1 Outline

Agenda: Uncertainty Principle

1. Weyl-Heisenberg Uncertainty Principle

2. Quantum mechanical interpretation

Last Time: Motivated by the boxcar function, whose Fourier transform is not summable, we introduced the space of square-integrable functions $L^2(\mathbb{R})$. We proved the Parseval-Plancherel theorem, which shows the Fourier transform preserves the L^2-inner product and is therefore an isometry (modulo a factor 2π). Using this result, we were able to extend the Fourier transform to $L^2(\mathbb{R})$ and thus make sense of the inverse Fourier transform of the boxcar function as a square-integrable function. The properties we proved also apply to this extension, with the remark that some equalities hold almost everywhere. We also defined the Fourier transform in higher dimensions by simply iterating the single-variable transform.

2 Weyl-Heisenberg's Uncertainty Principle

The uncertainty principle is commonly known in physics as saying that one cannot know simultaneously the position and the momentum of a particular with infinite precision. In fact, this statement is an implication of that mathematical observation that f and $\hat{f}$ cannot *both* be concentrated. This is something we can intuitively see from some examples. For instance, our calculations involving Gaussians (Lecture 1) revealed that increasing or decreasing the standard deviation in the time (or spatial) domain has the opposite effect in the frequency (or Fourier) domain. Fig. 1 illustrates this phenomenon. Taking this example to the extreme cases of 0 or infinite standard deviation gives the delta function in one domain and the constant function in the other. The former is completely localized, while the latter is completely delocalized.

Even without these examples, one can initially understand the uncertainty principle from the identity
$$F(t) = af(at) \quad \Leftrightarrow \quad \hat{F}(\omega) = \hat{f}(\omega/a).$$

That is, dilating (compressing) time is equivalent to compressing (dilating) frequency. What the uncertainty principle does is provide a precise, quantitative statement of this notion.
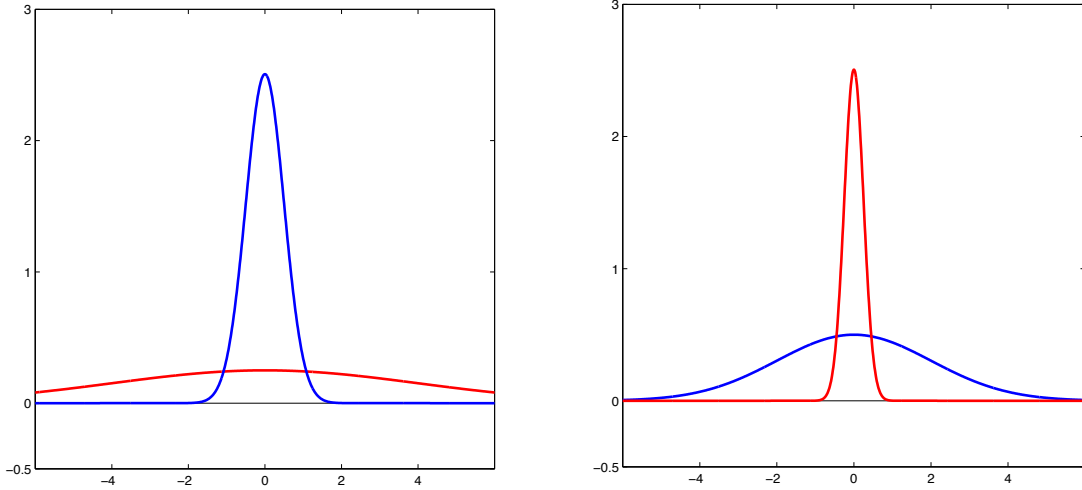
Time domain: The blue curve corresponds to a signal f_1 concentrated in time, whereas the red curve corresponds to a signal f_2 spread in time.

Frequency domain: The Fourier transform $\hat{f}_1$ is spread in frequency, whereas the Fourier transform $\hat{f}_2$ concentrates in frequency.

Figure 1: Visual evidence of the uncertainty principle

To state the result, we need to make a definition. For $f \in L^2(\mathbb{R})$, define the **spread** of f to be

$$\sigma^2(f) \triangleq \inf_{t_0 \in \mathbb{R}} \frac{1}{\|f\|^2} \int (t - t_0)^2 |f(t)|^2 \, dt.$$

We can interpret this quantity in probabilistic terms by noting that $p(t) = |f(t)|^2 / \|f\|^2$ is a probability density function, as

(i) $p(t) \geq 0$,

(ii) $\int p(t) \, dt = 1$.

Therefore, one can recognize $\sigma^2(f)$ as the variance of a random variable X with density $p(t)$, where the minimizing t_0 is exactly the mean

$$t_0 = \mathbb{E}X = \int t p(t) \, dt.$$

Then $\sigma(f)$ is, of course, the standard deviation of X. Similarly, $q(\omega) = |\hat{f}(\omega)|^2 / \|\hat{f}\|^2$ is the density of random variable Y having mean

$$\omega_0 = \mathbb{E}Y = \int \omega q(\omega) \, d\omega$$

and standard deviation $\sigma(\hat{f})$. With this notation, we can now state the uncertainty principle:

Theorem 1 (Weyl-Heisenberg Uncertainty Principle). *If $f \in L^2(\mathbb{R})$ is not identically 0, then*

(i) $\sigma(f)\sigma(\hat{f}) \geq 1/2.$

2

(ii) Equality holds if and only if f is a translation and modulation of a Gaussian function.

Thus, in terms of simultaneous time and frequency concentration, Gaussians are maximal with respect to mean squared-deviation.

Proof. Let us begin with some simplifying reductions. First, we may assume that X from above is centered ($\mathbb{E}X = 0$), since translations and modulations do not affect spread:

$$F = f(t + t_0) \quad \Rightarrow \quad \sigma^2(F) = \sigma^2(f) \quad \text{and} \quad \sigma^2(\hat{F}) = \sigma^2(e^{i\omega t_0} \hat{f}) = \sigma^2(\hat{f}).$$

Second, we may assume that Y is also centered, for the same reason:

$$F(t) = e^{-i\omega_0 t} f(t) \quad \Rightarrow \quad \sigma^2(F) = \sigma^2(f) \quad \text{and} \quad \sigma^2(\hat{F}) = \sigma^2(\hat{f}(\omega + \omega_0)) = \sigma^2(\hat{f}).$$

Third, f by $f/\|f\|$ does not change spread, and the density functions $p(t)$ and $q(\omega)$ are left unchanged. So we may assume f has norm 1. Numerically, our assumptions tell us

(i) $\int t|f(t)|^2 \, dt = \int \omega|\hat{f}(\omega)|^2 \, d\omega = 0$, and

(ii) $\int |f(t)|^2 \, dt = 1$.

Finally, smooth functions are dense in $L^2(\mathbb{R})$, and the Fourier transform preserves L^2-norm (modulo a factor of 2π). So it suffices to demonstrate the uncertainty principle for smooth, in particular differentiable, functions.

Now we will express $\sigma(\hat{f})$ in terms of f. Note that $f'(t) \xrightarrow{F} i\omega \hat{f}(\omega)$, and then the Parseval-Plancherel theorem yields

$$\sigma^2(\hat{f}) = \frac{1}{\|\hat{f}\|^2} \int \omega^2 |\hat{f}(\omega)|^2 \, d\omega = \frac{1}{2\pi} \int |i\omega \hat{f}(\omega)|^2 \, d\omega = \int |f'(t)|^2 \, dt,$$

where we used that $\|f\|^2 = 1$ implies $\|\hat{f}\|^2 = 2\pi$. Consequently, the uncertainty principle is equivalent to

$$\int t^2 |f(t)|^2 \, dt \int |f'(t)|^2 \, dt \geq \frac{1}{4}, \tag{1}$$

meaning it is also a statement about a trade-off between concentration and regularity. Namely, a function cannot simultaneously be very concentrated and have small derivatives (in the L^2 sense).

Although this may seem surprising, the uncertainty principle is nothing other than the Cauchy-Schwarz inequality, which states that for $g, h \in L^2(\mathbb{R})$,

$$\left| \int \overline{g(t)} h(t) \, dt \right| = |\langle g, h \rangle| \leq \|g\| \|h\| = \sqrt{\int |g(t)|^2 \, dt} \sqrt{\int |h(t)|^2 \, dt}.$$

Following Weyl's proof (see [2], p. 393), we assume that $|f(t)|^2$ decays faster than $1/t$ at infinity:

$$t|f(t)|^2 \to 0 \quad \text{as} \quad |t| \to \infty. \tag{2}$$

When $f' \in L^2(\mathbb{R})$, (2) holds; the verification is technical, and the interested reader can find it at the end of the proof. When $f' \notin L^2(\mathbb{R})$, we have $\sigma(\hat{f}) = \infty$, and so the uncertainty inequality certainly holds. Apply Cauchy-Schwarz to $g(t) = tf(t)$ and $h(t) = f'(t)$. Since $\|g\| = \sigma(f)$ and $\|h\| = \sigma(\hat{f})$, we have

$$\left| \int t\overline{f(t)} f'(t)\, dt \right| \leq \sigma(f)\sigma(\hat{f}).$$

We also have

$$\left| \int tf(t)\overline{f'(t)}\, dt \right| \leq \sigma(f)\sigma(\hat{f})$$

and so

$$\left| \int t\overline{f(t)} f'(t) + \int tf(t)\overline{f'(t)}\, dt \right| \leq 2\sigma(f)\sigma(\hat{f})$$

Now integration by parts gives

$$
\begin{aligned}
\int tf(t)\overline{f'(t)}\, dt + \int t\overline{f(t)} f'(t)\, dt &= \int t\frac{d}{dt}|f(t)|^2\, dt \\
&= t|f(t)|^2\big|_{-\infty}^{\infty} - \int |f(t)|^2\, dt \\
&= -1. \qquad \text{(from (2) and (ii))}
\end{aligned}
$$

We have thus shown

$$\sigma(f)\sigma(\hat{f}) \geq 1/2.$$

Equality holds in Cauchy-Schwarz if and only if $g \propto h$. So the uncertainty principle is satisfied with equality if and only if $tf(t) \propto f'(t)$. In this case, one can solve the differential equation

$$f'(t) = -\frac{t}{\sigma^2} f(t) \tag{3}$$

to see

$$f(t) = f(0)e^{-t^2/2\sigma^2}.$$

Since we allow translations, modulations, and rescalings (i.e. undoing our reductions from the beginning), in general we have

$$f(t) \propto e^{-i\omega_0 t} e^{-\frac{(t-t_0)^2}{2\sigma^2}}.$$

That is, f is a modulation and translation of a Gaussian function, as claimed. $\qquad\square$

Remark 1: Assume $f \in L^2(\mathbb{R})$ is differentiable, with $f' \in L^2(\mathbb{R})$. Verification of (2) proceeds as follows: Suppose against the claim that there exists $\epsilon > 0$ such that $|t||f(t)|^2 > \epsilon$ for arbitrarily large $|t|$. By possibly replacing $f(t)$ with $f(-t)$, we may assume this holds for arbitrarily large positive t. Since $f' \in L^2(\mathbb{R})$, we may choose t so large that

$$\int_t^\infty |f'(u)|^2\, du \leq \frac{1}{4}.$$

Applying Cauchy-Schwarz, we find that for any $t' \geq t$,

$$|f(t') - f(t)|^2 = \left| \int_t^{t'} f'(u)\, du \right|^2 \leq \left(\int_t^{t'} |f'(u)|\, du \right)^2 \leq (t' - t) \int_t^{t'} |f'(u)|^2\, du.$$

4

So for $t' \in [t, t + \epsilon/t]$,

$$|f(t') - f(t)|^2 \leq \frac{\epsilon}{t} \cdot \frac{1}{4}.$$

Hence

$$|f(t')| \geq |f(t)| - \frac{1}{2}\sqrt{\frac{\epsilon}{t}} > \frac{1}{2}\sqrt{\frac{\epsilon}{t}}$$

for such t'. It follows that

$$\int_t^\infty u^2 |f(u)|^2 \, du \geq \int_t^{t+\frac{\epsilon}{t}} u^2 |f(u)|^2 \, du > \int_t^{t+\frac{\epsilon}{t}} t^2 \cdot \frac{1}{4} \cdot \frac{\epsilon}{t} \, dt = \frac{\epsilon^2}{4}. \tag{4}$$

But $f \in L^2(\mathbb{R})$, and so

$$\lim_{t \to \infty} \int_t^\infty u^2 |f(u)|^2 \, du = 0.$$

Therefore, (4) is a contradiction.

Remark 2: If the RHS of (3) were given positive sign, $f(t)$ would be an exponential with positive exponent and thus not in $L^2(\mathbb{R})$.

Remark 3: We can verify directly that Gaussians satisfy the uncertainty principle with equality:

$$f(t) = e^{-t^2/2} \quad \implies \quad |f(t)|^2 = e^{-t^2} \quad \implies \quad \sigma^2(f) = \frac{1}{2},$$

and

$$\hat{f}(\omega) = \sqrt{2\pi} e^{-\omega^2/2} \quad \implies \quad |\hat{f}(\omega)|^2 = 2\pi e^{-\omega^2} \quad \implies \quad \sigma^2(\hat{f}) = \frac{1}{2}.$$

Indeed,

$$\sigma^2(f)\sigma^2(\hat{f}) = \frac{1}{4}.$$

3 Quantum mechanical interpretation

Let us consider the description of a simple physical system, such as an electron. From a classical point of view, the physical state of the electron at any instant is characterized by six quantities, that is, its three position coordinates, and its three momentum coordinates. In other words, to each time instant we associate a point in the **phase space**. Physical quantities of interest, such as the energy, can be measured from these quantities. However, in quantum mechanics the physical description is characterized by a **state vector**,

$$\textbf{State vector:} \quad |\psi\rangle \in \mathcal{H},$$

where $\mathcal{H}$ is an infinite-dimensional Hilbert space. The quantities of interest are represented by **observables**, which are symmetric operators acting on the Hilbert space

$$\textbf{Observable:} \quad A \in \mathcal{L}(\mathcal{H}), \quad A^* = A.$$

For instance, the **position** X and **momentum** P are observables. In classical mechanics, these only extract the position coordinates or momentum coordinates from a point in the phase space, but in quantum mechanics they are symmetric linear operators acting on $\mathcal{H}$.

Since they are symmetric, they can be diagonalized and have a complete set of eigenvectors and real eigenvalues. We can define

$$|x\rangle \quad \text{eigenvector of } X \text{ with eigenvalue } x, \text{ that is,} \quad X|x\rangle = x|x\rangle,$$

and

$$|p\rangle \quad \text{eigenvector of } P \text{ with eigenvalue } p, \text{ that is,} \quad P|p\rangle = p|p\rangle.$$

As usual, we let eigenvectors be normalized. Since $\{|x\rangle\}_x$ and $\{|p\rangle\}_p$ are both complete orthonormal sets, we can decompose the identity operator as

$$I = \int dx |x\rangle\langle x| = \int dp |p\rangle\langle p|.$$

Fundamental postulate of Quantum Mechanics: For an observable A with eigenstates $|a\rangle$ and a state $|\psi\rangle$, the outcome of measuring A is a with probability $|\langle a|\psi\rangle|^2$ and the system transitions to the state $|a\rangle$, that is,

$$|\psi\rangle \xrightarrow{\text{measure } A} |a\rangle \quad \text{with probability} \quad |\langle a|\psi\rangle|^2.$$

An experiment that shows this behavior is **Stern-Gerlach** experiment. For a description of this experiment and its significance, you can see `http://en.wikipedia.org/wiki/Stern-Gerlach_experiment`. The original paper by Otto Stern and Walther Gerlach (in German) can be found here `http://link.springer.com/article/10.1007%2FBF01326984`. Also the fundamental postulate is part of what is known as the **Copenhagen interpretation of quantum mechanics** (see `http://en.wikipedia.org/wiki/Copenhagen_interpretation#Principles`).

3.1 Uncertainty relations

Since the outcomes of measuring an obsevable are random variables, we can define **the expectation of an observable**. Let $|a\rangle$ denote the set of eigenvectors of the observable A and assume the system under consideration is on a state $|\psi\rangle$. Then its expected value is

$$\bar{a} = \langle A \rangle = \int da\, a |\langle a, \psi\rangle|^2$$

and its variance

$$\langle (\Delta A)^2 \rangle = \int da\, (a - \bar{a})^2 |\langle a, \psi\rangle|^2.$$

Heisenberg's uncertainty relation states that:

$$\langle (\Delta X)^2 \rangle \langle (\Delta P)^2 \rangle \geq (\hbar/2)^2 \tag{5}$$

with

$$\langle (\Delta X)^2 \rangle = \int dx\, (x - \bar{x})^2 |\langle x, \psi\rangle|^2$$
$$\langle (\Delta P)^2 \rangle = \int dp\, (p - \bar{p})^2 |\langle p, \psi\rangle|^2.$$

In other words, the product of standard deviations is at least $\hbar/2$, where $\hbar \approx 1.05 \, 10^{-34}$ is the reduced Planck constant $h/2\pi$.

As a remark, $\langle x, \psi \rangle$ is called the **the wave function** and often denoted by $\Psi(x)$. These are the coefficients of our state $|\psi\rangle$ in the basis of eigenvectors for X. Clearly $\int |\Psi(x)|^2 \, dx = 1$ and $|\Psi(x)|^2$ is usually interpreted as the probability finding the system (say, an electron) on an (infinitesimal) neighborhood of x.

To see why Heisenberg's uncertainty relation holds, we make use of this:

Claim 2. *If $|x\rangle$ is an eigenstate of X with eigenvalue x and, likewise, $|p\rangle$ is an eigenstate of P with eigenvalue p, then the dot product obeys*

$$\langle x, p \rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar}px}.$$

Define the wave function $\Phi(p)$ in momemtum space via $\Phi(p) = \langle p, \psi \rangle$. Since $|\psi\rangle = \int |x\rangle\langle x, \psi\rangle \, dx$, Φ and Ψ are related by

$$
\begin{aligned}
\underbrace{\Phi(p)}_{\text{wave function in momentum space}} &= \langle p | \psi \rangle \\[2em]
&= \int \langle x | \psi \rangle \langle p | x \rangle \, dx \\[1em]
&= \frac{1}{\sqrt{2\pi\hbar}} \int \underbrace{\Psi(x)}_{\text{wave function in position space}} e^{-\frac{i}{\hbar}px} \, dx.
\end{aligned}
$$

Similarly, we can write

$$
\begin{aligned}
\underbrace{\Psi(x)}_{\text{wave function in position space}} &= \langle x | \psi \rangle \\[2em]
&= \int \langle p | \psi \rangle \langle x | p \rangle \, dx \\[1em]
&= \frac{1}{\sqrt{2\pi\hbar}} \int \underbrace{\Phi(p)}_{\text{wave function in momentum space}} e^{\frac{i}{\hbar}px} \, dp.
\end{aligned}
$$

The Fourier transform arises naturally, and can be thought as **a change of variables between position space and momentum space**.

Hence, we can write uncertainties about position and momemtum as

$$\langle (\Delta X)^2 \rangle = \int (x - \bar{x})^2 |\Psi(x)|^2 \, dx$$

$$\langle (\Delta P)^2 \rangle = \int (p - \bar{p})^2 |\hat{\Psi}(p/\hbar)|^2 \, dp;$$

from here, Theorem 1 proved earlier gives Heisenberg's uncertainty relation (5).

3.2   Momemtum

We need to argue in favor of Claim 2, and in order to do this, we introduce the momemtum operator. We begin by considering a special class of maps acting on states, the **translation map** $U(a)$ for $a \in \mathbb{R}$ defined as

$$U(a)|x\rangle \mapsto |x + a\rangle.$$

This family of maps defines a group. We can consider its behavior for small δa, that is

$$U(\delta a) = I - \frac{i}{\hbar} P \delta a + o(\delta a^2).$$

The operator appearing on the first-order term of the above expansion is by definition **the momentum operator**. Roughly speaking, the momentum operator causes an infinitesimal displacement on a physical state. Informally,

$$U(\delta a)|\psi\rangle \quad \longrightarrow \quad \Psi(x) = \Psi(x) - \Psi'(x)\delta a + o(\delta a^2),$$

and consequently the momentum operator is such that $\Psi(x) \mapsto -i\hbar\frac{d}{dx}\Psi(x)$. Using our notation, its action on a state $|\psi\rangle$ is given by

$$P|\psi\rangle = -i\hbar \int \left(\frac{d}{dx'}\langle x'|\psi\rangle \right) |x'\rangle \, dx'.$$

With $|\psi\rangle = |p\rangle$ and using $P|p\rangle = p|p\rangle$, this gives

$$p|p\rangle = -i\hbar \int \left(\frac{d}{dx'}\langle x'|p\rangle \right) |x'\rangle \, dx'.$$

Taking the dot product with the bra $\langle x|$ gives

$$p\langle x|p\rangle = -i\hbar \int \left(\frac{d}{dx'}\langle x'|p\rangle \right) \langle x|x'\rangle \, dx' = -i\hbar\frac{d}{dx}\langle x|p\rangle.$$

This is a differential equation with solution

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar}px},$$

and consequently the eigenstates of the momentum represented in the position basis are complex exponentials with (spatial) frequency given by the momentum eigenvalue p.

3.3   Heisenberg's original formulation

In his 1927 paper [1], Heisenberg gave a slightly different formulation of all of this. First note that the mean outcome when applying A is given by[1]

$$\langle A \rangle = \langle \psi|A|\psi\rangle = \langle \psi|A \left(\int da|a\rangle\langle a| \right) |\psi\rangle = \int da\langle \psi|A|a\rangle\langle a|\psi\rangle = \int da\, a|\langle a, \psi\rangle|^2.$$

---

[1]In linear algebra, we would write $\langle \psi|A|\psi\rangle$ as $\psi^* A\psi$.

Set
$$\Delta A = A - \langle A \rangle I,$$
so that
$$\langle (\Delta A)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2,$$
is just the variance of the outcome as before. Last, the **commutator** or **Lie bracket** is
$$[A, B] = AB - BA,$$
which is a measure, in some sense, of how incoherent the observables A and B are.

Theorem 3 (Heisenberg (1927)). *For a state $|\psi\rangle$ and any observables A and B we have*
$$\langle (\Delta A)^2 \rangle \langle (\Delta B)^2 \rangle \geq \frac{1}{4} |\langle [A, B] \rangle|^2.$$

Proof. Not surprisingly, this is the Cauchy-Schwarz inequality in disguise. First, note that it suffices to assume that $\langle A \rangle = \langle B \rangle = 0$ (you do this part). Then setting $|u\rangle = A|\psi\rangle$ and $v = iB|\psi\rangle$, we have
$$\begin{aligned}
\langle (\Delta A)^2 \rangle \langle (\Delta B)^2 \rangle &= \langle \psi | A^2 | \psi \rangle \langle \psi | B^2 | \psi \rangle \\
&= \|u\|^2 \|v\|^2 \\
&\geq |\operatorname{Re}\{\langle u, v \rangle\}|^2 \\
&= \frac{1}{4} |\langle u, v \rangle + \langle v, u \rangle|^2 \\
&= \frac{1}{4} |\langle \psi | AB - BA | \psi \rangle|^2 \\
&= \frac{1}{4} |\langle [A, B] \rangle|^2.
\end{aligned}$$
$\square$

Claim 4. *We have*
$$[X, P] = i\hbar I.$$
Hence, $|\langle [X, P] \rangle|^2 = \hbar^2$, from where we obtain
$$\langle (\Delta X)^2 \rangle \langle (\Delta P)^2 \rangle \geq \frac{\hbar^2}{4}.$$

Proof. We compute the action of XP in the basis of eigenstates of X, and recall that $X|x\rangle = x|x\rangle$ and $P|\psi\rangle = -i\hbar \int \frac{d}{dx} \langle x, \psi \rangle |x\rangle \, dx$. On the one hand,
$$XP|\psi\rangle = -i\hbar \int x \frac{d}{dx} (\langle x, \psi \rangle) |x\rangle \, dx$$
while on the other
$$PX|\psi\rangle = -i\hbar \int \frac{d}{dx} (x\langle x, \psi \rangle) |x\rangle \, dx.$$
It thus follows that
$$[X, P]|\psi\rangle = i\hbar \int \left(-x\frac{d}{dx} \langle x, \psi \rangle + \frac{d}{dx} (x\langle x, \psi \rangle) \right) |x\rangle \, dx = i\hbar \int \langle x, \psi \rangle |x\rangle \, dx = i\hbar |\psi\rangle.$$
$\square$

References

[1] Werner Heisenberg. Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik. *Zeitschrift für Physik*, 43(3-4):172–198, 1927.

[2] Hermann Weyl. *The theory of groups and quantum mechanics*. Courier Corporation, 1950.