FUSRP Project Write-Up:

The Heisenberg Group and Uncertainty Principle in Mathematical Physics

Recep Çelebi, Kirk Hendricks, Matthew Jordan

August 29, 2015

Abstract

What is the relationship between Fourier analysis, quantum mechanics, and group theory? These important topics, all seemingly unrelated at the surface, are actually intimately related in a number of unexpected ways. One particularly interesting connection is via the *Heisenberg group*, which is surprisingly easy to define and understand, despite its far-reaching and deep applications. In this paper we will explore some properties of the Heisenberg group and the Fourier transform and introduce a selection of applications to quantum mechanics. We will assume undergraduate-level math background—very basic group theory and analysis. These topics were investigated as part of the Fields Undergraduate Summer Research Program 2015, under the supervision of Dr. Hadi Salmasian (University of Ottawa).

Contents

1 Fourier Series **2**
1.1 Periodic Functions 2
1.2 Orthogonality of Trigonometric Functions 2
1.3 The Complex-Valued Fourier Series 3

2 Extension to Hilbert Spaces: the Fourier Transform **4**
2.1 Plancherel Theorem 4

3 Applications to Quantum Mechanics **5**
3.1 The Hermite Polynomials 6

4 Lie Algebras **8**

5 Groups to Lie Groups **11**

6 From Lie Algebras to Lie Groups **13**
6.1 Exponential of Heisenberg Algebra 13

7 Digging Deep into the Heisenberg Algebra and Heisenberg Group **15**

8 Preliminary Definitions **18**

9 Unitary Dual **19**

10 The Heisenberg Group and its Unitary Dual **19**

11 Exploring the Schrödinger Representation **21**

12 Summary **23**

References **24**

1 Fourier Series

These first few sections will discuss the Harmonic analysis aspects of our project. We will begin with a brief overview of Fourier series and an introduction to the Fourier transform. After presenting some properties of the Fourier transform, we will prove Heisenberg's uncertainty priciple using two different methods.

1.1 Periodic Functions

The subject of Fourier analysis begins with the idea of a periodic function. A function f if periodic if there exists some $t \in \mathbb{R}$ which satisfies $f(x+t) = f(x)$.[1] Any linear combination of these periodic functions must also itself be a periodic function, as must any product or quotient, even if the periods of the functions are not the same.

1.2 Orthogonality of Trigonometric Functions

In a vector space V, an inner product is defined to have four properties, with three vectors $u, v, w \in V$ and scalar $c \in \mathbb{R}$,

1. $\langle u, v\rangle = \overline{\langle v, u\rangle}$
2. $\langle cu, v\rangle = c\langle u, v\rangle$
3. $\langle u, v+w\rangle = \langle u, v\rangle + \langle u, w\rangle$
4. $\langle u, u\rangle > 0$ for all $u \neq 0$ and $\langle u, u\rangle = 0$ for $u = 0$

Though the inner product is usually first introduced for finite-dimensional, real vector spaces (such as the dot product in $\mathbb{R}^n$), it can be extended to the infinite-dimensional vector space $L^2(\mathbb{R})$. The space $L^2(\mathbb{R})$ consists of all functions that satisfy the following:

$$\|f\|_2^2 := \int_{\mathbb{R}} |f(x)|^2\, dx \leq \infty.$$

The inner product of two functions $f, g \in L^2(\mathbb{R})$ is defined to be:

$$\langle f, g\rangle = \int_{\mathbb{R}} f(x)\overline{g(x)}\, dx\,. \tag{1.1}$$

The fact that we are defining the L^2 space over $\mathbb{R}$ simply means that the function takes in real inputs, though its output may still be complex. It is prudent to note here that the second function in the inner product definition is complex conjugated. This is necessary to satisfy the last condition of the inner product.

The idea of the space of functions being an inner product space is of paramount importance because it allows us to speak about the orthogonality of functions. Two functions are defined to be orthogonal if their inner product is equal to zero. In Fourier analysis, there are three very important orthogonality relations between basic trigonometric functions.

$$\langle \sin(2\pi nx), \sin(2\pi mx)\rangle = \int_0^1 \sin(2\pi nx)\sin(2\pi mx)dx = \int_0^1 \frac{\cos(2\pi(n-m)x) - \cos(2\pi(n+m)x)}{2} dx = \delta_{n,m} \tag{1.2}$$

$$\langle \cos(2\pi nx), \cos(2\pi mx)\rangle = \int_0^1 \cos(2\pi nx)\cos(2\pi mx)dx = \int_0^1 \frac{\cos(2\pi(n-m)x) + \cos(2\pi(n+m)x)}{2} dx = \delta_{n,m} \tag{1.3}$$

$$\langle \cos(2\pi nx), \sin(2\pi mx)\rangle = \int_0^1 \frac{\sin(2\pi(n+m)x) + \sin(2\pi(m-n)x)}{2} dx = 0 \tag{1.4}$$

[1] Note that by definition this T cannot be unique; any nonzero integer multiple of T must also satisfy this condition. ↑

for $n, m \in \mathbb{Z}$. Note that the Kronecker Delta function, represented by $\delta_{n,m}$, equals zero save when $n = m$, in which case it is equal to one. Notice also that since we are working over functions that are periodic on the interval $[0, 1)$, our integral is only over this interval, not over all real space, so these inner products are actually taken over $L^2([0, 1))$.

Armed now with the fact that the sines and cosines of different periods are always orthogonal to each other, consider then an orthonormal basis constructed of an infinite number of different sines and cosines, all of different period, to describe the space of all periodic functions on $[0, 1)$. Taking linear combinations of these basis functions allows us to represent a given function using trigonometric functions, and the expansion is called the *Fourier series*.

To be more precise, take $f(x) \in L^2([0, 1))$ and represent it as follows:

$$f(x) = \sum_{n=0}^{\infty} a_n \cos(2\pi nx) + \sum_{m=0}^{\infty} b_m \sin(2\pi mx).$$

This is simply any linear combination of sines and cosines with period equal to one. Now, if we take the inner product of both sides with with the k-th cosine frequency, $\cos(2\pi kx)$, we will get the expression,

$$\begin{aligned}\langle f(x), \cos(2\pi kx)\rangle &= \int_0^1 \left(\sum_{n=0}^{\infty} a_n \cos(2\pi kx)\cos(2\pi nx) + \sum_{m=0}^{\infty} b_m \cos(2\pi kx)\sin(2\pi nx) \right) dx \\ &= a_k\end{aligned}$$

Since the inner product of cosines with two different periods is zero, and that the inner product of cosines with sines is always zero. Now, since $\langle f(x), \cos(2\pi kx)\rangle = \int_0^1 f(x)\cos(2\pi kx)\,dx$, we now have a formula for a_n, and by a similar argument for b_m:

$$a_n = \int_0^1 f(x)\cos 2\pi nx dx \tag{1.5}$$

$$b_m = \int_0^1 f(x)\sin 2\pi mx dx. \tag{1.6}$$

Thus, for every function for which the integrals in equations (1.5) and (1.6) exist, there exists a Fourier series.

1.3 The Complex-Valued Fourier Series

Though in theory, the Fourier series is quite elegant, writing it out in terms of sines and cosines is really only used in the study of trigonometric polynomials. For several applications, it is best to think of the Fourier series as a complex valued sum. Consider, by Euler's equation,

$$f(x) = \sum_{n=0}^{\infty} a_n \cos 2\pi nx + b_n \sin 2\pi nx = \sum_{n=0}^{\infty} \frac{a_n}{2}\left(e^{i2\pi nx} + e^{-i2\pi nx}\right) + \frac{b_n}{2i}\left(e^{i2\pi nx} - e^{-i2\pi nx}\right) = \sum_{k=-\infty}^{\infty} c_k e^{2\pi ikx}$$

Through a method very similar to the ones we used to find a_n and b_m, it turns out that this is the expression for c_k:

$$c_k = \int_0^1 f(x)e^{-2\pi kx}\,dx.$$

The expression on the right is a far quicker way of writing the Fourier series, and one that lends itself more to the actual idea of the series: we are decomposing the function into an infinite superposition of waves of different frequency. It is a worthwhile exercise to show that, just like sines and cosines, exponentials of different frequencies are also orthogonal.[2]

[2] That is, show that $\langle e^{2\pi imx}, e^{2\pi inx}\rangle = \delta_{m,n}$. ↑

2 Extension to Hilbert Spaces: the Fourier Transform

The Fourier series is a very useful tool for restructuring a function in terms of an orthonormal basis of waves. It decomposes a function into a linear combination of waves of different frequencies, and the coefficient in front of each frequency in the series expansion (i.e. the a_n, b_n or the c_k) tells us the "strength" of each frequency. So, if we want to know how much the third harmonic of the cosine contributes to the overall function, we need only look at the magnitude of the coefficient in front of $\cos(2\pi(3)x)$. The magic of the Fourier series is that the frequencies can actually be indexed by integers. But what if we want to do the same process for a non-periodic function? Is there any way to know the strength of a given frequency in a function which does not repeat itself? The answer is yes, and it's made possible by the *Fourier transform*.

The Fourier transform of a function $f(x)$, denoted $\mathcal{F}\{f(x)\}(k)$ or $\hat{f}(k)$, is a new function that, given a real-valued frequency, tells us the strength of that frequency in the original function. For that reason, we speak of the Fourier transform as a representation of a function in frequency space. Formally, the Fourier series is defined as follows:

$$\mathcal{F}\{f(x)\}(k) = \langle f(x), e^{2\pi ikx}\rangle = \int_{\mathbb{R}} f(x)e^{-2\pi ikx}\,dx \tag{2.1}$$

(Remember of course that the inner product has always had a complex conjugation in its definition, even though this is the first time we are seeing it.) This Fourier transform has an inverse, which is simply

$$\mathcal{F}^{-1}\{f(k)\}(x) = \int_{\mathbb{R}} f(k)e^{2\pi ikx}dk \tag{2.2}$$

Delightfully, the Fourier transform preserves the norm of a function. That is, $\|\mathcal{F}f\|_2 = \|f\|_2$. This is a consequence of the Plancherel theorem, which will be proven shortly.[3] The Fourier transform relies on a property called *Pontryagin Duality* of locally compact groups, which states that there exist a canonical isomorphism between a locally compact group and its dual. This is the reason that the Fourier transform over $\mathbb{R}$ forms an isomorphism over $\mathbb{R}$, making the inverse transform possible and establishing the utility of the Fourier transform. Pontryagin Duality will be discussed more in a forthcoming section.

2.1 Plancherel Theorem

This theorem states that the L^2 norm of a function is equal to the L^2 norm of its frequency distribution/its Fourier transform. This theorem will become one of the most important tools we use later on. Its proof follows.

Theorem 2.1 (**Plancherel Theorem**). *For $f \in L^2(\mathbb{R})$ and its Fourier transform $\hat{f}$, $\langle f, f\rangle = \langle \hat{f}, \hat{f}\rangle$.*

Proof.

$$\begin{aligned}
\langle \hat{f}, \hat{f}\rangle = \int_{\mathbb{R}} \hat{f}(k)\overline{\hat{f}(k)}\,dk &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(x)e^{-2\pi ikx}\,dx \int_{\mathbb{R}} \overline{f(x')}e^{2\pi ikx'}\,dx'\,dk \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} f(x)\overline{f(x')}\int_{\mathbb{R}} e^{2\pi ik(x'-x)}\,dk\,dx'\,dx \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} f(x)\overline{f(x')}\delta(x'-x)\,dx'\,dx \\
&= \int_{\mathbb{R}} f(x)\overline{f(x)}\,dx \\
&= \langle f, f\rangle
\end{aligned}$$

[3] Other definitions of the transform and the Fourier series are acceptable, such as using e^{ikx} instead. However, the corresponding inner products must be defined over $[0, 2\pi)$ or $[-\pi, \pi)$ in these cases. This leads to a transformation that is either not unitary, or needs constant factors of $\frac{1}{\sqrt{2\pi}}$ to maintain the L^2 norm of a function over a transformation followed by an inverse transformation. ↑

□

Note that this proof relies upon the fact that the Fourier transform of the function $f(x) = 1$ is $\hat{f}(k) = \delta(k)$. Intuitively, this can be understood as the fact that the "zero-th frequency" is the only frequency present in the function $f(x) = 1$.

3 Applications to Quantum Mechanics

Heisenberg's Uncertainty Principle, normally stated as "the error in a position measurement multiplied by the error in a momentum measurement is bounded below," can seem very surprising at first glance. However, this uncertainty in measurement is a property of any wave in nature, a fact that was known long before quantum mechanics. Consider a small piece of wave. From this information, it is impossible to determine the frequency of the wave with absolute certainty. No matter how much it may look like an analytic function with a definite frequency on the interval given, this may only be an approximate solution. The longer our snippet of wave gets, the better we can approximate the frequency; however, to better approximate our frequency, we have given up information about the position of a "particle" "under" this wave. Since the frequency of a wavefunction is proportional to its momentum, we can see that the Heisenberg Uncertainty principle is not so much some existential axiom of the universe, but a simple property of any wave that follows from its definition. Thus, when quantum mechanics introduces the idea of wave-particle duality and says that everything has a wavefunction associated with it, it is apparent that we are moving towards a world where uncertainty, no matter the precision of our measurements, is guaranteed.

Once one knows the definition of the Fourier transform, this explanation becomes less hand-wavey and more concrete. The fact that the Fourier transform takes functions with small variances and transforms them to functions with large variances and vice versa means that the product of the area under a function and the area under its Fourier transform will always be bounded below by a positive constant.[4]

Now, to actually use some equations and prove something. The main ideas in the following proof are taken from [7].

Theorem 3.1 (**Heisenberg's Uncertainty Principle**). *Consider a differentiable function $f \in L^2(\mathbb{R})$ that vanishes at $\pm\infty$ faster than $x^{-\frac{1}{2}}$.*[5] *Set $\langle f, f\rangle = 1$. Set the mean of f, and therefore the mean of its Fourier transform, to zero. Then we can say,*

$$\|X(f(x))\|_2 \|P(f(x))\|_2 \geq \frac{1}{2} \tag{3.1}$$

Where

$$X(f(x)) = xf(x) \tag{3.2}$$

$$P(f(x)) = -if'(x) \tag{3.3}$$

Proof. First consider

$$\langle X(f), P(f)\rangle = \int_{\mathbb{R}} (xf(x))(\overline{if'(x)})dx$$

By an application of integration by parts, we have

$$i\int_{\mathbb{R}} xf(x)\overline{f'(x)}dx = ixf(x)\overline{f(x)}\Big|_{\mathbb{R}} - i\int_{\mathbb{R}} f(x)\overline{f(x)}dx - i\int_{\mathbb{R}} xf'(x)\overline{f(x)}dx$$

Since by assumption $f(x)$ vanishes faster than $x^{-\frac{1}{2}}$, the boundary term will disappear and, after taking the

[4] In quantum mechanics, this constant usually includes a fundamental constant of nature called the reduced Plank's constant, denoted $\hbar$. To keep things simple, we will set $\hbar = 1$, which is common practice in proofs of the uncertainty principle. ↑

[5] Note that f represents the wavefunction of quantum mechanics. ↑

modulus of both sides, we are left with

$$\|\langle X(f), P(f)\rangle\| \geq \mathrm{Re}\left[\int_{\mathbb{R}} ixf(x)\overline{f'(x)}dx\right] = \frac{1}{2}$$

Finally, we use the Cauchy-Schwartz inequality,[6]

$$\|X(f(x))\|_2\|P(f(x))\|_2 \geq \frac{1}{2} \tag{3.4}$$

□

However, we can continue with an application of the Plancherel Theorem to say something about the standard deviation of the wavefunction and its transform.

Corollary 3.1.1 (**Standard Deviation Representation**)**.** Note that,

$$\mathcal{F}(f'(x))(k) = \int_{\mathbb{R}} f'(x)e^{-2\pi ikx}dx = 2\pi ik\mathcal{F}(f(x))(k) \tag{3.5}$$

Thus taking a derivative is simply multiplication in frequency space (so the momentum operator is exactly the same as the position operator, just in frequency space!). Using Plancherel Theorem,

$$\begin{aligned}\left(\int_{\mathbb{R}} x^2 f(x)\overline{f(x)}dx\right)^{\frac{1}{2}}\left(\int_{\mathbb{R}} f'(x)\overline{f'(x)}dx\right)^{\frac{1}{2}} &= \left(\int_{\mathbb{R}} x^2 f(x)\overline{f(x)}dx\right)^{\frac{1}{2}}\|if'(x)\|_2 \\ &= \left(\int_{\mathbb{R}} x^2 f(x)\overline{f(x)}dx\right)^{\frac{1}{2}}\|i2\pi ik\hat{f}(k)\|_2\end{aligned}$$

This means that

$$\left(\int_{\mathbb{R}} x^2 f(x)\overline{f(x)}dx\right)^{\frac{1}{2}}\left(\int_{\mathbb{R}} k^2 \hat{f}(k)\overline{\hat{f}(k)}dk\right)^{\frac{1}{2}} \geq \frac{1}{4\pi} \tag{3.6}$$

So the standard Heisenberg's uncertainty principle also implies an inverse relationship between the standard deviation of any function and its transform.

The observant reader will have paused during the last proof when we assumed that the function f vanished sufficiently fast to send the boundary term to zero. A different, less direct proof avoids this sticking point and introduces the idea of the Hermite operator and Hermite polynomials, as seen in [5].

3.1 The Hermite Polynomials

First, let us define the differential operator H, called the *Hermite operator* like so,

$$H := -\frac{1}{4\pi^2}\frac{d^2}{dx^2} + x^2 \tag{3.7}$$

This operator is self-adjoint since it is a linear combination of even derivates. Since the eigenfunctions with distinct eigenvalues of self-adjoint operators are orthogonal over the L^2 inner product, we can define an orthonormal basis given by the eigenfunctions of H. These functions are given by the formula

$$h_k(x) = \frac{2^{\frac{1}{4}}}{\sqrt{k!}}\Big(-\frac{1}{\sqrt{2\pi}}\Big)^k e^{\pi x^2}\frac{d^k}{dx^k}e^{-2\pi x^2} \tag{3.8}$$

[6] Notice that this final statement can also be expressed in terms of the commutator: $\|X(f(x))\|_2\|P(f(x))\|_2 \geq \frac{1}{2}\Big|\big[P(f), X(f)\big]\Big|$. ↑

for non-negative integers k. These polynomials are incredibly useful since, in addition to being an orthonormal basis and eigenfunctions of the Hermite operator, they are also eigenfunctions of the Fourier transform. Note the eigenvalues are given by,

$$\begin{aligned} Hh_k &= \frac{2k+1}{2\pi} h_k \\ \mathcal{F}h_k &= (-i)^k h_k \end{aligned}$$

Heisenberg's Uncertainty Principle – An Alternate Proof. Define for any function f the mean, variance and standard deviation.

$$\begin{aligned} \text{Mean:} \quad \mu(f) &= \frac{1}{\|f\|_2^2} \int_{\mathbb{R}} x|f(x)|^2 dx \\ \text{Variance:} \quad \Delta^2(f) &= \int_{\mathbb{R}} |x-\mu(f)|^2 |f(x)|^2 dx \\ \text{Standard Deviation:} \quad \Delta(f) &= \sqrt{\Delta^2(f)} \end{aligned}$$

Using these definitions,

$$\begin{aligned} \langle Hf, f\rangle &= \int_{\mathbb{R}} -\frac{1}{4\pi^2} f''(x)\overline{f(x)}\, dx + \int_{\mathbb{R}} x^2 |f(x)|^2 dx \\ &= -\frac{1}{4\pi^2} f'(x)\overline{f(x)}\Big|_{\mathbb{R}} + \frac{1}{4\pi^2}\int_{\mathbb{R}} |f'(x)|^2 dx + \int_{\mathbb{R}} x^2|f(x)|^2 dx \\ &= \int_{\mathbb{R}} k^2 |\hat{f}(k)|^2 dk + \int_{\mathbb{R}} x^2 |f(x)|^2 dx \\ &= \int_{\mathbb{R}} (k-\mu(\hat{f}))^2 |\hat{f}(k)|^2 dk + \int_{\mathbb{R}} 2k\mu(\hat{f})|\hat{f}(k)|^2 dk - \int_{\mathbb{R}} \mu(\hat{f})^2 |\hat{f}(k)|^2 dk \\ &\quad + \int_{\mathbb{R}} (x-\mu(f))^2 |f(x)|^2 dx + \int_{\mathbb{R}} 2x\mu(f)|f(x)|^2 dx - \int_{\mathbb{R}} \mu(f)^2 |f(x)|^2 dx \\ &= \Delta^2(f) + \Delta^2(\hat{f}) + 2\mu(\hat{f})^2\|\hat{f}\|_2^2 - \mu(\hat{f})^2\|\hat{f}\|_2^2 + 2\mu(f)^2\|f\|_2^2 - \mu(f)^2\|f\|_2^2 \\ &= \Delta^2(f) + \Delta^2(\hat{f}) + \mu(\hat{f})^2\|\hat{f}\|_2^2 + \mu(f)^2\|f\|_2^2 \end{aligned}$$

Notice here that the boundary term disappears so long as $\|f\|$ is finite. That is, so long as $f \in L^2(\mathbb{R})$, avoiding the quickly decaying condition that we had in the earlier proof. Next, if we assign the means to be zero (which leaves the variance unchanged, since it's translationally invariant), we will have the expression

$$\begin{aligned} \Delta(f)^2 + \Delta(\hat{f})^2 &= \langle Hf, f\rangle && (3.9) \\ &= \int_{\mathbb{R}} H \sum_{k=0}^{\infty} \langle f, h_k\rangle h_k \sum_{k=0}^{\infty} \langle f, h_k\rangle \overline{h_k} dx && (3.10) \\ &= \sum_{k=0}^{\infty} \frac{2k+1}{2\pi} |\langle f, h_k\rangle|^2 && (3.11) \\ &\geq \frac{1}{2\pi} \sum_{k=0}^{\infty} |\langle f, h_k\rangle|^2 = \frac{1}{2\pi}\|f\|_2^2 && (3.12) \end{aligned}$$

Thus the sum of the variance of a function and its transform is bounded below. Now, since this formula is true for any function, we can define a new function g such that $g(x) = \frac{1}{\sqrt{\lambda}} f(\frac{x}{\lambda})$, where λ is an arbitrary constant. The Fourier transform of this g is $\hat{g}(k) = \sqrt{\lambda}\hat{f}(\lambda k)$. Plugging this information into our inequality

above,

$$\Delta(g)^2 + \Delta(\hat{g})^2 \geq \frac{1}{2\pi}\|g\|_2^2$$
$$\int_{\mathbb{R}} \frac{1}{\lambda}x^2 f\Big(\frac{x}{\lambda}\Big)^2 dx + \int_{\mathbb{R}} \lambda k^2 \hat{f}(k\lambda)^2 dk \geq \frac{1}{2\pi}\frac{1}{\lambda}\int_{\mathbb{R}} f(\frac{x}{\lambda})^2 dx$$
$$\lambda^2\Delta^2(f) + \frac{1}{\lambda^2}\Delta^2(\hat{f}) \geq \frac{1}{2\pi}\|f\|_2^2$$

And since this is true for all λ, then it is true for $\lambda = \sqrt{\frac{\Delta(\hat{f})}{\Delta(f)}}$, giving

$$\Delta(f)\Delta(\hat{f}) \geq \frac{1}{4\pi}\|f\|^2 \tag{3.13}$$

Notice that if we rescale the wavefunction to set $\|f\| = 1$, as we did before, then the above equation exactly equals (3.6), as expected. □

4 Lie Algebras

Having now proven the celebrated Heisenberg uncertainty principle, let's change gears and move into the more algebraic aspects of the project. In the coming sections, we will introduce the branch of abstract algebra called Lie theory, and describe our main object of study, the Heisenberg group.

We begin modestly, with a definition of an algebraic structure called an *algebra*. It is a relatively straightforward structure that consists of a vector space endowed with a special kind of binary operation called a *bilinear product*. Let us define each of these in turn.

Definition 4.1 (**Bilinear Product**)**.** Let A be a vector space defined over the field K. A bilinear product $B : A \times A \to A$ is a map that satisfies the following properties (for $x, y, z \in A$):

1. $B(x, y + z) = B(x, y) + B(x, z)$.
2. $B(x + y, z) = B(x, z) + B(y, z)$.
3. If c and d are scalars in K, then $B(c \cdot x, d \cdot y) = (cd)B(x, y)$.

It might now be clear where the name "bilinear product" comes from: each argument of the product is linear. Now the definition of an algebra:

Definition 4.2 (**Algebra over a Field**)**.** Let A be a vector space over a field K. Then A is an algebra if it is equipped with a bilinear product $B : A \times A \to A$.

Simple as that! An algebra is just a vector space with a bilinear product. One easy example of an algebra is the vector space $\mathbb{R}^3$ with the bilinear product given by the cross-product. Of course, there are plenty of others (including "boring" ones like $\mathbb{R}$ with multiplication).

Now that we know what algebras are, it's time to introduce the first **big** concept of the day: Lie algebras.

Definition 4.3 (**Lie Algebra**)**.** Let V be a vector space over a field K. Let $[\ ,\ ] : V \times V \to V$ be a bilinear product satisfying the following additional properties for all $x, y, z \in V$:

1. $[x, x] = 0$
2. $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$

We call the pair $\mathfrak{g} := (V, [\ ,\ ])$ a **Lie algebra**. The product $[\ ,\ ]$ is called the **Lie bracket** of $\mathfrak{g}$.

Using this definition, we can actually derive an interesting property about the Lie bracket. Applying property 1. to $[x + y, x + y]$, keeping in mind that the bracket is bilinear:

$$0 = [x + y, x + y] = [x, x + y] + [y, x + y]$$
$$0 = \cancel{[x, x]} + [x, y] + [y, x] + \cancel{[y, y]}$$
$$[x, y] = -[y, x]$$

And there you have it. We say that the Lie bracket is **skew-symmetric**.

Here are some examples of Lie algebras:

Example 4.4. Matrices.

The set $M_n(K)$ of all $n \times n$ matrices over the field K is a Lie algebra with the commutator defined as follows:

$$[A, B] = AB - BA.$$

It takes some toying around with matrix multiplication to show that this is a bilinear product that obeys the two additional properties in 4.3. We call this Lie algebra $\mathfrak{gl}_n$, where the $\mathfrak{gl}$ stands for "general linear."

Example 4.5. Zero-Trace Matrices.

Let $\mathfrak{sl}_n = \{A \in M_n(K) : \text{trace}(A) = 0\} \subset \mathfrak{gl}_n$. Then $\mathfrak{sl}_n$ is a Lie algebra under the same bracket:

$$[A, B] = AB - BA.$$

Let us quickly verify that this bracket is legit, i.e. that it indeed does map from $\mathfrak{sl}_n \times \mathfrak{sl}_n \to \mathfrak{sl}_n$. That is, let us show that if A and B both have trace 0, then so does $[A, B]$. The proof relies on two facts. First, that $\text{trace}(A+B) = \text{trace}(A) + \text{trace}(B)$, and second that $\text{trace}(AB) = \text{trace}(BA)$. We will not prove those facts here, but please trust that they are both true and easily Googleable.

So:

$$\text{trace}[A, B] = \text{trace}(AB - BA) = \text{trace}(AB) - \text{trace}(BA) = 0.$$

Success! This bracket is well-defined. The next step would be to show that the bracket is bilinear and obeys the properties from 4.3. Turns out that all of that is quite easy using the two properties of trace we mentioned above, so we won't bother right now.

Example 4.6. Linear Maps on a Vector Space.

If V is a vector space, let $A = \{f : V \to V \mid f \text{ is linear}\} = \text{End}(V)$.[7] This collection of linear mappings forms the Lie algebra $\mathfrak{gl}(V)$ with the bracket

$$[f, g] = f \circ g - g \circ f.$$

In these first two examples, you might see a theme a'brewin'. You might be thinking that there's something special going on with the bracket that looks like this:

$$[x, y] = xy - yx.$$

You are certainly right. But we have to wait a bit to discover why exactly this is so special. However, we will say that this type of bracket is called the **commutator** of x and y.

Example 4.7. Heisenberg Algebra.[8]

Take the vector space $\mathbb{R}^{2n+1}$, and represent each element of the space as follows:

$$(p_1, \ldots, p_n, q_1, \ldots, q_n, t) := (p, q, t).$$

Then this is a Lie algebra with bracket:

$$[(p, q, t), (p', q', t')] = (0, 0, pq' - qp'). \tag{4.1}$$

We often write this algebra as $\mathfrak{h}_n$ (the $\mathfrak{h}$ of course standing for Heisenberg).

We're almost done laying the groundwork with Lie algebras. There are only one thing left to do, at that's looking at how we can create mappings between one Lie algebra and another, or between a Lie algebra and another algebraic structure.

[7] "Endomorphism" is the fancy word for "linear map from a vector space to itself," hence the set $\text{End}(V)$. ↑

[8] Don't be too excited by the name yet. We'll see why it's called "Heisenberg" soon enough. ↑

Why would we even want to do this? Here's the idea: it's sometimes possible to view the elements of our Lie algebra as operators over a vector space. That is, each element of the algebra can be made to represent a function that acts on a vector space. Let's formalize this notion mathematically.

First, how can we map one Lie algebra to another?

Definition 4.8 (**Lie Algebra Homomorphism**)**.** Let $\mathfrak{g}$ and $\mathfrak{g}'$ be Lie algebras. Then a linear map $\varphi : \mathfrak{g} \to \mathfrak{g}'$ is a **Lie algebra homomorphism** if it obeys the following property for all $x, y \in \mathfrak{g}$:

$$\varphi([x, y]) = \big[\varphi(x), \varphi(y)\big] ,$$

where the bracket on the left is the bracket for $\mathfrak{g}$, and the one on the right is the bracket for $\mathfrak{g}'$.

Basically, in simpler terms, a homomorphism is a map that preserves the structure of the Lie bracket. Now that we have a notion of mappings between Lie algebras, we can now formalize the notion of Lie algebras representing operators on a vector space.

Definition 4.9 (**Lie Algebra Representation**)**.** Let $\mathfrak{g}$ be a Lie algebra and V a vector space. A **Lie algebra representation** is a Lie algebra homomorphism φ from the algebra $\mathfrak{g}$ to $\mathfrak{gl}(V)$, the set of all linear transformations over V. That is, if $x, y \in \mathfrak{g}$:

$$\begin{aligned} &\varphi : \mathfrak{g} \to \mathfrak{gl}(V) \\ &\varphi([x, y]) \quad = [\varphi(x), \varphi(y)] \\ &\qquad\qquad = \varphi(x)\varphi(y) - \varphi(y)\varphi(x) \end{aligned}$$

Up next is a fantastically relevant example of a representation, from the Heisenberg algebra. (Go back and reread Example 4.7 to remember how we defined the Heisenberg algebra)

Example 4.10. Representation of Heisenberg Algebra.

Consider the following mapping, $m : \mathfrak{h}_n \to M_{n+2}(\mathbb{R})$, given by:

$$m(p, q, t) = \begin{pmatrix} 0 & p_1 & \cdots & p_n & t \\ 0 & 0 & \cdots & 0 & q_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & q_n \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} . \tag{4.2}$$

The map m takes in an element of the Heisenberg algebra—which, you'll remember, is written in the shorthand $(p, q, t) = (p_1, \ldots, p_n, q_1, \ldots, q_n, t)$—and yields a matrix of size $(n + 2)^2$. "But wait!" I hear you say. "Wasn't a representation supposed to map from an algebra to a set of linear transformations on a vector space?"

You're totally right. But remember, matrices are themselves linear transformations: an $n \times n$ matrix is a linear operator on $\mathbb{R}^n$. So, the representation m defined in (4.2) maps from the Heisenberg algebra to a linear operator on $\mathbb{R}^{n+2}$.

Claim. *The map $m : \mathfrak{h}_n \to M_{n+2}(\mathbb{R})$ is a Lie algebra homomorphism.*

Proof. By the definition of Lie algebra homomorphism in 4.8, I need to show that

$$m([(p, q, t), (p', q', t')]) = [m(p, q, t), m(p', q', t')].$$

The bracket of the Heisenberg algebra was defined in (4.1) as

$$[(p, q, t), (p', q', t')] = (0, 0, pq' - qp').$$

Thus:

$$\begin{aligned}
m([(p,q,t),(p',q',t')]) &= m(0,0,pq'-qp') \\
&= \begin{pmatrix} 0 & 0 & \cdots & 0 & pq'-qp' \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & \cdots & 0 & pq' \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & \cdots & 0 & -q'p \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \\
&= m(p,q,t)m(p',q',t') - m(p',q',t')m(p,q,t) \\
&= [m(p,q,t), m(p',q',t')].
\end{aligned}$$

This completes the proof that the representation m is a homomorphism. □

The above chain of equalities uses the fact that that

$$m(p,q,t)m(p',q',t') = m(0,0,pq'). \tag{4.3}$$

To see this, simply write out the two matrices $m(p,q,t)$ and $m(p',q',t')$, multiply them, and notice that the only non-zero entry sits in the top-right corner and is precisely the dot-product between p and q'.

We have now said essentially all we need to say about algebras. We will now introduce the notion of groups and Lie groups in the coming section, on our way to showing how a Lie algebra can be turned into a Lie group.

5 Groups to Lie Groups

Let's start right from the beginning:

Definition 5.1 (**Group**)**.** Let G be a set and let $\circ$ be a binary operation. The pair $(G,\circ)$ is called a **group** if the following four axioms hold:

1. The set G is closed. (If $g_1, g_2 \in G$ then $g_1 \circ g_2 \in G$ too.)
2. The operation $\circ$ is associative. (If $g_1, g_2, g_3 \in G$, then $g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3$.)
3. The set G contains an identity element. (There exists $e \in G$ such that $e \circ g = g \circ e = e$ for any $g \in G$.)
4. Each element of G is invertible. (For any $g \in G$ there is $g^{-1} \in G$ such that $g \circ g^{-1} = g^{-1} \circ g = e$.)

Example 5.2. Automorphisms on a Vector Space.

If V is a vector space, let $G = \{f : V \to V | f \text{ is linear and invertible.}\} = \text{Aut}(V)$[9] and let $\circ$ be functional composition. Then $(G,\circ)$ is a group, which we call $GL(V)$.

This GL stands for the same "general linear" as before. When the vector space V in question is $\mathbb{R}^n$, the set of invertible linear operators on $\mathbb{R}^n$ is precisely the set of invertible matrices, or $GL_n(\mathbb{R})$. We often denote the set of $n \times n$ invertible matrices over a field K by $GL_n(K)$, though it's simply a special case of the general linear group.

It is part of standard linear algebra to show that this group satisfies all the group axioms. We will demonstrate only the closure axiom: given $f \in GL(V)$ and $g \in GL(V)$ I will show that $h := f \circ g \in GL(V)$.

[9] "Automorphism" is the fancy word for "linear, invertible map from a vector space to itself," hence the set $\text{Aut}(V)$. ↑

Well, for h to be in $GL(V)$ it needs to be invertible and linear. First, invertibility: let $h^{-1} = g^{-1} \circ f^{-1}$, which exists since both f and g are invertible because they're both in $GL(V)$. Now:

$$h^{-1} \circ h(x) = h^{-1} \circ (f \circ g(x)) = h^{-1} \circ f(g(x)) = g^{-1} \circ f^{-1}(f(g(x))) = g^{-1}(g(x)) = x.$$

and

$$h \circ h^{-1}(x) = h \circ g^{-1} \circ f^{-1}(x) = h \circ g^{-1}(f^{-1}(x)) = f \circ g(g^{-1}(f^{-1}(x))) = f(f^{-1}(x)) = x.$$

So, h is invertible. How about linear? Well,

$$h(x+y) = f \circ g(x+y) = f(g(x)+g(y)) = f(g(x)) + f(g(y)) = h(x) + h(y).$$

And there you have it: $h = f \circ g$ is invertible linear, so $f \circ g \in GL(V)$ and the first axiom is satisfied. The last three axioms are slightly less tedious (you're welcome).

Example 5.3 (Heisenberg Group)**.** Let G be the set of all vectors $(p_1, \ldots, p_n, q_1, \ldots, q_n, t) = (p, q, t) \in \mathbb{R}^{2n+1}$, and let $\circ$ be the operation given by:

$$(p, q, t) \circ (p', q', t') = \left(p + p', q + q', t + t' + \tfrac{1}{2}(pq' - qp')\right).$$

We call the group $(G, \circ)$ the **Heisenberg group**, denoted $\mathcal{H}$.

Let us quickly show that the group obeys the axioms. Closure is easy, because

$$\left(p + p', q + q', t + t' + \tfrac{1}{2}(pq' - qp')\right)$$

has the structure (n-tuple, n-tuple, real number). The operation is associative because so are addition and vector multiplication. Each element has an inverse: $(p, q, t)^{-1} = (-p, -q, -t)$. Finally, the identity is $(0, 0, 0)$.

Just like with algebras, it's nice to be able to map groups to one another. In particular, we're eventually going to want to represent groups as operators over a vector space (just like with algebras), so we will next define the notion of a group homomorphism.

Definition 5.4 (**Group Homomorphism**)**.** Let $(G, \circ_G)$ and $(H, \circ_H)$ be groups. Then a map $\varphi : G \to H$ is a **group homomorphism** if it obeys the following property for all $x, y \in G$:

$$\varphi(x \circ_G y) = \varphi(x) \circ_H \varphi(y).$$

Definition 5.5 (**Group Representation**)**.** Let $(G, \circ_G)$ be a group and V a vector space. A **group representation** is a homomorphism φ from G to $GL(V)$. That is, if $g_1, g_2 \in G$:

$$\varphi : G \to GL(V)$$
$$\varphi(g_1 \circ_G g_2) = \varphi(g_1) \circ \varphi(g_2)$$

Since these definitions are so similar to those for Lie algebras, we will skip giving examples.

One final thing we must define before we move on is the concept of a Lie group. Lie groups are incredibly important in mathematics and quantum mechanics, and form the backbone behind much of mathematical physics. Here is the definition:

Definition 5.6 (**Lie Group**)**.** A Lie group is a group that also has a manifold structure, on which the group multiplication and inversion operations are smooth—i.e. infinitely differentiable—maps.

If you're not familiar with the concept of a manifold, do not fret. There are simpler (though slightly more limiting) ways of defining Lie groups. The most important of these is the definition for matrix groups, because they are the most widely studied of the Lie groups out there.

Definition 5.7 (**Matrix Lie Group**)**.** A subgroup $H \subseteq GL_n(K)$ is called a **matrix Lie group** if the following property holds: if sequence of matrices $\{A_n\}$ in H converges to A, then either $A \in H$ or A is not invertible. An equivalent way of stating this is that H must be a closed subgroup of $GL_n(K)$.

The best examples of matrix Lie groups are the so-called *classical groups.* The following table summarizes a few examples of classical groups.

Example 5.8. Some Classical Groups.

Group Name	Definition
Orthogonal Group	$O(n) = \{A \in GL_n(\mathbb{R}) : AA^T = I\}$
Special Orthogonal Group	$SO(n) = \{A \in GL_n(\mathbb{R}) : AA^T = I \text{ and } \det(A) = 1\}$
Unitary Group	$U(n) = \{A \in GL_n(\mathbb{C}) : AA^\star = I\}$
Special Unitary Group	$SU(n) = \{A \in GL_n(\mathbb{C}) : AA^\star = I \text{ and } \det(A) = 1\}$

Representations of the above groups all describe fundamental symmetries of the physical universe, from planetary motion to special relativity to the spin of an electron. It's really quite magical stuff. The Heisenberg group is a Lie group, and its representation theory gives some insights into quantum mechanics. Before we discuss how this works, let's see how the Heisenberg group and Heisenberg algebra are actually related (besides by name).

6 From Lie Algebras to Lie Groups

Algebras are a nice structure, but groups are even better at describing the natural world. Group theory arises organically in the study of symmetries, so many problems in math and science can be well-understood using groups. The Heisenberg Group is no exception. Fortunately for us, there exists a mechanism that maps a Lie algebra directly to its corresponding Lie group. That handy little device is called the *exponential map.* We don't yet have the mathematical infrastructure to describe the exponential map in general terms, but we can still use the concept on our familiar matrices.

Definition 6.1 (**Matrix Exponential**)**.** Let A be an $n \times n$ matrix. Then the **matrix eponential** of A is defined as:

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} = I + A + \frac{1}{2}A^2 + \frac{1}{6}A^3 \dots$$

This formula ought to look familiar. It's precisely analogous to the Taylor Series expansion for the exponential function. Except this time, instead of a familiar number or variable in the exponent, it's a matrix. (We will sometimes write $\exp(A)$ instead of e^A.)

The matrix exponential is simply one small case of the more general phenomenon called an exponential map. But the matrix exponential still has great power in mapping a Lie algebra to its corresponding Lie group.

Back in the very lengthy Example 4.10 we saw that it's possible to represent the Heisenberg algebra via matrices. Exponentiating these matrices will lead us to a definition of the Heisenberg group.[10] So, here's the plan (inspired by [4]): We will first show how the matrix exponential works using a particular matrix. We will then show how we can take the exponential of $\mathfrak{h}_n$ in its entirety. Finally, we will show that the exponential of $\mathfrak{h}_n$ is actually a group. Here goes!

6.1 Exponential of Heisenberg Algebra

We can represent an element $(p, q, t) \in \mathfrak{h}_n$ by the matrix $m(p, q, t)$. Now, by (4.3) (or by simple matrix multiplication) we know that

$$m(p, q, t)^2 = m(0, 0, pq).$$

[10] Actually, it will lead us to two definitions of the Heisenberg group, which I will then show are equivalent to one another. ↑

What, then is $m(p,q,t)^3$? Well, it's simply:

$$\begin{aligned} m(p,q,t)^3 &= m(p,q,t)^2 m(p,q,t) \\ &= m(0,0,pq)m(p,q,t) \\ &= m(0,0,0). \end{aligned}$$

The zero matrix! Since multiplying any matrix by the zero matrix gives zero, we safely conclude that $m(p,q,t)^n = m(0,0,0)$ for $n \geq 3$.

Now, what is the matrix exponential of $m(p,q,t)$? Well, simply following the definition in 6.1 yields:

$$\begin{aligned} \exp(m(p,q,t)) &= I + m(p,q,t) + \frac{1}{2}m(p,q,t)^2 + \frac{1}{6}m(p,q,t)^3 + \dots \\ &= I + m(p,q,t) + \frac{1}{2}m(0,0,pq) \\ &= I + m(p,q,t+\tfrac{1}{2}pq) \end{aligned} \tag{6.1}$$

Now, to clear up the notation a little bit, let me introduce the following:

$$M(p,q,t) := I + m(p,q,t).$$

It's easy to imagine M as simply m with the diagonal "filled-in" with ones. Using this new definition and (6.1), we can write:

$$\exp(m(p,q,t)) = M(p,q,t+\tfrac{1}{2}pq).$$

Success! We have exponentiated an element of $\mathfrak{h}_n$. But remember, the goal is to exponentiate the *whole algebra*. Well, the following argument shows how that's possible:

$$\begin{aligned} \exp\{m(p,q,t) : p,q \in \mathbb{R}^n, t \in \mathbb{R}\} &= \{M(p,q,t+\tfrac{1}{2}pq) : p,q \in \mathbb{R}^n, t \in \mathbb{R}\} \\ &= \{M(p,q,t) : p,q \in \mathbb{R}^n, t \in \mathbb{R}\} \end{aligned}$$

I used the fact that the set of all things that look like $t + \frac{1}{2}pq$ is identical to the set of all t, since t can be any real number. Convince yourself of this!

We now have the following chain of mappings:

$$\{(p,q,t)\} \xrightarrow{m} \{m(p,q,t)\} \xrightarrow{\exp} \{M(p,q,t)\} \tag{6.2}$$

We call $\{M(p,q,t,)\}$ the **polarized Heisenberg group**, denoted $\mathcal{H}_{\text{pol}}$ with the group operation given by matrix multiplication:

$$M(p,q,t) \cdot M(p',q',t') = M(p+p', q+q', t+t'+pq'). \tag{6.3}$$

Actually, most of the time we drop the M and just define the polarized group like this:

$$\mathcal{H}_{\text{pol}} := \{(p,q,t) \in \mathbb{R}^{2n+1} : (p,q,t) \cdot (p,q,t') = (p+p', q+q', t+t'+\tfrac{1}{2}pq)\} \tag{6.4}$$

However, a slightly different definition is preferable. Instead of defining the group as the set of $M(p,q,t)$, let's use the set of all $\exp(m(p,q,t))$. Of course, these two things are equal, but the group operation we obtain from the latter is slightly different. Check it out:

$$\begin{aligned} \exp(m(p,q,t)) \cdot \exp(m(p',q',t')) &= M(p,q,t+\tfrac{1}{2}pq) \cdot M(p',q',t'+\tfrac{1}{2}p'q') \\ &= M(p+p', q+q', t+t'+\tfrac{1}{2}(pq+p'q')+pq') \\ &= \exp(m\big(p+p', q+q', t+t'+\tfrac{1}{2}(pq'-qp')\big)) \end{aligned} \tag{6.5}$$

That last line relies on the fact that

$$\begin{aligned}\exp(m(p+p', q+q', t+t'+\tfrac{1}{2}(pq'-qp'))) &= M(p+p', q+q', t+t'+\tfrac{1}{2}(pq'-qp')+\tfrac{1}{2}(p+p')(q+q')) \\ &= M(p+p', q+q', t+t'+ \\ &\qquad \tfrac{1}{2}pq' - \tfrac{1}{2}\cancel{(qp')} + \tfrac{1}{2}pq + \tfrac{1}{2}pq' + \tfrac{1}{2}p'q' + \tfrac{1}{2}\cancel{qp'} \\ &= M(p+p', q+q', t+t'+\tfrac{1}{2}(pq+p'q')+pq') \qquad \checkmark\end{aligned}$$

We have now obtained the operation for the Heisenberg group, found in (6.5). Just like with the polarized group, we drop the $\exp m$ and just retain the (p, q, t). As you can see, this precisely matches the original definition of the Heisenberg group back in Example 5.3. That is,

$$\mathcal{H} := \{(p,q,t) \in \mathbb{R}^{2n+1} : (p,q,t)\cdot(p',q',t') = (p+p', q+q', t+t'+\tfrac{1}{2}(pq'-qp'))\}$$

The relationship between the two group definitions are summarized in the table below.

Exponentials:	$\exp(m(p,q,t))$	$\cdot$	$\exp(m(p',q',t'))$	$=$	$\exp(m(p+p', q+q', t+t'+\tfrac{1}{2}(pq'-qp'))$
	$\Updownarrow$		$\Updownarrow$		$\Updownarrow$
Matrices:	$M(p,q,t)$	$\cdot$	$M(p',q',t')$	$=$	$M(p+p', q+q', t+t'+pq')$

In fact, it's possible to show that these two definitions are actually fully equivalent, since there is an isomorphism between them. Let $(\mathcal{H}, \circ)$ be the Heisenberg group, $(\mathcal{H}_{\text{pol}}, \circ_p)$ the polarized Heisenberg group, and $\varphi : \mathcal{H} \to \mathcal{H}_p$ a map between them. Define φ as follows:

$$\varphi(p,q,t) = (p,q,t+\tfrac{1}{2}pq).$$

Claim. *The map φ is an isomorphism between $\mathcal{H}$ and $\mathcal{H}_{pol}$.*

Proof. First, surjective: any element $(p,q,t) \in \mathcal{H}_{\text{pol}}$ is his by $(p,q,t-\frac{1}{2}pq) \in \mathcal{H}$.

Now, injective. If $\varphi(p,q,t) = \varphi(p',q',t')$, then $(p,q,t+\frac{1}{2}pq) = (p',q',t'+\frac{1}{2}p'q')$, and the only way this is true is if $p = p'$ and $q = q'$, which implies $t = t'$.

Finally, φ is a homomorphism:

$$\begin{aligned}\varphi((p,q,t)\circ(p',q',t')) &= \varphi(p+p', q+q', t+\tfrac{1}{2}(pq'-qp')) \\ &= (p+p', q+q', t+\tfrac{1}{2}pq' - \tfrac{1}{2}qp' + (p+p')(q+q')) \\ &= (p+p', q+q', t+\tfrac{1}{2}pq + t' + \tfrac{1}{2}p'q' + pq') \\ &= (p,q,t+\tfrac{1}{2}pq)\circ_p(p',q',t'+\tfrac{1}{2}p'q') \\ &= \varphi(p,q,t)\circ_p\varphi(p',q',t')\end{aligned}$$

□

7 Digging Deep into the Heisenberg Algebra and Heisenberg Group

Let's revisit the definition of the Heisenberg algebra. We said that is was made up of tuples in $\mathbb{R}^{2n+1}$ that looked something like:

$$(p_1, \ldots, p_n, q_1, \ldots, q_n, t) = (p, q, t).$$

We will now focus on the case where $n = 1$ in this definition. This means that our tuples sit in nice, familiar $\mathbb{R}^3$. If you'll indulge a small (and hopefully not too confusing) variable name change, let us write the elements of our $n = 1$ group as follows:

$$\mathfrak{h}_1 := \{(p,x,z) : p, x, z \in \mathbb{R}\}.$$

Hopefully you can already anticipate why we made that change. Our Lie algebra representation m showed how to view elements of this algebra as matrices, as follows:

$$(p, x, z) \xmapsto{m} \begin{pmatrix} 0 & p & z \\ 0 & 0 & x \\ 0 & 0 & 0 \end{pmatrix}$$

Now here's a question: how would we go about forming a basis for this Lie algebra? That is, can we obtain any member of $\mathfrak{h}_1$ through linear combinations of some elementary matrices? Well, here's a natural choice:

$$\begin{pmatrix} 0 & p & z \\ 0 & 0 & x \\ 0 & 0 & 0 \end{pmatrix} = p \underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{P} + x \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}}_{X} + z \underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{Z}$$

As you can see, we've defined the elementary matrices P, X, and Z, which serve as the basis of the Heisenberg algebra. Since these basis matrices are themselves elements of the Lie algebra, we can compute their brackets as follows:

$$[P, X] = Z$$
$$[P, Z] = 0$$
$$[X, Z] = 0$$

The relationship between the brackets of the three elements P, X, and Z is called the *canonical commutation relations* (CCR). This is the first we've seen of the CCR in this paper, so I will be careful to introduce them in an intuitive way (though there are many ways to explain it). At the level of the Lie algebras, the CCR tell us the degree to which two quantities fail to commute. Our two matrices P and X clearly do not commute, which is why their commutator is non-zero. All other pairs of matrices do—i.e. $PZ = ZP$—so their commutator is zero. Later on, we will see that the CCR points to a deeper truth about Fourier transforms and conjugate variables, but for now the matrix interpretation will suffice.

Next, let's see what happens when we exponentiate the basis matrices of the Lie algebra. Notice that $P^2 = 0, X^2 = 0$, and $Z^2 = 0$. Thus:

$$\exp(P) = I + P + \frac{1}{2}\cancel{P^2} + \ldots = M(1,0,0) := M(P)$$
$$\exp(X) = I + X + \frac{1}{2}\cancel{X^2} + \ldots = M(0,1,0) := M(X)$$
$$\exp(Z) = I + Z + \frac{1}{2}\cancel{Z^2} + \ldots = M(0,0,1) := M(Z)$$

This is a nice result, and is to be expected: exponentiating the basis matrices of the Heisenberg algebra gives the basis matrices of the Heisenberg group! Notice that these matrices also obey the canonical commutation relations, since by the commutator operation for groups:

$$\begin{aligned} \big[M(P), M(X)\big] &= M(P)M(X)M(P)^{-1}M(X)^{-1} \\ &= (1,0,0)(0,1,0)(-1,0,0)(0,-1,0) \\ &= (1,1,\tfrac{1}{2})(-1,-1,\tfrac{1}{2}) \\ &= (0,0,1) \\ &= M(Z) \end{aligned}$$

By similar methods, we can obtain $\big[M(P), M(Z)\big] = 0$ and $\big[M(X)M(Z)\big] = 0$. In an upcoming section, we will see the relationship between the CCR of the Heisenberg algebra and Heisenberg group and the operators

X and P introduced in section 3.

As a small preview of things to come, let us show how those operators that gave us the Uncertainty Principle also obey the CCR. Recall that we defined the operators X and P as follows: $X(f(x)) = x(f(x))$ and $P(f(x)) = -i\frac{d}{dx}f(x)$.

Let us also introduce a new operator Z, defined by $Z(f(x)) = -if(x)$. In summary, our operators are:

$$\begin{aligned} Xf &= xf \\ Pf &= -i\frac{d}{dx}f \\ Zf &= -if \end{aligned}$$

We will now calculate the commutators of these operators. First, X and P:

$$\begin{aligned} [P,X]f &= PXf - XPf \\ &= Pxf - X\left(-i\frac{d}{dx}f\right) \\ &= -\left(i\frac{d}{dx}(xf)\right) + ix\frac{d}{dx}f \\ &= -\big(i(f + xf')\big) + ixf' \\ &= \cancel{-ixf'} - if + \cancel{ixf'} \\ &= Zf \end{aligned}$$

Thus, we get the commutation relationship $[X,P] = Z$. Next, let's try $[X,Z]$:

$$\begin{aligned} [X,Z]f &= XZf - ZXf \\ &= Xif - Zxf \\ &= xif - ixf \\ &= 0 \end{aligned}$$

Finally, let's do $[P,Z]$:

$$\begin{aligned} [P,Z]f &= PZf - ZPf \\ &= Pif - Z(-i\hbar f') \\ &= -i(i\hbar f') - i(-i\hbar f') \\ &= 0 \end{aligned}$$

Thus we obtain the old commutation relations:

$$\begin{aligned} [P,X] &= Z \\ [P,Z] &= 0 \\ [X,Z] &= 0 \end{aligned}$$

In the next few sections, we will examine the representation theory of the Heisenberg group. Most importantly, we will see how a representation of $\mathcal{H}$ as operators on functions leads us right back to the position and momentum operators and this CCR.

8 Preliminary Definitions

The remainder of this paper will discuss the representation theory of the Heisenberg group.[11] In short, we want to represent elements of $\mathcal{H}$ as operators over the infinite-dimensional vector space $L^2(\mathbb{R})$. Since physical observables in quantum mechanics are represented as operators on $L^2(\mathbb{R})$, this representation is how we will make the connection between the Heisenberg group and physics. Though we have already seen the definition of a group and algebra representation, let us quickly revisit the definition in more abstract terms to get a better understanding.

Definition 8.1 (**Representation**). A representation π of a group G is a homomorphism from G to the group $GL(V)$ of invertible linear operators on V, where V is a nonzero complex vector space.

We refer to V as the *representation space* of π. If V is finite-dimensional, we say that π is finite-dimensional, and the *degree* of π is the dimension of V. Otherwise, we say that π is infinite-dimensional. We usually use the notation (π, V) when referring to a representation.

Remark. *The following definitional phrases of $GL(V)$ are equivalent: "the set of all invertible linear maps $T : V \to V$," "the set of all invertible linear operators on V," "the set of all automorphisms of V," and "the set of all bijective linear transformations $T : V \to V$."*

A representation π is called *unitary* if for every $g \in G$ the operator $\pi(g)$ is unitary on V, i.e., if

$$\langle \pi(g)(v), \pi(g)(w) \rangle = \langle v, w \rangle \quad \text{for all } v, w \in V, g \in G.$$

A closed subspace $W \subset V$ is called *invariant* for π if $\pi(g)W \subset W$ for every $g \in G$. The representation π is called *irreducible* if there is no proper closed invariant subspace, i.e., the only closed invariant subspaces are 0 and V itself. We can think of an irreducible representation as being made up of the most "basic" possible elements, and is therefore unable to be split into simpler pieces.

The next concept we need to introduce is that of a *character*:

Definition 8.2 (**Character**). Let A be an abelian group. A **character** of A is a continuous group homomorphism $\chi : A \to \mathbb{T}$ where $\mathbb{T}$ represents the unit torus, also called the circle group:

$$\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}.$$

It is important to note that $\mathbb{T}$ is isomorphic to both $\mathbb{R}/\mathbb{Z}$ and $\mathbb{R}/2\pi\mathbb{Z}$.

Proposition. *The characters of some important groups are given as follows:*

- *The characters of the group $\mathbb{Z}$ are given by $k \mapsto e^{2\pi i k x}$, where $x \in \mathbb{R}/\mathbb{Z}$.*
- *The characters of $\mathbb{R}/\mathbb{Z}$ are given by $x \mapsto e^{2\pi i k x}$, where $k \in \mathbb{Z}$.*
- *The characters of $\mathbb{R}$ are given by $x \mapsto e^{2\pi i x y}$, where $y \in \mathbb{R}$.*

Proof. We will only prove the first statement. Let $\varphi : \mathbb{Z} \to \mathbb{T}$ be a character. Then $\varphi(1) = e^{2\pi i x}$ for some $x \in \mathbb{R}/\mathbb{Z}$. So for an arbitrary $k \in \mathbb{Z}$ we get $\varphi(k) = \varphi(1 + 1 + ... + 1) = \varphi(1)^k = e^{2\pi i k x}$ due to group homomorphism. □

Let A be an locally compact and abelian group. The locally compact property is a formality that is beyond the scope of this paper, but simply means that the topology of A is sufficiently nice for Harmonic analysis. Let $\hat{A}$ denote the set of all characters of A. That is, $\hat{A}$ is the set of all homomorphisms from A to $\mathbb{T}$.

Lemma 1 (**Pontryagin Duality**). *$\hat{A}$ is an abelian group with the group operation given by the pointwise product*

$$\chi\eta(a) = \chi(a)\eta(a)$$

for all $\chi, \eta \in \hat{A}$ and $a \in A$. Moreover, $\hat{A}$ is called the dual group, or the Pontryagin dual, of A.

[11] All of the definitions and proofs in the remaining sections come from [2] and [3]. ↑

Proof. We need to show first that if $\chi, \eta \in \hat{A}$, then $\chi\eta \in \hat{A}$, meaning $\chi\eta$ is a continuous group homomorphism from A to $\mathbb{T}$. We will omit the continuity argument, but prove that $\chi\eta$ is a group homomorphism. Consider the following:

$$\begin{aligned}\chi\eta(ab) &= \chi(ab)\eta(ab) \\ &= \chi(a)\chi(b)\eta(a)\eta(b) \\ &= \chi(a)\eta(a)\chi(b)\eta(b) \\ &= \chi\eta(a)\chi\eta(b)\end{aligned}$$

And this completes the proof. □

The dual group of some of the important sets is given below.

- The dual group of $\mathbb{Z}$ is isomorphic to $\mathbb{R}/\mathbb{Z}$.
- The dual group of $\mathbb{T}$ is isomorphic to $\mathbb{Z}$.
- The dual group of $\mathbb{R}$ is isomorphic to $\mathbb{R}$.

The Pontryagin duality helps explain the nature of Fourier series and Fourier transforms. For example, Fourier series are used to represent periodic functions. For argument sake, let's assume we're dealing with a function $f(x)$ with period 2π. Then the domain of f is isomorphic to $\mathbb{R}/2\pi\mathbb{Z}$, or $\mathbb{T}$. The dual of $\mathbb{T}$ is $\mathbb{Z}$, which is why the Fourier coefficients c_k are indexed by integers. Similarly, if f is not periodic, and has domain $\mathbb{R}$ we can take its Fourier transform, which we can view as "continuously indexed" over the dual of $\mathbb{R}$, which is also $\mathbb{R}$. That is why we take an integral in a Fourier transform as opposed to a sum in a Fourier series.

9 Unitary Dual

Definition 9.1 (**Isomorphic/Unitarily equivalent representations**)**.** For a group G, two unitary representations (π, V_π) and (η, V_η) are called **isomorphic** or **unitarily equivalent** if there exists a unitary operator $T : V_\pi \to V_\eta$ such that

$$T \circ \pi(g) = \eta(g) \circ T$$

for every $g \in G$.

When two representations are unitarily equivalent, they are essentially indisinguishable as representations. That is, they differ by name alone. Since isomorphism is an equivalence relation on the class of unitary representations, we can consider the equivalence class $[\pi]$ for a representation $\pi : G \to GL(V)$. The class $[\pi]$ consists of all representations that are isomorphic to π. The set of all such equivalence classes is called the *unitary dual* of G, denoted by $\hat{G}$. This is summarized in the following definition:

Definition 9.2 (**Unitary Dual**)**.** The unitary dual of a group G is the collection of all equivalences clases of representations from G to $GL(V)$ for some fixed vector space V. That is:

$$\hat{G} = \{\pi \text{ irreducible unitary}\} / \text{isomorphy}.$$

10 The Heisenberg Group and its Unitary Dual

Recall that the Heisenberg group $\mathcal{H}$ is defined to be the group of real upper triangular 3×3 matrices with ones on the diagonal:

$$\mathcal{H} = \left\{ \begin{pmatrix} 1 & p & t \\ 0 & 1 & q \\ 0 & 0 & 1 \end{pmatrix} \middle| \; p, q, t \in \mathbb{R} \right\}$$

As we've seen, $\mathcal{H}$ is not an abelian group, though we can identify $\mathcal{H}$ with elements of $\mathbb{R}^3$ under the following group law:

$$(p,q,t)(p',q',z) = (p+p', q+q', t+t'+\tfrac{1}{2}(pq'-qp')).$$

The inverse of (p,q,t) is

$$(p,q,t)^{-1} = (-p,-q,-t).$$

The center of $\mathcal{H}$—i.e. the set of all elements in $\mathcal{H}$ that commute with all of $\mathcal{H}$— is $Z(\mathcal{H}) = \big\{(0,0,t) \mid t \in \mathbb{R}\big\}$. This can quickly be seen through the group operation. Since the center of $\mathcal{H}$ is isomorphic to $\mathbb{R}$, then the factor group $\mathcal{H}/Z(\mathcal{H})$ is unsurprisingly isomorphic to $\mathbb{R}^2$ through the following map:

$$\begin{aligned}\varphi &: \mathcal{H}/Z(\mathcal{H}) \to \mathbb{R}^2 \\ \varphi &: \big((p,q,t)Z(\mathcal{H})\big) \to (p,q)\end{aligned}$$

The map φ is surjective since any pair $(p,q) \in \mathbb{R}$ is mapped to by $(p,q,t)Z(\mathcal{H}) \in \mathcal{H}/Z(\mathcal{H})$. It is injective since if $(p,q) = (p',q') \in \mathbb{R}^2$, then their preimages under φ both belong to the same coset. Finally, φ is a homomorphism since.

$$\begin{aligned}\varphi\big((p,q,t)Z(\mathcal{H}) \cdot (p',q',t')Z(\mathcal{H})\big) &= \varphi\big((p+p', q+q', t+t'+\tfrac{1}{2}(pq'-q'p'))Z(\mathcal{H})\big) \\ &= (p+p', q+q') \\ &= (p,q) + (p',q') \\ &= \varphi\big((p,q,t)Z(\mathcal{H})\big) + \varphi\big((p',q',t')Z(\mathcal{H})\big)\end{aligned}$$

Thus, we can conclude that $\mathcal{H}/Z(\mathcal{H}) \cong \mathbb{R}^2$.

Let $\hat{\mathcal{H}}_0$ denote the subset of $\hat{\mathcal{H}}$ consisting of all equivalence classes $[\pi] \in \hat{\mathcal{H}}$ such that $\pi(h) = 1$ whenever h lies in the center of $\mathcal{H}$. Since $\mathcal{H}/Z(\mathcal{H}) \cong \mathbb{R}^2$, it follows that

$$\hat{\mathcal{H}}_0 = \widehat{\mathcal{H}/Z(\mathcal{H})} \cong \widehat{\mathbb{R}^2} \cong \hat{\mathbb{R}}^2,$$

and the latter can be identified with $\mathbb{R}^2$ in the following explicit way. Let $(a,b) \in \mathbb{R}^2$ and define a character

$$\begin{aligned}\chi_{a,b} &: \mathcal{H} \to \mathbb{T} \\ (x,y,z) &\mapsto e^{2\pi i(ax+by)}\end{aligned}$$

The identification is given by $(a,b) \mapsto \chi_{a,b}$. In particular, it follows that all representations in $\hat{\mathcal{H}}_0$ are one-dimensional. This observation indicates the importance of the behavior of the center under a representation. In addition, the equality of $\hat{\mathcal{H}}_0$ and $\widehat{\mathcal{H}/Z(\mathcal{H})}$ is given by the fact that the identity of $\mathcal{H}/Z(\mathcal{H})$ is just $Z(\mathcal{H})$.

Lemma 2. *Let (π, V_π) be an irreducible unitary representation of a locally compact group G. Let $Z(G) \subset G$ be the center of G. Then for every $z \in Z(G)$ the operator $\pi(z)$ on V_π is a multiple of the identity.*

We will not prove the lemma here, though we will point out an important consequence of the lemma: For each $[\pi] \in \hat{G}$ there is a character $\chi_\pi : Z(G) \to \mathbb{T}$ with $\pi(z) = \chi_\pi(z)\mathrm{Id}$ for every $z \in Z(G)$. This character χ_π is called the *central character* of the representation π.

For every character $\chi \neq 1$ of $Z(\mathcal{H})$, we will now construct an irreducible unitary representation of the Heisenberg group that has χ for its central character. So let $k \neq 0$ be a real number and consider the central character

$$\chi_k(0,0,t) = e^{ikt}.$$

For $(p,q,t) \in \mathcal{H}$ we define the operator $\pi_k(p,q,t)$ on $L^2(\mathbb{R})$ by

$$\pi_k(p,q,t)\varphi(x) = e^{i(qx+t)k}\varphi(x+p).$$

To show π_k is a unitary representation, we need to show $\langle \pi_k\varphi, \pi_k\psi\rangle = \langle \varphi, \psi\rangle$ for all $\varphi, \psi \in L^2(\mathbb{R})$.

$$\langle \pi_k(p,q,t)\varphi, \pi_k(p,q,t)\psi\rangle = \int_{\mathbb{R}} e^{i(qx+t)k}\varphi(x+p)\overline{e^{i(qx+t)k}\psi(x+p)}dx \tag{10.1}$$

$$= \int_{\mathbb{R}} \varphi(x+p)\overline{\psi(x+p)}dx \tag{10.2}$$

$$= \int_{\mathbb{R}} \varphi(x)\overline{\psi(x)}dx \tag{10.3}$$

$$= \langle \varphi, \psi\rangle \tag{10.4}$$

Another way to see that π_k is unitary is to observe that π is the product of three one-parameter unitary operators: rotation by q, multiplication by a complex number e^{itk}, and translation by p. More formally, these operators can be written as:

$$R(q)\varphi(x) = e^{iqx}\varphi(x)$$
$$\chi_k(t)\varphi(x) = e^{ikt}\varphi(x)$$
$$T(p)\varphi(x) = \varphi(x+p)$$

These operators will soon become very important in demonstrating the relationship between the representation theory of the Heisenberg group and the position and momentum operators from earlier. For now, though, let's make our way the crowning jewel of this section: the Stone-von Neumann theorem.

The Schwartz space $\mathcal{S}(\mathbb{R}^n)$ is the space of functions whose derivatives are rapidly decreasing. We have $\mathcal{S}(\mathbb{R}^n) \subseteq L^p(\mathbb{R}^n)$ for every $p \geq 1$ and $\mathcal{S}(\mathbb{R}^n)$ is stable under Fourier transform, i.e., if $f \in \mathcal{S}(\mathbb{R}^n)$ then $\mathcal{F}f \in \mathcal{S}(\mathbb{R}^n)$.

Theorem 10.1 (**Stone-von Neumann**). *For $k \neq 0$ the unitary representation π_k is irreducible. Every irreducible unitary representation of $\mathcal{H}$ with central character χ_t is isomorphic to π_t. It follows that*

$$\hat{\mathcal{H}} = \hat{\mathbb{R}}^2 \cup \{\pi_k : k \neq 0\}.$$

Proof. We will only prove irreducibility. Fix $t \neq 0$ and let us assume the contrary and suppose $V \subset L^2(\mathbb{R})$ be a closed non-zero subspace that is invariant under the set of operators $\pi_k(\mathcal{H})$. If $\varphi \in V$, then so is the function $\pi_k(-p,0,0)\varphi(x) = \varphi(x-p)$. Since V is closed, it therefore contains $\psi * \varphi(x) = \int_{\mathbb{R}} \psi(p)\varphi(x-p)dp$ for $\psi \in \mathcal{S} = \mathcal{S}(\mathbb{R})$. These convolution products are smooth functions, so V contains a smooth function $\varphi \neq 0$. One has

$$\pi_k(0,-q,0)\varphi(x) = e^{-iqkx}\varphi(x) \in V.$$

By integration it follows that for $\psi \in \mathcal{S}$ one has that $\hat{\psi}(kx)\varphi(x)$ lies in V. The set of possible functions $\hat{\psi}(kx)$ contains all smooth functions of compact support, as the Fourier transform is a bijection on the space of Schwartz functions to itself. Choose an open interval I, in which φ has no zero. It follows that $C_c^\infty(I) \subset V$. This space is dense in $L^2(\mathbb{R})$, so $V = L^2(\mathbb{R})$, which means that π_t is irreducible.

□

The Stone-von Neumann theorem suggests that there is, up to isomorphism, really only one "good" family of representations of the Heisenberg group on $L^2(\mathbb{R})$. By "good" I mean that the operators on $L^2(\mathbb{R})$ are irrediciuble, unitary, and act non-trivially on the center of $\mathcal{H}$, which means that they are suitable for use in quantum mechanics. In the following section, we will see why the uniqueness of this representation can give us some insights into the mathematical structure of the quantum world.

11 Exploring the Schrödinger Representation

The representation π_k is usually called the *Schrödinger Representation* of the Heisenberg group. Innocuous though it may seem, the Schrödinger representation actually contains a wealth of information. Recalling our

rotation, multiplication, and translation operators R, χ and T, we can write the Schrödinger representation as follows for $k = 1$:

$$\pi(p, q, t) = R(q)\chi(t)T(p) \tag{11.1}$$

Any element $(p, q, t) \in \mathcal{H}$ is made up of a combination of the basis matrices $(p, 0, 0)$, $(0, q, 0)$, and $(0, 0, z)$. So, let's consider the representaion π solely in the direction of these three basis matrices in turn. First, in the direction of $(p, 0, 0)$ we have that $\pi(p, 0, 0) = T(p)$.

To better understand the structure of $\pi(p, 0, 0)$, let's take its derivative at $p = 0$. The derivative of the translation operator when applied to a function $\varphi(x)$ is given by:

$$\begin{aligned}
\frac{\partial}{\partial p}\pi(p,0,0)\varphi(x)\bigg|_{p=0} = \frac{d}{dp}T(p)\varphi(x)\bigg|_{p=0} &= \lim_{\varepsilon\to 0}\frac{T(\varepsilon)-T(p)}{\varepsilon - p}\varphi(x)\bigg|_{p=0} \\
&= \lim_{\varepsilon\to 0}\frac{T(\varepsilon)\varphi(x) - T(0)\varphi(x)}{\varepsilon} \\
&= \lim_{\varepsilon\to 0}\frac{\varphi(x+\varepsilon)-\varphi(x)}{\varepsilon} \\
&= \frac{d}{dx}\varphi(x)
\end{aligned}$$

Thus, we obtain that the derivative of the translation operator is the differential operator $\frac{d}{dx}$.

Similarly, let's consider $\pi(0, q, 0) = R(q)$, which is the representation solely in the direction of the second basis matrix. Taking its derivative at $q = 0$ gives:

$$\begin{aligned}
\frac{\partial}{\partial q}\pi(0,q,0)\varphi(x)\bigg|_{q=0} = \frac{d}{dq}R(q)\varphi(x)\bigg|_{q=0} &= \lim_{\varepsilon\to 0}\frac{R(\varepsilon)-R(q)}{\varepsilon - p}\varphi(x)\bigg|_{q=0} \\
&= \lim_{\varepsilon\to 0}\frac{R(\varepsilon)\varphi(x) - R(0)\varphi(x)}{\varepsilon} \\
&= \lim_{\varepsilon\to 0}\frac{\varphi(x)(e^{i\varepsilon x}-1)}{\varepsilon} \\
&= \varphi(x)\lim_{\varepsilon\to 0}\frac{\cos(\varepsilon x) + i\sin(\varepsilon x)}{\varepsilon} \\
&= ix\varphi(x)
\end{aligned}$$

Thus, the derivative of the rotation operator is the operator ix, which multiplies a function by i times the function's argument.

You will notice that the two operators we've obtained by differentiating $R(q)$ and $T(p)$ look remarkably like our position and momentum operators from the discussion of Heisenberg's uncertainty principle. In fact, with a simple rescaling (multuplying by $-i$), we obtain that:

$$\begin{aligned}
-iR'(q)\varphi(x) &= x\varphi(x) = X(\varphi(x)) \\
-iT'(p)\varphi(x) &= -i\frac{d}{dx}\varphi(x) = P(\varphi(x))
\end{aligned}$$

According to the Stone-von Neumann theorem, since the Schrödinger representation is unique, these two operators are the only valid representations of the Heisenberg group on $L^2(\mathbb{R})$. What this means, more importantly, is that X and P are the only operators on $L^2(\mathbb{R})$ that satisfy the canonical commutation relationship (up to isomorphism), since the CCR is the defining structure of the Heisenberg algebra and Heisenberg group. Everything has now come full circle! The one-ness of everything we have discussed will be summarized in the next section.

12 Summary

We have covered a lot of ground in this project write-up. From introducing the uncertainty princpiple to the Hermite functions, and from Lie algebras to the Schrödinger representation, it is often easy to lose sight of the connections between these seemingly disparate areas of mathematics. However, in this final summary, we will hopefully make clear that everything discussed here is intimately related, and can be expressed in terms of a small number of elegant, consise mathematical statements.

The "lowest level" object we discussed was the Heisenberg algebra, a structure generated by two elementary matrices. Exponentiating this algebra gave a Lie group known as the Heisenberg group, made up of all real upper-triangular matrices with ones along the main diagonal. A look into the representation theory of the Heisenberg group revealed the Schrödinger representation, which by the Stone-von Neumann theorem is the only family of irreducuble unitary representations of the Heisenberg group, up to isomorphism. Taking derivatives of the Schrödinger representation along the basis matrices of the Heisenberg group yielded the position and momentum operators of quantum mechanics.

These position and momentum operators are conjugates of each other under Fourier transform. By the Heisenberg uncertainty principle, there is a fundamental limit on how accurately one can simultaneously specify two quantities which are Fourier conjugates. This means that the universe precludes knowing precisely both the momentum and position of a particle. Our proof of the uncertainty principle used the Hermite operator and its orthonormal eigenfunctions, the Hermite functions. This Hermite operator is in fact the Laplacian of the Schödinger representation, which points to why it so readily proved the theorem.

This last statement requires further qualification, because it's particularly important. Recall that in Euclidean space, the Laplacian is simply the differential operator applied twice in the direction of each basis vector. In $\mathbb{R}^2$, for example, the Laplacian operator is $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$. In the Schrödinger representation, the the derivative along the p direction was $P = -i\frac{d}{dx}$. Along the x basis vector, the derivative is $X = x$. Applying these operators twice gives

$$X^2 + P^2 = x^2 - \frac{d^2}{dx^2} = H.$$

All of the above can be summarized in this sentences. The Schrödinger representation of the Heisenberg group uniquely determines two operators that satisfy the CCR over $L^2(\mathbb{R})$, and its Laplacian can help prove that those operators obey Heisenberg's uncertainty principle.

Thank you for reading!

References

[1] Michael G. Cowling and John F. Price, *Bandwidth Versus Time Concentration: The Heisenberg-Pauli-Weyl Inqauality*, SIAM Journal on Mathematical Analysis **15** (1984), 151 – 165.

[2] Anton Deitmar, *A First Course in Harmonic Analysis*, 2 ed., Springer-Verlag, New York, 2005.

[3] Anton Deitmar and Siegfried Echterhoff, *Principles of Harmonic Analysis*, Springer-Verlag, New York, 2009.

[4] Gerald B. Folland, *Harmonic Analysis in Phase Space*, Annals of Mathematics Studies, Princeton University Press, 1989.

[5] Philippe Jaming, *Uncertainty Principles for Orthonormal Bases*, ArXiv Mathematics e-prints (2006).

[6] Brad Osgood, *Lecture Notes for EE 261: The Fourier Transform and its Applications*, Electrical Engineering Department, Stanford University.

[7] Alladi Sitaram, *Uncertainty Principles and Fourier Analysis*, Resonance (1999), 20 – 23.

[8] Sundaram Thangevlu, *An Introduction to the Uncertainty Principle: Hardy's Teorem on Lie Groups*, Springer Science & Business Media, Birkhäuser Boston, 2004.