

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

2013-05-10

## 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  $\varphi$  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  $\varphi$ . 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  $\varphi$  in an appropriate norm<sup>1</sup>—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:

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

---

<sup>1</sup>This 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.

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

$$\begin{aligned} \mathbb{N} &:= \{0, 1, 2, \dots\}, & \alpha &= (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n \text{ is a multi-index,} \\ |\alpha| &:= \alpha_1 + \dots + \alpha_n, & \partial_x^\alpha &:= \frac{\partial^{|\alpha|}}{\partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n}} \quad (\text{we use } \partial^\alpha \text{ when the variable is implicit),} \\ \text{and } x^\beta &:= x_1^{\beta_1} \dots x_n^{\beta_n} \text{ for } x \in \mathbb{R}^n, \beta \in \mathbb{N}^n. \end{aligned}$$

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

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

◇

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  $\varphi_k \xrightarrow{\mathcal{S}(\mathbb{R}^n)} \varphi$ .

◇

**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  $\lim_{k \rightarrow \infty} U(\varphi_k) = U(\varphi)$ . ◇

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,

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

LECTURE 5, 2013-03-05.

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

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

## 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  $\varphi$ , also written  $\mathcal{F}(\varphi)$ , as:

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

LECTURE 6, 2013-03-07.

**Lemma (“Your life in Fourier land depends on it.”).**

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

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

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

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

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

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

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

**Proof.** Proof of (2.2):

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

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):

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

Proof of the second part of (2.5):

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

*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):

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

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

**Proof.**

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

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

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

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

$$\begin{aligned}
 \mathcal{F}(\partial_{\mathbf{x}}^\alpha (\mathbf{k}^\beta \varphi))(\mathbf{k}) &= i^{|\alpha|+|\beta|} \mathbf{k}^\alpha \partial_{\mathbf{k}}^\beta \widehat{\varphi}(\mathbf{k}) && \text{by applying the first part} \\
 &&& \text{of (2.5) } |\alpha| \text{ times and the} \\
 &&& \text{second part } |\beta| \text{ times.} \\
 \forall \alpha, \beta, |\mathbf{k}^\alpha \partial_{\mathbf{k}}^\beta \widehat{\varphi}(\mathbf{k})| &\leq \frac{1}{(2\pi)^{\frac{n}{2}}} \|\partial_{\mathbf{x}}^\beta (\mathbf{x}^\beta \varphi)\|_1 < \infty && \text{from (2.2).}
 \end{aligned}$$

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

$$\begin{aligned}
 \check{\widehat{\varphi}}(\mathbf{x}) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \widehat{\varphi}(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}} \, d\mathbf{k} \\
 &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \varphi(\mathbf{y}) e^{-i\mathbf{k} \cdot \mathbf{y}} \, d\mathbf{y} \, e^{i\mathbf{k} \cdot \mathbf{x}} \, d\mathbf{k}
 \end{aligned}$$

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  $(2\pi)^{-n} \int_{\mathbb{R}^n} e^{i\mathbf{k} \cdot (\mathbf{x}-\mathbf{y})} \, d\mathbf{k}$  by  $\delta(\mathbf{x}-\mathbf{y})$ . “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."

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

"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.*

$$\begin{aligned}\frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \hat{\varphi}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} e^{-\epsilon \frac{|\mathbf{k}|^2}{2}} d\mathbf{k} &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(\mathbf{y}) e^{-i\mathbf{y}\cdot\mathbf{k}} d\mathbf{y} e^{i\mathbf{k}\cdot\mathbf{x}} e^{-\epsilon \frac{|\mathbf{k}|^2}{2}} d\mathbf{k} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(\mathbf{y}) \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})} e^{-\epsilon \frac{|\mathbf{k}|^2}{2}} d\mathbf{k} d\mathbf{y}\end{aligned}$$

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.

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

We therefore get:

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

□

### 3 Seeking eigenfunctions of the Fourier transform

"The main ideas of many things are right here." We seek to find the solutions  $(\varphi, \lambda)$  of

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

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  $n = 1$ .

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

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:

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

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  $[Q, P] = i1$ . Multiplying by some  $\hbar \in \mathbb{R}$ , redefine  $(Pf)(x) := \frac{\hbar}{i} \frac{d}{dx}f(x)$ , we get  $[Q, P] = \hbar i1$ .

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  $e^x e^y$  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

$$(e^{itP} f)(x) = f(x+t) \quad \text{as } (e^{it(\frac{1}{i} \frac{d}{dx})}) f(x) \stackrel{\text{formally}}{=} \sum_{l=0}^{\infty} \frac{t^l}{l!} \frac{d^l}{dx^l} f(x),$$

the Taylor series of  $f(x+t)$  at  $x$ .

$$e^{itP} e^{isQ} = e^{ist} e^{isQ} e^{itP}$$

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 subgroup generated by the orbit of  $W$  under the action of  $F_X$  on itself by conjugation,  $^g s := gsg^{-1}$ , so we can write

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

"Out of this trivial stuff comes something deep."

LECTURE 8, 2013-03-14.

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

“A faithful representation of a group is just changing variables.”  $\rho$  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.”

LECTURE 9, 2013-03-19.  
There are official notes for the material covered in this lecture.

The dual group only works for abelian groups. The way a bourgeois mathematician thinks of the classical phase space  $\mathbb{R}^n \times \mathbb{R}^n$  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: *Ἡ φανταχτερή γλῶσσα τῶν ἀστικῶν μαθηματικῶν*<sup>2</sup>

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  $\mathbf{q} : \mathbb{R} \rightarrow G$  describes the evolution of the generalised coordinates of the system with time, the generalised velocity at time  $t$ ,  $\dot{\mathbf{q}}(t)$ , lies in the tangent space  $T_{\mathbf{q}(t)} G$ —recall that the tangent space of  $g$  at  $\mathbf{x}$  is defined as the set of derivatives<sup>3</sup>  $\dot{\gamma}(0)$  at 0 of smooth curves  $\gamma : \mathbb{R} \rightarrow G$  with  $\gamma(0) = \mathbf{x}$ .

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  $T_{\mathbf{q}} G$  and  $T_{\varphi_q(\mathbf{q})} = T_{\mathbb{1}} G$ , by

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

where  $\dot{\gamma}(0)$  is an element of  $T_{\mathbf{q}} G$ , so  $\gamma(0) = \mathbf{q}$  and  $\gamma_0 := \varphi_q \circ \gamma$  verifies  $\gamma_0(0) = \mathbb{1}$ , therefore  $\dot{\gamma}_0(0) \in T_{\mathbb{1}} G =: \mathfrak{g}$ . 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,  $\dot{\mathbf{q}}(t) \in T_{\mathbf{q}(t)} G$ ,

<sup>2</sup>The fancy language of bourgeois mathematicians.

<sup>3</sup>One 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  $\gamma$  in the set of curves through  $\mathbf{x}$  with  $\gamma(0) = \mathbf{x}$  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  $\mathbf{x}$  onto  $\mathbb{R}^n$ . It can then be shown that this is independent of  $f$ .

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 dual<sup>4</sup>  $\widehat{G}$  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 space<sup>5</sup>  $(X, \Sigma, \mu)$  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  $X$ ,  $L^2(X) := \{\varphi \in \mathbb{C}^X \mid \|\varphi\|_2 < \infty\}$ , 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  $\mu$  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  $\mu$  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  $\mathcal{S}(\mathbb{R}^n) \subseteq L^2(\mathbb{R}^n)$ .

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

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

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

**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  $\rho$  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  $\mu$ , i.e.,  $\mu(S^1) = 1$ . In other words,  $\widehat{\varphi}(j) = \langle e^{ij \cdot}, \varphi \rangle$  is the projection of  $\varphi$  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  $\rho$ -invariant, as  $\rho$  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  $\varphi$  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  $\rho$  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  $\rho$  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  $\rho$ -invariant, where  $\rho$  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 measure<sup>6</sup>—the representation  $\rho$  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

<sup>6</sup>The 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  $\mu$  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  $\widetilde{\varphi}(\mathbf{x}) := \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{\varphi}(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}} d\mathbf{k}$ , i.e., for  $\mu = \mathcal{L}^n$ ,  $\widehat{\mu} = \frac{1}{(2\pi)^n} \mathcal{L}^n$ .

$\mathbf{k} \mapsto e^{2\pi i \mathbf{k} \cdot}$  is sometimes used, as it allows for  $\mu = \hat{\mu} = \mathcal{L}^n$ —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  $\lambda \mapsto e^{2\pi i \lambda}$ , with  $\widehat{\mathbb{R}^n} \ni e^{2\pi i \lambda} : \mathbf{k} \mapsto e^{2\pi i \lambda \mathbf{k}}$ , 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  $\rho$  decomposes as

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

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

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

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  $\mathbf{M}$  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.”

LECTURE 10, 2013-03-21.

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

## 5 All the eigenfunctions of the Fourier transform

LECTURE 11, 2013-03-26. Recall:

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

$$\mathcal{F}^{-1}(\psi)(\mathbf{x}) := \check{\psi}(\mathbf{x}) := \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \psi(\mathbf{x}) e^{i\mathbf{k} \cdot \mathbf{x}} d\mathbf{k}, \quad \psi \in \mathcal{S}(\mathbb{R}^n),$$

$$\mathcal{F} \mathcal{F}^{-1} = \mathcal{F}^{-1} \mathcal{F} = \mathbb{1},$$

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

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:

$$A := \frac{1}{\sqrt{2}} \left( x + \frac{d}{dx} \right),$$

$$A^\dagger := \frac{1}{\sqrt{2}} \left( x - \frac{d}{dx} \right).$$

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

$$A_x := \frac{1}{\sqrt{2}} \left( x + \frac{d}{dx} \right),$$

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

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

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

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 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.

We also define  $H := A^\dagger A$ ,  $h_n(x) := (A^\dagger)^n h_0$  for  $n \geq 1$ . We have  $h_n \in \mathcal{S}(\mathbb{R})$  for  $n \geq 0$ .  $A$ ,  $A^\dagger$  and  $H$  all map  $\mathcal{S}(\mathbb{R})$  to itself. Now we compute: LECTURE 12, 2013-03-28.

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

This yields the following proposition.

**Proposition.**

$$[A, A^\dagger] = 1 \tag{5.1}$$

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

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

$$\langle h_m, h_n \rangle = \delta_{m,n}, \tag{5.4}$$

$$A h_n = \sqrt{n} h_{n-1}, \tag{5.5}$$

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

$$H h_n = n h_n, \quad n \geq 0, \tag{5.7}$$

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

$$\widehat{A_x \varphi} = i A_k \widehat{\varphi}, \tag{5.9}$$

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

**Proof.** Proof of (5.2) by induction:

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

“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  $\langle h_m, h_n \rangle = \frac{1}{\sqrt{m!}} \frac{n(n-1) \dots 1}{\sqrt{n!}} \langle 0, h_0 \rangle = 0$  if  $m > n$ . If  $m = n$ , we have  $\langle h_n, h_n \rangle = \frac{1}{\sqrt{n!}} \frac{n(n-1) \dots 1}{\sqrt{n!}} \langle h_0, h_0 \rangle = 1$ . 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

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

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

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

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  $\sup_{x \in \mathbb{R}} |h_n(x)|$ .

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

## 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  $(q_1, q_2) \in \Omega = \mathbb{R}^2$ .

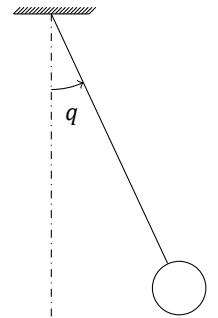


Figure 1. The configuration space  $\Omega = \mathbb{R}$  of a simple pendulum. Here  $q \in \mathbb{R}$ , as the pendulum can do a full turn—or several—around its pivot.

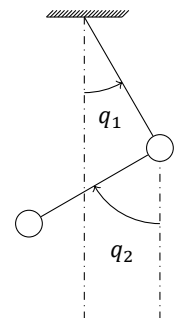


Figure 2. The configuration space  $\Omega = \mathbb{R}^2$  of a double pendulum, where  $q = (q_1, q_2) \in \mathbb{R}^2$ .

**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

$$q_i \in C^\infty(\Omega \times \mathbb{R}^n, \mathbb{R})$$

$$\Omega \times \mathbb{R}^n \ni (\mathbf{q}, \mathbf{p}) \mapsto q_i \in \mathbb{R}.$$

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

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

**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  $d\mathbf{s} = \begin{pmatrix} \frac{\partial s_j}{\partial x_k} \end{pmatrix}$  satisfies

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

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*  $\text{Sp}(2n, \mathbb{R}) := \{\mathbf{A} \in \text{GL}(2n, \mathbb{R}) \mid \mathbf{A}^\top = -J\mathbf{A}^{-1}J\}$ .  $\diamond$

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

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

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

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”<sup>7</sup>, 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

<sup>7</sup>*Gone with the Wind*, 1939, directed by Victor Fleming.

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  $h(q, p) = \frac{1}{2}(\frac{p^2}{m} + m\omega^2 q^2)$ , where  $(q, p) \in \mathbb{R} \times \mathbb{R}$ , 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.

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

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

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

You can check the Jacobi identity and all that. □

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

**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:

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

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

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<sup>8</sup>. Unless you can find Waldo, it's just a meaningless complicated picture. Finding Waldo gives it its meaning. Actually let's call the Hamiltonian  $W$ <sup>9</sup>. 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.

**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:*

$$\begin{aligned} P_k, Q_k &\in \text{Hermitian}(\mathcal{S}(\mathbb{R}^n)) \\ \frac{i}{\hbar} [P_k, P_l] &= 0, \\ \frac{i}{\hbar} [Q_k, Q_l] &= 0, \\ \frac{i}{\hbar} [P_k, Q_l] &= \delta_{kl} \mathbb{1}. \end{aligned}$$

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

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

<sup>8</sup>This 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.

<sup>9</sup>This notation was—fortunately—not actually used during the lecture, so we stick to  $H$ .

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:

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

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  $\mathbf{q} \wedge \mathbf{p} = (q_2 p_3 - q_3 p_2, q_3 p_1 - q_1 p_3, q_1 p_2 - q_2 p_1)^T$ . 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

$$\begin{aligned} h(p, q) &:= \frac{1}{2} \left( \frac{p^2}{m} + m\omega^2 q^2 \right), \\ H &= h \left( \frac{\hbar}{i} \frac{d}{dq}, q \right) = \frac{1}{2} \left( \frac{1}{m} \left( \frac{\hbar}{i} \frac{d}{dq} \right)^2 + m\omega^2 q^2 \right) \\ &= \frac{1}{2} \left( -\frac{\hbar^2}{m} \frac{d^2}{dq^2} + m\omega^2 q^2 \right). \end{aligned}$$

Now we look for the eigenvalues  $E$ ,  $H\psi = E\psi$ , for  $\psi \in \mathcal{S}(\mathbb{R})$ .

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. Write

$$x := \left( \frac{m\omega}{\hbar} \right)^{\frac{1}{2}} q,$$

we get

$$\frac{d}{dx} = \frac{dq}{dx} \frac{d}{dq} = \left( \frac{\hbar}{m\omega} \right)^{\frac{1}{2}} \frac{d}{dq}.$$

We now expand the eigenvalue equation from above,

$$(H\psi)(q) = E\psi(q) \Leftrightarrow -\frac{\hbar^2}{2m} \frac{d^2}{dq^2} \psi(q) + \frac{m\omega^2}{2} q^2 \psi(q) = E\psi(q).$$



Let

$$f(x) := \psi\left(\left(\frac{\hbar}{m\omega}\right)^{\frac{1}{2}} x\right).$$

The eigenvalue equation becomes

$$\begin{aligned}\frac{1}{2}\left(-\omega\hbar\frac{d^2}{dx^2} + \omega\hbar x^2\right)f(x) &= Ef(x), \\ \frac{1}{2}\left(-\frac{d^2}{dx^2} + x^2\right)f(x) &= \frac{E}{\hbar\omega}f(x).\end{aligned}$$

“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. What do I need to do?” *Somebody*<sup>10</sup> answers with a complicated circumlocution for ‘subtract one’. You should write novels. This is like Percy Bysshe Shelley when he talks about a skylark:

### To a Skylark

HAIL to thee, blithe spirit!  
Bird thou never wert—  
That from heaven or near it  
Pourest thy full heart  
In profuse strains of unpremeditated art.  
[...]

—Percy Bysshe Shelley.

I just say ‘it’s a skylark’. Here I just want the formula. We add a  $-1$  in the last expression:

$$\frac{1}{2}\left(-\frac{d^2}{dx^2} + x^2 - 1\right)f(x) = \left(\frac{E}{\hbar\omega} - \frac{1}{2}\right)f(x).$$

We rewrite that as

$$\textcolor{red}{H}f(x) = \lambda f(x), \text{ where } \lambda = \frac{E}{\hbar\omega} - \frac{1}{2}.$$

The red  $\textcolor{red}{H}$  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:

$$\begin{aligned}\textcolor{red}{H}h_n &= nh_n, \text{ where} \\ h_n &= \frac{1}{\sqrt{n!}}(A^\dagger)^n h_0, \\ h_0(x) &= \frac{1}{\sqrt[4]{\pi}}e^{-\frac{1}{2}x^2}.\end{aligned}$$

So here,

$$\begin{aligned}n &= \frac{E}{\hbar\omega} - \frac{1}{2}, \\ E_n &= \hbar\omega\left(n + \frac{1}{2}\right),\end{aligned}$$

and we have found the energy levels of the quantum harmonic oscillator. *The poet 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.

<sup>10</sup>He will henceforth be referred to as ‘the poet’.

Let us look at the wavefunction  $\varphi$ . It's normalised so that

$$\int_{\mathbb{R}} |\psi(q)|^2 dq = 1,$$

and we interpret this as the probability

$$P_{\psi}[a < q < b] = \int_a^b |\psi(q)|^2 dq.$$

There's no reason why it should work, but it does. The other thing is

$$P_{\psi}[a < k < b] = \int_a^b |\widehat{\psi}(k)|^2 dk.$$

What is the Heisenberg uncertainty principle? Define

$$\begin{aligned} E_{\psi}[A] &:= \langle \psi, A \psi \rangle, \\ \text{Var}_{\psi}[A] &:= E_{\psi}[(A - E_{\psi}[A]\mathbb{1})^2]. \end{aligned}$$

Here's the punchline:

$$\text{Var}_{\psi}[Q] \text{Var}_{\psi}[P] \geq \frac{\hbar^2}{4}.$$

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  $\rho_h$ . 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. *The professor being extremely sleep-deprived, he ends the lecture 25 minutes early, leading to widespread cheering and applause.*

LECTURE 17, 2013-04-23.

You've talked at length about  $\text{Mat}_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$ ,  $\text{Mat}_n(\mathbb{C}) \cong \mathbb{C}^{n^2}$ .  $n \times n$  matrices on  $\mathbb{K}$  are just elements of  $\mathbb{K}^{n^2}$ . The only reason we write them as square arrays is to equip these spaces with a—polynomial—multiplication,

$$(AB)_{ij} := \sum_{k=1}^n A_{ik} B_{kj}.$$

Manifolds are just  $\mathbb{R}^d$  or  $\mathbb{C}^d$  if what you are doing has anything to do with reality.

Let  $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ . We have the continuous *determinant* function,

$$\det : \text{Mat}_n(\mathbb{K}) \rightarrow \mathbb{K}.$$

Define

$$\begin{aligned} \text{GL}(n, \mathbb{K}) &:= \{\mathbf{g} \in \text{Mat}_n(\mathbb{K}) \mid \det \mathbf{g} \neq 0\} \\ &= \text{Mat}_n(\mathbb{K}) \setminus \underbrace{\{\mathbf{g} \in \text{Mat}_n(\mathbb{K}) \mid \det \mathbf{g} = 0\}}_{\text{closed as det is continuous and } \{0\} \text{ is closed.}} \end{aligned}$$

$GL(n, \mathbb{K})$  is an open subset of  $\mathbb{K}^{n^2}$ . Now let us look at the multiplication and inversion.

$$\begin{aligned} GL(n, \mathbb{K}) \times GL(n, \mathbb{K}) &\xrightarrow{c^\infty \text{ map}} GL(n, \mathbb{K}) \\ (g, h) &\mapsto gh \\ GL(n, \mathbb{K}) &\xrightarrow{c^\infty \text{ map}} GL(n, \mathbb{K}) \\ g &\mapsto \left( \frac{1}{\det g} ((-1)^{i+j} A_{ij}(g)) \right)_{(i,j) \in \{1, \dots, n\}} \end{aligned}$$

where  $A_{ij}(g)$  is the determinant of  $g$  with the  $i$ th row and  $j$ th column cut out. Thus  $(GL(n, \mathbb{K}), \text{matrix multiplication})$  is a group with smooth operations. What you do when you talk about Lie groups is the same, except you start with completely abstract manifolds. As physicists, it's unlikely that you'll work with Lie groups that are not very closely related to the following concept.

**Definition (Matrix Lie group).** A matrix Lie group  $G$  is a closed subgroup of  $GL(n, \mathbb{K})$  under the topology generated by the norm

$$\|g\| := \sqrt{\text{tr } gg^\dagger} = \left( \sum_{i,j=1}^n |g_{ij}|^2 \right)^{\frac{1}{2}}. \quad \diamond$$

You already know an example of one from last week, the continuous Heisenberg group

$$\text{Heisenberg}_n = \left\{ \left( \begin{array}{c|ccc|c} 1 & r_1 & \dots & r_n & t \\ \hline \mathbf{0} & & \mathbb{1}_n & & \begin{smallmatrix} s_1 \\ \vdots \\ s_n \end{smallmatrix} \\ \hline 0 & & \mathbf{0} & & 1 \end{array} \right) \mid \begin{array}{l} \mathbf{r}, \mathbf{s} \in \mathbb{R}^n, \\ t \in \mathbb{R} \end{array} \right\}.$$

Why is it closed? Look at the map

$$\mathbb{R}^{(n+2)^2} \supseteq [\mathbf{r}, \mathbf{s}; t] \mapsto (r_1, \dots, r_n, s_1, \dots, s_n, t) \in \mathbb{R}^{2n+1}.$$

Do I have to prove that it's continuous? The poet:—Isn't it open?—"Well, that's the problem with education. Does anyone remember the definition of connected?" *The poet answers with the definition of path-connected.*—"There was another definition, one with open sets and so on. What are the subsets of  $\mathbb{R}^n$  that are both open and closed?" Somebody else:—The half open, half closed intervals. *The professor whistles "La Vie en rose".*—"Sorry, nothing personal. I had to wait for my blood pressure to go down."

An open set  $\Omega$  of  $\mathbb{R}^n$  is a set such that every point is contained in an open ball inside  $\Omega$ . A closed set of  $\mathbb{R}^n$  is a set in which every Cauchy sequence converges. Here our group verifies both, so it is both open and closed in  $\mathbb{R}^{(n+2)^2}$ , and thus it is connected.

Now we look at a Lie group at the very bottom of the food chain. This is the plankton in the sea of algebraic objects. The reason we talk about it though is that to understand spin, you have to eat plankton sandwiches with  $SU(2)$  inside. And you have to look at the representations. That's what we need this representation crap for.

**Definition (Special unitary group of degree 2).** We define the group of unitary  $2 \times 2$  matrices with determinant 1,

$$GL(2, \mathbb{C}) \supseteq SU(2) := \{g \in GL(2, \mathbb{C}) \mid g^\dagger g = 1, \det(g) = 1\}. \quad \diamond$$

It is a subgroup of  $GL(2, \mathbb{C})$ , as

$$\begin{aligned} gh \in SU(2) &\Rightarrow (gh)^\dagger gh = h^\dagger g^\dagger gh = \mathbb{1} \\ &\Rightarrow gh \in SU(2), \\ g \in SU(2) &\Rightarrow g^\dagger g = \mathbb{1} \\ &\Rightarrow g^{-1} = g^\dagger \\ &\Rightarrow (g^{-1})^\dagger = g \\ &\Rightarrow (g^{-1})^\dagger g^{-1} = \mathbb{1} \\ &\Rightarrow g^{-1} \in SU(2). \end{aligned}$$

It is closed, as  $h \mapsto h^\dagger$  is a continuous function, and thus

$$\begin{aligned} \lim_{j \rightarrow \infty} g_j &= g \\ \Rightarrow \lim_{j \rightarrow \infty} g_j^\dagger &= g^\dagger \\ \Rightarrow \lim_{j \rightarrow \infty} g_j^\dagger g_j &= g^\dagger g = \mathbb{1}, \end{aligned}$$

and the determinant is a continuous function too. So you have a nice closed subgroup of  $GL(2, \mathbb{C})$  looking at you. Now I don't know how you feel about things looking at you, but it's looking at you.

“Do you remember what a compact set is?”—Any cover by open sets has a finite subcover.—“This is correct, but with this definition you have to look at all open covers, and check whether they have a finite subcover, and we don't have time for that. There is a nicer property, named after two Germans—actually it was one German and one Frenchman<sup>11</sup>.” The poet answers “closed and bounded”<sup>12</sup>, though he got the answer from someone else.—“You should be a novelist. You know how to incorporate ideas from others”. So  $SU(2)$  is a compact matrix Lie group, as

$$g \in SU(2) \Rightarrow \|g\| = \sqrt{\text{tr } gg^\dagger} = \sqrt{\text{tr } \mathbb{1}_2} = \sqrt{2}.$$

The fundamental reason we study compact sets is that it's like being finite without being finite. For instance you can extract a finite subcover from any open cover.

Now, after an hour of being yelled at and patronised, and all sorts of other wonderful things, we now have a compact matrix Lie group. We could have picked the circle group, but this has a nice property that the circle group doesn't have. The poet:—The circle group is abelian, this one isn't.—“You see? leave it to a poet”.

LECTURE 18, 2013-04-25.

Okay, so we have  $SU(2) \subseteq GL(2, \mathbb{C})$ , which is a compact matrix group. We inflicted upon you all sorts of things about finite groups. What is it about these groups that make them richer? Let's look at  $GL(n, \mathbb{K}) \subseteq \text{Mat}_n(\mathbb{K}) \cong \mathbb{K}^{n^2}$ . Now one thing that you've learned about is paths.

**Definition (Smooth path).** A smooth path in  $GL(n, \mathbb{K})$  is a smooth map

$$\gamma : (a, b) \rightarrow GL(n, \mathbb{K}). \quad \diamond$$

<sup>11</sup>This is the Heine-Borel theorem, after Heinrich Eduard Heine (1828/1881) and Félix Édouard Justin Émile Borel (1871/1956). Note that in France this is called the *théorème de Borel-Lebesgue* after the aforementioned Borel and an additional Frenchman, Henri Léon Lebesgue (1875/1941). We have found no Russians associated with this theorem, but interestingly, the theorem is called *лемма Гейне — Бореля* there, so it is only a lemma. Many mathematicians have actually contributed to this theorem, so taking those mentioned by *Wikipedia*, it could be called the Borel-Cousin-Dirichlet-Heine-Pincherle-Lebesgue-Schoenflies-Weierstrass theorem—that's 1 Italian, 3 French, 3 Prussians and Dirichlet, who was born in Napoleonic France in 1805 and became *de facto* Prussian in 1815. Do not confuse this Heine with Christian Johann Heinrich Heine (1797/1856), another poet.

<sup>12</sup>This is true in  $\mathbb{R}^n$ . For a general metric space, replace “closed and bounded” by “complete and totally bounded”.

**Definition (Tangent vector).** A tangent vector  $X$  to the matrix Lie group  $G$  at  $1 \in G$  is

$$X = \left. \frac{d}{dt} \gamma \right|_{t=0}$$

for some smooth path  $\gamma(t) \in G$ ,  $a < t < b$  and  $\gamma(0) = 1$ .  $\diamond$

The idea is the following. The group is a complex thing that you can't visualise, so I draw a complex thing that you can't visualise (Figure 3). You study it locally instead, by looking at the tangent space. You pick a curve lying in  $G$ , and as it lies in  $G$ —which it can do—it lies in  $GL(n, \mathbb{K})$ , so you can differentiate it,

$$\frac{d}{dt} g = \left( \frac{d}{dt} \gamma_{ij}(t) \right) \in \text{Mat}_n(\mathbb{K}).$$

We now define the Lie algebra of  $G$ .

**Definition (Lie algebra of a matrix Lie group).** Lie algebra of  $G$  is its tangent space at the identity,

$$\text{Lie}(G) := \{X \in \text{Mat}_n(\mathbb{K}) \mid X \text{ tangent to } G \text{ at } 1\}. \quad \diamond$$

The reason mathematicians like this is that everything you know, linear algebra, geometry, analysis—actually you don't need much analysis for this—conspire, legally, to make this object. Now we have a definition, so we see how it applies to our example  $SU(2)$ . Consider a path in  $SU(2)$ ,

$$\begin{aligned} \gamma : ]-\epsilon, \epsilon[ &\rightarrow SU(2), \\ \gamma(0) &= 1_2, \\ \gamma(t) &= \begin{pmatrix} \gamma_{11}(t) & \gamma_{12}(t) \\ \gamma_{21}(t) & \gamma_{22}(t) \end{pmatrix} \quad \text{for } -\epsilon < t < \epsilon. \end{aligned} \quad (6.1)$$

As the path lies in  $SU(2)$ ,

$$\begin{aligned} 1_2 &= \gamma(t)\gamma(t)^\dagger, \\ 0 &= \frac{d}{dt}(\gamma(t)\gamma(t)^\dagger) && \text{—'Who do you gotta call?'—Leibniz.} \\ &= \left( \frac{d}{dt} \gamma(t) \right) \gamma(t)^\dagger + \gamma(t) \left( \frac{d}{dt} (\gamma(t)^\dagger) \right) \\ &= \dot{\gamma}(t) \gamma(t)^\dagger + \gamma(t) \left( \frac{d}{dt} \gamma(t) \right)^\dagger. \end{aligned}$$

In particular, at  $t = 0$ ,

$$\begin{aligned} 0 &= \dot{\gamma}(0) \gamma(0)^\dagger + \gamma(0) \dot{\gamma}(0)^\dagger && \text{—'What's } \gamma(0)?\text{'—Zero.—'Zero! I} \\ &= \dot{\gamma}(0) + \dot{\gamma}(0)^\dagger. && \text{understand poetic license, I like} \\ &&& \text{poetic license, but saying a one is} \\ &&& \text{a zero is too much poetic license.} \end{aligned}$$

So we have shown

$$X \in \text{Lie}(SU(2)) \Rightarrow X + X^\dagger = 0 \quad (6.2)$$

The professor scrambles for a pointing stick in order to show the equalities he wants to refer to. “They removed it, they didn't want me to threaten you with it—ah! here it is.” This (6.1) is not a linear condition, *this* (6.2) is the infinitesimal version of that—and that's the fundamental idea in calculus: you take a nonlinear condition and you approximate it linearly. We have another nonlinear condition:

$$1 = \det \gamma(t) \Rightarrow 0 = \frac{d}{dt} \det \gamma(t). \quad (6.3)$$

For the moment, don't be a mathematician. You don't need to know how to differentiate determinants. You just differentiate this one, like a physicist—or a not so

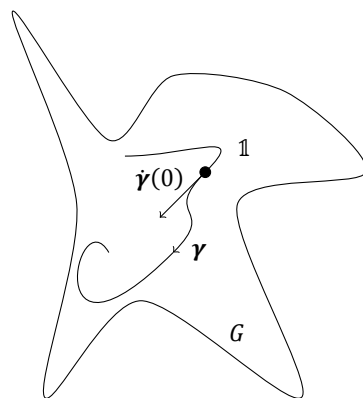


Figure 3. A complex thing that you can't visualise.

bourgeois and pedantic mathematician. A mathematician, at least one with a certain kind of education, would stop dead here. They'd develop a whole new theory of differentiating matrices, with their own journal, and they'd have conjectures, and prove them, and go on to prove things nobody wants to hear about.

$$\begin{aligned}
 0 &= \frac{d}{dt} \det \gamma(t) \\
 &= \frac{d}{dt} (\gamma_{11}(t)\gamma_{22}(t) - \gamma_{21}(t)\gamma_{12}(t)) && \text{—'Who do you gotta call?'} \\
 &= \dot{\gamma}_{11}(t)\gamma_{22}(t) + \gamma_{11}(t)\dot{\gamma}_{22}(t) && \text{—Leibniz.} \\
 &\quad - \dot{\gamma}_{12}(t)\gamma_{21}(t) - \gamma_{12}(t)\dot{\gamma}_{21}(t).
 \end{aligned}$$

At  $t = 0$ ,

$$0 = \left. \frac{d}{dt} \det \gamma(t) \right|_{t=0} = \dot{\gamma}_{11}(t) + \dot{\gamma}_{22}(t),$$

So the infinitesimal version of (6.3) is

$$\text{tr } \dot{\gamma}(0) = 0.$$

These two conditions on the tangent vectors motivate the following definition.

**Definition.** We call  $\mathfrak{su}(2)$  the set of skew-Hermitian  $2 \times 2$  complex matrices with trace 0,

$$\mathfrak{su}(2) := \{X \in \text{Mat}_2(\mathbb{C}) \mid X + X^\dagger = 0, \text{tr } X = 0\}. \quad \diamond$$

So far we have shown

$$X \in \text{Lie}(\text{SU}(2)) \Rightarrow X \in \mathfrak{su}(2).$$

But we don't know if there *are* any tangent vectors yet. We haven't looked at any actual smooth curves. —“How do we find smooth curves?”—Exponentiate.

If you tell me you have never heard about this I'll call you all liars, because when I tried to do the properties of the exponential map last semester you told me you had already seen that. I know, last semester was a long time ago, you don't remember anything.

Speaking of not remembering anything. Two Harvard professors—there are lots of Harvard professors, but I'm only going to talk about two of them—published a paper<sup>13</sup> saying that if your debt reaches 90 % of your GDP, your economy would slow down by 0.1 %—to show that there is no gender bias, one was of the masculine persuasion, and the other one was of the feminine persuasion. This has been used to justify austerity policies all over. A graduate student<sup>14</sup> in econometrics—to economists, that's like street sweeping—couldn't reproduce the results, and asked them some of their data. So they sent their Excel spreadsheet—that's mathematics for economists—and the student noticed basic arithmetic mistakes, and that Australia, Tasmania and New Zealand<sup>15</sup> were left out—and minus 0.1 % became plus 2.2 %. You could have done that too. Things are the same in mathematics and physics. Maybe I'll give you a list of danger signs when picking who to write a paper with. If they have been doing the same things for the past 15 years, *isch nöt guet!* Every time you publish a paper in one of the top journals in economics—up to 4 papers—your salary goes up by 25000 Swiss francs. It doesn't matter whether it's wrong. The science behind that is “not worth a tub of hot spit”<sup>16</sup>.—Do we have to take a break?

As the gentleman said here, “exponentiate”. Let  $X \in \mathfrak{su}(2)$ . We have

$$\begin{aligned}
 X + X^\dagger &= 0, \\
 \text{tr } X &= 0.
 \end{aligned}$$

<sup>13</sup>Carmen Reinhart and Kenneth Rogoff, Growth in a Time of Debt, in *American Economic Review: Papers & Proceedings*, 100.

<sup>14</sup>Thomas Herndon, Political Economy Research Institute, University of Massachusetts Amherst.

<sup>15</sup>The list of excluded countries is actually Denmark, Canada, Belgium, Austria and Australia.

<sup>16</sup>This is a reference to a quote by John Nance Garner on the vice-presidency of the United States. The original quote is: “not worth a bucket of warm piss”.

*Pointing to the poet*—“You told me you had done the properties of the exponential map.”—So it’s my fault.—“We’ve got to have a fall guy. It’s like *the Maltese Falcon*<sup>17</sup>.”

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!},$$

$$e^{-A} = (e^A)^{-1}.$$

Define the path

$$\gamma_X(t) := e^{tX}.$$

We have, with  $X \in \mathfrak{su}(2)$ ,

$$\begin{aligned}\gamma_X(t)^\dagger &= (e^{tX})^\dagger = e^{(tX)^\dagger} \\ &= e^{-tX} = (e^{tX})^{-1} = (\gamma_X(t))^{-1}, \\ \det \gamma_X(t) &= \det e^{tX} = e^{\operatorname{tr} tX} \\ &= e^{t \operatorname{tr} X} = 1.\end{aligned}$$

So we have the following FACT:  $\forall t \in \mathbb{R} \quad \forall X \in \mathfrak{su}(2) \quad \gamma_X(t) \in \mathrm{SU}(2)$ . Moreover,  $\gamma_X(0) = \mathbb{1}$ , so

$$X = \left. \frac{d}{dt} \gamma_X \right|_{t=0} \in \operatorname{Lie}(\mathrm{SU}(2)).$$

We have proven the following proposition.

**Proposition.** *The Lie algebra of  $\mathrm{SU}(2)$  is the set of skew-Hermitian  $2 \times 2$  complex matrices with trace 0,*

$$\operatorname{Lie}(\mathrm{SU}(2)) = \mathfrak{su}(2). \quad \square$$

I have called  $\mathfrak{su}(2)$  a “Lie algebra”. Where’s the bracket? It’s the commutator.

**Proposition.**  $(\mathfrak{su}(2), [\cdot, \cdot])$  is a Lie algebra.

**Proof.** Let  $X, Y \in \mathfrak{su}(2)$ .

$$\begin{aligned}\operatorname{tr}[X, Y] &= \operatorname{tr} XY - \operatorname{tr} YX = 0. \\ [X, Y]^\dagger &= (XY - YX)^\dagger \\ &= Y^\dagger X^\dagger - X^\dagger Y^\dagger \\ &= (-Y)(-X) - (-X)(-Y) \\ &= [Y, X] = -[X, Y].\end{aligned} \quad \square$$

[The poet] just proved it’s a real vector space, without even knowing it. Consider two tangent vectors  $\dot{\gamma}(0)$  and  $\dot{\gamma}'(0)$  at  $\mathbb{1}$ . Then we can construct a smooth path whose tangent vector at  $\mathbb{1}$  is any real linear combination of those.  $\gamma(at)$  just means that we go slower or faster along the path  $\gamma$ . —“Who do you gotta call?”—Chain rule; but it’s not a name.—“Just call it Chain’s rule.”

$$\left. \frac{d}{dt} (\gamma(at) \gamma'(bt)) \right|_{t=0} = a \dot{\gamma}(0) + b \dot{\gamma}'(0).$$

You may ask, “why is the commutator,  $[\dot{\gamma}(0), \dot{\gamma}'(0)] = \dot{\gamma}(0) \dot{\gamma}'(0) - \dot{\gamma}'(0) \dot{\gamma}(0)$ , the Lie bracket?”. How do we get a minus in the derivative? The poet:—The guy behind me whispers ‘quotient’.—“It’s a girl.”—Oh! I’m sorry. *He turns around to see who was whispering.* Evidently.—“You’re a real poet. You had me worried for a moment. Poets are supposed to know that stuff.” So, the derivative of a quotient is

$$\frac{d}{dx} f(x)^{-1} = (-1) \frac{f'(x)}{f(x)^2}. \quad \text{‘There’s no analogy, just differentiate the damn function!’}$$

<sup>17</sup> *The Maltese Falcon*, 1941, directed by John Huston.

The *group* commutator of  $\gamma(t)$  and  $\gamma'(t)$  is  $\gamma(t)\gamma'(t)\gamma(t)^{-1}\gamma'(t)^{-1}$ . If we differentiate it, we get

$$\left. \frac{d}{dt} \gamma(t)\gamma'(t)\gamma(t)^{-1}\gamma'(t)^{-1} \right|_{t=0} = \dot{\gamma}(0) + \dot{\gamma}'(0) - \dot{\gamma}(0) - \dot{\gamma}'(0) = 0,$$

so that doesn't work. We have to differentiate twice. The reason is that we're just taking the first term in the expansion of  $\log(\gamma(t)\gamma'(t)\gamma(t)^{-1}\gamma'(t)^{-1})$ , and we want more terms.

$$\log(e^{sA}e^{sB}) = \log(1 + (e^{sA}e^{sB} - 1))$$

Now we plug this into the Taylor series of the logarithm, which is—*after several attempts at getting the signs right*—

$$\log(1+x) = -\sum_{k=1}^{\infty} \frac{(-x)^k}{k}.$$

The moment you memorise you're dead, unless you're an actor like him, or you're counting cards in Vegas. But if you count cards and they catch you, you're not in good shape either. We get, for  $A, B$  in the Lie algebra,

$$\begin{aligned} \log(e^A e^B) &= A + B + \frac{1}{2}[A, B] + \dots \\ \log(e^A e^B e^{-A} e^{-B}) &= [A, B] + \dots \end{aligned}$$

Here's the question, ladies and gentlemen.

**Proposition.**  $\exp : \mathfrak{su}(2) \rightarrow \mathrm{SU}(2)$  is

- (A) *injective*;
- (B) *surjective*;
- (C) *bijective*;
- (D) *all of the above*.

*What do you pick? What can you hit with the exponential map? Quiet man:— $\mathbb{C}^*$ . So it's surjective. It can't be bijective, so it's (B).*

**Proof.** [...] “Where do the eigenvalues of a unitary operator lie? Please, after a whole year”—On the unit circle.—“Thank God.”  $\square$

Next time we'll do something more fun. We'll talk about spin, and we'll talk about quarks.

This will involve a lot of calculations, you're not going to like it. I'll tell you why you'll love it.

—“Common, you keep mixing real and complex. If  $X$  is skew-Hermitian,  $X^\dagger = -X$ . What is  $(iX)^\dagger$ ?” *The professor whistles “la Vie en rose”.*

$$(iX)^\dagger = iX$$

Now these groups should appeal to physicists.  $\mathbb{R}^3$  is both *what* and *what*? It's both homogeneous, the laws of physics are the same in Boston and in Zürich, and isotropic. What does this have to do with groups? What it looks like at the origin is what it looks like everywhere: multiplication by  $g$  is an invertible smooth map.

The first thing you do with  $\mathrm{SU}(2)$  is we look at the electron. Let's just have an electron sitting there in the Universe. We turn everything else off. How many degrees of freedom do we have? Two, I've yelled at you last semester about this, trying to get *some* response.

$$\begin{pmatrix} \psi_\uparrow \\ \psi_\downarrow \end{pmatrix} \in \mathbb{C}^2$$

LECTURE 19, 2013-04-30.

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



—“Is that how you write it?”—Everybody has their own notation.—“Okay, so I’ll use my own.”

$$\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{C}^2,$$

where the up and down states are respectively

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

If you rotate the spin space—by applying an element of  $SU(2)$  to it—it doesn’t change anything. It’s a law of physics. —“Have you ever heard of spin one in physics?”—We only talked about spin one half, although we mentioned that some exotic particles had other spins.—“Has anyone heard of Helium 3?”—It’s a liquid.—“It depends on the temperature. You may be a poet, but the temperature has to come into the equation. Do you know about Percy Bysshe Shelley? Has anyone heard of *Frankenstein*? It was written by Mary Wollstonecraft Shelley. This Mary Shelley had a husband, Percy Bysshe Shelley, who was a poet, and who wrote about the Moon:

### The Moon

AND, like a dying lady lean and pale,  
Who totters forth, wrapp’d in a gauzy veil,  
Out of her chamber, led by the insane  
And feeble wanderings of her fading brain,  
The moon arose up in the murky east  
A white and shapeless mass.

Art thou pale for weariness  
Of climbing heaven and gazing on the earth,  
Wandering companionless  
Among the stars that have a different birth,  
And ever changing, like a joyless eye  
That finds no object worth its constancy?

—Percy Bysshe Shelley.

So you can write poetry about modern science—‘the stars that have a different birth’ was modern science at the time—but you’ve got to have temperature as part of the equation.—I’m not here for poetry.—“You may not be here for that, but maybe that’s what you are, like this basketball player<sup>18</sup>.”

Spin one has three dimensions,

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{C}^3.$$

The fact of the matter is, what you want is an irreducible representation  $\rho_n$  of  $SU(2)$  of dimension  $n \geq 2$ . Now there are two things we need. We need that there is an irreducible representation, and we need to combine them. In economics, they have equilibria. The problem is, they have too many, so they don’t know which one they’re talking about. And that’s the problem with string theory. The trouble with string theory is, it’s like economics: they have too many ground states.

—“You go to that guy, who’s half Swiss and half German and who likes to calculate”—Well, only sometimes.—“You’re only half Swiss and half German part of the time? So we have some sort of quantum European here.”

<sup>18</sup>This is a reference to NBA player Jason Collins, who recently revealed he is gay. “I didn’t set out to be the first openly gay athlete playing in a major American team sport. But since I am, I’m happy to start the conversation.”

Has anybody seen *the four hundred blows*<sup>19</sup>? Well they dress with scarfs like him—it's not a happy film, by the way.

Now if we only have one quark in the Universe—we turn everything else off—it has inside it *three* degrees of freedom, *up*, *down* and *strange*.

$$\begin{pmatrix} u \\ d \\ s \end{pmatrix} \in \mathbb{C}^3, \text{ the flavour space.}$$

The poet:—Why up, down and strange? Why not give them greek names, so it looks important and fancy?—“Gell-Mann came up with these names. He was—is he still alive?”—He is still alive.—“he is an obstreperous guy. Do you know what obstreperous means? If you don't, look it up.”

Heisenberg—*he's* dead, like Franco—considered a part of that, and he called it *isospin*. It described the symmetry between up and down, and so it used SU(2). Indeed, if you have  $\mathbf{U} \in \text{SU}(2)$ , then

$$\left( \begin{array}{cc|c} \mathbf{U} & & 0 \\ & & 0 \\ \hline 0 & 0 & 1 \end{array} \right)$$

is in SU(3).

The law of physics is that the antiquark lives in the dual flavour space  $\mathbb{C}^{3*}$ . The flavour space is a representation  $\mathbf{3}$  of SU(3), and the dual flavour space is the dual representation  $\bar{\mathbf{3}}$ . So if we have a quark-antiquark pair—a meson—, it lies in the tensor product  $\mathbb{C}^3 \otimes \mathbb{C}^{3*}$ , or  $\mathbf{3} \otimes \bar{\mathbf{3}}$ . Now the dimension of this representation is  $3 \times 3 = 9$ , and it decomposes into the trivial—one-dimensional—representation  $\mathbf{1}$ , and another—eight-dimensional—irreducible representation, which we call  $\mathbf{8}$ , that is,  $\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}$ . Gell-Mann came up with this, and he called it the eightfold way. I hope I have convinced you we need to look at this stuff. This is not some bourgeois pedantic mathematician stuff. When I do it next Friday you cannot complain. You can leave the room, but you cannot complain.

LECTURE 20, 2013-05-02.

Alright, here we are. I'm going to talk to you about  $\mathfrak{su}(2)$  and SU(2) at a rate at which you can absorb the information—I know that's a slow rate. I'm going to give you a lecture today, just using the blackboard and talking.

$$\begin{aligned} \text{SU}(2) &:= \{g \in \text{GL}(2, \mathbb{C}) \mid gg^\dagger = 1, \det g = 1\} \\ \mathfrak{su}(2)(2) &:= \{x \in \mathfrak{gl}(2, \mathbb{C}) \mid x^\dagger + x = 0, \text{tr } x = 0\} \end{aligned}$$

We have a little lemma.

**Lemma.**

$$\frac{i}{\sqrt{2}}\sigma_k = \frac{i}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \frac{i}{\sqrt{2}}\sigma_k = \frac{i}{\sqrt{2}} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \frac{i}{\sqrt{2}}\sigma_k = \frac{i}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

is an orthonormal basis for  $(\mathfrak{su}(2), \langle \cdot, \cdot \rangle)$ , where  $\langle \mathbf{A}, \mathbf{B} \rangle := \text{tr}(\mathbf{A}\mathbf{B}^\dagger)$ .

**Proof.** Calculate

$$\begin{aligned} \left\langle \frac{i\sigma_k}{\sqrt{2}}, \frac{i\sigma_k}{\sqrt{2}} \right\rangle &= \text{tr} \left( \frac{i}{\sqrt{2}}\sigma_k \frac{-i}{\sqrt{2}}\sigma_k^\dagger \right) \\ &= \frac{1}{2} \text{tr } \sigma_k^2 = \frac{1}{2} \text{tr} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1, \end{aligned}$$

etc. □

So you have an orthonormal basis. It's nice to have an orthonormal basis.

**Definition.**  $\text{Ad}(g)x := gxg^{-1}$ , for  $g \in \text{SU}(2)$  and  $x \in \mathfrak{su}(2)$ . Two by two matrix, two by two matrix, two by two matrix. It's well defined. ◊

<sup>19</sup>*Les quatre cents coups*, 1959, directed by François Truffaut.

**Proposition.**  $\mathbf{Ad} : \mathrm{SU}(2) \rightarrow \mathrm{U}(\mathfrak{su}(2)) := \mathrm{U}((\mathfrak{su}(2), \langle \cdot, \cdot \rangle))$  is an irreducible representation of  $\mathrm{SO}(2)$ .

**Proof.** Let us prove  $\mathbf{Ad}(\mathbf{g}) \in \mathrm{U}(\mathfrak{su}(2))$ :

$$\begin{aligned} \langle \mathbf{Ad}(\mathbf{g})x, \mathbf{Ad}(\mathbf{g})y \rangle &= \mathrm{tr}((\mathbf{Ad}(\mathbf{g})x)(\mathbf{Ad}(\mathbf{g})y)^\dagger) && \mathbf{g} \text{ and } \mathbf{g}^{-1} \text{ kill each other off—no,} \\ &= \mathrm{tr}(\mathbf{g}x\mathbf{g}^{-1}(\mathbf{g}y\mathbf{g}^{-1})^\dagger) && \text{let's use a more poetic term. They} \\ &= \mathrm{tr}(\mathbf{g}x\mathbf{g}^{-1}(\mathbf{g}^{-1})^\dagger y^\dagger \mathbf{g}^\dagger) && \text{merge into the identity. Then we} \\ &= \mathrm{tr}(\mathbf{g}x\mathbf{g}^{-1}\mathbf{g}y^\dagger \mathbf{g}^{-1}) && \text{bring the } \mathbf{g}^{-1} \text{ to the front, and} \\ &= \mathrm{tr}(xy^\dagger) = \langle x, y \rangle. && \text{again, they merge into the iden-} \\ & && \text{tity.} \end{aligned}$$

I hope you can handle all that merging. There's something I didn't do! I need to show  $\mathbf{Ad}(\mathbf{g})x \in \mathfrak{su}(2)$ . —“I have to show that it somehow *metamorphoses* into  $\mathfrak{su}(2)$ .”—Maybe *mirrors* would be a good metaphor.—“No. I don't like mirrors.”

$$\begin{aligned} (\mathbf{Ad}(\mathbf{g})x)^\dagger &= (\mathbf{g}x\mathbf{g}^{-1})^\dagger \\ &= (\mathbf{g}^{-1})^\dagger x^\dagger \mathbf{g}^\dagger \\ &= \mathbf{g}x^\dagger \mathbf{g}^{-1} = -\mathbf{g}x\mathbf{g}^{-1}. \end{aligned}$$

The trace is *ewig entlang Konjugationsklassen*—it's *much* easier to be a bad poet than a bad mathematician. This is a basic fact, as our algebraic conscience would tell us. It's as basic as “you don't eat soup with a fork”. Isn't it? *Claude nods*. It's good to have an algebraic conscience, because I don't have one.

Now we prove that it's a representation,  $\mathbf{Ad}(\mathbf{g})\mathbf{Ad}(\mathbf{h})x = \mathbf{Ad}(\mathbf{gh})x$  for all  $\mathbf{g}, \mathbf{h} \in \mathrm{SU}(2)$  and  $x \in \mathfrak{su}(2)$ .

$$\begin{aligned} \mathbf{Ad}(\mathbf{g})(\mathbf{Ad}(\mathbf{h})x) &= \mathbf{g}(\mathbf{h}x\mathbf{h}^{-1})\mathbf{g}^{-1} \\ &= (\mathbf{gh})x(\mathbf{gh})^{-1} = \mathbf{Ad}(\mathbf{gh})x, \\ \mathbf{Ad}(\mathbf{1})x &= x. \end{aligned}$$

□

**Proposition.**  $(\mathfrak{su}(2), \langle \cdot, \cdot \rangle) \cong (\mathbb{R}^3, \cdot)$ . The isomorphism is given by

$$\begin{aligned} \mathbb{R}^3 \ni \xi &\mapsto \xi_1 \frac{i}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \xi_3 \frac{i}{\sqrt{2}} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \xi_2 \frac{i}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \frac{i}{\sqrt{2}} \begin{pmatrix} \xi_1 & \xi_2 - i\xi_3 \\ \xi_2 + i\xi_3 & -\xi_1 \end{pmatrix}. \end{aligned}$$

**Proof.** The poet:—We already encountered this formula last year.—“That's right, but I certainly didn't expect you to remember it.” *the professor holds on to a table and breathes deeply*. “Sorry. I was beginning to hyperventilate. I certainly didn't expect you to remember this.”—Isn't this also the measurement operator of a Stern-Gerlach box?—“A what? [...] Yes, that's right.”

$$\left\langle \frac{i}{\sqrt{2}} \xi \cdot \sigma, \frac{i}{\sqrt{2}} \lambda \cdot \sigma \right\rangle = \xi_1 \lambda_1 + \xi_2 \lambda_2 + \xi_3 \lambda_3.$$

□

$$(\mathbf{Ad}(\mathbf{g}))_{kl} = \left\langle \mathbf{Ad}(\mathbf{g}) \frac{i}{\sqrt{2}} \sigma_k, \frac{i}{\sqrt{2}} \sigma_l \right\rangle$$

The poet:—I'm not sure about the order of the  $k$  and  $l$  I dictated you.—“Let's just check! Here's how you do linear algebra.”

$$\mathbf{Ad}(\mathbf{g})x = \mathbf{Ad}(\mathbf{g}) \sum_{k=1}^3 \langle x, \phi_k \rangle \phi_k.$$

I don't want to write  $\frac{i}{\sqrt{2}}$  all the time, so I defined

$$\phi_k := \frac{i}{\sqrt{2}} \sigma_k.$$

—“Algebraic conscience, do you agree? I'm just projecting onto an orthogonal basis.” *Claude looks confused* “I'll just give him a few seconds to give us the go-ahead.” *Claude points out a mistake.* “It's good to have an algebraic conscience. I don't have an algebraic conscience—I don't have a conscience at all, but that's another matter.”

$$\mathbf{Ad}(g)x = \sum_{k=1}^3 \langle x, \phi_k \rangle \mathbf{Ad}(g)\phi_k.$$

Let us look at  $\mathbf{Ad}(g)\phi_k$ .

$$\mathbf{Ad}(g)\phi_k = \sum_{l=1}^3 \langle \mathbf{Ad}(g)\phi_k, \phi_l \rangle \phi_l.$$

Substituting,

$$\mathbf{Ad}(g)x = \sum_{k,l=1}^3 \langle \mathbf{Ad}(g)\phi_k, \phi_l \rangle \langle x, \phi_k \rangle \phi_l.$$

Algebraic conscience, do you agree? You see, I'm doing this to show that you don't have to memorise this. If you memorise, you always make a mistake. We have proven:

$$\mathbf{Ad}(g) = (\langle \mathbf{Ad}(g)\phi_l, \phi_k \rangle).$$

Moreover,

$$\begin{aligned} \mathbf{Ad}(g) \mathbf{Ad}(g)^\dagger &= \mathbb{1}_{\mathfrak{su}(2)}. \\ \langle \mathbf{Ad}(g)\phi_l, \phi_k \rangle &= \langle \phi_k, \mathbf{Ad}(g)\phi_l \rangle \\ &= \langle (\mathbf{Ad}(g))^\dagger \phi_k, \phi_l \rangle \\ &= \langle \mathbf{Ad}(g)^{-1} \phi_k, \phi_l \rangle \\ &= \langle \mathbf{Ad}(g^{-1}) \phi_k, \phi_l \rangle \dots \end{aligned}$$

—“I want to show it's real. It should be real, right?” Claude:—The vector space  $\mathfrak{su}(2)$  is a real vector space.  $\mathbf{Ad}(g)$  is a transformation of it.—“We have to take a break, right?”

So what I want to show is that  $\mathbf{Ad} \in \mathrm{SO}((\mathfrak{su}(2), \langle \cdot, \cdot \rangle), 3)$ , where  $(\mathfrak{su}(2), \langle \cdot, \cdot \rangle)$  is a real inner product space.  $\langle x, y \rangle = \mathrm{tr}(xy^\dagger) = -\mathrm{tr}(xy)$  is real linear in both variables.  $\mathbf{Ad}(g)x := gxg^{-1}$  is a real linear transformation of the real inner product space  $(\mathfrak{su}(2), \langle \cdot, \cdot \rangle)$ .

Claude:—We checked that it's an orthogonal matrix. So we *know* that.—“Yes. We checked  $\langle \mathbf{Ad}(g)x, \mathbf{Ad}(g)y \rangle = \langle x, y \rangle$  for  $g \in \mathrm{SU}(2)$ ,  $x, y \in \mathfrak{su}(2)$ , in other words,  $\mathbf{Ad}(g) \in \mathrm{O}((\mathfrak{su}(2), \langle \cdot, \cdot \rangle))$ . We checked that in complete detail.”

The poet is right, it's irreducible because it has no invariant subspace. We now know—hey, how do you like that, ‘we now know’, all you have to do is to add a ‘k’ in front of the ‘n’—that  $\mathrm{SU}(2) \ni g \mapsto \mathbf{Ad}(g) \in \mathrm{O}((\mathfrak{su}(2), \langle \cdot, \cdot \rangle))$ .

We still have to prove that it's in  $\mathrm{SO}((\mathfrak{su}(2), \langle \cdot, \cdot \rangle))$ . We already know that it's orthogonal, so we only have to prove that its determinant is 1. As it's orthogonal, its determinant can only be 1 or  $-1$ . Now the determinant is a continuous function. Can a continuous function take both the values 1 and  $-1$ , and nothing in between?—No.—“Well, that's not quite true. The domain has to have a property...”—Connected.—“Exactly. Is  $\mathrm{SU}(3)$  connected? In the notes from last semester...”

The poet:—They are not available online.—“Yes they are”<sup>20</sup>. In *Finite Dimensional Quantum Mechanics*, on page eight, there is a section called “Unitary Groups”, and there, you have the fact that it’s connected. Namely, you have—in an exercise—the fact that  $SU(2)$  is homeomorphic to the 3-sphere  $S^3$ . As physicists, you don’t quite know what the 3-sphere is, but you know what the 2-sphere looks like: it’s round, and it’s connected.”

We have proven the following.

**Proposition.** *The map  $\mathbf{Ad} : SU(2) \rightarrow SO(\mathfrak{su}(2), \langle \cdot, \cdot \rangle) \cong SO(3)$  is a surjective homomorphism, and*

$$\begin{aligned} g \in \mathbf{Ad}^{-1}(R) &\Rightarrow -g \in \mathbf{Ad}^{-1}(R), \\ \forall R \in SO(3) \quad |\mathbf{Ad}^{-1}(R)| &= 2, \end{aligned}$$

that is, it is a double covering.

**Proof.** We have done the fact that the determinant is one. How do we prove it’s surjective? When we have proven that, we want to know how many people go to each rotation, *wieviele blätter diese Überlagerung hat*. So we have to go door-to-door and ask. How the hell are we going to prove that? Or how are we bloody well going to prove that?—you see, in the thirties, ‘bloody well’ was really vulgar. Now if you say that, you just sound like a bad British actor. Look at the maps

$$\begin{aligned} \mathbf{a} \in \mathfrak{su}(2) &\mapsto e^{\mathbf{a}} \in SU(2), \\ \mathbf{a} \in \mathfrak{su}(2) &\mapsto \mathbf{Ad}(e^{\mathbf{a}}) \in SO(3). \end{aligned}$$

—“How do you characterise a rotation?”—You have to have an axis and angle.—“Exactly. It can be an axis of evil, but you need an axis and an angle. Now this will amuse you. What is the definition of  $\mathbf{Ad}(e^{\mathbf{a}})\mathbf{A}$ ?” The poet:— $e^{\mathbf{a}}\mathbf{A}(e^{\mathbf{a}})^{-1} = \mathbf{A}e^{\mathbf{a}}e^{-\mathbf{a}} = \mathbf{A}$ —“Does this commute? The poet tells me it commutes. Is there an anti-poet?”—I have the reason why it commutes.—“You are the reason why it commutes?”—No, I *have* the reason.

$\mathbf{a}$  is the axis of the special orthogonal linear transformation  $\mathbf{Ad}(e^{\mathbf{a}})$ .  $\mathbf{Ad}(e^{\phi\mathbf{a}})$  has axis  $\mathbf{a}$ . As  $\phi$  changes, the angle has to change. Pick a  $\mathbf{b} \in \mathfrak{su}(2)$  perpendicular to  $\mathbf{a}$  such that  $|\mathbf{b}| = 1$ . We have

$$\begin{aligned} \langle \mathbf{a}, \mathbf{b} \rangle &= 0, \\ \langle \mathbf{b}, \mathbf{b} \rangle &= 1, \\ \langle \mathbf{b}, \mathbf{Ad}(e^{\phi\mathbf{a}}\mathbf{b}) \rangle &= \cos \alpha. \end{aligned}$$

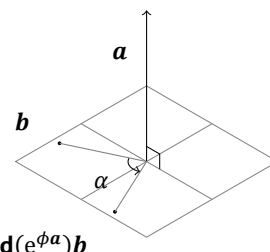
—“What is  $\alpha$  as a function of  $\phi$ ?” The poet:—Traditionally,  $\phi$  is the angle, so the angle must be  $\phi$ .—“You’re making the *Candide* argument. Pangloss said “the reason we have noses is to keep up our eyeglasses”<sup>21</sup>. But speaking *culturally*  $\alpha$  could be  $\phi$  or  $2\phi$ . It’s a little proposition that it’s just  $\phi$ .” Claude:—I think it’s  $2\phi$ .—“I think this is an excellent exercise”. So there is a beautiful map from the beautiful poetic space  $SU(2)$  onto the space  $SO(3)$  where we, as sinners, all live.  $\square$

I learned yesterday that there was no lecture on Thursday, so I’ll try to do something that fits in the hour. Let

$$X \in \mathfrak{so}(3) := \{X \in \mathfrak{gl}(3, \mathbb{R}) \mid X^T + X = 0\}.$$

We have  $x_{ji} = -x_{ij}$ , and thus  $x_{ii} = 0$ , so

$$\begin{aligned} (e^X)^T &= e^{X^T} = e^{-X} = (e^X)^{-1}, \\ \det(e^X) &= e^{\text{tr} X} = e^0 = 1, \text{ so we have proven} \\ \mathfrak{so}(3) \ni X &\mapsto e^X \in SO(3). \end{aligned}$$



**Figure 4.**  $\mathbf{Ad}(e^{\phi\mathbf{a}})$  is a rotation around  $\mathbf{a}$  of angle  $\alpha$ . What is  $\alpha$  as a function of  $\phi$ ?

LECTURE 21, 2013-05-07.

<sup>20</sup>The lecture notes from last semester are still available online at <http://www.math.ethz.ch/~gruppe5/group5/lectures/mmp/hs12/Files/>.

<sup>21</sup>« Il est démontré, disait-il, que les choses ne peuvent être autrement : car tout étant fait pour une fin, tout est nécessairement pour la meilleure fin. Remarquez bien que les nez ont été faits pour porter des lunettes ; aussi avons-nous des lunettes. [...] »—Voltaire, *Candide, ou l'Optimisme*.

You were probably told that this is an infinitesimal rotation—*das ist eine infinitesimale Drehung*—but we are going to prove it. For  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ , define

$$[\mathbf{x}, \mathbf{y}]_{\mathbb{R}^3} := \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix},$$

$$\mathbf{L}_x := \begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix}.$$

We have

$$\mathbf{L}_x^\top = -\mathbf{L}_x.$$

**Lemma.**  $\mathbb{R}^3 \ni \mathbf{x} \mapsto \mathbf{L}_x \in \mathfrak{so}(3)$  is an isomorphism of Lie algebras,  $\mathbf{L}_{[\mathbf{x}, \mathbf{y}]_{\mathbb{R}^3}} = [\mathbf{L}_x, \mathbf{L}_y]$ , and you can easily check that it's surjective and blah blah blah blah.

**Proof.** You just have to prove that the kernel is empty. You had something about that last year, when you were younger and things were still partially commutative. —“They did that last year, didn't they?” Claude:—Yes, yes.—“Should we do that again?”—No, no!—“I hope that this rings a bell, or at least a—what's it called, the thing from Proust, *À la recherche du temps perdu*”—A *madeleine*.—“Yes, exactly, a *madeleine*. That's in *Du côté de chez Swann*.<sup>22</sup>” □

—“Do you agree? You look like you don't agree.”—It's morning.—“It depends on your job. I see you're not a nocturnal animal. If you want, you can come back at two, I'll be there. I don't think [Claude] will come though. Will you?” Claude:—Nooo...

$$\mathbf{L}_x \mathbf{y} = [\mathbf{x}, \mathbf{y}]_{\mathbb{R}^3}$$

$$\mathbf{L}_x \mathbf{x} = 0$$

$$e^{\mathbf{L}_x} = \sum_{l \geq 0} \frac{\mathbf{L}_x^l}{l!}$$

$$e^{\mathbf{L}_x} \mathbf{x} = \sum_{l \geq 0} \frac{\mathbf{L}_x^l}{l!} \mathbf{x} = \mathbb{1} \mathbf{x} = \mathbf{x}, \text{ that is, } e^{\mathbf{L}_x} \text{ is a rotation around } \mathbf{x}.$$

$$\mathbf{L}_x^2 = \begin{pmatrix} -(x_3^2 + x_2^2) & x_1 x_2 & x_1 x_3 \\ x_1 x_2 & -(x_1^2 + x_3^2) & x_2 x_3 \\ x_1 x_3 & -x_2 x_3 & -(x_1^2 + x_2^2) \end{pmatrix}$$

$$\mathbf{L}_x^3 = -|\mathbf{x}|^2 \mathbf{L}_x$$

and by induction

$$\mathbf{L}_x^{2m+1} = (-|\mathbf{x}|^2)^m \mathbf{L}_x \quad \text{for } m \geq 0,$$

$$\mathbf{L}_x^{2m} = (-|\mathbf{x}|^2)^{m-1} \mathbf{L}_x^2 \quad \text{for } m \geq 1.$$

So, we sum.

$$\begin{aligned} e^{\mathbf{L}_x} &= \mathbb{1} + \sum_{m \geq 1} \frac{1}{(2m)!} \mathbf{L}_x^{2m} + \sum_{m \geq 0} \frac{1}{(2m+1)!} \mathbf{L}_x^{2m+1} \\ &= \mathbb{1} + \sum_{m \geq 1} \frac{(-|\mathbf{x}|^2)^{m-1}}{(2m)!} \mathbf{L}_x^2 + \sum_{m \geq 0} \frac{(-|\mathbf{x}|^2)^m}{(2m+1)!} \mathbf{L}_x \end{aligned}$$

<sup>22</sup>“Et dès que j'eus reconnu le goût du morceau de madeleine trempé dans le tilleul que me donnait ma tante (quoique je ne susse pas encore et dusse remettre à bien plus tard de découvrir pourquoi ce souvenir me rendait si heureux), aussitôt la vieille maison grise sur la rue, où était sa chambre, vint comme un décor de théâtre s'appliquer au petit pavillon donnant sur le jardin, qu'on avait construit pour mes parents sur ses derrières (ce pan tronqué que seul j'avais revu jusque-là); et avec la maison, la ville, la Place où on m'envoyait avant déjeuner, les rues où j'allais faire des courses depuis le matin jusqu'au soir et par tous les temps, les chemins qu'on prenait si le temps était beau.”—Marcel Proust, *À la recherche du temps perdu*.

What is the Taylor series of the sine? You think it's not worth knowing that stuff, but it is. *Somebody says*  $x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$ . Just give me the series!

$$\begin{aligned}\sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \\ &= \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m+1}}{(2m+1)!}\end{aligned}$$

So we get

$$e^{L_x} = \mathbb{1} - \frac{1}{|x|^2} \sum_{m \geq 1} \frac{(-|x|^2)^m}{(2m)!} L_x^2 + \frac{1}{|x|} L_x \sin |x|.$$

Now what is the Taylor series of the cosine? You don't remember the series, but you seem to be pretty good at remembering the first terms. *The professor whistles "La Vie en rose" while writing the series.*

$$\begin{aligned}\cos x &= 1 - \frac{x^2}{2!} + \dots \\ &= \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m}}{2m!}\end{aligned}$$

We get

$$e^{L_x} = \mathbb{1} + \frac{1 - \cos |x|}{|x|^2} L_x^2 + \frac{\sin |x|}{|x|} L_x.$$

So we have this very nice formula—and we don't need Euler angles. The calculation is so simple that going to the North pole doesn't bring you anything. We have  $x \cdot (x_2, -x_1, 0)^\top$ , and  $(e^{L_x}(x_2, -x_1, 0)) \cdot (x_2, -x_1, 0) = |(x_2, -x_1, 0)|^2 \cos \alpha$ . Once we have the angle  $\alpha$ , we have the axis and the angle, so we have a rotation. Why didn't I take the vector product with an arbitrary vector? I want to actually calculate with this. We will need the following vectors:

$$\begin{aligned}L_x \begin{pmatrix} x_2 \\ -x_1 \\ 0 \end{pmatrix} &= \begin{pmatrix} x_3 x_1 \\ x_2 x_3 \\ -x_1^2 - x_2^2 \end{pmatrix}, \\ L_x^2 \begin{pmatrix} x_2 \\ -x_1 \\ 0 \end{pmatrix} &= \begin{pmatrix} -x_2 x_3^2 - x_2(x_1^2 + x_2^2) \\ x_3^2 x_1 + x_1(x_1^2 + x_2^2) \\ -x_1 x_2 x_3 + x_1 x_2 x_3 \end{pmatrix}, \\ e^{L_x} \begin{pmatrix} x_2 \\ -x_1 \\ 0 \end{pmatrix} &= \begin{pmatrix} x_2 \\ -x_1 \\ 0 \end{pmatrix} + \frac{\sin |x|}{|x|} \begin{pmatrix} x_3 x_1 \\ x_2 x_3 \\ -x_1^2 - x_2^2 \end{pmatrix} + \frac{1 - \cos |x|}{|x|^2} \begin{pmatrix} -x_2 |x|^2 \\ x_1 |x|^2 \\ 0 \end{pmatrix}.\end{aligned}$$

The poet:—Couldn't you just give us the result? It's a  $3 \times 3$  matrix, we can do it ourselves, if we want.—“If you *want*. But you won't do it. So once in your life, do some calculations. I know it's painful. No pain, no gain.”

$$\begin{aligned}\left( e^{L_x} \begin{pmatrix} x_2 \\ -x_1 \\ 0 \end{pmatrix} \right) \cdot \begin{pmatrix} x_2 \\ -x_1 \\ 0 \end{pmatrix} &= x_1^2 + x_2^2 + \frac{1 - \cos |x|}{|x|^2} (-x_2^2 |x|^2 - x_1^2 |x|^2) \\ &= x_1^2 + x_2^2 + (1 - \cos |x|)(-x_1^2 + x_2^2) \\ &= (x_1^2 + x_2^2) \cos |x|\end{aligned}$$

—“You see, the point is I didn't prepare for this calculation.”—Evidently.—“We now know that  $|x|$  is the angle, goodbye, farewell, have a nice holiday.”

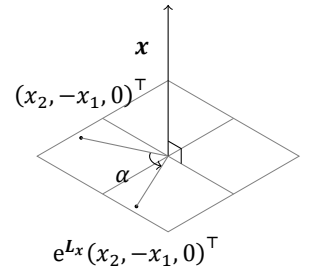


Figure 5.  $e^{L_x} \in \text{SO}(3)$  is a rotation around  $x$ . What is the angle  $\alpha$ ?

## Colophon



This document is available online at <http://goo.gl/m6YNz> (QR-code beside).

It was typeset using the  $\text{\LaTeX}$  engine by Jonathan Kew and Khaled Hosny, version 0.9998, published under the MIT License, and written in the  $\text{\LaTeX}$  markup language.

The serif text is typeset in the Linux Libertine font family, version 5.3, and the sans serif text in the Linux Biolinum font family, version 5.3, both published by the Libertine Open Fonts Project under the GNU General Public License with the font exception and under the Open Font License. The mathematical expressions are typeset in the Cambria Math font family, version 5.93, designed under the direction of Jelle Bosma (Monotype Imaging Incorporated) and Ross Mills (Tiro Typeworks), and licensed by Microsoft Corporation.

Version control was done using the Git source code management system by Linus Torvalds *et alii*. The document is hosted on the GitHub hosting service provided by GitHub Incorporated.

## Bug reporting

The reader is kindly invited to report any falsehoods, typos, unclarities, etc. For the sake of everybody's sanity, please follow the following guidelines when reporting an issue:



- check whether the issue still is present in the latest version of the document, and whether it has already been reported on the “issues” page of the repository <http://goo.gl/XNQRk> (QR-code beside);
- if you have a GitHub account, file an issue at the above page and properly label it as
  - a *bug* if it is a typo, falsehood or any other issue which is unambiguously an error;
  - an *enhancement* if it is a request for clarification, a missing part of a lecture or any other addition.



Please file separate issues for unrelated issues found in the document, and include a concise yet descriptive title;

- if you do not have a GitHub account—and are not willing to create one—, send the report to [egg.robin.1eroy+MMPII@gmail.com](mailto:egg.robin.1eroy+MMPII@gmail.com) (see QR-code beside). Note that bugs reported by email will take longer to process than those reported via GitHub;
- the report should include:
  - the relevant page(s) and lecture number(s);
  - the build date and time found at the bottom of this page;
  - a description of the erroneous version, with enough context to find the issue;
  - a clear description of how the issue should be fixed.

An example of a bug report could be the following:

### **Typo: “them macho wait” p. 19**

Page 19, lecture 14, build date 2013-04-18T08:47:21+02:00, “remember, them macho wait as long as possible” should read “remember, the macho waits as long as possible”.