

Notes from the Methods of Mathematical Physics II lectures of 2013-02-19/05-30 by Prof. Dr. Eugene Trubowitz

Robin Leroy and David Nadlinger

2013-04-18

1 *True Lies: An introduction to distribution theory*

In many areas of physics, the Dirac delta “function” is an important tool, e.g., for describing localized phenomena. While it is usually described in vague terms as “only making sense under an integral”, physicists tend to still write identities such as

LECTURE 4, 2013-02-28.

$$\Delta \frac{1}{|\mathbf{x}|} = -4\pi \delta(\mathbf{x}),$$

and rely on their intuition and previous mistakes to distinguish situations in which they can safely apply this calculus from the ones where that would lead to incorrect results.

How can we formalize this? Note that the operation of taking the integral of the Dirac delta multiplied with another function φ can be considered as a map from that function to a scalar in \mathbb{C} . Thus, we could consider

$$\delta_{\mathbf{x}}(\varphi) = \int_{\mathbb{R}^n} \delta(\mathbf{x} - \mathbf{y}) \varphi(\mathbf{y}) \, d\mathbf{y}$$

to be a functional, and more specifically an element of the dual space of whatever class of functions we can allow for φ . But the question is: what restrictions do we need on this space of functions if we want to find objects like the Dirac delta in its dual space?

Actually, we want these functionals to be in the continuous dual of our space, that is, we want them to be functionals U such that $U(\varphi_k) \rightarrow U(\varphi)$ for all sequence of functions $(\varphi_k)_{k \in \mathbb{N}}$ that converge to φ in an appropriate norm¹ — we want our space to be a Banach space. Intuition tells us that the smaller, i.e., more restricted, a space of functions is, the stronger the conditions on $\varphi_k \rightarrow \varphi$ must be in order for it to be complete, and so the bigger its continuous dual can be, as we only need $U(\varphi_k) \rightarrow U(\varphi)$ for the sequences $(\varphi_k)_{k \in \mathbb{N}}$ that converge in that space. Thus, we will investigate a space of functions $\varphi(\mathbf{x}) \in C^\infty(\mathbb{R}^n)$ that vanish quickly as $|\mathbf{x}| \rightarrow \infty$ (“converge with a vengeance”), hoping to find the Dirac delta in its continuous dual.

Definition. First, a few definitions used throughout this course:

$$C^\infty(\mathbb{R}^n) := \{\varphi : \mathbb{R}^n \rightarrow \mathbb{C} \mid \forall \alpha \in \mathbb{N}^n, \partial^\alpha \varphi \in C^0(\mathbb{R}^n)\}$$

is the space of *smooth functions* from \mathbb{R}^n to \mathbb{C} , where

$$\mathbb{N} := \{0, 1, 2, \dots\}, \quad \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n \text{ is a } \textit{multi-index},$$

$$|\alpha| := \alpha_1 + \dots + \alpha_n, \quad \partial_{\mathbf{x}}^\alpha := \frac{\partial^{|\alpha|}}{\partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n}} \quad (\text{we use } \partial^\alpha \text{ when the variable is implicit}),$$

$$\text{and } \mathbf{x}^\beta := x_1^{\beta_1} \dots x_n^{\beta_n} \text{ for } \mathbf{x} \in \mathbb{R}^n, \beta \in \mathbb{N}^n.$$

¹This simply means that U is a continuous map from the function space endowed with its norm topology to \mathbb{R} endowed with the standard topology.

Definition (Schwartz space). Now we define the *Schwartz space*

$$\mathcal{S}(\mathbb{R}^n) := \{\varphi \in C^\infty(\mathbb{R}^n) \mid \forall \alpha, \beta \in \mathbb{N}^n, \|\varphi\|_{\alpha, \beta} < \infty\}, \text{ where } \|\varphi\|_{\alpha, \beta} := \sup_{x \in \mathbb{R}^n} |x^\alpha \partial_x^\beta \varphi(x)|. \quad \diamond$$

Note that $\varphi \in \mathcal{S}(\mathbb{R}^n)$ implies $\lim_{|x| \rightarrow \infty} (1 + |x|^2)^k \partial_x^\alpha \varphi(x) = 0$, as

$$\begin{aligned} (1 + |x|^2)^m \partial_x^\alpha \varphi(x) &= \sum_{j=0}^m \binom{m}{j} |x|^{2j} \partial_x^\alpha \varphi(x) \\ &= \sum_{j=0}^m \binom{m}{j} \left(\sum_{i=1}^n x_i^2 \right)^j \partial_x^\alpha \varphi(x) \\ &= \sum_{j=0}^m \binom{m}{j} \sum_{|\alpha|=j} \left(\prod_{i=1}^n x_i^{2\alpha_i} \right) \partial_x^\alpha \varphi(x) \end{aligned}$$

and $\sup_{x \in \mathbb{R}^n} |x^{2\alpha} \varphi(x)| < \infty$. It immediately follows:

Lemma. If $\varphi \in \mathcal{S}(\mathbb{R}^n)$, then

$$\forall k \geq 0 \quad \forall \alpha \in \mathbb{N}^n \quad \exists C_\alpha \in \mathbb{R} \quad |\partial_x^\alpha \varphi(x)| < \frac{C_\alpha}{(1 + |x|^2)^k}. \quad \square$$

Definition (Convergence in $\mathcal{S}(\mathbb{R}^n)$). The sequence $(\varphi_k)_{k \in \mathbb{N}} \in (\mathcal{S}(\mathbb{R}^n))^{\mathbb{N}}$ converges to $\varphi \in \mathcal{S}(\mathbb{R}^n)$ if and only if

$$\forall \alpha, \beta \in \mathbb{N}^n \quad \lim_{k \rightarrow \infty} \|\varphi_k - \varphi\|_{\alpha, \beta} = 0.$$

We use the notation $\varphi_k \xrightarrow{\mathcal{S}(\mathbb{R}^n)} \varphi$. \diamond

Definition (Tempered distributions). Note that $\mathcal{S}(\mathbb{R}^n)$ is a complex vector space. We define that U is an element of its continuous dual $\mathcal{S}'(\mathbb{R}^n)$, the space of *tempered distributions*, if and only if:

1. U is a linear map from $\mathcal{S}(\mathbb{R}^n)$ to \mathbb{C} , i.e., $U \in \mathcal{S}'(\mathbb{R}^n)$
2. U is continuous on $\mathcal{S}(\mathbb{R}^n)$, i.e. if $\varphi_k \xrightarrow{\mathcal{S}(\mathbb{R}^n)} \varphi$ then $\lim_{k \rightarrow \infty} U(\varphi_k) = U(\varphi)$. \diamond

Let us now consider some examples. First, define the space of locally integrable functions as follows.

Definition (Locally integrable functions).

$$L_{\text{loc}}^1(\mathbb{R}^n) := \{f : \mathbb{R}^n \rightarrow \mathbb{C} \mid \forall x \in \mathbb{R}^n, \exists \epsilon > 0, \int_{B_\epsilon x} |f(y)| \, dy < \infty\} \quad \diamond$$

Now let $f \in L_{\text{loc}}^1(\mathbb{R}^n)$. If additionally $\sup_{x \in \mathbb{R}^n} (1 + |x|^2)^{-s} |f(x)| < \infty$ for some $s > 0$ (e.g. $f(x) \sum_{|\beta| \leq s} c_\beta x^\beta$), then

$$U_f(\varphi) := \int_{\mathbb{R}^n} f(x) \varphi(x) \, dx$$

is a tempered distribution.

Definition (Dirac delta distribution). Another example is the *Dirac delta distribution* mentioned above, with

$$\delta(\varphi) := \varphi(0), \quad \delta_x(\varphi) := \varphi(x). \quad \diamond$$

Distributions have the nice property that we can always differentiate them.

Definition (Derivative of a distribution). Let $U \in \mathcal{S}'(\mathbb{R}^n)$. Then $\partial^\alpha U$ is the distribution defined by

$$(\partial^\alpha U)(\varphi) := (-1)^{|\alpha|} U(\partial^\alpha \varphi) \quad \text{for } \varphi \in \mathcal{S}(\mathbb{R}^n). \quad \diamond$$

This is natural, as

$$U_{\partial^\alpha f}(\varphi) = \int_{\mathbb{R}^n} (\partial^\alpha f(x)) \varphi(x) \, dx = (-1)^{|\alpha|} \int_{\mathbb{R}^n} f(x) \partial^\alpha \varphi(x) \, dx$$

by integrating by parts. Let's find the derivative of $|\cdot|$ in the distribution world:

$$\begin{aligned} \text{Abs}(\varphi) &:= \int_{-\infty}^{\infty} |x| \varphi(x) \, dx \\ \left(\frac{d}{dx} \text{Abs} \right)(\varphi) &= -\text{Abs} \frac{d\varphi}{dx} \\ &= -\int_0^{\infty} x \varphi'(x) \, dx - \int_{-\infty}^0 (-x) \varphi'(x) \, dx \\ &= -[x\varphi(x)]_0^{\infty} + \int_0^{\infty} \varphi(x) \, dx + [x\varphi(x)]_{-\infty}^0 - \int_{-\infty}^0 \varphi(x) \, dx \\ &= \int_{-\infty}^{\infty} \text{sgn}(x) \varphi(x) \, dx = U_{\text{sgn}}(\varphi), \\ \text{where } \text{sgn}(x) &:= \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases} \quad \text{is the sign function.} \end{aligned}$$

We define $\text{Sgn}(\varphi) := \int_{-\infty}^{\infty} \text{sgn}(x) \varphi(x) \, dx$ and write $\frac{d}{dx} \text{Abs} = \text{Sgn}$. This corresponds to $\frac{d}{dx} |x| = \text{sgn}(x)$ in physicist-speak.

What about $\Delta \frac{1}{|x|}$? By defining the *Newtonian distribution*

$$N(\varphi) := \int_{\mathbb{R}^3} \frac{1}{|x|} \varphi(x) \, dx,$$

we get

$$\Delta N(\varphi) = \sum_{j=1}^3 \frac{\partial^2}{\partial x_j^2} N(\varphi) = \sum_{j=1}^3 N\left(\frac{\partial^2 \varphi}{\partial x_j^2}\right) = \dots = -4\pi \varphi(0),$$

in other words,

$$\Delta N = -4\pi \delta.$$

Definition (Inner product on $\mathcal{S}(\mathbb{R}^n)$). Let $\varphi, \psi \in \mathcal{S}(\mathbb{R}^n)$. We define:

LECTURE 5, 2013-03-05.

$$\langle \psi, \varphi \rangle := \int_{\mathbb{R}^n} \bar{\psi}(x) \varphi(x) \, dx. \quad \diamond$$

2 General properties of the Fourier transform

Definition (Fourier Transform). For $\varphi \in \mathcal{S}(\mathbb{R}^n)$, we define the *Fourier transform* $\widehat{\varphi}$ of φ , also written $\mathcal{F}(\varphi)$, as:

$$\mathcal{F}(\varphi)(\mathbf{k}) := \widehat{\varphi}(\mathbf{k}) := \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(\mathbf{x}) e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x}. \quad \diamond$$

LECTURE 6, 2013-03-07. **Lemma** (“Your life in Fourier land depends on it.”).

$$\begin{aligned} \widehat{\varphi + \psi} &= \widehat{\varphi} + \widehat{\psi}, \\ \widehat{\lambda \varphi} &= \lambda \widehat{\varphi} \end{aligned} \quad (2.1)$$

$$|\widehat{\varphi}(\mathbf{k})| \leq \frac{1}{(2\pi)^{\frac{n}{2}}} \|\varphi\|_1 < \infty, \quad \|\varphi\|_1 := \int_{\mathbb{R}^n} |\varphi(\mathbf{x})| d\mathbf{x} \quad (2.2)$$

$$\varphi(\widehat{\lambda \cdot})(\mathbf{k}) = \frac{1}{|\lambda|^n} \widehat{\varphi}\left(\frac{\mathbf{k}}{\lambda}\right), \quad \lambda \neq 0 \quad (2.3)$$

$$\widehat{T_{\mathbf{y}}\varphi}(\mathbf{x}) = e^{i\mathbf{k} \cdot \mathbf{y}} \widehat{\varphi}(\mathbf{k}), \quad (T_{\mathbf{y}}\varphi)(\mathbf{x}) := \varphi(\mathbf{x} + \mathbf{y}) \quad (2.4)$$

$$\frac{\partial \widehat{\varphi}}{\partial x_j}(\mathbf{k}) = i k_j \widehat{\varphi}(\mathbf{k}), \quad (2.5)$$

$$\widehat{x_j \varphi}(\mathbf{k}) = i \frac{\partial \widehat{\varphi}}{\partial k_j}(\mathbf{k})$$

$$\int_{\mathbb{R}^n} \widehat{\varphi}(\mathbf{y}) \psi(\mathbf{y}) d\mathbf{y} = \int_{\mathbb{R}^n} \varphi(\mathbf{y}) \widehat{\psi}(\mathbf{y}) d\mathbf{y} \quad (2.6)$$

Proof. Proof of (2.2):

$$\begin{aligned} |\widehat{\varphi}(\mathbf{k})| &= \left| \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(\mathbf{x}) e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x} \right| \\ &\leq \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} |\varphi(\mathbf{x}) e^{-i\mathbf{k} \cdot \mathbf{x}}| d\mathbf{x} = \frac{1}{(2\pi)^{\frac{n}{2}}} \underbrace{\int_{\mathbb{R}^n} |\varphi(\mathbf{x})| \underbrace{|e^{-i\mathbf{k} \cdot \mathbf{x}}|}_1 d\mathbf{x}}_{\|\varphi\|_1} \end{aligned}$$

Proof of existence of the integral $\|\varphi\|_1$:

$$\begin{aligned} \int_{\mathbb{R}^n} |\varphi(\mathbf{x})| d\mathbf{x} &= \int_{\mathbb{R}^n} (1 + |\mathbf{x}|^2)^{\frac{s}{2}} |\varphi(\mathbf{x})| \frac{d\mathbf{x}}{(1 + |\mathbf{x}|^2)^{\frac{s}{2}}} \\ &\leq \|\varphi\|_{s,0} \int_{\mathbb{R}^n} \frac{1}{(1 + |\mathbf{x}|^2)^{\frac{s}{2}}} d\mathbf{x} \\ &= \|\varphi\|_{s,0} \int_0^\infty \frac{r^{n-1}}{(1 + r^2)^{\frac{s}{2}}} dr < \infty \quad \text{for } s \geq n + 1. \end{aligned}$$

“(2.1) I will not do!” Proof of (2.3):

$$\begin{aligned}
 \widehat{\varphi(\lambda \cdot)} &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(\lambda x) e^{-ik \cdot x} dx \\
 &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(y) e^{-i \frac{1}{\lambda} k \cdot y} d\left(\frac{y}{\lambda}\right) \quad \text{where } x = \frac{1}{\lambda} y. \\
 &= \frac{1}{|\lambda|^n} \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(y) e^{-i \frac{k}{\lambda} \cdot y} dy \\
 &\quad \underbrace{\hspace{10em}}_{\widehat{\varphi}\left(\frac{k}{\lambda}\right)}
 \end{aligned}$$

Proof of the second part of (2.5):

$$\begin{aligned}
 \widehat{x_j \varphi}(x) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} x_j \varphi(x) e^{-ik \cdot x} dx \\
 &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(x) i \frac{\partial}{\partial k_j} e^{-ik \cdot x} dx \quad \text{as } \frac{\partial}{\partial k_j} e^{-ik \cdot x} = \frac{\partial}{\partial k_j} e^{-i \sum_{r=0}^n k_r x_r} = -i x_j e^{-ik \cdot x} \\
 &= i \frac{\partial}{\partial k_j} \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(x) e^{-ik \cdot x} dx \\
 &= i \frac{\partial \widehat{\varphi}}{\partial k_j}(k)
 \end{aligned}$$

The professor starts whistling some elevator music while waiting for a student to write down the proof before he can wipe the board. When the student is done, the professor notices that there still is some room left on the blackboard and starts writing the rest there, without erasing anything.

Proof of (2.6):

$$\begin{aligned}
 \int_{\mathbb{R}^n} \widehat{\varphi}(y) \psi(y) dy &= \int_{\mathbb{R}^n} \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(x) e^{-iy \cdot x} dx \psi(y) dy \\
 &= \int_{\mathbb{R}^n} \varphi(x) \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix \cdot y} \psi(y) dy dx \quad \text{As } \int_{\mathbb{R}^{2n}} |\varphi(x) e^{-iy \cdot x} \psi(y)| dx dy < \infty, \text{ we} \\
 &\quad \text{can use Fubini's theorem.} \\
 &= \int_{\mathbb{R}^n} \varphi(y) \widehat{\psi}(y) dy \quad \square
 \end{aligned}$$

Proposition. The Fourier transform $\widehat{\cdot}$ is a continuous bijective linear map from the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ to itself. $\widehat{\widehat{\varphi}} = \varphi$, where $\widehat{\psi}(x) := (2\pi)^{-n/2} \int_{\mathbb{R}^n} \psi(k) e^{ik \cdot x} dx$, read “unhat”, “bird”, “seagull”, or whatever you like.

Proof.

$$\begin{aligned}
 \varphi \in \mathcal{S}(\mathbb{R}^n) &\Rightarrow x^\alpha \partial_x^\beta \varphi \in \mathcal{S}(\mathbb{R}^n) \\
 \mathcal{F}(x^\alpha \partial_x^\beta \varphi)(k) &= i^{|\alpha|} \partial_k^\alpha (\partial_x^\beta \widehat{\varphi})(k) \quad \text{from the second equality} \\
 &= i^{|\alpha|} \partial_k^\alpha (i^{|\beta|} k^\beta \widehat{\varphi})(k) \quad \text{in (2.5), applied } |\alpha| \text{ times.} \\
 &= i^{|\alpha|+|\beta|} \partial_k^\alpha (k^\beta \widehat{\varphi})(k)
 \end{aligned}$$

“I should have done it the other way around. [...] Let’s try it the other way around.”
We want:

$$\forall \alpha, \beta, \sup_{k \in \mathbb{R}^n} |k^\alpha \partial_k^\beta \widehat{\varphi}(k)| < \infty$$

As that implies $\widehat{\varphi} \in \mathcal{S}(\mathbb{R}^n)$.

$$\mathcal{F}(\partial_x^\alpha (k^\beta \varphi))(k) = i^{|\alpha|+|\beta|} k^\alpha \partial_k^\beta \widehat{\varphi}(k)$$

by applying the first part
of (2.5) $|\alpha|$ times and the
second part $|\beta|$ times.

$$\forall \alpha, \beta, |k^\alpha \partial_k^\beta \widehat{\varphi}(k)| \leq \frac{1}{(2\pi)^{\frac{n}{2}}} \|\partial_x^\beta (x^\alpha \varphi)\|_1 < \infty \quad \text{from (2.2).}$$

We now know $\widehat{\varphi} \in \mathcal{S}(\mathbb{R}^n)$. Let us prove $\widetilde{\varphi} = \varphi$. “When you do the wrong thing I’m gonna scream loudly.”

$$\begin{aligned} \widetilde{\varphi}(x) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \widehat{\varphi}(k) e^{ik \cdot x} dk \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \varphi(y) e^{-ik \cdot y} dy e^{ik \cdot x} dk \end{aligned}$$

We don’t interchange the integrals here because that would lead to ugly calculations. “If you paint the walls before you start building, it’s not a good idea.” At this point, somebody suggests replacing $(2\pi)^{-n} \int_{\mathbb{R}^n} e^{ik \cdot (x-y)} dk$ by $\delta(x-y)$. “It’s plausible! — WHY? [...] Ah I said it, so it’s plausible.” However, we would need to do some nasty calculations in order to do this. “How are we going to do it so fast that you don’t get bored, and yet in enough detail that he’s convinced? Be sneaky.”

$$\begin{aligned} \widetilde{\varphi} &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \widehat{\varphi}(k) e^{ik \cdot x} \underbrace{\lim_{\epsilon \downarrow 0} e^{-\epsilon \frac{|k|^2}{2}}}_{\substack{1 \text{ written in} \\ \text{some other way}}} dk \\ &= \lim_{\epsilon \downarrow 0} \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \widehat{\varphi}(k) e^{ik \cdot x} e^{-\epsilon \frac{|k|^2}{2}} dk \end{aligned}$$

“Don’t worry.” This is actually a one-liner using Lebesgue’s dominated convergence theorem.

The 10-minute break ends with the loud noise of a metallic pointing stick hitting the desk.

$$\begin{aligned} \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \widehat{\varphi}(k) e^{ik \cdot x} e^{-\epsilon \frac{|k|^2}{2}} dk &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(y) e^{-iy \cdot k} dy e^{ik \cdot x} e^{-\epsilon \frac{|k|^2}{2}} dk \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(y) \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{ik \cdot (x-y)} e^{-\epsilon \frac{|k|^2}{2}} dk dy \end{aligned}$$

Recall: “every path leads to Rome and every Gaussian is in $\mathcal{S}(\mathbb{R}^n)$.” Also, the last foot of a dactylic hexameter is always a spondee.

$$\frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\epsilon \frac{|k|^2}{2}} e^{i(y-x) \cdot k} dk = \frac{e^{-\frac{1}{2\epsilon} |x-y|^2}}{\epsilon^{\frac{n}{2}}}$$

We therefore get:

$$\begin{aligned}
 \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \widehat{\varphi}(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}} e^{-\epsilon \frac{|\mathbf{k}|^2}{2}} d\mathbf{k} &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(\mathbf{y}) \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{i\mathbf{k} \cdot (\mathbf{x} - \mathbf{y})} e^{-\epsilon \frac{|\mathbf{k}|^2}{2}} d\mathbf{k} d\mathbf{y} \\
 &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(\mathbf{y}) \frac{e^{-\frac{1}{2\epsilon} |\mathbf{x} - \mathbf{y}|^2}}{\epsilon^{\frac{n}{2}}} d\mathbf{y} \\
 &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(\mathbf{y}\sqrt{\epsilon} + \mathbf{x}) \frac{e^{-\frac{1}{2\epsilon} |\mathbf{y}\sqrt{\epsilon}|^2}}{\epsilon^{\frac{n}{2}}} d(\mathbf{y}\sqrt{\epsilon}) \quad \text{“Epsilons everywhere!”} \\
 &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(\mathbf{x} + \mathbf{y}\sqrt{\epsilon}) e^{-\frac{1}{2} |\mathbf{y}|^2} d\mathbf{y} \quad \text{“Now we can paint.”} \\
 &\xrightarrow{\epsilon \downarrow 0} \frac{1}{(2\pi)^{\frac{n}{2}}} \varphi(\mathbf{x}) \int_{\mathbb{R}^n} e^{-\frac{1}{2} |\mathbf{y}|^2} d\mathbf{y} = \varphi(\mathbf{x}). \quad \square
 \end{aligned}$$

3 Seeking eigenfunctions of the Fourier transform

“The main ideas of many things are right here.” We seek to find the solutions (φ, λ) of

$$\widehat{\varphi} = \lambda \varphi.$$

We shall find all of them. “Not one is going to get away.” Actually, we have already got one:

$$e^{-\frac{1}{2} |\mathbf{x}|^2} = e^{-\frac{1}{2} |\mathbf{x}|^2},$$

so we know that 1 is an eigenvalue. How do we find the others? “Here you actually have to have an idea.”

Idea Find an $H \in \text{End}(\mathcal{S}(\mathbb{R}^n))$ such that $[H, \widehat{\cdot}] = 0$. Then they have the same eigenspaces, so we just need to find the eigenfunctions of H (see *Finite Dimensional Quantum Mechanics*).

Let $n = 1$.

$$\left. \begin{aligned} \widehat{(x^2 \varphi)}(k) &= -\frac{d^2}{dk^2} \widehat{\varphi}(k) \\ \widehat{\left(\frac{d^2}{dx^2} \varphi\right)}(k) &= -k^2 \widehat{\varphi}(k) \end{aligned} \right\} \quad \begin{array}{l} \text{From the lemma your life} \\ \text{depends on.} \end{array}$$

Subtracting the second equality from the first one above:

$$\mathcal{F}\left(\left(-\frac{d^2}{dx^2} + x^2\right)\varphi\right)(k) = \left(-\frac{d^2}{dk^2} + k^2\right)\mathcal{F}\varphi(k).$$

We smell a harmonic potential. “Don’t get your fingers too close, they sometimes bite.”

Let $H := \frac{1}{2}\left(-\frac{d^2}{dx^2} + x^2 - 1\right)$. Then $[H, \widehat{\cdot}] = 0$. We multiplied by $\frac{1}{2}$ so that it looks even more like a harmonic potential (it is the Hamiltonian of the quantum harmonic oscillator). Why did we add -1 ? Define

$$\begin{aligned} A &:= \frac{1}{\sqrt{2}} \left(\frac{d}{dx} + x \right), \\ A^\dagger &:= \frac{1}{\sqrt{2}} \left(-\frac{d}{dx} + x \right). \end{aligned}$$

We then have $H = A^\dagger A$. If a and b commute, then we have $(a - b)(a + b) = a^2 - b^2$, but here x and $\frac{d}{dx}$ don't quite commute, hence the -1 .

LECTURE 7, 2013-03-12.

Let us look at these operators in more detail: define $(Qf)(x) := xf(x) \in \mathcal{S}(\mathbb{R})$, $(Pf)(x) := \frac{d}{dx}f(x) \in \mathcal{S}(\mathbb{R})$ for $f \in \mathcal{S}(\mathbb{R})$. We saw that $[Q, P] = 1$. We also have:

$$\begin{aligned} \langle \varphi, P\psi \rangle &= \int_{\mathbb{R}} \overline{\varphi}(x) \frac{d}{dx} \psi(x) dx \\ &= [\overline{\varphi}\psi]_{-\infty}^{\infty} - \int_{\mathbb{R}} \frac{d}{dx} \varphi(x) \psi(x) dx \\ &= -\langle P\varphi, \psi \rangle, \end{aligned} \quad \begin{aligned} [\overline{\varphi}\psi]_{-\infty}^{\infty} &= 0 \text{ as } \varphi \in \mathcal{S}(\mathbb{R}), \\ \psi &\in \mathcal{S}(\mathbb{R}), \text{ and therefore} \\ \overline{\varphi}\psi &\in \mathcal{S}(\mathbb{R}). \end{aligned}$$

so P is skew symmetric. What is better than skew symmetric? Self adjoint is better. How do we make a skew symmetric operator self adjoint? We add an i . Namely, redefine $(Pf)(x) := \frac{1}{i} \frac{d}{dx}f(x)$, we then have $\langle \varphi, P\psi \rangle = \langle P\varphi, \psi \rangle$. Q is obviously self adjoint. We now have $[Q, P] = i\mathbb{1}$. Multiplying by some $\hbar \in \mathbb{R}$, redefine $(Pf)(x) := \frac{\hbar}{i} \frac{d}{dx}f(x)$, we get $[Q, P] = \hbar i\mathbb{1}$.

If we let $\hbar := \hbar$, the conditions verified by P and Q are the conditions for infinite matrices \mathbf{P} and \mathbf{Q} in the “catechism of matrix mechanics” in *Finite Dimensional Quantum Mechanics*. “Being something a little more, and something a lot less than cockroaches, we are curious.” What we have been doing here is not the ideal way of doing this: we have to pick a vector space, in this case $\mathcal{S}(\mathbb{R}^n)$, last year the space of infinite matrices. As we learned in linear algebra, we gain a deeper understanding if we do not use coordinate systems. “We are going to study the pure essence” of this. We are going to replace that by an abstract group, which lives in Plato's cave, and when we need to calculate, we will look at a representation. “The group is as close to God as you can get.”

Definition (Heisenberg group). We define the *Heisenberg group*, introduced by Hermann Weyl.

$$\left(H := \left\{ \begin{pmatrix} 1 & r & t \\ 0 & 1 & s \\ 0 & 0 & 1 \end{pmatrix} \mid r, s, t \in \mathbb{R} \right\}, \text{ordinary matrix multiplication} \right) \quad \diamond$$

You know that if you take e^{xy} on reals numbers, you get e^{x+y} , but that doesn't work for matrices. Hermann Weyl's idea is the following: we look at the identities

$$\begin{aligned} (e^{itP}f)(x) &= f(x+t) & \text{as } (e^{it(\frac{1}{i}\frac{d}{dx})})f(x) &\stackrel{\text{formally}}{=} \sum_{l=0}^{\infty} \frac{t^l}{l!} \frac{d^l}{dx^l} f(x), \\ & & \text{the Taylor series of } f(x+t) &\text{at } x. \\ e^{itP}e^{isQ} &= e^{ist}e^{itP} \end{aligned}$$

Write $X := e^{iQ}$ and $Y := e^{iP}$, $X^r Y^s = e^{irs} Y^s X^r$, $Z := (e^i)^t \mathbb{1}$. We will generate a group with these three elements. “I'll think up another way of doing this; it's something worth doing twice.” How do we preserve all the ideas, and yet get rid of all the infinite dimensional vector spaces? We study a finite dimensional analog, the finite Heisenberg group H_n , in which we use $\mathbb{Z}/n\mathbb{Z}$ instead of \mathbb{R} . And we'll use that little group to do all sorts of experiments. No one will complain about what we do to it.

Let us start again. “Those of you who have done the exercises, and those of you who have the determination to go to Prof. Willwacher's — ‘Willwacher.’ I like that name. — lecture, we are going to do that again, but on a concrete example.” We will find all the conjugacy classes. “We're going to hunt them down!”

“I'll see you on Thursday, and I expect there'll be fewer people.”

4 The finite Heisenberg group

The professor whistles “Singin’ in the Rain” while waiting for the lecture to start.

In its soul, the finite Heisenberg group has captured the essence of calculating something in quantum mechanics. The point of abstraction is to remove all that gets in the way of communing with the thing in and of itself. “We’re gonna go deep into Plato’s cave today.”

“You write commutators to see the level of how much it doesn’t commute (*sic*).”

“Conjugacy is where it’s at, in the language of the 1960s.” This is what the entire linear algebra lecture is about. For instance, the Jordan normal form gives you a unique representative of a conjugacy class. “You can dig down here and spend years in this hole. Or you can go up. I’ll go up, I’ll let you dig down.”

$X := [1, 0; 0]$, $Y := [0, 1; 0]$, $Z := [0, 0; 1]$ generate the finite Heisenberg group. With $Z = XYX^{-1}Y^{-1}$, X, Y generate it. “Jacobi? — no, no, no. No Jacobi identity.”

The free group on one element is isomorphic to $(\mathbb{Z}, +)$. The normal subgroup of F_X generated by W is the smallest normal subgroup which contains W . This is well-defined, as F_X is such a normal subgroup, and the intersection of normal subgroups is a normal subgroup, so the normal subgroup generated by W is

$$\bigcap_{\substack{N \trianglelefteq F_X \\ W \subseteq N}} N.$$

Note that this is also equal to the orbit of W under the action of G on its subsets by conjugation, ${}^g S := gSg^{-1}$, so we can write

$$\langle X|W \rangle = F_X / {}^G W = F_X / \bigcap_{\substack{N \trianglelefteq F_X \\ W \subseteq N}} N.$$

“Out of this trivial stuff comes something deep.”

“A faithful representation of a group is just changing variables.” ρ is a faithful complex irreducible representation of the finite group G if and only if $\sum_{g \in G} |\chi_\rho(g)|^2 = |G|$.

“I will see those of you who actually do come back next Tuesday.”

The dual group only works for abelian groups. The way a bourgeois mathematician thinks of the classical phase space $\mathbb{R}^n \times \widehat{\mathbb{R}^n}$ is as the cotangent bundle $\mathbb{R}^n \times (\mathbb{R}^n)^*$. “If you would have the energy to learn ancient Greek, and you talked about the real housewives of Atlanta, it would be a waste.” It is not sufficient to have a fancy language, you actually have to say something.

We started this lecture with the finite Fourier transform. This is a special case of what we discuss here, with $\mathbb{Z}/n\mathbb{Z}$ as the abelian group A . This is perfect pedagogy: after only eight months, we come back to the beginning. Everything I’ve done so far is just that for various groups. When the group is infinite, e.g., the unit circle for Fourier series, you have to do analysis. “[Claude] doesn’t go there. I do.”

4.1 SUPPLEMENTAL: *Η φανταχτερή γλῶσσα τῶν ἀστικῶν μαθηματικῶν*²

Assume the space G of generalised coordinates of a physical system is a Lie group, i.e., a group which is also a differentiable manifold, and in which composition and inversion are smooth maps. Then if $q : \mathbb{R} \rightarrow G$ describes the evolution of the generalised coordinates of the system with time, the generalised velocity at time t , $\dot{q}(t)$, lies in the tangent space $T_{q(t)} G$ — recall that the tangent space of g at x is defined as the set of derivatives³ $\dot{\gamma}(0)$ at 0 of smooth curves $\gamma : \mathbb{R} \rightarrow G$ with $\gamma(0) = x$.

²The fancy language of bourgeois mathematicians.

³One has to be careful here: unless G is a submanifold of \mathbb{R}^n , the derivative is not well-defined. In that case however, $\dot{\gamma}(0)$ can be defined as the equivalence class of γ in the set of curves through x with $\gamma(0) = x$ under the relation $\gamma_1 \sim \gamma_2 \Leftrightarrow (f \circ \gamma_1)'(0) = (f \circ \gamma_2)'(0)$, where f is a differentiable chart of a neighbourhood of x onto \mathbb{R}^n . It can then be shown that this is independent of f .

LECTURE 8, 2013-03-14.

There are official notes for the material covered in this lecture.

LECTURE 9, 2013-03-19.

There are official notes for the material covered in this lecture.

The ordered pair $(\mathbf{q}(t), \dot{\mathbf{q}}(t))$ lies in the *tangent bundle* TG , that is,

$$(\mathbf{q}(t), \dot{\mathbf{q}}(t)) \in TG := \bigcup_{\mathbf{x} \in G} \{\mathbf{x}\} \times T_{\mathbf{x}}G.$$

Informally, TG is the disjoint union of all the tangent spaces of the manifold G . An element $(\mathbf{x}, \mathbf{v}) \in TG$ represents a tangent vector \mathbf{v} to G at \mathbf{x} . Physically, this is a state of the system, namely the one with coordinates \mathbf{x} and velocity \mathbf{v} .

As G is a Lie group, composition on the left — or right — by any element is a diffeomorphism. In particular, for $q \in G$, $\varphi_q : G \rightarrow G, \mathbf{x} \mapsto q^{-1}\mathbf{x}$ is a diffeomorphism, and so it induces a canonical linear isomorphism $d\varphi_q(\mathbf{q})$ between the tangent spaces T_qG and $T_{\varphi_q(q)} = T_{\mathbb{1}}G$, by

$$d\varphi_q(\mathbf{q})\dot{\mathbf{r}}(0) = d\varphi_q(\mathbf{r}(0))\dot{\mathbf{r}}(0) = (\varphi_q \circ \mathbf{r})'(0) = \dot{\mathbf{r}}_0(0),$$

where $\dot{\mathbf{r}}(0)$ is an element of T_qG , so $\mathbf{r}(0) = \mathbf{q}$ and $\mathbf{r}_0 := \varphi_q \circ \mathbf{r}$ verifies $\mathbf{r}_0(0) = \mathbb{1}$, therefore $\dot{\mathbf{r}}_0(0) \in T_{\mathbb{1}}G =: \mathfrak{g}$. By definition, \mathfrak{g} is the Lie algebra $\text{Lie}(G)$ associated with G . It follows from this isomorphism that the tangent bundle of G is *parallelisable*, i.e.,

$$TG = \bigcup_{\mathbf{x} \in G} \{\mathbf{x}\} \times T_{\mathbf{x}}G \cong \bigcup_{\mathbf{x} \in G} \{\mathbf{x}\} \times T_{\mathbb{1}}G = G \times T_{\mathbb{1}}G = G \times \mathfrak{g},$$

where the isomorphism is canonical. Recall from Hamiltonian mechanics that for a Lagrangian $\mathcal{L} : TG \times \mathbb{R} \rightarrow \mathbb{R}, (\mathbf{q}, \mathbf{v}, t) \mapsto \mathcal{L}(\mathbf{q}, \mathbf{v}, t)$, the Hamiltonian is defined as the Legendre transform of \mathcal{L} ,

$$\mathcal{H} := \sum_i v_i \frac{\partial \mathcal{L}}{\partial v_i} - \mathcal{L} = \sum_i v_i p_i - \mathcal{L},$$

where $p_i := \frac{\partial \mathcal{L}}{\partial v_i}$. The *generalised momentum* $\mathbf{p} : \mathbf{v} \mapsto \sum_i v_i p_i$ is therefore a linear form on the space $\mathfrak{g} \ni \mathbf{v}$. Using $\mathbf{q}(t)$ as the evolution of the system as above, $\dot{\mathbf{q}}(t) \in T_{\mathbf{q}(t)}G$, and so $\mathbf{p}(t) \in (T_{\mathbf{q}(t)}G)^* =: T_{\mathbf{q}(t)}^*G$, the *cotangent space* of G at \mathbf{q} , and the tuples (\mathbf{q}, \mathbf{p}) lie in the *cotangent bundle* T^*G ,

$$T^*G := \bigcup_{\mathbf{q} \in G} \{\mathbf{q}\} \times T_{\mathbf{q}}^*G.$$

As the tangent spaces of G are canonically isomorphic to each other, so are their duals, so we can write

$$T^*G = G \times \mathfrak{g}^*.$$

The dual of the Lie algebra, \mathfrak{g}^* , is the momentum space. The Hamiltonian is a function from the cotangent bundle (and time) to the reals, $\mathcal{H} : T^*G \times \mathbb{R} = G \times \mathfrak{g}^* \times \mathbb{R} \rightarrow \mathbb{R}$.

The Pontryagin dual⁴ \widehat{G} is the set of homomorphisms from G to $U(1)$. Compare this with the vector space dual V^* , which is the set of linear maps from V to the underlying field. One can see \widehat{G} as a sort of non-linear analogue to V^* . We will see that for $G = \mathbb{R}^n$, we have a meaningful isomorphism between \widehat{G} and \mathfrak{g}^* . This will also give a motivation for defining the Hamiltonian on the *cotangent* bundle and not on the tangent bundle, answering the question: *why should the generalised momenta not lie in the same space as the velocities?*

4.2 SUPPLEMENTAL: Examples of representations from Raisa Galimova's exercise class

Let G be a group acting on a measure space⁵ (X, Σ, μ) by the left action $f : G \rightarrow X^X, g \mapsto f_g$, that is, we denote by $f_g(x)$ the action of $g \in G$ on $x \in X$, and $f_g \circ f_h = f_{gh}$. Define the space of square-integrable functions on X , $L^2(X) := \{\varphi \in X^{\mathbb{C}} \mid \|\varphi\|_2 < \infty\}$,

⁴The Pontryagin dual is named after Лев Семёнович Понтрягин (1908/1988).

⁵Here X is a set, $\mu : X \rightarrow [0, \infty]$ is a measure on X and $\Sigma \subseteq 2^X$ is the σ -algebra of μ -measurable subsets.

where $\|\varphi\|_2 := \langle \varphi, \varphi \rangle$, $\langle \psi, \varphi \rangle := \int_X \bar{\psi} \varphi \, d\mu$. Note that if X is finite, and μ is the counting measure,

$$\begin{aligned}\langle \psi, \varphi \rangle &= \int_X \bar{\psi} \varphi \, d\mu = \sum_{x \in X} \bar{\psi}(x) \varphi(x), \\ \forall \varphi \in \mathbb{C}^X \quad \|\varphi\|_2 &= \int_X |\varphi|^2 \, d\mu = \sum_{x \in X} |\varphi|^2 < \infty,\end{aligned}$$

so this is consistent with the definitions from the lecture for a finite group A , namely $L^2(A) = \mathbb{C}^A$, and $\langle \psi, \varphi \rangle = \sum_{g \in A} \bar{\psi}(g) \varphi(g)$. If $X = \mathbb{R}^n$, and μ is the Lebesgue measure,

$$\begin{aligned}\langle \psi, \varphi \rangle &= \int_X \bar{\psi} \varphi \, d\mu = \int_{\mathbb{R}^n} \bar{\psi}(x) \varphi(x) \, dx, \\ \|\varphi\|_2 &= \int_X |\varphi|^2 \, d\mu = \int_{\mathbb{R}^n} |\varphi(x)|^2 \, dx,\end{aligned}$$

and again we recover the definition we gave for the inner product on $L^2(\mathbb{R}^n) \subseteq L^2(\mathbb{R}^n)$.

Now consider the representation $\rho : G \rightarrow \text{GL}(L^2(X))$, $g \mapsto \rho_g$ of G on $L^2(X)$ defined by $(\rho_g(\varphi))(x) = \varphi(f_g(x))$.

Fourier Series and S^1 . If $G = S^1 = X$ is the circle group, acting on itself by addition (here we use the definition $S^1 = \mathbb{R}/2\pi\mathbb{Z}$), $L^2(X) = L^2(S^1)$ decomposes under ρ in a way which is dictated by Fourier series. Indeed, we have for $\varphi \in L^2(S^1)$, $j \in \mathbb{Z}$ — see *The Discrete Fourier Transform*, page 8, though note that the inner product was antilinear in the second argument back then, whereas it is antilinear in the first one this semester:

$$\widehat{\varphi}(j) = \frac{1}{2\pi} \int_0^{2\pi} \varphi(\theta) e^{-ij\theta} \, d\theta = \int_{S^1} \varphi e^{-ij \cdot} \, d\mu = \langle e^{ij \cdot}, \varphi \rangle,$$

where $e^{ij \cdot} := \theta \mapsto e^{ij\theta}$, $e^{ij \cdot} \in \widehat{S^1} \cong \mathbb{Z}$ and where we take a normed μ , i.e., $\mu(S^1) = 1$. In other words, $\widehat{\cdot}(j) = \langle e^{ij \cdot}, \cdot \rangle$ is the projection of φ onto the subspace $V_j := \text{span}\{e^{ij \cdot}\}$ of $L^2(S^1)$. Note that the $e^{ij \cdot}$ are exactly the elements of $\widehat{S^1}$. Moreover, V_j is ρ -invariant, as ρ acts by translation on $\varphi \in L^2(S^1)$, namely $(\rho_\alpha(\varphi))(\theta) = \varphi(\alpha + \theta)$, and hence $\forall \lambda e^{ij \cdot} \in V_j$, $(\rho_\alpha(\lambda e^{ij \cdot})) = \lambda e^{ij(\alpha + \cdot)} = \lambda e^{ij\alpha} e^{ij \cdot} \in V_j$. As it is one-dimensional, V_j is irreducible. We also have, with $\langle e^{ij \cdot}, e^{ik \cdot} \rangle = \widehat{e^{ik \cdot}}(j) = \delta_{jk}$, $V_j \perp V_k$ for $k \neq j$. By Parseval's identity, $\sum_{j \in \mathbb{Z}} \widehat{\varphi}(j) e^{ij \cdot}$ converges absolutely to φ for all $\varphi \in L^2(S^1)$, so

$$L^2(S^1) = \bigoplus_{j \in \mathbb{Z}} V_j = \bigoplus_{\chi \in \widehat{S^1}} \text{span } \chi.$$

To summarise, the translation action ρ is a unitary representation of S^1 on $L^2(S^1)$ with respect to the inner product $\langle \cdot, \cdot \rangle$, it decomposes in the irreducible representations V_j which are the spans of individual elements $e^{ij \cdot}$ of $\widehat{S^1}$ and $\widehat{\cdot}(j)$ is the projector onto V_j . Since $\widehat{S^1} \cong \mathbb{Z}$, and the elements of $\widehat{S^1}$ are the basis onto which functions are projected by the Fourier transform, it makes sense to think of the projector as $\widehat{\cdot}(e^{ij \cdot})$ instead, with $\widehat{\varphi}(\chi) := \langle \chi, \varphi \rangle$ for $\chi \in \widehat{S^1}$.

The Fourier transform and \mathbb{R} . If $G = \mathbb{R}^n = X$ acts by addition on itself, the decomposition of $L^2(X) = L^2(\mathbb{R}^n)$ under ρ is described by the Fourier transform. As above, we have for $\varphi \in L^2(\mathbb{R}^n)$, $\mathbf{k} \in \mathbb{R}^n$,

$$\widehat{\varphi}(\mathbf{k}) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(\mathbf{x}) e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x} = \int_{\mathbb{R}^n} \varphi e^{-i\mathbf{k} \cdot} d\mu = \langle e^{i\mathbf{k} \cdot}, \varphi \rangle$$

by taking $\mu = \frac{1}{(2\pi)^{\frac{n}{2}}} \mathcal{L}^n$, where \mathcal{L}^n is the Lebesgue measure on \mathbb{R}^n . Again, we see that the $e^{i\mathbf{k} \cdot} := \mathbf{x} \mapsto e^{i\mathbf{k} \cdot \mathbf{x}}$ are exactly the elements of $\widehat{\mathbb{R}^n} \cong \mathbb{R}^n$, and that the $V_{\mathbf{k}} := \text{span}\{e^{i\mathbf{k} \cdot}\}$ are ρ -invariant, where ρ acts by translation and are orthogonal. The projector onto $V_{\mathbf{k}}$ is $\widehat{\cdot}(\mathbf{k})$. As for $\varphi \in L^2(\mathbb{R}^n)$,

$$\widetilde{\varphi}(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \widehat{\varphi}(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}} d\mathbf{k} \varphi(\mathbf{x}) = \int_{\mathbb{R}^n} \widehat{\varphi} e^{i\mathbf{x} \cdot} d\widehat{\mu} = \varphi(\mathbf{x})$$

for almost all \mathbf{x} — $\widehat{\mu}$ is the dual measure⁶ —, the representation ρ of \mathbb{R} on $L^2(\mathbb{R})$ by translation decomposes in irreducible representations as

$$L^2(S^1) = \bigoplus_{\mathbf{k} \in \mathbb{R}^n} V_{\mathbf{k}} = \bigoplus_{\chi \in \widehat{\mathbb{R}^n}} \text{span } \chi.$$

Once again, it makes sense to think of the projectors as $\widehat{\cdot}(e^{i\mathbf{k} \cdot})$, with $\widehat{\varphi}(\chi) := \langle \chi, \varphi \rangle$ for $\chi \in \widehat{\mathbb{R}^n}$, as the isomorphism $\widehat{\mathbb{R}^n} \cong \mathbb{R}^n$ is not unique. Actually, the isomorphism $\mathbf{k} \mapsto e^{2\pi i \mathbf{k} \cdot}$ is sometimes used, as it allows for $\mu = \widehat{\mu} = \mathcal{L}^n$ — there are no constant factors before the integrals. Using this isomorphism, the Gaussian $e^{-\frac{1}{2}|\mathbf{x}|^2}$ is no longer self-dual, i.e., the same as its Fourier transform, but $e^{-\pi|\mathbf{x}|^2}$ is instead.

This also gives us an isomorphism between the momentum space $T_0^* \mathbb{R}^n = (\mathbb{R}^n)^*$ and the Pontryagin dual $\widehat{\mathbb{R}^n}$, namely $\boldsymbol{\lambda} \mapsto e^{2\pi i \boldsymbol{\lambda} \cdot}$, with $\widehat{\mathbb{R}^n} \ni e^{2\pi i \boldsymbol{\lambda} \cdot} : \mathbf{k} \mapsto e^{2\pi i \boldsymbol{\lambda} \cdot \mathbf{k}}$, which, in contrast to the isomorphism from \mathbb{R}^n to $\widehat{\mathbb{R}^n}$, does not depend on the choice of an inner product. It therefore makes sense to think of momenta as lying in the dual of the tangent space, and not in the tangent space. Thus the phase space should indeed be $T^* \mathbb{R}^n$, and not $T \mathbb{R}^n$.

Spherical harmonics and $SO(3)$. If $G = SO(3)$ acts by rotations on $X = S^2$, the representation ρ decomposes as

$$L^2(S^2) = \bigoplus_{\ell=1}^{\infty} \mathbf{H}_{\ell},$$

where the irreducible \mathbf{H}_{ℓ} are defined as follows: \mathbf{H}_{ℓ} is the set of all homogeneous harmonic polynomials of degree ℓ on S^2 , i.e.,

$$\mathbf{H}_{\ell} = \left\{ f : S^2 \rightarrow \mathbb{C}, (x, y, z)^{\top} \mapsto f(x, y, z) \mid f(x, y, z) = \sum_{\alpha+\beta+\gamma=\ell} c_{\alpha\beta\gamma} x^{\alpha} y^{\beta} z^{\gamma} \wedge \Delta f = 0 \right\}.$$

LECTURE 10, 2013-03-21.

There are official notes for the material covered in this lecture.

The dual of a group is a sort of nonlinear version of the dual of a vector space.

We switched from an additive to a multiplicative notation. “Algebraists... you have to be careful. They’re sensitive.”

“Representation theory is like a bubble on Wall Street.” We have, obviously, $T \frac{1}{\sqrt{n}} = \frac{1}{\sqrt{n}}$. “Does nobody know about Elmo?” Now, by taking this eigenvector and applying M

⁶The dual measure $\widehat{\mu}$ is the measure we need to get the expressions to look the same when written in terms of measures. We chose μ in such a way that $\mu = \widehat{\mu}$, but if we had defined a non-unitary version, e.g. $\widehat{\varphi}(\mathbf{k}) := \int_{\mathbb{R}^n} \varphi(\mathbf{x}) e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x}$, we would need to define the inverse Fourier transform accordingly as $\widetilde{\psi}(\mathbf{x}) := \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \psi(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}} d\mathbf{k}$, i.e., for $\mu = \mathcal{L}^n$, $\widehat{\mu} = \frac{1}{(2\pi)^n} \mathcal{L}^n$.

over and over again, — in a perfectly legal manner — we get all the other eigenvectors! All the representation theory, all the modules, all the things they talk about... it's just this. "[Claude's] entire thesis is just a super-duper version of this." You don't have to do anything, you just talk. This is the French influence.

"We're going to come to the point where you see the difference between [Claude] and me." We will need not one, but *two* proofs for the Stone-Von Neumann theorem.

"Why am I not happy [about Proof 1] whereas he is happy? [...] I'm only going to talk about mathematical reasons." Man cannot live on algebra alone. What happens as n goes to infinity? The proof doesn't converge. He doesn't care.

The problem is phase transitions: the symmetries change, and you cannot see that with algebra. For instance, for superconductivity, you have to have enough stuff. You cannot just have an explicit n . "Suddenly, we're going to go back to the Fourier transform."

"That's what's good about these algebraists. They're tenacious."

"We have a new Pope now. You cannot confuse the Father, the Son and the Holy Ghost. If you do, you are going to get in trouble. If you confuse eigenvectors and eigenvalues, you're also going to get into trouble. Not the same kind of trouble though."

"If you don't come back next week I understand."

5 All the eigenfunctions of the Fourier transform

Recall:

LECTURE 11, 2013-03-26.

$$\begin{aligned}\mathcal{F}(\varphi)(\mathbf{k}) &:= \widehat{\varphi}(\mathbf{k}) := \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(\mathbf{x}) e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x}, & \varphi \in \mathcal{S}(\mathbb{R}^n), \\ \mathcal{F}^{-1}(\psi)(\mathbf{x}) &:= \check{\psi}(\mathbf{x}) := \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \psi(\mathbf{x}) e^{i\mathbf{k} \cdot \mathbf{x}} d\mathbf{k}, & \psi \in \mathcal{S}(\mathbb{R}^n), \\ \mathcal{F} \mathcal{F}^{-1} &= \mathcal{F}^{-1} \mathcal{F} = \mathbb{1}, \\ \mathcal{F} \left(\frac{d^2}{dx^2} + x^2 \right) (k) &= \left(\frac{d^2}{dk^2} + x^2 \right) \mathcal{F}(\varphi)(k).\end{aligned}$$

So we have this very nice guy here: $\frac{d^2}{dx^2} + x^2$, and having gone to *Elementarschule*, you'd think

$$\frac{d^2}{dx^2} + x^2 \stackrel{?}{=} \left(x - \frac{d}{dx} \right) \left(x + \frac{d}{dx} \right).$$

But the *real* identity they should teach you at school — life is short, the sooner you get going, the better — is

$$(a - b)(a + b) = a^2 - b^2 + [a, b].$$

That's the difference between childhood and adulthood: ab is no longer equal to ba , and it causes a lot of problems. Here $\left[\frac{d}{dx}, x \right] = 1$. We now know $\left(x - \frac{d}{dx} \right) \left(x + \frac{d}{dx} \right) = -\frac{d^2}{dx^2} + x^2 - 1$. This will be around long after strings are seen as a strange cult of the 21st century.

We defined:

$$\begin{aligned}A &:= \frac{1}{\sqrt{2}} \left(x + \frac{d}{dx} \right), \\ A^\dagger &:= \frac{1}{\sqrt{2}} \left(x - \frac{d}{dx} \right).\end{aligned}$$

We wrote “dagger”. This makes sense, as;

$$\begin{aligned}
 \langle \psi, A \varphi \rangle &= \frac{1}{\sqrt{2}} \int_{\mathbb{R}} \bar{\psi}(x) \left(\frac{d}{dx} + x \right) \varphi(x) dx \\
 &= \frac{1}{\sqrt{2}} \int_{\mathbb{R}} \bar{\psi}(x) \frac{d}{dx} \varphi(x) dx + \frac{1}{\sqrt{2}} \int_{\mathbb{R}} \overline{(x\psi(x))} \varphi(x) dx \\
 &= \int_{\mathbb{R}} \overline{\left(\frac{1}{\sqrt{2}} \left(-\frac{d}{dx} + x \right) \psi(x) \right)} \varphi(x) dx \\
 &= \langle A^\dagger \psi, \varphi \rangle.
 \end{aligned}$$

Write

$$\begin{aligned}
 A_x &:= \frac{1}{\sqrt{2}} \left(x + \frac{d}{dx} \right), \\
 A_x^\dagger &:= \frac{1}{\sqrt{2}} \left(x - \frac{d}{dx} \right),
 \end{aligned}$$

and similarly for k instead of x . Observe:

$$\begin{aligned}
 \mathcal{F}(A_x \varphi)(k) &= i A_k \widehat{\varphi}(k), \\
 \mathcal{F}(A_x^\dagger \varphi)(k) &= i A_k^\dagger \widehat{\varphi}(k), \\
 A_x e^{-\frac{1}{2}x^2} &= \left(\frac{d}{dx} + x \right) e^{-\frac{1}{2}x^2} = 0.
 \end{aligned}$$

define

$$h_0(x) := \frac{e^{-\frac{1}{2}x^2}}{\sqrt[4]{\pi}}$$

so that $\langle h_0, h_0 \rangle = 1$. We can now perform the following induction:

$$\begin{aligned}
 A h_0 &= 0, \\
 \widehat{h_0} &= h_0, \\
 \widehat{A_x^\dagger h_0}(k) &= (-i) A_k^\dagger h_0(k) \text{ using the vital lemma,} \\
 \mathcal{F}((A_x^\dagger)^n h_0)(k) &= (-i)^n A_k^\dagger \mathcal{F}((A_x^\dagger)^{n-1} h_0)(k) = (-i)^n (A_k^\dagger)^n h_0(k).
 \end{aligned}$$

In other words, the $(A^\dagger)^n h_0$ — with an A^\dagger , not an A ; “a dagger can be the difference between life and death!” — are eigenfunctions of the the Fourier transform. We will need to prove that these are orthogonal and form a complete set, and we will have an orthogonal basis of $\mathcal{S}(\mathbb{R})$. “I’ll do some coherent states. I like coherent states.” Next time we shall prove:

$$\begin{aligned}
 [A, A^\dagger] &= 1, \\
 [A, (A^\dagger)^n] &= n(A^\dagger)^{n-1}.
 \end{aligned}$$

This is the only real idea in group representations: you find some commutation relations, and you generate all the eigenvectors and eigenvalues.

$n \geq 0$. A , A^\dagger and H all map $\mathcal{S}(\mathbb{R})$ to itself. Now we compute:

$$\begin{aligned}
 2[A, A^\dagger]\varphi &= 2A A^\dagger \varphi - 2A^\dagger A \varphi \\
 &= \left(\frac{d}{dx} + x\right)\left(-\frac{d}{dx} + x\right)\varphi - \left(-\frac{d}{dx} + x\right)\left(\frac{d}{dx} + x\right)\varphi \\
 &= \left(\frac{d}{dx} + x\right)(\varphi' + x\varphi) - \left(-\frac{d}{dx} + x\right)(\varphi' + x\varphi) \\
 &= (\varphi'' + (x\varphi)' + (-x\varphi') + x^2\varphi) - (\varphi'' - (x\varphi)' + x\varphi' + x^2\varphi) \\
 &= 2(x\varphi)' - 2x\varphi' = 2\varphi
 \end{aligned}$$

This yields the following proposition.

Proposition.

$$[A, A^\dagger] = 1 \quad (5.1)$$

$$[A, (A^\dagger)^n] = n(A^\dagger)^{n-1}, \quad n \geq 1, \quad (5.2)$$

$$\widehat{h_n} = (-i)^n h_n, \quad n \geq 0, \quad (5.3)$$

$$\langle h_m, h_n \rangle = \delta_{m,n}, \quad (5.4)$$

$$A h_n = \sqrt{n} h_{n-1}, \quad (5.5)$$

$$A^\dagger h_n = \sqrt{n+1} h_{n+1}, \quad (5.6)$$

$$H h_n = n h_n, \quad n \geq 0, \quad (5.7)$$

$$h_n \in \mathcal{S}(\mathbb{R}), \quad n \geq 0, \quad (5.8)$$

$$\widehat{A_x \varphi} = i A_k \widehat{\varphi}, \quad (5.9)$$

$$\widehat{A_x^\dagger \varphi} = (-i) A_k^\dagger \widehat{\varphi}. \quad (5.10)$$

Proof. Proof of (5.2) by induction:

$$[A, A^\dagger] = 1, \quad (n = 1),$$

$$\begin{aligned}
 [A, (A^\dagger)^{n+1}] &= A(A^\dagger)^{n+1} - (A^\dagger)^{n+1} A \\
 &= A(A^\dagger)^{n+1} - A^\dagger(A(A^\dagger)^n - n(A^\dagger)^{n-1}) \\
 &= A(A^\dagger)^{n+1} - A^\dagger A(A^\dagger)^n + n(A^\dagger)^n
 \end{aligned}$$

“Ah, there must be an easier way to show this, it can't be that hard. Let's just reboot this argument:”

$$\begin{aligned}
 &A(A^\dagger)^n - (A^\dagger)^n A = n(A^\dagger)^{n-1} \\
 \Rightarrow &A(A^\dagger)^{n+1} - (A^\dagger)^n A A^\dagger = n(A^\dagger)^n \\
 \Rightarrow &A(A^\dagger)^{n+1} - (A^\dagger)^n (\mathbb{1} + A^\dagger A) = n(A^\dagger)^n \\
 \Rightarrow &A(A^\dagger)^{n+1} - (A^\dagger)^{n+1} A - (A^\dagger)^n = n(A^\dagger)^n.
 \end{aligned}$$

Proof of (5.4). Now we are going to do something quickly, where algebra really shines: you do nothing but apply the same basic identities over and over, yet get something useful out. Let us assume $m > n \geq 0$, because we already did the case where both are 0.

$$\begin{aligned}
 \langle h_m, h_n \rangle &= \frac{1}{\sqrt{m!}} \frac{1}{\sqrt{n!}} \langle (A^\dagger)^m h_0, (A^\dagger)^n h_0 \rangle \\
 &= \frac{1}{\sqrt{m!}} \frac{1}{\sqrt{n!}} \langle (A^\dagger)^{m-1} h_0, A(A^\dagger)^n h_0 \rangle \\
 &= \frac{1}{\sqrt{m!}} \frac{1}{\sqrt{n!}} \langle (A^\dagger)^{m-1} h_0, ((A^\dagger)^n A + n(A^\dagger)^{n-1}) h_0 \rangle \\
 &= \frac{1}{\sqrt{m!}} \frac{n}{\sqrt{n!}} \langle (A^\dagger)^{m-1} h_0, (A^\dagger)^{n-1} h_0 \rangle \\
 &= \frac{1}{\sqrt{m!}} \frac{n(n-1)}{\sqrt{n!}} \langle (A^\dagger)^{m-2} h_0, (A^\dagger)^{n-2} h_0 \rangle \\
 &= \dots = \frac{1}{\sqrt{m!}} \frac{n(n-1) \dots 1}{\sqrt{n!}} \langle (A^\dagger)^{m-n} h_0, h_0 \rangle.
 \end{aligned}$$

As A^\dagger kills h_0 — it is called the *annihilation operator*, in German the *Vernichtungsoperator* —, we get $\langle h_m, h_n \rangle = \frac{1}{\sqrt{m!}} \frac{n(n-1)\cdots 1}{\sqrt{n!}} \langle 0, h_0 \rangle = 0$ if $m > n$. If $m = n$, we have $\langle h_n, h_n \rangle = \frac{1}{\sqrt{n!}} \frac{n(n-1)\cdots 1}{\sqrt{n!}} \langle h_0, h_0 \rangle = 1$. In other words, $\{h_n\}$ is an orthogonal family.

Proof of (5.5):

$$\begin{aligned} A h_n &= \frac{1}{\sqrt{n!}} A (A^\dagger)^n h_0 \\ &= \frac{1}{\sqrt{n!}} ((A^\dagger)^n A + n(A^\dagger)^{n-1}) h_0 \\ &= \frac{n}{\sqrt{n!}} (A^\dagger)^{n-1} h_0 = \sqrt{n} h_{n-1}. \end{aligned}$$

The other ones are even easier; I'm not doing them, you are getting bored.

Now for something fun, the proof of (5.7). First, note that

$$H h_0 = A^\dagger A h_0 = 0.$$

With what we already have, we can complete the proof elegantly — but then, I have always maintained that elegance is for tailors:

$$H h_n = A^\dagger (A h_n) = A^\dagger \sqrt{n} h_{n-1} = \sqrt{n} \sqrt{n} h_n = n h_n. \quad \square$$

Now, ladies and gentlemen, your jaws should be dropping. We have these beautiful identities:

$$\begin{aligned} \hat{h}_n &= (-i)^n h_n, & n \geq 0, \\ H h_n &= n h_n, & n \geq 0. \end{aligned}$$

So, “out of nowhere” we now have eigenvectors with eigenvalues 0, 1, 2, etc.!

“Whenever I think of algebra, I think about Lawrence of Arabia. It's clean.” *The professor asks if anyone in the audience has actually heard about T. E. Lawrence, and goes on to discuss how the David Lean film features a scene widely known in the film world as an example for the match cut technique.* “If you are really interested, I can just tell you about film and skip this [points to the blackboard] boring crap. I know something about it.” *At this point the lecture has extended five minutes into the break, and people are beginning to leave.* “You should leave the moment you feel lied to. That's a basic thing in all human relationships.”

Now, what would make this even better? Let us show that we are done, that we have all the eigenvectors — in other words, that our orthogonal family is a basis. “That, is, you can't get away from it. You can run, but you can't get away.”

Proposition. *If $\varphi \in \mathcal{S}(\mathbb{R})$ and $\langle \varphi, h_n \rangle = 0$, $n \geq 0$, then $\varphi \equiv 0$.*

Proof. We have $x = \frac{A+A^\dagger}{\sqrt{2}}$. We have:

$$\begin{aligned} \langle \varphi, x h_0 \rangle &= \frac{1}{\sqrt{2}} \langle \varphi, A h_0 \rangle + \frac{1}{\sqrt{2}} \langle \varphi, A^\dagger h_0 \rangle \\ &= \frac{1}{\sqrt{2}} \langle \varphi, h_1 \rangle = 0. \end{aligned}$$

What do we do? We only need one word: Repeat.

$$\begin{aligned} \langle \varphi, x^2 h_0 \rangle &= \frac{1}{2} \langle \varphi, (A + A^\dagger)^2 h_0 \rangle \\ &= \frac{1}{2} \langle \varphi, (A A + A^\dagger A + A A^\dagger + A^\dagger A^\dagger) h_0 \rangle \\ &= \frac{1}{2} \langle \varphi, A A^\dagger h_0 + A^\dagger A^\dagger h_0 \rangle \\ &= \frac{1}{2} \langle \varphi, (A^\dagger A + \mathbb{1}) h_0 + A^\dagger h_1 \rangle \\ &= \frac{1}{2} \langle \varphi, h_0 + \sqrt{2} h_2 \rangle = 0. \end{aligned}$$

And now? Again: Repeat.

$$\begin{aligned}\langle \varphi, x^3 h_0 \rangle &= \left(\frac{1}{\sqrt{2}} \right)^3 \langle \varphi, (A + A^\dagger)^3 h_0 \rangle \\ &= \dots = 0.\end{aligned}$$

By induction, we can show $\langle \varphi, x^n h_0 \rangle = 0$ for $n \geq 0$. It follows

$$\begin{aligned}\sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} \langle \varphi, x^n h_0 \rangle &= 0 \\ \Rightarrow \left\langle \varphi, \left(\sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} \right) h_0 \right\rangle &= 0, \quad \forall k \in \mathbb{R}.\end{aligned}$$

which, if we look carefully, we can recognize as a Fourier transform:

$$\begin{aligned}0 &= \langle \varphi, e^{-ikx} h_0 \rangle \\ &= \sqrt{2\pi} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \overline{\varphi}(x) h_0(x) e^{-ikx} dx, \quad \forall k \in \mathbb{R}.\end{aligned}$$

Note that $\overline{\varphi}(x) h_0(x) \in \mathcal{S}(\mathbb{R})$. It follows:

$$\begin{aligned}\forall k \in \mathbb{R} \quad \widehat{\overline{\varphi} h_0}(k) &= 0 \\ \Rightarrow \forall x \in \mathbb{R} \quad \overline{\varphi}(x) h_0(x) &= 0 \quad (\text{as shown previously}) \\ \Rightarrow \forall x \in \mathbb{R} \quad \varphi(x) &= 0. \quad (h_0(x) \neq 0, \forall x \in \mathbb{R})\end{aligned}$$

So $\varphi \equiv 0$. □

We now know that the h_n form a basis of $\mathcal{S}(\mathbb{R})$. But we are greedy. We really want to write

$$\psi \in \mathcal{S}(\mathbb{R}) \stackrel{?}{=} \sum_{n=0}^{\infty} \langle \psi, h_n \rangle h_n(x) := \lim_{N \rightarrow \infty} \sum_{n=0}^N \langle \psi, h_n \rangle h_n(x).$$

It is exceedingly boring to prove this, I'm sure you don't have the patience. But there is something I *can* do:

$$\begin{aligned}|\langle \psi, h_n \rangle| &= \left| \int_{\mathbb{R}} \overline{\psi}(x) h_n(x) dx \right| \\ &\leq \int_{\mathbb{R}} |\overline{\psi}(x)| |h_n(x)| dx \\ &= \int_{\mathbb{R}} |\overline{\psi}(x)| h_n(x) dx \\ &\leq \|\psi\|_2 \underbrace{\|h_n\|_2}_{=1}.\end{aligned}$$

But in fact, we can do better:

$$\begin{aligned}\langle \psi, h_n \rangle &= \frac{1}{n^p} \langle \psi, H^p h_n \rangle \\ &= \frac{1}{n^p} \langle H^p \psi, h_n \rangle.\end{aligned}$$

So,

$$\begin{aligned} |\langle \psi, h_n \rangle| &\leq \frac{1}{n^p} |\langle \psi, H^p h_n \rangle| \\ &\leq \frac{1}{n^p} \|H^p \psi\|_2. \end{aligned}$$

Concluding, we have the following

Lemma. *If $\psi \in \mathcal{S}(\mathbb{R})$, then $|\langle \psi, h_n \rangle| \leq \frac{\|H^p \psi\|_2}{n^p}$, $n \geq 1$, $p \geq 0$.* □

LECTURE 13, 2013-04-09.

Proposition. *The $h_n(x)$, $n \geq 0$ form an orthonormal basis of $\mathcal{S}(\mathbb{R})$. More precisely, let*

$$\mathcal{S}(\mathbb{N}) := \left\{ (s_0, s_1, s_2, \dots) \mid s_j \in \mathbb{C} \text{ and } \forall p, \exists c_p, |s_j| \leq c_p \frac{1}{(1+j)^p} \right\}.$$

Then,

$$\begin{aligned} \mathcal{S}(\mathbb{N}) &\rightarrow \mathcal{S}(\mathbb{R}) \\ s &\mapsto \sum_{n=0}^{\infty} s_n h_n(x) \end{aligned}$$

is a continuous, bijective linear map with a continuous inverse. □

“Whatever somebody means by an orthonormal basis for an infinite dimensional vector space, this is about as good as it gets.” In proving the theorem, which is an exercise, we will have to do something like

$$\left| \sum_{n \geq 0} s_n h_n(x) \right| \leq \sum_{n \geq 0} |s_n| |h_n(x)|$$

— and remember, the macho waits as long as possible before bringing the $|\cdot|$ in. Compare this to what we had for Fourier series:

$$\left| \sum_{j=-\infty}^{\infty} \widehat{\varphi}(j) e^{ijx} \right| \leq \sum_{j=-\infty}^{\infty} |\widehat{\varphi}(j)| \underbrace{|e^{ijx}|}_{=1}$$

For the h_n , we are not so lucky, and actually have to estimate $\sup_{x \in \mathbb{R}} |h_n(x)|$.

$$\begin{aligned}
 |h_n(x)| &= |(-i)^{-n} \widehat{h_n}(x)| \\
 &\leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |h_n(y)| \, dy \\
 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{1+y^2}} \sqrt{1+y^2} |h_n(y)| \, dy \\
 &\leq \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{\infty} \frac{1}{1+y^2} \, dy \right)^{\frac{1}{2}} \left(\int_{-\infty}^{\infty} (1+y^2) h_n^2(y) \, dy \right)^{\frac{1}{2}} \\
 &= \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{\infty} \frac{1}{1+y^2} \, dy \right)^{\frac{1}{2}} (\langle h_n, (1+y^2) h_n \rangle)^{\frac{1}{2}} \\
 &= \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{\infty} \frac{1}{1+y^2} \, dy \right)^{\frac{1}{2}} \left(\langle h_n, h_n \rangle + \frac{1}{2} \langle h_n, (A + A^\dagger)^2 h_n \rangle \right)^{\frac{1}{2}} \\
 &\leq \frac{1}{\sqrt{2\pi}} \underbrace{\left(\int_{-\infty}^{\infty} \frac{1}{1+y^2} \, dy \right)^{\frac{1}{2}}}_{< \infty, \text{ independent of } n} \sqrt{2(n+1)}
 \end{aligned}$$

6 Quantising the harmonic oscillator

“Now, we come into the good stuff”. Once upon a time, there was a classical Hamiltonian:

$$h(q, p) := \frac{1}{2} \left(\frac{p^2}{m} + m\omega^2 q^2 \right)$$

Do you remember anything from classical mechanics? “So, Thursday: from the classical to the quantum world and beyond.”

6.1 Classical Mechanics

What’s classical mechanics? We start with a number $n \geq 1$, the number of *degrees of freedom*. It does not change while you keep discussing that system. Then, we have a set $\Omega \subseteq \mathbb{R}^n$, open and connected, called the *classical configuration space*. For instance, we can consider a pendulum, with $q \in \Omega = \mathbb{R}$ the angle. Or we could have a joint in the pendulum, yielding $(q_1, q_2) \in \Omega = \mathbb{R}^2$.

Definition (Classical phase space). The *classical phase space* is defined as $\Omega \times \mathbb{R}^n$, where $(\mathbf{q}, \mathbf{p}) \in \Omega \times \mathbb{R}^n$. \diamond

“If you don’t measure anything, then you’re doing string theory, and your argument is: there must be supersymmetries, because it’s so beautiful.”

Definition (Classical observables). The set of *classical observables* is the algebra of smooth, real valued functions on $\Omega \times \mathbb{R}^n$, namely $C^\infty(\Omega \times \mathbb{R}^n, \mathbb{R})$. \diamond

An example of a classical observable is

$$\begin{aligned}
 q_i &\in C^\infty(\Omega \times \mathbb{R}^n, \mathbb{R}) \\
 \Omega \times \mathbb{R}^n &\ni (\mathbf{q}, \mathbf{p}) \mapsto q_i \in \mathbb{R}.
 \end{aligned}$$

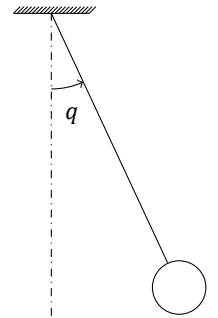


Figure 1. The configuration space $\Omega = \mathbb{R}$ of a simple pendulum. Here $q \in \mathbb{R}$, as the pendulum can do a full turn — or several — around its pivot.

LECTURE 14, 2013-04-11.

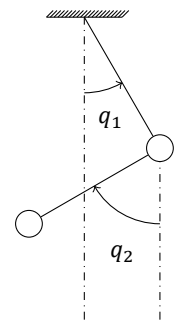


Figure 2. The configuration space $\Omega = \mathbb{R}^2$ of a double pendulum, where $\mathbf{q} = (q_1, q_2) \in \mathbb{R}^2$.

Somebody complained about complex-valued functions not being real enough, so we use real-valued functions. “You could ask me: why smooth? — Why smooth?” You can’t measure something that isn’t regular. “No human being can work with something other than a polynomial anyway.”

Definition (Poisson bracket). Finally, define the *Poisson bracket* of ψ, φ as

$$\{\psi, \varphi\}(\mathbf{q}, \mathbf{p}) := \sum_{j=1}^n \frac{\partial \psi}{\partial q_j} \frac{\partial \varphi}{\partial p_j} - \frac{\partial \psi}{\partial p_j} \frac{\partial \varphi}{\partial q_j},$$

for $\psi, \varphi \in C^\infty(\Omega \times \mathbb{R}^n)$. ◇

Proposition (Ur-something of classical mechanics). $(C^\infty(\Omega \times \mathbb{R}^n), \{\cdot, \cdot\})$ is a Lie algebra, das heißt, $\{\cdot, \cdot\}$ is bilinear, skew-symmetric and satisfies the Jacobi identity. □

Definition (Classical mechanical system). A classical mechanical system on $\Omega \times \mathbb{R}^n$ is a (Hamiltonian) total energy function

$$h(\mathbf{q}, \mathbf{p}) \in C^\infty(\Omega \times \mathbb{R}^n). \quad \diamond$$

Definition (Evolution of a classical observable). The evolution in time f of an observable $\psi \in C^\infty(\Omega \times \mathbb{R}^n)$, given at $t = 0$ by $f(\mathbf{q}, \mathbf{p}, 0) = \psi(\mathbf{q}, \mathbf{p})$, is determined by the *Liouville equation*

$$\frac{d}{dt} f = \{f, h\}. \quad \diamond$$

The conservation of energy immediately follows:

$$\frac{d}{dt} h = \{h, h\} = 0.$$

This is all of classical mechanics. We can certainly “pimp” the theoretical structure behind it further by introducing symplectic manifolds, Poisson algebras, and so on. But so far, this could have all been done by mathematicians. Where does the physicist come in? The art and the genius lies in finding the correct h . In doing so, at least when describing a real system you see around you, you more or less establish a law of nature.

Starting from this theoretical perspective, as opposed to doing a history course and starting with Newton and $F = ma$, could be described as the modern view, because it brings the symmetries of the theory to the front. From what we have, we can now derive Hamilton’s equations of motion:

$$\begin{aligned} \frac{d}{dt} q_j &= \{q_j, h\} \\ &= \sum_{k=1}^n \left(\frac{\partial}{\partial q_k} q_j \right) \frac{\partial h}{\partial p_k} - \left(\frac{\partial}{\partial q_k} p_j \right) \frac{\partial h}{\partial q_k} \\ &= \frac{\partial h}{\partial p_j} \\ \frac{\partial}{\partial t} p_j &= -\frac{\partial h}{\partial q_j}. \end{aligned}$$

How can we now describe a “natural” transformation between two phase spaces? If the map

$$\begin{aligned} \mathbf{s} : \Omega \times \mathbb{R}^n &\rightarrow \bar{\Omega} \times \mathbb{R}^n \\ (\mathbf{s}^* \psi)(\mathbf{q}, \mathbf{p}) &= \psi(\mathbf{s}(\mathbf{q}, \mathbf{p})) \in C^\infty(\Omega \times \mathbb{R}^n) \end{aligned}$$

verifies

$$\begin{aligned} \psi &\in C^\infty(\bar{\Omega} \times \mathbb{R}^n) \\ \{s^* \psi, s^* \varphi\}(\mathbf{q}, \mathbf{p}) &= \{\psi, \varphi\}(\mathbf{s}(\mathbf{q}, \mathbf{p})) = s^* \{\psi, \varphi\} \quad \forall \psi, \varphi \in C^\infty(\bar{\Omega} \times \mathbb{R}^n), \end{aligned}$$

then it is called *symplectic* — or “canonical” in 19th century math slang. Why is this important? It allows us to define when two systems, i.e., two hamiltonians and the associated phase spaces, are physically equivalent.

Let me just show you one thing. \mathbf{s} is a symplectic isomorphism if and only if $\mathrm{d}\mathbf{s} = \left(\frac{\partial s_j}{\partial x_k}\right)$ satisfies

$$(\mathrm{d}\mathbf{s}(\mathbf{q}, \mathbf{p})\mathbf{v}) \cdot (\mathbf{J} \mathrm{d}\mathbf{s}(\mathbf{q}, \mathbf{p})\mathbf{w}) = \mathbf{v} \cdot (\mathbf{J}\mathbf{w}), \text{ where } \mathbf{J} := \begin{pmatrix} \mathbf{0} & \mathbb{1}_n \\ -\mathbb{1}_n & \mathbf{0} \end{pmatrix},$$

for all $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{w} \in \mathbb{R}^n$ and $(\mathbf{q}, \mathbf{p}) \in \Omega \times \mathbb{R}^n$.

Definition (Symplectic group). And if we really want to, we can now also define the group of *symplectic matrices* $\mathrm{Sp}(2n, \mathbb{R}) := \{\mathbf{A} \in \mathrm{GL}(2n, \mathbb{R}) \mid \mathbf{A}^\top = -\mathbf{J}\mathbf{A}^{-1}\mathbf{J}\}$. \diamond

Note that “it’s even in the Heisenberg stuff”, as $\mathrm{Sp}(2, \mathbb{R}) = \mathrm{SL}(2, \mathbb{R})$. With this definition, the above characterisation of a symplectic isomorphism can be reformulated to $\mathrm{d}\mathbf{s} \in \mathrm{Sp}(2n, \mathbb{R})$.

What is the simplest possible motion? Let’s take an h_0 that only depends on \mathbf{p} , not on \mathbf{q} . In that case we have

$$\begin{aligned} \dot{q}_j &= \frac{\partial h_0}{\partial p_j}(\mathbf{p}) \quad \Rightarrow \quad \mathbf{q} = \frac{\mathrm{d} h_0}{\mathrm{d} \mathbf{p}} t + \mathbf{q}(0) \\ \dot{p}_j &= \frac{\partial h_0}{\partial q_j}(\mathbf{p}) = 0 \Rightarrow \mathbf{p} = \mathbf{p}(0) \\ \mathbf{v} &:= \dot{\mathbf{q}} = \frac{\partial h}{\partial \mathbf{p}} \end{aligned}$$

That’s a straight line. There’s nothing simpler than a straight line, except no line at all — but I’m not going to get into discussions of nothingness, though it’s fashionable these days. There are entire books written about nothingness, and how string theory explains everything about nothingness...

Now suppose somebody over there in Oberwinterthur has a fancy Hamiltonian $h(\mathbf{q}, \mathbf{p})$. How do you know it’s not a straight line, and that you are not being charged non-straight line prices for a straight line? You want to pull back, to find \mathbf{s} such that $\mathbf{s}^*h = h_0$. You heard about the Hamilton-Jacobi equation, right? And probably your mind went fuzzy at this time, and you heard about maps, and maybe generating functions, and this and that... There is an \mathbf{s} if and only if $\{\varphi \mid \{\varphi, h\} = 0\}$ has an Abelian subalgebra \mathfrak{A} of dimension n such that

$$\begin{aligned} \mathfrak{A} &= \{I_1, \dots, I_n\}, \\ \{I_j, I_k\} &= 0, \\ \{I_j, h\} &= 0. \end{aligned}$$

Setting things up like this, the logical structure is crystal clear. This is something a mathematician could do. You may say, as a physicist, “frankly, my dear, I don’t give a damn,”⁷ but the reason this structure is interesting is because we have beautiful examples. So, if mathematicians are coming up with all different kinds of beautiful structures, the question is — and I am putting this in very low terms — Are they not just playing with themselves? With this you can predict where Pluto or Neptune is. String theory has predicted nothing, and it may never.

In the case $h(\mathbf{q}, \mathbf{p}) = \frac{1}{2}(\frac{p^2}{n} + m\omega^2 q^2)$, where $(\mathbf{q}, \mathbf{p}) \in \mathbb{R} \times \mathbb{R}$, we have the harmonic oscillator. And that’s classical mechanics. Now on to quantum mechanics.

6.2 Quantum Mechanics

How do we start? There is a number $n \geq 1$ of *quantum degrees of freedom*. Then, $\Omega \subseteq \mathbb{R}^n$ is the *quantum configuration space*, like the classical configuration space above. If you had fallen asleep at the beginning of the previous hour, and you thought Q.M. was just a misspelling of classical mechanics, you might think the lecture had just started. Now it becomes different.

⁷*Gone with the Wind*, 1939, directed by Victor Fleming.

Definition (Quantum phase space). The *quantum phase space* is $\mathcal{S}(\mathbb{R}^n)$. \diamond

Now we need an algebra of quantum observables.

Definition (Quantum observables). The *quantum observables* are the elements of

$$\text{QO} := \left\{ A \in \text{End}(\mathcal{S}(\mathbb{R}^n)) \mid \begin{array}{l} A \text{ is a continuous linear map such that} \\ \forall \psi, \varphi \in \mathcal{S}(\mathbb{R}^n), \langle A\psi, \varphi \rangle = \langle \psi, A\varphi \rangle \end{array} \right\}. \quad \diamond$$

Let me give you two examples of observables:

$$\begin{aligned} (Q_j \psi)(\mathbf{q}) &:= q_j \psi(\mathbf{q}), \\ (P_j \psi)(\mathbf{q}) &:= \frac{\hbar}{i} \frac{\partial}{\partial q_j} \psi(\mathbf{q}). \end{aligned}$$

Proposition. $(\text{QO}, \frac{i}{\hbar}[A, B])$ is a Lie algebra.

Proof. We check that QO is closed under the bracket $\frac{i}{\hbar}[\cdot, \cdot]$: let A, B be self-adjoint. Then

$$\begin{aligned} \left(\frac{i}{\hbar}[A, B] \right)^\dagger &= -\frac{i}{\hbar}(AB - BA)^\dagger \\ &= -\frac{i}{\hbar}(B^\dagger A^\dagger - A^\dagger B^\dagger) \\ &= -\frac{i}{\hbar}(BA - AB) = \frac{i}{\hbar}[A, B]. \end{aligned}$$

You can check the Jacobi identity and all that. \square

Definition (Quantum system). A *quantum system* is an $H \in \text{QO}$. \diamond

Definition (Evolution of a quantum observable). The evolution of the observable A is determined by

$$\dot{A} = \frac{i}{\hbar}[A, H]. \quad \diamond$$

“That’s it. That’s quantum mechanics.”

LECTURE 15, 2013-04-16.

Okay! So where were we, ladies and gentlemen? When I last saw you, we talked about quantum mechanics. We have:

$$\begin{array}{c} \text{configuration space} \\ \left(\underbrace{n}_{\substack{\uparrow \\ \text{degrees of} \\ \text{freedom}}}, \underbrace{\mathbb{R}^n}_{\substack{\downarrow \\ \text{classical phase space}}}, \underbrace{(\mathcal{S}(\mathbb{R}^n), \langle \cdot, \cdot \rangle)}_{\substack{\text{classical phase space}}}, \underbrace{\text{Hermitian}(\mathcal{S}(\mathbb{R}^n))}_{\substack{\text{quantum observables} \\ \text{Lie algebra structure}}}, \underbrace{\frac{i}{\hbar}[\cdot, \cdot]}_{\substack{\text{Lie algebra structure}}}) \end{array}$$

$$\frac{d}{dt} A = \frac{i}{\hbar}[A, H]$$

We’re doing this the French way. This is a 5-tuple, the way Bourbaki would write it. To them, mathematics is just long strings of symbols.

Such a logical structure is of no interest if you can’t find H , unless there is a way of picking H s that describe quantum systems, not just because you believe they do, but because they do actually predict things that you can verify. It’s like finding Waldo.⁸ Unless you can find Waldo, it’s just a meaningless complicated picture. Finding Waldo gives it its meaning. Actually let’s call the Hamiltonian W .⁹ Finding Waldo is *not* the Bourbaki way of doing things. I’ll show you two books any physicist will tell you about: *Find classical Waldo*, and *Find quantum Waldo*. *Find classical Waldo* you already had a lecture about last semester.

Recall the two quantum observables from the last lecture.

⁸This is a reference to Martin Handford’s *Where’s Wally?*, a series of children’s books in which a character (Wally) is hidden on every page, published in the United Kingdom from 1987 onwards. It was published as *Where’s Waldo?* in the United States, and as *Wo ist Walter?* in Germany.

⁹This notation was — fortunately — not actually used during the lecture, so we stick to H .

Definition (Momentum and position operators).

$$(P_k \varphi)(\mathbf{q}) := \frac{\hbar}{i} \frac{\partial}{\partial q_k} \varphi(\mathbf{q}) \text{ is the momentum operator,}$$

$$(Q_k \varphi)(\mathbf{q}) := q_k \varphi(\mathbf{q}) \text{ is the position operator.}$$

Lemma. These operators have the following properties:

$$P_k, Q_k \in \text{Hermitian}(\mathcal{S}(\mathbb{R}^n))$$

$$\frac{i}{\hbar} [P_k, P_l] = 0,$$

$$\frac{i}{\hbar} [Q_k, Q_l] = 0,$$

$$\frac{i}{\hbar} [P_k, Q_l] = \delta_{kl} \mathbb{1}.$$

Proof. The cases where the commutator is 0 are obvious. We prove the last equality,

$$\begin{aligned} \frac{i}{\hbar} \left[\frac{\hbar}{i} \frac{d}{dq}, q \right] \varphi &= \frac{d}{dq} (q\varphi) - q \frac{d}{dq} \varphi \\ &= \varphi = \mathbb{1}\varphi. \end{aligned} \quad \square$$

Now we do something *really* weird. We find Waldo in the classical case — this may be hard, but we find Waldo hiding amidst the electric and magnetic fields —, and then we quantise him. Namely, for a classical Waldo $h(\mathbf{q}, \mathbf{p})$, replace any q_k by Q_k and any p_k by P_k .

$$h(\mathbf{q}, \mathbf{p}) \xrightarrow{\text{Quantise Waldo}} H(Q_1, \dots, Q_n, P_1, \dots, P_n)$$

If you believe that... you'll believe anything. As an example of quantisation, consider the following:

$$h(\mathbf{q}, \mathbf{p}) = \frac{1}{2m} \mathbf{p}^2 + V(\mathbf{q}) \Rightarrow H(\mathbf{Q}, \mathbf{P}) = \frac{1}{2m} \sum_{j=1}^n P_j^2 + V(Q_1, \dots, Q_n),$$

where we have for a continuous function f on \mathbb{R} and an operator A with spectral decomposition

$$A = \sum_{k=1}^r \lambda_k \Pi_k,$$

$$f(A) = \sum_{k=1}^r f(\lambda_k) \Pi_k.$$

Why? This is obvious if f is a polynomial, because $\Pi_k \Pi_l = \delta_{kl} \Pi_k$. Recall the Weierstrass approximation theorem: any continuous function can be approximated by polynomials, and the approximation is uniform, so by taking the limit, $f(A) = \sum_{k=1}^r f(\lambda_k) \Pi_k$.

This works because we have a strict Apartheid between the P s and Q s. They have to sit on different sides of the bus. For instance, look at the harmonic oscillator:

$$h(p, q) = \frac{1}{2} \left(\frac{p^2}{m} + m\omega^2 q^2 \right)$$

$$\Rightarrow H(Q, P) = \frac{1}{2} \left(-\frac{\hbar^2}{m} \frac{d^2}{dq^2} + m^2 \omega^2 q^2 \right).$$

Note that there are no qp terms in the classical Hamiltonian. But if we do not have this Apartheid in the classical Hamiltonian, if Waldo lives in the United States after the civil rights movement — or even after Obama was elected —, we may have $q_i p_j$ in

the Hamiltonian — say with $\mathbf{q} \wedge \mathbf{p} = (q_2 p_3 - q_3 p_2, q_3 p_1 - q_1 p_3, q_1 p_2 - q_2 p_1)^\top$. And then this commutes in the classical case, but not in the quantum case, so we have a problem: do we write $Q_i P_j$ or $P_j Q_i$? Somehow, if you have a physically interesting Hamiltonian, it just happens that you can fiddle with it to get rid of those $q_i p_j$. This may sound like a ridiculous story like *Samichlaus*, but it works. That's all I can tell you. Do you know about Pauling? He got the Nobel Peace Prize — and also the Nobel Prize in Chemistry for something he did on the side. Well Pauling used that sort of bullshit to calculate the angle in the water molecule, and he got the right answer.

So do you understand the point? There is a logical structure which is neat, and there is this weird recipe to quantise Waldo, and it works. I can't tell you more, and nobody can. Actually, there is a slightly more satisfying recipe, using these W s — hey, more Waldos! — these Weyl operators we talked about. We'll come back to this later. It's a messy story, and it doesn't even have a good end. Somehow or other, quantum mechanics and gravity don't mix. There was a time when string theory was supposed to quantise gravity. Twenty years ago there was a long list of all the things string theory was supposed to do. Now the list looks very different, and it's much shorter; none of the things that used to be on that list are on it anymore. String theory has changed its *curriculum vitae*.

LECTURE 16, 2013-04-18.

The last time I saw you we were looking for Waldo. We're going to look for Waldo some more. We were looking at EQ1. The easiest case is when we have this strict Apartheid between q and p , and of course this does not happen all that much in real life — thank God for that. EQ2 There is one more thing that I should do to bring us back to the nostalgic world of our childhood. There is one more thing that I should do, and all will be right in the world. We add -1 in the last expression: EQ3 The red H is the one you know — I ran out of H s. The fact that I picked red has no political overtones whatsoever. This is Switzerland, red's okay. It's on the flag. Recall: EQ4. So here, EQ5, so we have found the energy levels of the quantum harmonic oscillator. *Somebody asks whether this has something to do with blackbody radiation.* “Most physicists aren't even aware that they're lying,” but you *really* need quantum field theory for that.

MISSING TEXT It's normalised so that EQ6, and we interpret this as the probability EQ7. There's no reason why it should work, but it does. The other thing is EQ8. What is the Heisenberg uncertainty principle? Define EQ9. Here's the punchline: EQ10. What does this mean? The tinier the variance, the more the expected value is expected, so the more you know about Q , the less you know about P . I'm not going to prove this, you'll do it in quantum mechanics next year.

I'm now going to show you a more rational way of quantising Waldo. This is still a story, but while — unless you're really gullible — you stopped believing in *Samichlaus* at the age of 4. This you might believe until 10.

From here onwards there are official notes for the material covered in the lecture.

Somebody asks about the dominated convergence theorem, which is used in the proof of irreducibility of ρ_h . It's just you, right? No, he wants to do that too. So next Tuesday, I'll bite the bullet and we'll talk about the Lebesgue dominated convergence theorem.

The biggest mistake people do when they decide to do research is to read a book. That's just as bad as jumping into a crevice, except if you fall into a crevice, some Austrian might find you in 3000 years and you'll end up in a museum. If you read a book, you're really finished.