prove this linearity, the proof itself is quite straightforward: just express everything in terms of a connection and compute. $\square$

9.4. **Proof of Duistermaat-Heckman by deformation to fixed points.** This is due to Deligne-Verligne. For our symplectic manifold (M, ω),

$$\Omega^\bullet(M) = \bigoplus_{0 \le i \le 2d} \Omega^i(M).$$

The idea is to take a one-parameter deformation from $\int e^{itH}\omega^n$ to an integral localized at the critical points.

We introduce a deformation of the usual differential $\eta \mapsto d\eta$. For $X = X_H$, we define

$$d_X \eta = d\omega + X \lrcorner \eta.$$

Note that this is weird: it doesn't preserve the grading on $\Omega^\bullet$ because d raises the degree by 1 and the $X \lrcorner$ decreases it by 1.

The operator d_X is no longer a differential: we have

$$\begin{aligned} d_X^2(\eta) &= d(\eta + X \lrcorner \eta) + X \lrcorner (d\eta + X \lrcorner \eta) \\ &= d(X \lrcorner \eta) + X \lrcorner d\eta \\ &= \mathscr{L}_X \eta. \end{aligned}$$

So we see that the obstruction to $d_X^2 = 0$ is the Lie derivative. This implies that $d_X^2 = 0$ on forms which are fixed by the flow X_H, i.e. forms in $\Omega^*(M)^{S_1}$.

We want to write down an example of a form which is closed for d_X. The symplectic form ω is *not* closed, since

$$d_X \omega = \underbrace{d\omega}_{=0} + X \lrcorner \omega.$$

We claim that $X_H \lrcorner \omega = dH$: indeed, we have

$$X_H \lrcorner \omega(Y) = \omega(X_H, Y) = dH(Y).$$

Therefore, $d_X H = dH$, so

$$d_X(\omega - H) = 0$$

i.e. $\omega - H$ is d_X-closed.

Exercise 9.4.1. Check that the form $\omega - H$ is fixed by S^1. (This is basically a tautology once you remember how we defined everything.)

Recall that d is a derivation in the sense that

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^m \omega \wedge d\eta.$$

The same is true for $X \lrcorner$, hence also for d_X. This implies that $(\omega - H)^\ell$ is also killed by d_X.

Similarly, expressions like

$$e^{(\omega - H)} = e^{-H} e^{\omega}$$

interpreted as a formal series (which truncates at $\dim M$), are also d_X-closed. (Note that since ω is in even degree, there are no worries about commutativity vs anti-commutativity.)

Definition 9.4.2. We define the functional

$$\int_M : \Omega^*(M) \to \mathbf{C}$$

which simply integrates the top-degree component:

$$\sum_{i=0}^{\dim M} \omega_i \mapsto \int_M \omega_{\dim M}.$$

Remark 9.4.3. Stokes' Theorem says that $\int_M d\eta = 0$ if M is closed, and the same is true for d_X, since $X \lrcorner$ can't produce any terms in top degree. (This will imply that the integral descends to a functional on cohomology.)

Lemma 9.4.4. *Supopse $d_X \alpha = 0$ and $\gamma \in (\Omega^1)^{S^1}$. Set $\beta = d_X \gamma \in \Omega^0 \oplus \Omega^2$. Then*

$$\int_M \alpha e^{-s\beta} = \int_M \alpha \quad \textit{for all } s \in \mathbf{R}.$$

Remark 9.4.5. We'll apply this with $s \to \infty$ and a non-negative function β, so that the integrand $\alpha e^{-s\beta}$ concentrates around the zeros of β.

Proof. Using that $d_X \alpha = 0$ and $d_X \gamma = 0$, we have

$$\frac{d}{ds}(\alpha e^{-s\beta}) = \alpha \beta e^{-s\beta} = d_X(\alpha \gamma e^{-s\beta}).$$

□

We want to set this up so that β vanishes precisely at the critical points of H. To do this, choose an S^1-invariant Riemannian metric g on M. Use g to turn X_H into a 1-form γ, defined by

$$\gamma(Y) = g(X_H, Y).$$

Then

$$d_X\gamma = d\gamma + \underbrace{X_H \lrcorner \gamma}_{g(X_H,X_H)}.$$

The function $g(X_H, X_H) \geq 0$ and vanishes exactly when $X_H = 0$, which is equivalent to $H = 0$. The other piece $d\gamma$ is a 2-form, so it's nilpotent. Lemma 9.4.4 implies

$$\begin{aligned}\int_M \alpha &= \int_M \alpha e^{-sg(X_H,X_H)} e^{-s\,d\gamma} \\ &= \int_M \alpha \underbrace{e^{-sg(X_H,X_H)}}_{\text{function}} \sum_{j=1}^{d} \frac{(-s)^j}{j} \underbrace{(d\gamma)^j}_{2j-\text{form}}\end{aligned}$$

Near a critical point, this looks like

$$\sum_{j=1}^{d} \frac{(-s)^j}{j!} \underbrace{\int (\alpha \cdot (d\gamma)^j) e^{-s\text{Gaussian}}}_{\sim s^{-d}}$$

Letting $\dim M = 2d$, the only term that survives as $s \to \infty$ is $j = d$. So

$$\int_M \alpha = \frac{(-s)^d}{d!} \int_M \alpha e^{-sg(X_H,X_H)} (d\gamma)^d. \tag{9.4.1}$$

On the right side of (9.4.1), the only term that contributes is α_0. On the left side of (9.4.1), the only term that contributes is α_{2d}. Also, the right side localizes over the critical points as

$$\sum_{P \text{ critical point}} \alpha_0(P) \left[\lim_{s\to\infty} \frac{(-s)^d}{d!} \int_M \alpha e^{-sg(X_H,X_H)} (d\gamma)^d \right].$$

Probably this implies the result by pure thought, because once we know the result only depends on the values at the critical points we can probably deduce that it must be the stationary phase asymptotic.

However, we're actually going to compute the answer. So we want to know the local structure near the critical points. The first step is to linearize the S^1-action. S^1 fixes any critical point P (since the S^1 action is generated by the flow associated to dH), so S^1 acts on $T_P M$.

Lemma 9.4.6. *A symplectic vector space with a (linear) S^1-action is a sum of*

$$(\mathbf{R}^2, dx \wedge dy, \theta \in S^1 \text{ rotates by } n\theta).$$

Proof. Just linear algebra - left as exercise. $\square$

So we can choose coordinates $x_1, \ldots, x_d$ and $y_1, \ldots, y_d$ centered at P such that the action of θ on S^1 is given by a block diagonal matrix with blocks

$$R_{n_i\theta} = \begin{pmatrix} \cos n_i\theta & -\sin n_i\theta \\ \sin n_i\theta & \cos n_i\theta \end{pmatrix}$$

plus quadratic terms. The vector field X_H is then

$$X_H = \sum_{i=1}^{d} n_i(y_i \partial_{x_i} - x_i \partial_{y_i}) + (\text{higher order terms}).$$

From the form of ω, we see that (by Example 8.2.2)

$$H = \frac{1}{2}\sum n_i(x_i^2 + y_i^2) + (\text{higher order terms}).$$

We choose an S^1-invariant metric g such that $\partial_{x_i}, \partial_{y_i}$ are an orthonormal basis at P. This tells us that

$$g(X_H, X_H) = \sum n_i^2(x_i^2 + y_i^2) + (\text{higher order terms}).$$

Finally, the form γ was obtained by dualizing the vector field, so by the form of g we just replace the ∂'s with d's:

$$\gamma = \sum n_i(y_i\, dx_i - x_i\, dy_i) + (\text{higher order terms}).$$

(No n_i vanishes because we assumed that the critical points are non-degenerate.) Then

$$d\gamma = \sum -2n_i(dx_i \wedge dy_i) + (\text{higher order terms}).$$

We can ignore the higher order terms in the limit $s \to \infty$.

The conclusion is that (after a little manipulation)

$$\begin{aligned}\int_M \alpha &= \sum_{P \text{ critical point}} \alpha_0(P)\left[\lim_{s\to\infty}(-s)^d \prod_{i=1}^{d}(-2n_i)\underbrace{\int_M e^{-s\sum n_i^2(x_i^2+y_i^2)}\, dx_i\, dy_i}_{\prod_{i=1}^{d}\frac{2\pi}{2n_i s}}\right] \\ &= \sum_P \alpha_0(P)(2\pi)^d \prod_{i=1}^{d}\frac{1}{n_i}\end{aligned}$$

Now apply this to $\alpha = e^{it(H-\omega)}$, which we saw earlier was a closed form. The left hand side becomes

$$\begin{aligned}\int_M e^{itH}\left(\sum_j \frac{(-i\omega)^j}{j!}\right) &= \int_M e^{itH}(-it)^d \frac{\omega^d}{d!} \\ &= (2\pi)^d \sum_P e^{itH(P)} \prod \frac{1}{n_i(P)}\end{aligned}$$

so the conclusion is then that

$$\int e^{itH}\frac{\omega^d}{d!} = \frac{(2\pi)^d}{t^d}\sum_P e^{itH(P)}\prod_{k=1}^{d}\frac{i}{n_k(P)}. \tag{9.4.2}$$

9.5. **Equivariant cohomology.** The Duistermaat-Heckman theorem inspired Atiyah-Bott's work on equivariant cohomology [AB84], which we discuss next.

In the preceding subsection we introduced the deformation

$$d_X\omega = d\omega + X\lrcorner\omega.$$

We showed that if $d_X\alpha = 0$, then

$$\int_M \alpha = \sum_{P \text{ crit point}} \alpha(P)c_P.$$

To be clear, writing $\alpha = \sum \alpha_i$ the left hand side only depends on $\alpha_{\dim M}$ and the right hand side only depends on α_0. This makes sense because the condition $d_X\alpha = 0$ forces the different components to be related in some way.

For $\alpha \in \ker(d_X \text{ on } \Omega^*(M)^{S^1})$, the evaluation

$$\alpha \mapsto \alpha(P_i)$$

for a critical point P_i depends only on the cohomology class of α. To see this, we just have to show that the evaluation map vanishes on d_X-boundaries. The component d raises degree and thus always vanishes under the evaluation. Since X vanishes at the critical points, for any 1-form β we have $X\beta(P_i) = 0$, so the evaluation map also kills any term of the form $X\lrcorner(-)$. Thus we have a map

$$\frac{\ker(d_X \text{ on } \Omega^*(M)^{S^1})}{\operatorname{im}(d_X \text{ on } \Omega^*(M)^{S^1})} \to \bigoplus_{P_i} \mathbf{R}. \tag{9.5.1}$$

This is actually an isomorphism. If you believe this, then the Duistermaat-Heckman formula becomes very reasonable because $\int_M \alpha$ is also a function on cohomology, so it can be expressed as a sum over the fixed points P_i of some function of $\alpha(P_i)$.

Suppose S^1 acts freely on a compact manifold M. How can we compute $H^*(M/S^1)$ using differential forms on M? De Rham theory says that

$$H^*(M/S^1) = H^*(\Omega_{M/S^1}).$$

If we have a differential form on M/S^1, we can pull it back to a differential form on M, and it will land in the S^1-invariants.

Example 9.5.1. This isn't an isomorphism, as one can see in the simple example $M = Y \times S^1$. In this case, the is an invariant form $d\theta$ on S^1, which we can also pull back to M. So

$$\Omega^*(M)^{S^1} = \pi^*\Omega_Y^* \oplus (\pi^*\Omega_Y^*) \wedge d\theta.$$

To get rid of the extra part, we impose the condition that $X\lrcorner\omega = 0$. We now have an isomorphism

$$\{\text{diff. forms on } M/S^1\} \xrightarrow{\sim} \{\omega \in \Omega^*(M)^{S^1} : X\lrcorner\omega = 0\}.$$

Remark 9.5.2. We record for later use that if $X\lrcorner\omega = 0$, then $\omega = X\lrcorner\omega'$. To prove this, use a partition of unity to reduce to the product case.

The conclusion is that *if S^1 acts freely on M*, then

$$H^*(M/S^1) = H^*(\{\omega \in \Omega^*(M)^{S^1} : X \lrcorner \omega = 0\}). \tag{9.5.2}$$

But if the S^1-action isn't free, then neither side is well-behaved. What we mean by this is that a given a map $f : M \to M'$ that commutes with the S^1-action and is a homotopy equivalence, it doesn't necessarily induce an isomorphism for the quotient by S^1 on either side of 9.5.2.

You can fix the left hand side by using *equivariant cohomology* $H^*_{S^1}(M)$, obtained by replacing M by a homotopy-equivalent space on which S^1 acts freely.

We want to talk about how you can fix the right hand side correspondingly. The process of taking invariant forms for a group action behaves well. The problem is with the condition $X \lrcorner \omega = 0$. The construction of taking levelwise kernel on a complex is not well-behaved.

Note that $X \lrcorner (X \lrcorner \omega) = 0$. So the map $\epsilon \colon \omega \mapsto X \lrcorner \omega$ gives an action of the algebra $A := \mathbf{R}[\epsilon]/\epsilon^2$ on the complex $\Omega^*(M)^{S^1}$ (shifting degrees). In these terms, we can think of things killed by $X \lrcorner$ as

$$\{\omega \colon X \lrcorner \omega = 0\} = \mathrm{Hom}_A(A/\epsilon, \Omega(M)^{S^1}).$$

This clarifies what we should do. This construction is bad because A/ϵ is not a projective A-module, and the way to correct it is to replace it by a projective resolution:

$$P := \ldots \xrightarrow{\epsilon} A \xrightarrow{\epsilon} A \xrightarrow{\epsilon} A \to \ldots \xrightarrow{\epsilon} A.$$

Then $\mathrm{Hom}(P, \Omega^*(M)^{S^1})$ is the total complex of the double complex

$$\Omega^*(M)^{S^1} \xrightarrow{\epsilon} \Omega^*(M)^{S^1} \xrightarrow{\epsilon} \Omega^*(M)^{S^1} \xrightarrow{\epsilon} \ldots \tag{9.5.3}$$

(We'll write this out explicitly shortly.)

Note that if the S^1-action is free then $X \lrcorner = 0 \implies \omega = X \lrcorner \omega'$ (Remark 9.5.2), so (9.5.3) is quasi-isomorphic to

$$\ker(\epsilon \text{ on } \Omega^*(M)^{S^1}) = \{\omega \in \Omega^*(M)^{S^1} : X \lrcorner \omega = 0\}.$$

The double complex (9.5.3) is

$$\begin{tikzcd}
\vdots & \vdots & \vdots & \\
\Omega^3(M)^{S^1} \arrow[r, "X \lrcorner"] \arrow[u, phantom] & \Omega^2(M)^{S^1} \arrow[r] & \Omega^*(M)^{S^1} \arrow[r] & \ldots \\
\Omega^2(M)^{S^1} \arrow[r, "X \lrcorner"] \arrow[u, "d"] & \Omega^1(M)^{S^1} \arrow[r] \arrow[u, "d"] & \Omega^*(M)^{S^1} \arrow[r] \arrow[u, "d"] & \ldots \\
\Omega^1(M)^{S^1} \arrow[r, "X \lrcorner"] \arrow[u, "d"] & \Omega^0(M)^{S^1} \arrow[r] \arrow[u, "d"] & 0 \arrow[r] \arrow[u, "d"] & \ldots \\
\Omega^0(M)^{S^1} \arrow[r, "X \lrcorner"] \arrow[u, "d"] & 0 \arrow[r] \arrow[u, "d"] & 0 \arrow[r] \arrow[u] & \ldots
\end{tikzcd}$$

The total complex is then

$$\Omega^0 \xrightarrow{d_X} \Omega^1 \xrightarrow{d_X} \Omega^0 \oplus \Omega^2 \xrightarrow{d_X} \Omega^1 \oplus \Omega^3 \to \dots \tag{9.5.4}$$

Theorem 9.5.3. *The complex* (9.5.4) *computes* $H^*_{S^1}(M)$.

The Cartan model for equivariant cohomology. We can rewrite the double complex as follows. Consider the space of polynomial functions

$$\mathbf{R} \to (\Omega^*)^{S^1}$$

graded by $\deg(t^i\omega_j) = 2i + j$. For example, the terms of degree 3 are $t\omega_1 + \omega_3$. The differential of a polynomial P has value at $t \in \mathbf{R}$ being $d_{tX}P(t)$. (More intrinsically, $\mathbf{R} = \mathrm{Lie}(S^1)$.) This is the *Cartan model.* Think of it as a family of differential forms in $\Omega^*(M)^{S^1}$ indexed by $\mathbf{R} \cdot X$, and the derivative at t is the derivative of the form using the vector field tX.

Theorem 9.5.4. *The inclusion* $M^{S^1} \hookrightarrow M$ *induces an almost isomorphism*

$$H^*_{S^1}(M) = H^*_{S^1}(M^{S^1}).$$

The meaning of "almost" is that both sides are modules over $\mathbf{R}[t]$*, and it's an isomorphism after inverting* t.

At $t = 1$, you get the isomorphism (9.5.2). Atiyah-Bott traced through it to prove the Duistermaat-Heckman theorem. It's basically an issue of finding that the constants are related to the normal bundles at the fixed points.

10. The Kirillov character formula for U_n

10.1. **The Weyl character formula for U_n.** First we'll compute the characters of the irreducible representations of U_n, following Weyl's classical proof.

Let $T \subset U_n$ be the subgroup of diagonal matrices

$$\begin{pmatrix} z_1 = e^{i\theta_1} & & & \\ & z_2 = e^{i\theta_2} & & \\ & & \ddots & \\ & & & z_n = e^{i\theta_n} \end{pmatrix}$$

The irreducible representations of T are all of the form

$$\begin{pmatrix} z_1 & & & \\ & z_2 & & \\ & & \ddots & \\ & & & z_n \end{pmatrix} \mapsto \prod z_i^{m_i} =: \underline{z}^{\underline{m}}.$$

Since any unitary matrix can be diagonalized, to describe the character it suffices to describe the character restricted to T.

$$\chi_V \begin{pmatrix} z_1 & & & \\ & z_2 & & \\ & & \ddots & \\ & & & z_n \end{pmatrix} = \sum_{\underline{m}\in\mathbf{Z}^n} c_{\underline{m}} \underline{z}^{\underline{m}}, \quad c_{\underline{m}} \in \mathbf{Z}_{\geq 0}.$$

Lemma 10.1.1. *For any f conjugacy-invariant on U_n,*

$$\int_{U_n} f = \frac{1}{n!} \int_T f \cdot \prod_{i<j} |z_i - z_j|^2 s$$

Proof. Compute the Jacobian of the conjugation map

$$\begin{aligned} U_n/T \times T &\to U_n \\ (g,t) &\mapsto g t g^{-1}. \end{aligned}$$

□

We know that if V is irreducible, then

$$\langle \chi_V, \chi_V \rangle = 1$$

so Lemma 10.1.1 implies

$$\int_T \left| \sum c_{\underline{m}} \underline{z}^{\underline{m}} \prod_{i<j} (z_i - z_j) \right|^2 = n!.$$

Note that $c_{\underline{m}} \underline{z}^{\underline{m}}$ is symmetric under permutation and $\prod_{i<j}(z_i - z_j)$ is antisymmetric, so

$$\sum c_{\underline{m}} \underline{z}^{\underline{m}} \prod_{i<j} (z_i - z_j) =: \sum d_{\underline{m}} \underline{z}^{\underline{m}}$$

is anti-symmetric. So

$$\int_T \left| \sum d_{\underline{m}} z^{\underline{m}} \right|^2 = \sum |d_{\underline{m}}|^2 = n!.$$

Since $d_{\underline{m}} \in \mathbf{Z}$ and is anti-symmetric, we conclude that

$$\sum d_{\underline{m}} z^{\underline{m}} = \pm \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \underline{z}^{\sigma(\underline{m}_0)} \quad \text{for some } \underline{m}_0 \in \mathbf{Z}^n.$$

This implies that

$$\chi_V(\underline{z}) = \pm \frac{\sum_{\sigma \in S_n} \underline{z}^{\sigma(\underline{m}_0)}}{\prod (z_i - z_j)} \tag{10.1.1}$$

Remark 10.1.2. The fact that characters are complete is sufficient to imply that every m_0, up to permutation, occurs.

To summarize, for every $m_1 > m_2 > \ldots > m_n$, with $m_i \in \mathbf{Z}$, there exists an irreducible representation V of U_n with character (10.3.1). We rewrite (10.3.1) in a different way, interpreting both the numerator and denominator as a determinant. The numerator

$$\sum_{\sigma \in S_n} \underline{z}^{\sigma(\underline{m}_0)} = \det(z_i^{m_j})$$

and the denominator is manifestly a Vandermonde determinant. So we have

$$\chi(\underline{z}) = \frac{\det(z_i^{m_j})}{\prod (z_j - z_i)}$$

Remark 10.1.3. One feature that makes the character theory simple is that there is an abelian subgroup that meets every conjugacy class. This can never happen in a finite group: by counting arguments you can show that no proper subgroup meets every conjugacy class.

If we write

$$\chi_V = \sum_{\underline{k} \in \mathbf{Z}^n} c_{\underline{k}} \underline{z}^{\underline{k}}$$

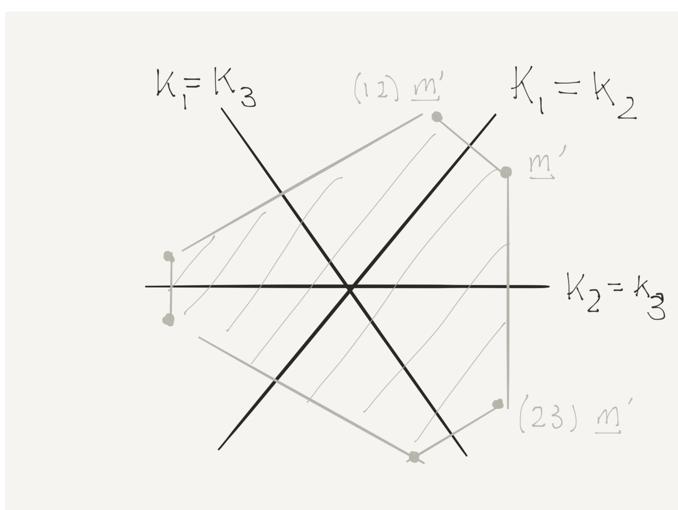
which $\underline{k}$ occur? (These are called the *weights* of V.)

This question is answered by highest weight theory. It says that $\underline{m}' = (m_1 - (n-1), m_2 - (n-2), \ldots, m_n)$ occurs, with coefficient $c_{\underline{m}'} = 1$. This (and its orbit under S_n) is the *highest weight*, which means that if $c_{\underline{k}} \neq 0$ then $\|\underline{k}\| \leq \|\underline{m}'\|$.

Example 10.1.4. We consider U_3. Then $V \longleftrightarrow \underline{m} = (m_1 > m_2 > m_3)$ has highest weight $\underline{m}' = (m_1 - 2, m_2 - 1, m_3)$. All $k = (k_1, k_2, k_3)$ with $c_{\underline{k}} \neq 0$ have

$$\sum k_i' = (m_1 - 2) + (m_2 - 1) + m_3$$

which you can prove by looking at the action of the center.



Fact 10.1.5. If $c_{\underline{k}} \neq 0$, then $\underline{k}$ lies in the shaded hexagon, which is the convex hull of $(\sigma \underline{m}' : \sigma \in S_3)$.

How do the $c_{\underline{k}}$ behave inside this hexagon? We'll come back and discuss this after we do the Kirillov character formula.

10.2. **The Kirillov character formula.**

We want to establish a formula

$$\chi_V \cdot \sqrt{j} = \mathrm{FT}\Big(\frac{1}{d!}\left(\frac{\omega}{2\pi}\right)^d \text{ on } \mathcal{O}\Big)$$

for some orbit $\mathcal{O}$ of G on $\mathfrak{g}^*$, which we will determine.

Remark 10.2.1. In the nilpotent case all orbits occur, but in this case only "discretely many" do.

First we're going to finally fulfill a promise made long ago (Fact 1.2.1) about a formula for $j(X)$. Recall that $j(X)$ is the Jacobian at X of the exponential map $\exp\colon \mathfrak{g} \to G$, with reference to the Lebesgue measure on $\mathfrak{g}$ and the right invariant measure on G.

Lemma 10.2.2. *We have*

$$j(X) = \det\left(\frac{e^{\mathrm{ad}\, X} - 1}{\mathrm{ad}\, X}\right) = \prod_{\lambda \text{ eigenvalue of } \mathrm{ad}\, X} \frac{e^{\lambda} - 1}{\lambda}.$$

In other words, we have an endomorphism $\mathrm{ad}(X)\colon \mathfrak{g} \to \mathfrak{g}$ and we can apply $\frac{e^z - 1}{z}$ to this endomorphism.

Proof. We need to compute the derivative of the exponential map. This means that we should study $\exp(X + \delta X)$, where δX is a small element of $\mathfrak{g}$. (The symbol δX is unrelated to X.) You can write out a power series for this, but it's a little confusing because they don't commute. However, they commute to first order, so

$$\exp(X + \delta X) = \lim_{N\to\infty} \exp\left(\frac{X + \delta X}{N}\right)^N = \lim_{N\to\infty} \left(\exp\left(\frac{X}{N}\right)\exp\left(\frac{\delta X}{N}\right)\right)^N$$

We can expand out the term in the limit as

$$\begin{aligned}
\left(\exp\left(\frac{X}{N}\right)\exp\left(\frac{\delta X}{N}\right)\right)^N &= \underbrace{e^{X/N}}_{=:g} e^{\delta X/N} e^{X/N} e^{\delta X/N} \dots \\
&= (g e^{\delta X/N} g^{-1})(g^2 e^{\delta X/N} g^{-2}) \dots (g^N e^{\delta X/N} g^{-N}) g^N \\
&= e^{\mathrm{Ad}(g)\delta X/N} e^{\mathrm{Ad}(g^2)\delta X/N} \dots e^{\mathrm{Ad}(g^N)\delta X/N} e^X \\
&= \exp\left(\sum \frac{\mathrm{Ad}(g^i)\delta X}{N} + O(|\delta X|^2)\right) e^X \\
&\xrightarrow{N\to\infty} \exp\left(\int_0^1 \mathrm{Ad}(e^{tX})\delta X + O(|\delta X|^2)\right) e^X
\end{aligned}$$

Finally,

$$\int_0^1 \mathrm{Ad}(e^{tX})\delta X = \int_0^1 e^{t\,\mathrm{ad}(X)}\delta X = \frac{e^{\mathrm{ad}\, X} - 1}{\mathrm{ad}\, X}\delta X.$$

□

Remark 10.2.3. The significance of using a *right*-invariant Haar measure was the e^X that popped out at the right end.

For $U_n = \{A \mid A\overline{A}^T = \mathrm{Id}_n\}$, we have

$$\mathrm{Lie}(U_n) = \mathfrak{u}_n = \{A \mid A + \overline{A}^T = 0\} = \{iX \mid X \text{ hermitian}\}.$$

If

$$X = i\begin{pmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_n \end{pmatrix}$$

the eigenvalues of $\mathrm{ad}\, X : \mathfrak{u}_n \to \mathfrak{u}_n$ are $\{\theta_i - \theta_j \mid 1 \le i, j \le n\}$. Therefore

$$j(X) = \prod_{1 \le k, l \le n} \frac{e^{i(\theta_k - \theta_l)} - 1}{i(\theta_k - \theta_l)}$$

This will mostly (but not entirely) cancel the denominator of the Weyl character formula.

$$\begin{aligned} j(X) &= \prod_{1 \le k, l \le n} \frac{e^{i(\theta_k - \theta_l)} - 1}{i(\theta_k - \theta_l)} \\ &= \prod_{k<l} \frac{e^{i(\theta_k - \theta_l)} - 1}{i(\theta_k - \theta_l)} \frac{e^{i(\theta_l - \theta_k)} - 1}{i(\theta_l - \theta_k)} \\ &= \prod_{k<l} e^{-i(\theta_k + \theta_l)} \left[\frac{e^{i\theta_k} - e^{i\theta_l}}{i(\theta_k - \theta_l)} \right]^2 \end{aligned}$$

So the conclusion is that

$$j(X) = \underbrace{\prod_k e^{-i(n-1)\theta_k}}_{(\det e^X)^{-(n-1)}} \prod_{l<k} \left[\frac{e^{i\theta_k} - e^{i\theta_l}}{i(\theta_k - \theta_l)} \right]^2$$

Recall that

$$\chi(\underline{z}) = \frac{\det(z_l^{m_k})}{\prod_{k<l}(z_l - z_k)}$$

where z_k corresponds to $e^{i\theta_k}$.

$$\chi(e^X)\sqrt{j(X)} = (\det e^X)^{-(n-1)/2} \frac{\det e^{i m_k \theta_l}}{\prod_{k<l} i(\theta_k - \theta_l)}.$$

The square root makes the global validity of the formula concerning, but near the identity there is no confusion.

We now prepare to apply Duistermaat-Heckman. Write

$$\mathfrak{u}_n = \mathrm{Lie}(U_n) = \{iA \mid A \text{ Hermitian}\}.$$

We identify $\mathfrak{u}_n^* = \{B \text{ hermitian}\}$, the space of Hermitian matrices, via

$$(iA \mapsto \mathrm{Tr}(AB)) \leftarrow B.$$

In other words, the pairing on $\mathfrak{u}_n \times \mathfrak{u}_n^*$ is

$$\langle iA, B\rangle = \mathrm{Tr}(AB).$$

Definition 10.2.4. For $\underline{\lambda} = (\lambda_1, \dots, \lambda_n) \in \mathbf{R}^n$, the orbit $\mathscr{O}_{\underline{\lambda}}$ is the conjugacy class of

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \in \mathfrak{u}_n^*.$$

In other words (diagonalizability of Hermitian matrices) $\mathscr{O}_{\underline{\lambda}}$ is the set of all Hermitian matrices with eigenvalues $\lambda_1, \dots, \lambda_n$.

We want to compute the Fourier transform of the symplectic measure for $X \in \mathfrak{u}_n$.

$$\mathrm{FT}_{\mathscr{O}_{\underline{\lambda}}}(X) = \int_{\xi \in \mathscr{O}_{\underline{\lambda}}} \frac{1}{d!}\left(\frac{\omega}{2\pi}\right)^d e^{i\langle X, \xi\rangle}.$$

The function on $\mathscr{O}_{\underline{\lambda}}$ given by $\xi \mapsto \langle \xi, X\rangle$ generates a flow, which is the vector X^* field on $\mathscr{O}_{\lambda}$ associated to X. (Another way to say this is that G acts on $\mathscr{O}_{\underline{\lambda}}$, and the flow is the one generated by the action in the direction of X.)

Take

$$X = i\begin{pmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_n \end{pmatrix}$$

The vector field X^* generates an S^1-action if $(\theta_1, \dots, \theta_n) = \theta(m_1, \dots, m_n)$ with $m_i \in \mathbf{Z}$. We need to assume that the m_i are distinct in order to apply Duistermaat-Heckman, which will ultimately correspond to the representation-theoretic requirement on the weights.

The action of $\theta \in S^1$ is conjugation by

$$\begin{pmatrix} e^{im_1\theta} & & & \\ & e^{im_2\theta} & & \\ & & \ddots & \\ & & & e^{im_k\theta} \end{pmatrix}$$

Duistermaat-Heckman says that (9.4.2)

$$\mathrm{FT}_{\mathscr{O}_{\underline{\lambda}}}(X) = \sum_{S^1\text{-fixed points } P \in \mathscr{O}_{\underline{\lambda}}} \frac{e^{i\langle X, P\rangle}}{\prod (in_\alpha \theta)}$$

where the n_α come from the action of S^1 on $T(P)$:

$$(T(P), \omega, S^1) = \bigoplus (\mathbf{R}^2, \mathrm{std}, \theta \text{ rotates by } n_\alpha\theta).$$

Under our assumption that $\theta_i \neq \theta_j$, i.e. $m_i \neq m_j$, the S^1-fixed points are the diagonal Hermitian matrices in $\mathcal{O}_{\underline{\lambda}}$, which is just the orbit under permutation:

$$\left\{ \begin{pmatrix} \lambda_{\sigma(1)} & & \\ & \ddots & \\ & & \lambda_{\sigma(n)} \end{pmatrix} : \sigma \in S_n \right\}.$$

So

$$e^{i\langle X,P\rangle} = e^{i\sum \lambda_{\sigma(k)}\theta_k}.$$

Now,

$$X = i\begin{pmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_n \end{pmatrix} = i\theta \begin{pmatrix} m_1 & & \\ & \ddots & \\ & & m_n \end{pmatrix}$$

so $\{n_\alpha\}$ at P_σ is $\{m_{\sigma(l)} - m_{\sigma(k)} \colon k < l\}$. So

$$\begin{aligned} \mathrm{FT}_{\mathcal{O}_{\underline{\lambda}}}(X) &= \sum_{\sigma \in S_n} \frac{\mathrm{sgn}(\sigma) e^{i\sum \lambda_{\sigma(k)}\theta_k}}{i\prod_{k<l}(m_l - m_k)\theta} \\ &= \frac{\det(e^{i\lambda_k\theta_l})_{1\le k,l\le n}}{\prod_{k<l} i(\theta_l - \theta_k)}. \end{aligned}$$

(We proved this for a dense set of θ's, but it follows for all of them by continuity.)

Remark 10.2.5. The interesting quaitative features of this computation were that the $\sqrt{j}$ almost cancels the denominator in the Weyl Character Formula, and the sum over S_n corresponds bijectively to fixed points.

The **conclusion** is that

$$\chi_{(m_1,\dots,m_n)}\sqrt{j}(e^X) = \mathrm{FT}_{\mathcal{O}_{\underline{\lambda}}}(X)$$

where

$$\lambda = \begin{pmatrix} m_1 & & & \\ & m_2 & & \\ & & \ddots & \\ & & & m_n \end{pmatrix} - \frac{n-1}{2}\,\mathrm{Id}.$$

Remark 10.2.6. Note that for V the trivial representation, $(m_1,\dots,m_n) = (n-1, n-2, 0, \dots, 0)$ so $\mathcal{O}_V$ is *not* $\{0\}$.

So we've seen that the irreducible characters of U_n corresponds to *certain* coadjoint orbits, namely the orbits of

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}, \quad \lambda_1 > \lambda_2 > \ldots > \lambda_n, \quad \lambda_i \in \frac{n-1}{2} + \mathbf{Z}.$$

10.3. **Applications.** This supports the idea that the same picture holds for U_n as for nilpotent groups, namely that there is a basis for V indexed by balls of symplectic volume 1 on $\mathcal{O}_V$, with the ball B centered at ξ being an approximate eigenvector in the sense that

$$e^X v_B \approx e^{i\langle \xi, X\rangle} v_B.$$

This heuristic picture lets us guess things about representation theory.

10.3.1. *Weights for U_3.* Let V be the irreducible representation indexed by $(m_1 > m_2 > m_3)$.

Then the character of V is

$$\chi_V \begin{pmatrix} z_1 & & \\ & z_1 & \\ & & z_3 \end{pmatrix} = \sum_{\underline{k} \in \mathbf{Z}^3} c_{\underline{k}} \underline{z}^{\underline{k}}.$$

Here $c_{\underline{k}}$ is the multiplicity of the character

$$\begin{pmatrix} z_1 & & \\ & z_1 & \\ & & z_3 \end{pmatrix} \mapsto z_1^{k_1} z_2^{k_2} z_3^{k_3}$$

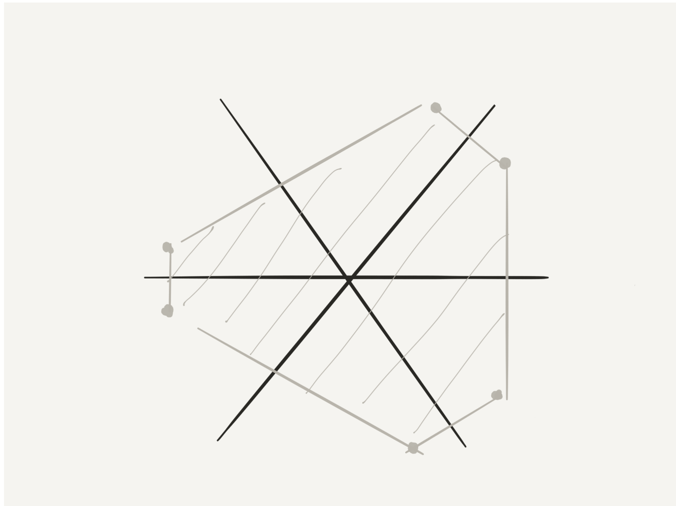
in the restriction of V to the diagonal subgroup.

We want to guess what $c_{\underline{k}}$ is. The picture says that V corresponds to an orbit $\mathcal{O}_V \subset \mathfrak{u}_3^*$. Let T be the diagonal (maximal) torus and $\mathfrak{t} \in \mathfrak{u}_3$ be its Lie algebra. As we saw, if V has weights (m_1, m_2, m_3) then $\mathcal{O}_V$ consists of hermitian matrices with eigenvalues $(m_1 - 1, m_2 - 1, m_3 - 1)$.

A ball of volume 1 corresponds to an approximate eigenvector for U_n, hence also for T. So $V|_T$ corresponds to projecting $\mathcal{O}$ to $\mathfrak{t}^*$. Identifying $\mathfrak{u}_3^*$ with Hermitian matrices, the projection $\mathfrak{u}_3^* \to \mathfrak{t}^*$ sends a hermitian matrix to the diagonal. So the image consists of the diagonals of 3×3 hermitian matrices with fixed eigenvalues $(\mu_1, \ldots, \mu_n)$. This was worked out by Schur: it is the convex hull of permutations of

$$\begin{pmatrix} \mu_{\sigma(1)} & & \\ & \mu_{\sigma(2)} & \\ & & \ddots \end{pmatrix}$$

For $n = 3$, the picture is



This suggests that the weights $c_{\underline{k}}$ are non-zero exactly when $\underline{k}$ is in the convex hull $(m_{\sigma(1)} - 1, m_{\sigma(2)} - 1, m_{\sigma(3)} - 1)$. This is almost true, but the shifts are not quite right.

We can also estimate how the $c_{\underline{k}}$ are varying. We expect $c_{\underline{k}}$ to be approximately the volume of the pre-image of $\underline{k}$ in $\mathscr{O}$. To calculate the volume, we can push forward the symplectic measure on $\mathscr{O}$ to $\mathfrak{t}^*$. By the same principle as in Duistermaat-Heckman, this measure is piecewise polynomial of the form $L(\underline{k}) \cdot$ Lebesgue where $L(\underline{k})$ is linear.

Exercise 10.3.1. Analyze the tensor product of two irreducibles of SO_3. If $\dim V = 2n+1$ and $\dim W = 2m + 1$, then $\mathscr{O}_V$ is the sphere of radius $n + 1/2$ and $\mathscr{O}_W$ is the sphere of radius $m + 1/2$ in $\mathfrak{so}_3^* = \mathbf{R}^3$. Consider taking the convolution of measures on $\mathscr{O}_V + \mathscr{O}_W$, corresponding to $\mathfrak{so}_3 \hookrightarrow \mathfrak{so}_3 \times \mathfrak{so}_3$ and $\mathfrak{so}_3^* \times \mathfrak{so}_3^* \to \mathfrak{so}_3^*$.

10.3.2. *The semiclassical heuristic.* Suppose G is a compact Lie group, acting on a compact manifold X. We have a decomposition of G-representations $L^2(X) = \bigoplus V$. Which representations V occur and with what multiplicity?

Think back to the case $X = \mathbf{R}$. The theory of pseudo-differential operators suggested that we can index a basis for $L^2(\mathbf{R})$ by a decomposition of $\mathbf{R}^2$. (The two directions in this $\mathbf{R}^2$ are the "position" and "frequency" in X.) We can think of $\mathbf{R}^2$ as $T^*\mathbf{R}$.

We expect a similar relationship between $L^2(X)$ and the symplectic manifold T^*X. Given a ball B in T^*X centered at $(x \in X, \xi \in T_x^*)$, we should get a corresponding approximate eigenfunction $f_B \in L^2(X)$ which is localized near x, with frequency localized near ξ. Think of f_B as $e^{i\langle \xi, x\rangle}$. This suggests that for $Y \in \mathfrak{g}$ near 0, giving $Y_x \in T_x X$,

$$e^Y f_B \approx e^{i\langle \xi, Y_x\rangle} f_B.$$

The association $Y \mapsto Y_x$ defines a map

$$\mathfrak{g} \to T_x X$$

or dually

$$T_x^* X \to \mathfrak{g}^*$$

which we can promote to a G-equivariant map

$$\Phi\colon T^*X \to \mathfrak{g}^*.$$

In other words, f_B is an approximate eigenvector for G, with eigenvalue given by $\Phi(x,\xi)$. This suggests the following ("semiclassical") heuristic. We consider pushing forward the symplectic volume via Φ_*. Since the map Φ is G-equivariant, the result will be a G-invariant measure, so we can decompose it into measures on the G-orbits:

$$\Phi_*(\text{symplectic measure}) = \int_{\mathscr{O}} \nu_{\mathscr{O}}\, d\mu(\mathscr{O}). \tag{10.3.1}$$

The orbit $\mathscr{O}$ should correspond to a representation V, and the multiplicity of $\pi_{\mathscr{O}}$ should have something to do with $\mu(\mathscr{O})$.

This is obviously can't be right in general, because if G is compact then the representations appearing in $L^2(G)$ are discrete, while the right side of (10.3.1) is continuous - so we need to discretize it in some way.

Example 10.3.2. Consider $G = U_n$ acting on $X = \mathbf{CP}^{n-1}$. The stabilizer of $[0,0,\dots,0,1]$ is

$$H = \begin{pmatrix} \boxed{U_{n-1}} & \\ & \boxed{U_1} \end{pmatrix} \cong U_{n-1} \times U_1.$$

We can view $X = G/H$. So (by Peter-Weyl)

$$L^2(X) = \bigoplus_{\pi \text{ irred.}} (\dim \pi^H)\pi.$$

At x_0, we can identify $T_{x_0}X = \mathfrak{g}/\mathfrak{h}$, hence

$$T^*_{x_0}X = \{\lambda \in \mathfrak{g}^* \colon \lambda|_{\mathfrak{h}} = 0\} = \mathfrak{h}^\perp.$$

The map

$$\Phi\colon T^*_{x_0}X \to \mathfrak{g}^*$$

is then identified with the inclusion

$$\mathfrak{h}^\perp \hookrightarrow \mathfrak{g}^*.$$

We've been identify $\mathfrak{g} = \mathfrak{u}_n^*$ with $n \times n$ hermitian matrices. In these terms,

$$\mathfrak{h}^\perp = \left\{ \begin{pmatrix} 0 & \dots & 0 & Y_1 \\ 0 & \dots & 0 & \vdots \\ 0 & \dots & 0 & Y_{n-1} \\ \overline{Y}_1 & \dots & \overline{Y}_{n-1} & 0 \end{pmatrix} \right\}$$

so $\Phi(T^*X)$ is the union of the G-conjugates of the $\mathfrak{h}^\perp$. What are the eigenvalues of

$$\begin{pmatrix} 0 & \dots & 0 & Y_1 \\ 0 & \dots & 0 & \vdots \\ 0 & \dots & 0 & Y_{n-1} \\ \overline{Y}_1 & \dots & \overline{Y}_{n-1} & 0 \end{pmatrix}?$$

Clearly many of them are 0, except 2. Since the trace is 0, and the sum of the squares of the matrix entries will be the norm of the matrix squared, which is

$$||Y|| = \sqrt{\sum |Y_i|^2}$$

we see that the eigenvalues are

$$(||Y||, 0, \dots, 0, -||Y||).$$

Since the map $\mathfrak{g}^* \to \mathfrak{g}^*/G$ takes a matrix to its eigenvalues, we conclude that the composite map

$$T^*X \to \mathfrak{g}^* \to \mathfrak{g}^*/G$$

has image $\{(r, 0, \dots, 0, -r)\}$. We put the Lebesgue measure on $\mathbf{C}^{n-1} \cong \mathbf{R}^{2n-2}$ and push it forward to $\mathbf{R}$ via the map

$$(x_1, \dots, x_{2n-2}) \mapsto \sqrt{\sum x_i^2}$$

to obtain a measure proportional to $r^{2n-3}\,dr$. Let $\mathcal{O}(r) \subset \mathfrak{g}^*$ be the orbit mapping to $(r, 0, \dots, 0, -r)$. The dimension of this orbit is $4n-6$, so the symplectic volume is homogeneous of degree $2n-3$. Since

$$\Phi_*(\text{symplectic measure}) = c \int_r \nu_{\mathcal{O}(r)}\,dr$$

this suggests something like each $\pi_{\mathcal{O}(r)}$ occurs with the same multiplicity.

In fact

$$L^2(U_n/U_{n-1} \times U_1) = \bigoplus V_{\underline{m}=(m_1,\dots,m_n)}$$

where $\underline{m'} = (m_1 - (n-1), \dots, m_n) = (k, 0, \dots, 0, -k)$.

11. Classical mechanics

11.1. **Newton's laws of motion.** Suppose you have a system of n particles in $\mathbf{R}^3$, with position vectors $(\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n)$, interacting according to a potential $V(\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n)$. Newton's law says

$$m_i \frac{d^2}{dt^2}\vec{q}_i = -\nabla_{\vec{q}_i} V(\vec{q}_1, \dots, \vec{q}_n). \tag{11.1.1}$$

Example 11.1.1. If the ith particle has mass m_i, and the potential is all from gravity, the potential function is

$$V(\vec{q}_1, \dots, \vec{q}_n) = -G \sum_{i \neq j} \frac{m_i m_j}{|\vec{q}_i - \vec{q}_j|}.$$

11.2. **Hamilton's reformulation.** Hamilton observed that you can rewrite this as a flow on a symplectic manifold. One advantage of this is to highlight symmetries that would otherwise be obscured. We introduce the *momentum*

$$\vec{p}_i = m_i \frac{d\vec{q}_i}{dt}.$$

The equations (11.2.1) become

$$\vec{p}_i = m_i \frac{d\vec{q}_i}{dt} \tag{11.2.1}$$

$$\frac{d\vec{p}_i}{dt} = -\nabla_{\vec{q}_i} V(\vec{q}_1, \dots, \vec{q}_n). \tag{11.2.2}$$

Write

$$x = (\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n)$$
$$p = (\vec{p}_1, \vec{q}_2, \dots, \vec{p}_n)$$

Introduce the "Hamiltonian"

$$H(p,q) := \sum \frac{|\vec{p}_i|^2}{2m_i} + V(\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n).$$

Using this we can rewrite (11.2.1) as

$$\frac{d\vec{p}_i}{dt} = -\frac{\partial H}{\partial x_i} \tag{11.2.3}$$

$$\frac{d\vec{x}_i}{dt} = \frac{\partial H}{\partial p_i} \tag{11.2.4}$$

Think of $(p,x) \in \mathbf{R}^{3n} \times \mathbf{R}^{3n}$ as a symplectic manifold with differential form $\sum dp_i \wedge dx_i$.

Recall that we have a Poisson bracket

$$\{f,g\} = \sum \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial x_i} - \frac{\partial f}{\partial x_i}\frac{\partial g}{\partial p_i}.$$

This allows us to rewrite (11.2.3) as

$$\frac{d\vec{p}_i}{dt} = -\frac{\partial H}{\partial x_i} = \{H, p_i\} \tag{11.2.5}$$

$$\frac{d\vec{x}_i}{dt} = \frac{\partial H}{\partial p_i} = \{H, x_i\}. \tag{11.2.6}$$

Recall the meaning of these Poisson brackets. A function H on a symplectic manifold M gives a flow X_H. For $f \in C^\infty(M)$,

$$X_H f = \pm\{H, f\}.$$

So the equations tells us that the time evolution of p and x is given by the Hamiltonian flow X_H associated to H. In particular, H is constant along the flow of X_H, corresponding to conservation of energy.

11.3. **Quantization.** We'll describe the quantization of this system.

In classical mechanics, the state (position and momentum) is represented by a vector in $\mathbf{R}^{6n}$, which we view as $T^*\mathbf{R}^{3n}$. The quantum state is represented by a wavefunction $\psi \in L^2(\mathbf{R}^{3n})$ with $||\psi|| = 1$, which encodes the probability density that a particle is at position $(x_1, \dots, x_n)$.

In classical mechanics observables are represented by a function F on the phase space $T^*\mathbf{R}^{3n}$. The corresponding quantum observable is represented by an operator $\mathrm{Op}(F)$,with $\langle \mathrm{Op}(F)\psi, \psi\rangle$ being the expected value of F.

Classical	Quantum
State $(x, p) \in T^*(\mathbf{R}^{3n})$	State $\psi \in L^2(\mathbf{R}^{3n})$, with $\|\|\psi\|\| = 1$
Observable F on $T^*\mathbf{R}^{3n}$	Operator $\mathrm{Op}(F)$ on $L^2(\mathbf{R}^{3n})$

We remind you how to quantize an observable, i.e. how to construct $\mathrm{Op}(F)$ from F. Let $n = 1$, so the potential is

$$V(x) = \frac{-1}{\sqrt{x_1^2 + x_2^2 + x_3^2}}.$$

Then $L^2(\mathbf{R})$ is the quantization of $\mathbf{R}^2 = T^*\mathbf{R}$. We wrote down a rule Op from functions f on $T^*\mathbf{R}$ to operators. The function x was sent to multiplication by x. The function p was sent to $i\frac{d}{dx}$.

What about xp? You might want to send it to $x \cdot i\frac{d}{dx}$, but this is not self-adjoint, while operators corresponding to observables should be self-adjoint since their eigenvalues are real numbers. So you might try $\frac{1}{2}(\widehat{x}\widehat{p} + \widehat{p}\widehat{x})$, which is the Weyl-twisted operator Op^W, corresponding to a different splitting of the Heisenberg group (Remark 6.3.1).

We also had the Weil representation

$$\mathrm{Lie}(\mathrm{SL}_2\,\mathbf{R}) \to \text{Operators on } L^2(\mathbf{R})$$

and there was a map $\mathrm{Lie}(\mathrm{SL}_2\,\mathbf{R})$ to vector fields on $T^*\mathbf{R}$.

What about a vector field Y on $T^*\mathbf{R}$ not coming from $\mathrm{Lie}(\mathrm{SL}_2\,\mathbf{R})$? We want to attach to Y an operator A_Y such that $A^* = -A$. We're going to follow an analogue of our story for the Weil representation; the analogue of the defining equation (6.1.1) is

$$[A_Y, \mathrm{Op}(f)] = \mathrm{Op}(Yf).$$

Suppose $Y = X_H$ for some Hamiltonian $H \in C^\infty(T^*\mathbf{R})$, i.e. Y preserves the symplectic form. Then $\mathrm{Op}(Yf) = \mathrm{Op}(-\{H, f\})$. Recall that

$$[\mathrm{Op}(f), \mathrm{Op}(g)] = i\,\mathrm{Op}(\{f, g\}) + (\text{higher order terms}).$$

This suggests that $A_Y = i\,\mathrm{Op}(H)$ when $Y = X_H$ associated to H.

This tells us that X_H can be thought of as $i\,\mathrm{Op}(H)$ to first order, which is correct when you restrict to quadratic H.

Finally, we discuss one more aspect of quantization. The time evolution associated to a Hamiltonian H is the flow of the vector field X_H. In quantum mechanics, time evolution is governed by the *time-dependent Schrödinger equation*

$$\frac{\partial \psi}{\partial t} = i\,\mathrm{Op}(H)\psi.$$

Classical	Quantum
State $(x,p)\in T^*(\mathbf{R}^{3n})$	State $\psi \in L^2(\mathbf{R}^{3n})$, with $\|\psi\|=1$
Observable F on $T^*\mathbf{R}^{3n}$	Operator $\mathrm{Op}(F)$ on $L^2(\mathbf{R}^{3n})$
Time evolution (flow of X_H)	Schrödinger equation $\frac{\partial\psi}{\partial t} = i\,\mathrm{Op}(H)\psi$

Suppose you're looking for an observable that doesn't evolve with time. In particular, if $\psi_0 \in L^2(\mathbf{R}^{3n})$ is an eigenfunction of $\mathrm{Op}(H)$, with

$$\mathrm{Op}(H)\psi_0 = E_0\psi_0$$

then $\psi(t) = \psi_0 \cdot e^{iE_0 t}$ is a "stationary" solution to the Schrödinger equation.

In other words, the quantum behavior of the system is governed by the spectrum of $\mathrm{Op}(H)$ acting on $L^2(\mathbf{R}^{3n})$.

11.4. **The Kepler problem.** Consider a singular particle in $\mathbf{R}^3$ with potential

$$V(\vec{q}) \propto \frac{-1}{\|\vec{q}\|}.$$

The periodic nature of planetary orbits is a special feature of this (inverse square) potential. If you change the exponent slightly, that would fail. In the Hamiltonian formulation, the symplectic manifold is $T^*\mathbf{R}^3$s, with

$$H = p_1^2 + p_2^2 + p_3^2 - \frac{1}{\sqrt{x_1^2 + x_2^2 + x_3^2}}.$$

There are some conserved quantities of this flow: the energy H, and the angular momentum vector L_x, L_y, L_z. So this flow should be constrained to a 6—4 = 2-dimensional space. But in fact the flows are 1-dimensional, which suggests that there are more invariants. We'll construct K_x, K_y, K_z, which together with the L's generates Lie(SO_4). The quantization of this system is SO_4 acting on $L^2(\mathbf{R}^3 - \{0\})$.

We'll work in the 2D case, considering $T^*(\mathbf{R}^2)$, so

$$H = p_1^2 + p_2^2 - \frac{1}{\sqrt{x_1^2 + x_2^2}}.$$

If we believe that the orbits are closed, then we should be able to find 3 invariants preserved by the flow. We'll write down 3 conserved quantities.

First, note that the whole system is stable under rotation. Letting $R \in SO_2$, the operation

$$(x_1, x_2) \mapsto R(x_1, x_2)$$
$$(p_1, p_2) \mapsto R(p_1, p_2)$$

preserves H. The corresponding vector field is $(x_1\partial_{x_2} - x_2\partial_{x_1}) + (p_1\partial_{p_2} - p_2\partial_{p_1})$. Since this preserves the symplectic form, it is the flow of some $J \in C^\infty(X)$. The vector field X_J of the flow is defined by

$$dJ(Y) = \omega(X_J, Y),$$

so $J = x_1 p_2 - p_1 x_2$. Then $dH(J)$ is the rate of change of H along the J-flow, which is 0 since the flow of J is rotation. This tells us that $\{J, H\} = 0$. By symmetry, the change of J along the H flow is 0, i.e. J is conserved. This J is angular momentum.

More generally, F is conserved if and only if $\{F, H\} = 0$.

So we have the two conserved quantities

$$J = x_1 p_2 - x_2 p_1$$
$$H = p_1^2 + p_2^2 - \frac{1}{\sqrt{x_1^2 + x_2^2}}.$$

Now we have to do something not entirely formal (that depends on inverse square). We study the level set $\{T^*\mathbf{R}^2 \colon J = J_0\}$. This is preserved by rotation, so it's preserved by the S^1-action, so we can take the quotient $\{T^*\mathbf{R}^2 \colon J = J_0\}/S^1$. This is an example of symplectic reduction. It inherits a symplectic form by "descending" the restriction of the symplectic form on $T^*\mathbf{R}^2$.

Let's introduce some coordinates on this quotient. If (x_1, x_2, p_1, p_2) has $J = J_0$, then we can rotate so that $x_2 = 0$. Since $x_1 p_2 - x_2 p_1 = J_0$, this tells us that $p_2 = J_0/x_1$. So our point is then $(x_1, 0, p_1, J_0/x_1)$, which is controlled by the coordinates $r := x_1$ and $p := p_1$. The Hamiltonian is then

$$H = p^2 + \frac{J_0^2}{r^2} - \frac{1}{r} = p^2 + \left(\frac{J_0}{r} - \frac{1}{2J_0}\right)^2 - \frac{1}{4J_0^2}.$$

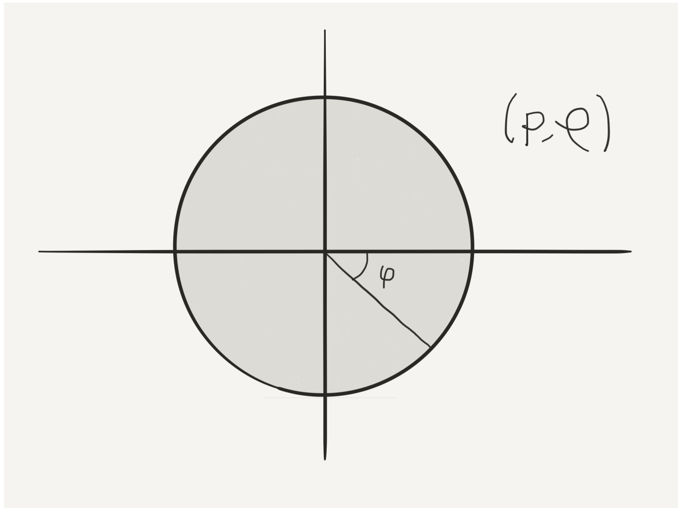
Let $\rho = \frac{J_0}{r} - \frac{1}{2J_0}$, so

$$H = p^2 + \rho^2 - \text{constant}.$$

The symplectic form in these coordinates is $dp \wedge dr = \frac{r^2}{J_0} d\rho \wedge dp$. In (p, ρ)-space, the particle moves along circles of the form $p^2 + \rho^2 = \text{constant}$. The factor $\frac{r^2}{J_0}$ just rescales the speed of the flow. (If the symplectic form doubles then the flow halves speed, so the rate of change of the angle is $2J_0/r^2$ instead of 1. We should be writing things in terms of ρ instead of r, but they determine each other and the formulas are cleaner in terms of r.)

Although the orbit is closed in (p, ρ)-space, upstairs in the level set $\{T^*\mathbf{R}^2 \colon J = J_0\}$ (before the S^1-quotient) you might have precession. This is typical for central potentials. But there is a coincidence that the angular speed of the particle in these coordinates is the same as the original one.

To summarize, in (ρ, p)-space the orbit is a circle. Letting φ be the angle *clockwise* from the x-axis, the angular speed is $d\varphi/dt = 2J_0/r^2$.



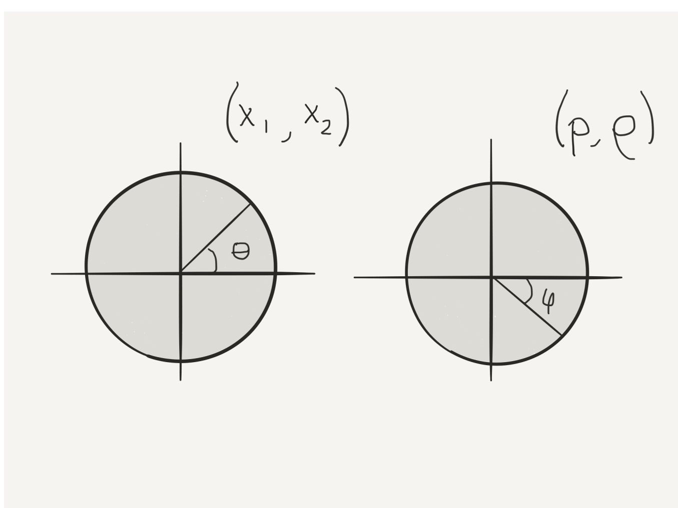
In (x_1, x_2)-space, letting θ denote the angle *clockwise* from the x-axis at$x_2 = 0$, we have $d\theta/dt = x_2'/x_1$. By Hamilton's equations (11.2.3) we have

$$x\frac{\partial H}{\partial p_2} = 2p_2.$$

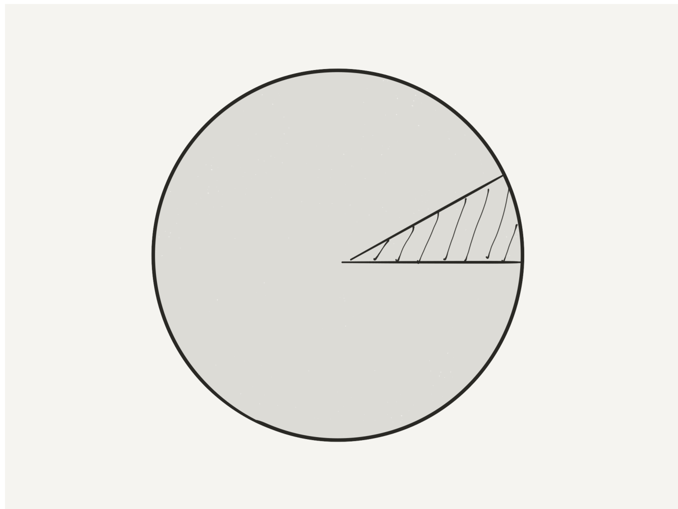
Since $J = x_1p_2 - x_2p_1$, simplifies at $x_2 = 0$ to x_1p_2, we conclude that

$$\frac{x_2'}{x_1} = \frac{2J_0}{x_1^2} = \frac{2J_0}{r^2}.$$

So the miracle is that *the angles are changing at the same speed.*



This is a reformulation of one of Kepler's laws: $r^2 d\theta/dt$ is the rate of sweeping out area, which is constant in time.



We emphasize that $d\theta/dt$ and $d\varphi/dt$ are not constant, since r depends on t (the actual orbits are not spherical), but they are exactly opposite. This leads to a time invariance that we now explain.

Call R_θ a counterclockwise rotation by θ. Since we have just seen that in time t the (p,ρ) vector rotates by $R_{-\theta(t)}$, we have that $R_{\theta(t)}(p(t),\rho(t))$ *is actually constant in time.* We now compute the consequences of this observation. We have

$$(p,\rho) = \left(p, \frac{J_0}{r} - \frac{1}{2J_0}\right) = \left(p, \frac{J_0}{r}\right) - \left(0, \frac{1}{2J_0}\right)$$

The vector $\left(p, \frac{J_0}{r}\right)$ was by obtained by definition from rotating the momentum vector (p_1, p_2) so that $x_2 = 0$, which is a rotation by $-\theta$. Therefore $\left(p, \frac{J_0}{r}\right) = R_{-\theta}(p_1, p_2)$, so

$$R_\theta(p,\rho) = (p_1, p_2) - R_\theta\left(0, \frac{1}{2J_0}\right).$$

Now we unravel this by explicitly computing the entries of R_θ. We have

$$R_\theta = \begin{pmatrix} x_1/r & -x_2/r \\ x_2/r & x_1/r \end{pmatrix}$$

Substituting this into the constancy of $R_\theta(p,\rho)$ in time, we find that

$$\boxed{\left(p_1 + \frac{1}{2J_0}\frac{x_2}{r}, p_2 - \frac{1}{2J_0}\frac{x_1}{r}\right) \text{ is time-independent.}}$$

Exercise 11.4.1. Double check this by computing the Poisson brackets.

Define

$$B = J p_1 + \frac{1}{2}\frac{x_2}{r}$$
$$A = J p_2 - \frac{1}{2}\frac{x_1}{r}$$

The vector (A, B) is called the *Lenz vector*. One can check that $\{H, A\} = \{H, B\} = 0$.

Now A, B, J, H are all invariant under the flow X_H. These are all functions on $T^*\mathbf{R}^2$, which is 4-dimensional, yet we found a 1-dimensional orbit. This means that we should really have 3 "independent" invariants, so there needs to be a relation. Indeed, we compute

$$\begin{aligned} A^2 + B^2 &= J^2(p_1^2 + p_2^2) + \frac{1}{4} + J \underbrace{\frac{p_1 x_2 - x_1 p_2}{4}}_{-J/r} \\ &= J^2\left(p_1^2 + p_2^2 - \frac{1}{r}\right) + \frac{1}{4} \\ &= J^2 H + \frac{1}{4} \end{aligned}$$

which we rewrite as

$$\left(\frac{A}{\sqrt{-H}}\right)^2 + \left(\frac{B}{\sqrt{-H}}\right)^2 + J^2 = \frac{-1}{4H}.$$

(We've written things this way because we want to quantize them in a moment.)

Remark 11.4.2. Unbounded orbits correspond to H positive, so the closed orbits correspond to negative H.

Since $\{H, A\} = \{H, B\} = 0$, the Jacobi identity implies that $\{A, B\}$ also has 0 bracket with H. Since we already have enough nvariants, this should be a function of the ones we already have. In fact, $\{A, B\} = HJ$. Similarly, $\{J, A\} = -B$ and $\{J, B\} = A$.

We introduce new coordinates $\alpha = A/\sqrt{-H}$, $\beta = B/\sqrt{-H}$. In these terms

$$\{\alpha, \beta\} = -J$$
$$\{J, \alpha\} = \beta$$
$$\{\beta, J\} = -\alpha$$

which is a presentation of Lie(SO_3), with the Casimir

$$\alpha^2 + \beta^2 + J^2 = -\frac{1}{4H}.$$

(The calculations are simplified by the fact that H Poisson-commutes with everything.)

Remark 11.4.3. In 3 dimensions we would have gotten SO_4, etc. If H were positive, we would have gotten SO(2, 1) instead.

This means that at least infinitesimally, SO_3 acts on $T^*(\mathbf{R}^2 - \{0\})$. In fact this action does not exponentiate to SO_3, because of the non-compactness of the space. So it does extend after compactifying.

Remark 11.4.4. One way to think about the compactification is that you can map S^2 by stereographic projection to $\mathbf{R}^2$, so you have $T^*S^2 \to T^*\mathbf{R}^2$ away from ∞, and there is an SO_3-action on T^*S^2. This goes back to Fock, and is explained in [Mos]. (We have given a highly oversimplified account.)

We have defined a map

$$T^*(\mathbf{R}^2 - \{0\})^{H<0} \xrightarrow{(\alpha,\beta,H)} \mathrm{Lie}(\mathrm{SO}_3)^*.$$

In this picture the orbits of energy H correspond to spheres of radius squared is $-1/4H$. These are the coadjoint orbits.

We now pass to the quantum story. As we know from the Kirillov character formula, not all coadjoint orbits correspond to representations. The question is: what are the eigenvalues of the Schrödinger operator

$$\mathrm{Op}(H) = -\Delta - \frac{1}{4} \text{ on } \mathbf{R}^2?$$

We have to replace everything by operators. Recall that $\mathrm{Op}(x_1)$ is multiplication by x_1, (p_1) is $i\partial_1$, and

$$\mathrm{Op}(H) = \mathrm{Op}\left(p_1^2 + p_2^2 - \frac{1}{\sqrt{x_1^2 + x_2^2}}\right) = -\Delta - \frac{1}{r}.$$

What about A, B, J? Let $\widehat{A} = \mathrm{Op}(A)$, etc. To compute their quantizations we need to symmetrize:

$$\begin{aligned}
\mathrm{Op}(A) &= \frac{1}{2}[\mathrm{Op}(J)\mathrm{Op}(p_1) + \mathrm{Op}(p_1)\mathrm{Op}(J)] + \frac{1}{2}\frac{x_2}{r} \\
\mathrm{Op}(B) &= \frac{1}{2}[\mathrm{Op}(J)\mathrm{Op}(p_2) + \mathrm{Op}(p_2)\mathrm{Op}(J)] - \frac{1}{2}\frac{x_1}{r}.
\end{aligned}$$

These are all differential operators on $\mathbf{R}^2 - \{0\}$.

Miraculously, they turn out to satisfy almost the exact same relations as their classical versions.

Fact 11.4.5. We have

$$\begin{aligned}
[\widehat{J}, \widehat{A}] &= -i\widehat{B}, \\
[\widehat{B}, \widehat{J}] &= -\widehat{A}, \\
[\widehat{A}, \widehat{B}] &= i\widehat{H}\widehat{J},
\end{aligned}$$

and

$$\widehat{A}^2 + \widehat{B}^2 = \widehat{J}^2\widehat{H} + \frac{1}{4}(1 + \widehat{H}).$$

Assume that you can make sense of $1/\sqrt{-\widehat{H}}$. (This is true in ≥ 3 dimensions. It has to do with self-adjointness; these operators are formally self-adjoint but you also need some convergence conditions for functional analysis.) Then we can make sense of the operators

$$\alpha = i\widehat{A}/\sqrt{-\widehat{H}},$$
$$\beta = i\widehat{B}/\sqrt{-\widehat{H}},$$
$$\gamma = i\widehat{J}.$$

Then α, β, γ generate $SO_3(\mathbf{R})$, and

$$\alpha^2 + \beta^2 + \gamma^2 = -\frac{1+\widehat{H}^{-1}}{4}.$$

The space $L^2(\mathbf{R}^2 - \{0\})^{\widehat{H}<0}$ carries a representation of $\mathrm{Lie}(SO_3)$, which you can actually lift to SO_3. So it decomposes as

$$L^2(\mathbf{R}^2 - \{0\})^{\widehat{H}<0} \cong \bigoplus \underbrace{V_{2n+1}}_{\dim 2n+1}$$

(possibly with multiplicities). By explicit computation,

$$\alpha^2 + \beta^2 + \gamma^2 \text{ acts by } n(n-1) \text{ on } V_{2n+1}.$$

This implies that the $\widehat{H}$-eigenvalues E_n all satisfy

$$-\frac{(1+E_n^{-1})}{4} = n(n-1) \implies E_n = \frac{-1}{(2n-1)^2}.$$

The Kepler Problem in 3 dimensions. In three dimensions you instead get energy levels $1/n^2$. This matches with the known energy levels of the hydrogen atoms, so the theory seems to work!

In summary, you have SO_4 or $SO_{3,1}$ acting on the 3d Kepler problem. In fact there is a larger symmetry group $SO_{4,2}$ acting, after compactification, that preserves the symplectic form, and interpolates between these two cases.

11.5. **Summary.** For the Kepler problem in 3 dimensions, the classical mechanics are controlled by the phase space $T^*(\mathbf{R}^2 - \{0\})$ with Hamiltonian

$$H = \sum p_i^2 - \frac{1}{\sqrt{x_1^2 + x_2^2 + x_3^2}}.$$

The periodic orbits led us to discover extra conserved quantities, which generated $\mathrm{Lie}(SO_4)^*$.

The quantum version is controlled by the Hamiltonian

$$\mathrm{Op}(H) = -\Delta - \frac{1}{\sqrt{x_1^2 + x_2^2 + x_3^2}}$$

acting on $L^2(\mathbf{R}^3)$. The Hamiltonian operator has a *degenerate* spectrum. The eigenvalues are $-\frac{1}{4n^2}$ with multiplicity n^2. In fact the states correspond to n^2-dimensional representations of SO_4.

12. Quantization

We have seen so far that for G nilpotent or compact, there is a correspondence from irreducible representations π of G to coadjoint orbits $\mathscr{O}_\pi \subset (\operatorname{Lie} G)^*$, which is described by the Kirillov character formula.

We now address the questions:

(1) Which orbits $\mathscr{O}$ occur?
(2) Can we give a recipe for making $\pi_{\mathscr{O}}$ from $\mathscr{O}$? (More generally, can we make a Hilbert space from a symplectic manifold (M,ω)?)

12.1. **Quantizable orbits.** We address (1) first.

Example 12.1.1. Let's start by considering a very simple Lie group: $G = U(1) = S^1$. Since G is abelian, its adjoint action on $\operatorname{Lie}(G)$ is trivial, so the G-orbits on $\operatorname{Lie}(G)^*$ are just points. The corresponding representations are 1-dimensional, and the correspondence sends

$$\chi \mapsto \frac{1}{i} d\chi \colon \operatorname{Lie}(G) \to \mathbf{R}.$$

The elements of $\operatorname{Lie}(G)^*$ occurring are the ones that exponentiate to G, i.e. which vanish on $\ker \log$.

In general (for any Lie group G), if $\mathscr{O} = G\cdot\lambda$ for some $\lambda \in \operatorname{Lie}(G)^*$, we have seen (§7.3) that λ defines a Lie algebra homomorphism

$$\operatorname{Lie}(G_\lambda) \to \mathbf{R}$$

where G_λ is the stabilizer of λ. In view of Example 12.1.1, this suggests:

Guess 12.1.2. The orbit $\mathscr{O} = G\cdot\lambda$ corresponds to a representation when you can lift the character

$$i\lambda \colon \operatorname{Lie}(G_\lambda) \to i\mathbf{R}$$

to a homomorphism $\widetilde{\lambda} \colon G_\lambda \to S^1$.

Let's see how this guess plays out for the unitary group U_n. Recall that, identifying $\operatorname{Lie} U_n$ with hermitian matrices, the relevant $\mathscr{O}$ were orbits of

$$\begin{pmatrix} m_1 & & \\ & \ddots & \\ & & m_n \end{pmatrix} + \frac{n-1}{2}\operatorname{Id}$$

where $m_1 > \ldots > m_n \in \mathbf{Z}$. We'll see that Guess 12.1.2 leads to an answer that is sort of close to being correct, but for instance it doesn't know about the $\frac{n-1}{2}\operatorname{Id}$.

Take $\lambda \in \operatorname{Lie}(U_n)^*$ corresponding to

$$\lambda \longleftrightarrow \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

and without loss of generality assume that

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \ldots.$$

The stabilizer of λ is then

$$G_\lambda \cong U_{n_1} \times U_{n_2} \times \ldots \times U_{n_k}$$

The only possibility for $\widetilde{\lambda}$ is $(\det g_1)^{\lambda_1}(\det g_2)^{\lambda_2}\ldots(\det g_k)^{\lambda_k}$. So the guess is that the relevant $\mathscr{O}$-orbits are those of the characters

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \quad \lambda_1 \geq \ldots \geq \lambda_n; \lambda_i \in \mathbf{Z}.$$

This is wrong if n is even, because it is off by $\frac{n-1}{2}\,\mathrm{Id}$, and also because we miss the strict inequalities.

We now describe an improvement due to Duflo. The problem is that even when we try to quantize a vector space, we really get something like $\mathrm{Mp}(V)$ instead of $\mathrm{Sp}(V)$. So considering these double-cover issues is really unavoidable. Since G_λ fixes $\lambda \in \mathscr{O}$, we have a map $G_\lambda \to \mathrm{Aut}(T_\lambda\mathscr{O})$. But in fact G_λ preserves the symplectic structure (clear from inspection) and so lands in $\mathrm{Sp}(T_\lambda\mathscr{O})$

$$G_\lambda \to \mathrm{Sp}(T_\lambda\mathscr{O}).$$

This symplectic group has a metaplectic cover $\mathrm{Mp}(T_\lambda\mathscr{O})$. Does this split, i.e. lift to $G_\lambda \to \mathrm{Mp}(T_\lambda\mathscr{O})$?

$$\begin{tikzcd} & \mathrm{Mp}(T_\lambda\mathscr{O}) \arrow[d] \\ G_\lambda \arrow[r] \arrow[ur, dashed] & \mathrm{Sp}(T_\lambda\mathscr{O}). \end{tikzcd}$$

We can rephrase this as follows. Let $\widetilde{G}_\lambda$ be the pullback cover to G_λ, so concretely

$$\widetilde{G}_\lambda = \{g \in G_\lambda, \text{ lift to } \mathrm{Mp}(T_\lambda\mathscr{O}) \text{ of its image in } \mathrm{Sp}(T_\lambda\mathscr{O})\}.$$

This gives us a double cover

$$0 \to \{\pm 1\} \to \widetilde{G}_\lambda \to G_\lambda \to 0.$$

The existence of a lift $G_\lambda \to \mathrm{Mp}(T_\lambda\mathscr{O})$ is equivalent to the existence of a lift $G_\lambda \to \widetilde{G}_\lambda$.

Guess 12.1.3 (Duflo)**.** The orbit $\mathscr{O} = G \cdot \lambda$ corresponds to a representation when you can lift the character λ to a character of $\widetilde{G}_\lambda$ such that $\widetilde{\lambda}(-1) = -1$.

Let's see how this plays out for the unitary group. If

$$\lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \in \mathrm{Lie}(U_n)^*, \quad \lambda_1 > \ldots > \lambda_n,$$

then

$$G_\lambda = \underbrace{(S^1) \times \ldots \times (S^1)}_{n}$$

and

$$\widetilde{G}_\lambda = \{g \in G_\lambda, z \in S^1 : z^2 = (\det g)^{n-1}\}$$

In this case Guess 12.1.3 correctly picks out integral λ_i for odd n, and half-integral λ_i for even n. However, it still doesn't pick out the strict inequalities.

The two guesses are equivalent when the cover $\widetilde{G}_\lambda$ is split. In the case of the unitary group, this happens when n is odd. From now one *we assume that* $\widetilde{G}_\lambda$ *is split.*

12.2. **Konstant's recipe for quantization.** We now turn to (2), the problem of making $\pi_{\mathcal{O}}$ from $\mathcal{O}$, under the assumption above. We'll give a recipe of Kostant.

We are given a character $\widetilde{\lambda}\colon G_\lambda \to S^1$ with derivative $i\lambda$. Then $\widetilde{\lambda}$ gives a line bundle (or S^1-bundle) $\mathscr{L}_\lambda$ on $G/G_\lambda = \mathcal{O}$. Explicitly, the total space of $\mathscr{L}_\lambda$ is

$$\{(g,z)\in G\times \mathbf{C}\}/(g,z)\sim (gh,\widetilde{\lambda}(h)^{-1}z) \text{ for } h\in G_\lambda.$$

The sections are $f\colon G\to \mathbf{C}$ with $f(gh) = \widetilde{\lambda}(h)^{-1}f(g)$, which you can recognize as $\operatorname{Ind}_{G_\lambda}^G \widetilde{\lambda}$. However, this is still too big to be the right representation.

Example 12.2.1. For $G = \text{Heis}$, and λ on a planar orbit, we have $G_\lambda = Z(G)$. Therefore the induced representation $\operatorname{Ind}_{G_\lambda}^G \widetilde{\lambda}$ is $L^2(\mathbf{R}^2)$, but the correct quantization is $L^2(\mathbf{R})$.

We will make a candidate for $\pi_{\mathcal{O}}$ inside the sections of our line bundle.

Lemma 12.2.2. *The Chern class* $c_1(\mathscr{L}_\lambda) \in H^2(\mathcal{O},\mathbf{R})$ *is the class of* $[\frac{\omega}{2\pi}] \in H^2(M,\mathbf{R})$. *In particular,* $[\frac{\omega}{2\pi}]$ *is integral.*

Remark 12.2.3. This is significant - for instance if M is compact then it implies that the volume of M will be integral, which is desirable because that volume is supposed to be related to the dimension of the representation.

Outline of proof. One way to compute c_1 is as the curvature of a connection. So we'll make a connection on $\mathscr{L}$.

Every $X \in \operatorname{Lie}(G)$ gives a vector field X^* on $\mathcal{O}$. We will specify the connection by telling you how to differentiate along X^*. For f a section of $\mathscr{L}$, equivalently $f\colon G\to\mathbf{C}$ satisfying $f(hg) = \widetilde{\lambda}(h)^{-1}f(g)$, we define a connection ∇ by

$$\nabla_{X_*}f(g) = X_l f - i\lambda(g^{-1}Xg)f.$$

Here

$$X_l f(g) := \frac{d}{dt}|_{t=0} f(e^{-itX}g).$$

This connection is also G-equivariant.

You can compute the curvature:

$$\nabla_{X_*}\nabla_{Y_*} - \nabla_{Y_*}\nabla_{X_*} - \nabla_{[X_*,Y_*]}.$$

We won't write out the details, but you explicitly compute that the curvature at $e\in\mathcal{O}$ sends $X_*, Y_* \mapsto -i\lambda([X,Y])$. Recall that the symplectic form is $(X_*, Y_*)\mapsto \lambda([X,Y])$. The Chern class is then given by

$$[i\frac{(\text{curvature})}{2\pi}] = [\frac{\omega}{2\pi}].$$

$\square$

Let's highlight some properties of the connection ∇ that we didn't use above. The connection is defined by

$$\nabla_{X_*} f(g) = X_l f - i\lambda(g^{-1} X g) f.$$

This is G-equivariant, meaning that for $\gamma \in G$ we have

$$\nabla_{\gamma X_*}(\gamma f) = \gamma(\nabla_{X_*} f).$$

Also, for sections f, g of $\mathscr{L}$, $f\overline{g}$ is a function on $\mathscr{O} = G/G_\lambda$, and

$$X_*(f\overline{g}) = (\nabla_X f)\overline{g} + f\overline{(\nabla_X g)}$$

i.e. *the connection is compatible with a hermitian structure on* $\mathscr{L}$. (The concrete consequence is that parallel transport preserves the hermitian inner product.)

In summary, we have for $\lambda \in \mathrm{Lie}(G)^*$:

- The orbit $\mathscr{O} = G.\lambda$, with a canonical symplectic form ω,
- $\widetilde{\lambda}\colon G_\lambda \to S^1$ with $d\widetilde{\lambda} = \lambda$, corresponding to a line bundle $\mathscr{L} \to \mathscr{O}$ with $c_1(\mathscr{L}) = [\frac{\omega}{2\pi}]$.

Algebra	Geometry
λ, $\mathscr{O} = G.\lambda$	(M, ω)
$\widetilde{\lambda}\colon G_\lambda \to S^1$	$\mathscr{L}$, $c_1(\mathscr{L}) = [\frac{\omega}{2\pi}]$.

Recall that for a nilpotent group N, we constructed $\pi_{\mathscr{O}}$ using a polarization, i.e. a subgroup $G_\lambda \subset Q \subset G$ such that $\mathfrak{q}$ is Lagrangian for the symplectic form ω on $\mathrm{Lie}(G)/\mathrm{Lie}(G_\lambda)$. Given this, we make the induced representation $\mathrm{Ind}_Q^G \widetilde{\lambda}$. This turns out to be identified with sections of $\mathscr{L}$ which are flat, i.e. $\nabla = 0$, along right Q-orbits, and this will lead us to a better construction in general.

A polarization is a subgroup Q fitting into $G_\lambda \subset Q \subset G$ which is maximal isotropic with respect to $\lambda([X, Y])$. Geometrically, the map $\mathscr{O} = G/G_\lambda \to G/Q$ is a family of Lagrangians filling out M.

In the nilpotent case the quantization was the induced representation

$$\mathrm{Ind}_Q^G \lambda = \{f : G \to \mathbf{C} \colon f(gq) = f(g)\widetilde{\lambda}(q)^{-1}\}.$$

We can view this as sections of $\mathscr{L}$ which are flat along the fibers of $\mathscr{O} \to G/Q$. To make this precise, we formalize the notion of a "real polarization".

Definition 12.2.4. A *real polarization* on (M, ω) is a smooth assignment $x \in M \rightsquigarrow$ Lagrangian $L_x \subset T_x M$ which is integrable, in the sense that if X, Y are such that $X_x, Y_x \in L_x$ then $[X, Y]_x \in L_x$.

Given a real polarization $\{L_x\}$ we can form

$$\{f \in \Gamma(\mathscr{L}) \colon \nabla_X f(x) = 0 \text{ when } X_x \in L_x\}$$

is a first approximation to the quantization of (M, ω). However we need to give this a Hilbert space structure. One candidate is

$$(f, g) = \int_{\mathscr{O}} f\overline{g}.$$

This isn't quite right because of the intervention of the metaplectic group, so we need to correct it by some sort of twist.

Remark 12.2.5. This doesn't work for unitary groups, or even for SO_3. We'll modify it to something that works better, but everything we describe in this section is a template, which does not work universally.

Algebra	Geometry
λ, $\mathcal{O} = G.\lambda$	(M, ω)
$\widetilde{\lambda}: G_\lambda \to S^1$	$\mathscr{L}$, $c_1(\mathscr{L}) = [\frac{\omega}{2\pi}]$.
Q	real polarization
$\operatorname{Ind}_Q^G \lambda$	sections constant along Lagrangian leaves

Kostant's fix was to introduce a notion of a *complex polarization.*

Definition 12.2.6. A *complex polarization* on (M, ω) is a smooth assignment of a complex Lagrangian $L_x \subset T_x M^{\mathbf{C}}$ for every $x \in M$, which is integrable.

What does this look like? We have

$$T_x^{\mathbf{C}} = T_x \oplus i T_x.$$

If L_x is in general position, then it will be the graph of some function Φ:

$$L_x = \{v + i\Phi(v) : \Phi\}.$$

Let's explicate the Lagrangian condition. It implies

$$\omega(v + i\Phi(v), w + i\Phi(w)) = 0 \text{ for all } v, w \in T_x^{\mathbf{C}}.$$

The real and imaginary components are

$$\omega(v, w) - \omega(\Phi v, \Phi w) = 0$$
$$\omega(\Phi v, w) + \omega(v, \Phi w) = 0.$$

This tells us that $\omega(v, w) + \omega(\Phi^2 v, w) = 0 \implies \Phi^2 = -1$. (This also follows from L_x being a complex vector space.) In other words Φ^2 defines a complex structure on T_X, and

$$\omega(\Phi v, \Phi w) = \omega(v, w).$$

So generically (under a transversality condition) L_x arises from a complex structure on T_X, and

$$\omega(J v, J w) = \omega(v, w).$$

Another way to say this is that a complex structure on T_X induces a decomposition $T_x^{\mathbf{C}} \cong T_x \oplus \overline{T_x}$, and $L_x = \overline{T_x}$.

The integrability assumption implies that this is an integrable complex structure, i.e. M is a complex manifold. Assume that it's further Kähler, i.e. Im $\omega(J v, v) > 0$ for all $v \in T_x$.

The line bundle $\mathscr{L} \to M$ also has a holomorphic structure, (i.e. we can choose transition functions that are not just smooth but holomorphic.) Indeed, obstruction to endowing a complex line bundle with a holomorphic structure is that its Chern class is type $(1, 1)$, which is ensured by the Kähler assumption ensures that.

The connection

$$\nabla_X : \mathscr{L} \to \mathscr{L} \otimes \Omega^1$$

has $(0,1)$-component $\mathscr{L} \to \mathscr{L} \otimes \Omega^{(0,1)}$ defining the $\overline{\partial}$-operator. (This is not a formality.)

What does it mean for a section of $\mathscr{L}$ to be "flat along the Lagrangian direction"? It means exactly that they are holomorphic sections of $\mathscr{L}$. So Kostant's quantization is holomorphic sections $\Gamma(M,\mathscr{L})$.

Algebra	Geometry
λ, $\mathscr{O} = G.\lambda$	(M,ω)
$\widetilde{\lambda}\colon G_\lambda \to S^1$	$\mathscr{L}$, $c_1(\mathscr{L}) = [\frac{\omega}{2\pi}]$.
Q	real polarization
$\operatorname{Ind}_Q^G \lambda$	sections constant along Lagrangian leaves
J	complex polarization (e.g. coming from Kähler structure)
$\Gamma(M,\mathscr{L})$	holomorphic sections

In the nilpotent case, the quantization of an orbit comes from a real polarization. In the compact case, the quantization of an orbit comes from a complex polarization. The general case of a semisimple Lie group will be a hybrid between the real polarization (Lagrangian) and complex polarization (Kähler) pictures.

We mentioned earlier that the presence of the metaplectic group forces a modification of this discussion.

Example 12.2.7. Consider the representations of SU_2. The quantizable coadjoint orbits $\mathscr{O} = G.\lambda$ are spheres of radius $n+1/2$, with

$$\int \frac{\omega}{2\pi} = 2n+1$$

hence are associated to a $2n+1$-dimensional representation V_{2n+1}.

The sphere has a complex structure, making it $\mathbf{CP}^1$. The symplectic form ω is a multiple of Fubini-study, namely $(2n+1)c_1(\mathscr{O}(1))$. You can check that $\mathscr{L} = \mathscr{O}(2n+1)$. In this case the recipe gives $\Gamma(\mathbf{CP}^1, \mathscr{O}(2n+1)) = \mathrm{Sym}^{2n+2}\mathbf{C}^2$, which is $2n+2$-dimensional. That's not quite right.

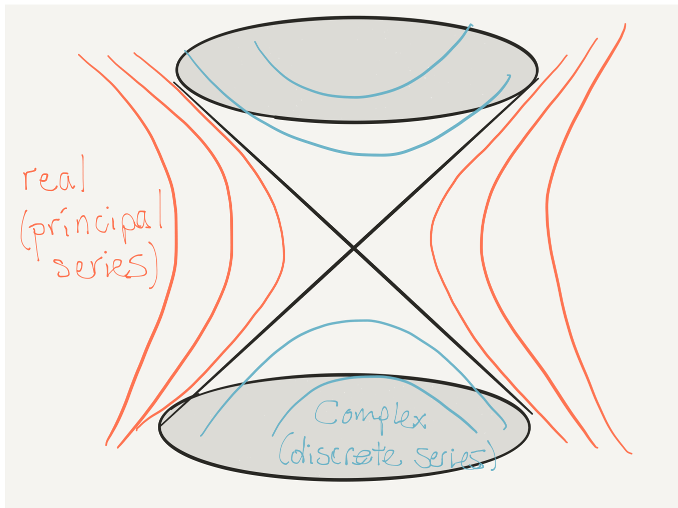
The necessary modification comes from the metaplectic group. We need to twist by the square root of K. In this case $K_{\mathbf{CP}^1} = \mathscr{O}(-2)$, so $\sqrt{K} = \mathscr{O}(-1)$. Another good thing about this is that it gives you a more canonical inner product: the product of two sections of $\sqrt{K}$ is an honest differential, which can be integrated. What does this square root have to do with a lifting to the metaplectic group? For Riemannian manifolds a square root of K corresponds to a spin structure; an analogue of spin for symplectic manifolds is metaplectic structure.

(This twist by $\sqrt{K}$ is necessary even for real polarizations.)

12.3. **Representations of** $\mathrm{SL}_2\,\mathbf{R}$**.** To conclude, we'll use this picture to describe the representations of $\mathrm{SL}_2\,\mathbf{R}$. This is a non-compact Lie group, with many infinite-dimensional representations. Let's think about its coadjoint orbits: $\mathrm{Lie}(\mathrm{SL}_2)$ is the 3-dimensional vector space

$$\begin{pmatrix} x & y \\ z & -x \end{pmatrix}$$

which has the invariant quadratic form given by det. The co-adjoint orbits are the level sets of this quadratic form, which are hyperboloids and paraboloids.



The hyperboloids are ruled by lines, which are the real polarizations. (There are two rulings, which give isomorphic representations.) That gives a family of representations called the *principal series*, which satisfy Kirillov's character formula. All the hyperboloid orbits comes from representations, so they form a *continuous* part of the spectrum $\widehat{\mathrm{SL}_2\,\mathbf{R}}$.

The paraboloids correspond to complex polarizations. (They are realized as holomorphic or antiholomorphic sections, depending on whether they point up or down.) In this paraboloids which come from representations are discrete, and the corresponding representations are called the *discrete series*. The discrete series are quantized by Kähler forms (taking L^2 sections).

So $\widehat{\mathrm{SL}_2\,\mathbf{R}}$ has principal and discrete series, the trivial representation, and one more mysterious class: the complementary series.

References

[AB84] Atiyah, M. F.; Bott, R. The moment map and equivariant cohomology. Topology 23 (1984), no. 1, 1–28.

[Dix96] Dixmier, Jacques. "Les C^*-algébres et leurs représentations". (French) [[C^*-algebras and their representations]] Reprint of the second (1969) edition. Les Grands Classiques Gauthier-Villars. [Gauthier-Villars Great Classics] Éditions Jacques Gabay, Paris, 1996. 403 pp. ISBN: 2-87647-013-6.

[ES76] Eckmann, J.-P.; Sénér, R. The Maslov-WKB method for the (an-)harmonic oscillator. Arch. Rational Mech. Anal. 61 (1976), no. 2, 153–173.

[LV80] Lion, Gérard; Vergne, Michèle. The Weil representation, Maslov index and theta series. Progress in Mathematics, 6. Birkhäuser, Boston, Mass., 1980. vi+337 pp. ISBN: 3-7643-3007-4.

[Mos] Moser, J. Regularization of Kepler's problem and the averaging method on a manifold. Comm. Pure Appl. Math. 23 1970 609–636.