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

2013-03-27

Lecture 4, 2013-02-28.

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

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

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

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

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

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

Actually, we want these functionals to be in the continuous dual of our space, that is, we want them to be functionals U such that $U(\varphi_k) \to U(\varphi)$ for all sequence of functions $(\varphi_k)_{k \in \mathbb{N}}$ that converge to φ in an appropriate norm[1] — 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 $\varphi_k \to \varphi$ must be in order for it to be complete, and so the bigger its continuous dual can be, as we only need $U(\varphi_k) \to U(\varphi)$ for the sequences $(\varphi_k)_{k \in \mathbb{N}}$ that converge in that space. Thus, we will investigate a space of functions $\varphi(\boldsymbol{x}) \in C^\infty(\mathbb{R}^n)$ that vanish quickly as $|\boldsymbol{x}| \to \infty$ ("converge with a vengeance"), hoping to find the Dirac delta in its continuous dual.

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

$C^\infty(\mathbb{R}^n) := \left\{ \varphi : \mathbb{R}^n \to \mathbb{C} \mid \forall \boldsymbol{\alpha} \in \mathbb{N}^n, \partial^{\boldsymbol{\alpha}}\varphi \in C^0(\mathbb{R}^n) \right\}$ *is the space of* smooth functions *from* $\mathbb{R}^n$ *to* $\mathbb{C}$, *where*

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

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

and $\boldsymbol{x}^{\boldsymbol{\beta}} := x_1^{\beta_1} \cdots x_n^{\beta_n}$ *for* $\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{\beta} \in \mathbb{N}^n$.

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

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

---

[1]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.

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

$$\left(1+|x|^2\right)^m \partial_x{}^\alpha \varphi(x) = \sum_{j=0}^{m}\binom{m}{j}|x|^{2j}\partial_x{}^\alpha \varphi(x)$$

$$= \sum_{j=0}^{m}\binom{m}{j}\left(\sum_{i=1}^{n} x_i^2\right)^j \partial_x{}^\alpha \varphi(x)$$

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

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

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

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

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*

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

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

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

Definition (Locally integrable functions).

$$\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n) := \left\{f : \mathbb{R}^n \to \mathbb{C} \ \middle| \ \forall x \in \mathbb{R}^n, \exists \epsilon > 0, \int_{B_\epsilon x}|f(y)|\,\mathrm{d}y < \infty\right\}$$

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

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

is a tempered distribution.

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

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

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*

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

This is natural, as

$$U_{\partial^\alpha f}(\varphi) = \int_{\mathbb{R}^n} (\partial^\alpha f(\boldsymbol{x}))\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = (-1)^{|\alpha|}\int_{\mathbb{R}^n} f(\boldsymbol{x})\partial^\alpha\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \qquad \text{by integrating by parts.}$$

Let's find the derivative of $|\cdot|$ in the distribution world:

$$\mathrm{Abs}(\varphi) := \int_{-\infty}^{\infty} |x|\varphi(x)\,\mathrm{d}x$$

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}\,\mathrm{Abs}\right)(\varphi) = -\,\mathrm{Abs}\,\frac{\mathrm{d}\varphi}{\mathrm{d}x}$$

$$= -\int_0^\infty x\varphi'(x)\,\mathrm{d}x - \int_{-\infty}^0 (-x)\varphi'(x)$$

$$= -[x\varphi(x)]_0^\infty + \int_0^\infty \varphi(x)\,\mathrm{d}x + [x\varphi(x)]_{-\infty}^0 - \int_{-\infty}^0 \varphi(x)\,\mathrm{d}x$$

$$= \int_{-\infty}^\infty \mathrm{sgn}(x)\varphi(x)\,\mathrm{d}x = U_{\mathrm{sgn}}(\varphi),$$

$$\text{where } \mathrm{sgn}(x) := \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases} \quad \text{is the sign function.}$$

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

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

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

we get

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

in other words,

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

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

$$\langle\psi,\varphi\rangle := \int_{\mathbb{R}^n} \overline{\psi}(\boldsymbol{x})\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

2   General properties of the Fourier transform

Lemma ("Your life in Fourier land depends on it.").

$$\widehat{\varphi + \psi} = \widehat{\varphi} + \widehat{\psi},$$
$$\widehat{\lambda \varphi} = \lambda \widehat{\varphi} \tag{1}$$

$$|\widehat{\varphi}(\boldsymbol{k})| \le \frac{1}{(2\pi)^{\frac{n}{2}}} \|\varphi\|_1 < \infty, \qquad\qquad \|\varphi\|_1 := \int_{\mathbb{R}^n} |\varphi(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} \tag{2}$$

$$\widehat{\varphi(\lambda \cdot)}(\boldsymbol{k}) = \frac{1}{|\lambda|^n} \widehat{\varphi}\left(\frac{\boldsymbol{k}}{\lambda}\right), \qquad\qquad \lambda \ne 0 \tag{3}$$

$$\widehat{T_{\boldsymbol{y}}\varphi}(\boldsymbol{x}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{y}} \, \widehat{\varphi}(\boldsymbol{k}), \qquad\qquad (T_{\boldsymbol{y}}\varphi)(\boldsymbol{x}) := \varphi(\boldsymbol{x} + \boldsymbol{y}) \tag{4}$$

$$\widehat{\frac{\partial \varphi}{\partial x_j}}(\boldsymbol{k}) = \mathrm{i}k_j \widehat{\varphi}(\boldsymbol{k})$$
$$\widehat{x_j \varphi}(\boldsymbol{k}) = \mathrm{i}\frac{\partial \widehat{\varphi}}{\partial k_j}(\boldsymbol{k}) \tag{5}$$

$$\int_{\mathbb{R}^n} \widehat{\varphi}(\boldsymbol{y})\psi(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} = \int_{\mathbb{R}^n} \varphi(\boldsymbol{y})\widehat{\psi}(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} \tag{6}$$

Proof. Proof of (2):

$$|\widehat{\varphi}(\boldsymbol{k})| = \left| \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(\boldsymbol{x})\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \, \mathrm{d}\boldsymbol{x} \right| \le \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \left|\varphi(\boldsymbol{x})\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\right| \, \mathrm{d}\boldsymbol{x} = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} |\varphi(\boldsymbol{x})| \underbrace{\left|\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\right|}_{1} \, \mathrm{d}\boldsymbol{x}$$

$$\underbrace{}_{\|\varphi\|_1}$$

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

$$\int_{\mathbb{R}^n} |\varphi(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^n} \left(1 + |\boldsymbol{x}|^2\right)^{\frac{s}{2}} |\varphi(\boldsymbol{x})| \frac{\mathrm{d}\boldsymbol{x}}{\left(1 + |\boldsymbol{x}|^2\right)^{\frac{s}{2}}}$$

$$\le \|\varphi\|_{s,0} \int_{\mathbb{R}^n} \frac{1}{\left(1 + |\boldsymbol{x}|^2\right)^{\frac{s}{2}}} \, \mathrm{d}\boldsymbol{x}$$

$$= \|\varphi\|_{s,0} \int_0^{\infty} \frac{r^{n-1}}{(1 + r^2)^{\frac{s}{2}}} \, \mathrm{d}r < \infty \qquad\qquad \text{for } s \ge n + 1.$$

"(1) I will not do!" Proof of (3):

$$\widehat{\varphi(\lambda\,\cdot\,)} = \frac{1}{(2\pi)^{\frac{n}{2}}} \int\limits_{\mathbb{R}^n} \varphi(\lambda x) e^{-i k \cdot x} \, d x$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} \int\limits_{\mathbb{R}^n} \varphi(y) e^{-i \frac{1}{\lambda} k \cdot y} \, d\left(\frac{y}{\lambda}\right) \qquad \text{where } x = \frac{1}{\lambda} y.$$

$$= \frac{1}{|\lambda|^n} \underbrace{\frac{1}{(2\pi)^{\frac{n}{2}}} \int\limits_{\mathbb{R}^n} \varphi(y) e^{-i \frac{k}{\lambda} \cdot y} \, d y}_{\widehat{\varphi}\left(\frac{k}{\lambda}\right)}$$

Proof of the second part of (5):

$$\widehat{x_j \varphi}(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int\limits_{\mathbb{R}^n} x_j \varphi(x) e^{-i k \cdot x} \, d x$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} \int\limits_{\mathbb{R}^n} \varphi(x) i \frac{\partial}{\partial k_j} e^{-i k \cdot x} \, d x \qquad \text{as } \frac{\partial}{\partial k_j} e^{-i k \cdot x} = \frac{\partial}{\partial k_j} e^{-i \sum_{r=0}^{n} k_j x_j} = -i x_j e^{-i k \cdot x}$$

$$= i \frac{\partial}{\partial k_j} \frac{1}{(2\pi)^{\frac{n}{2}}} \int\limits_{\mathbb{R}^n} \varphi(x) e^{-i k \cdot x} \, d x$$

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

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 6:

$$\int\limits_{\mathbb{R}^n} \widehat{\varphi}(y) \psi(y) \, d y = \int\limits_{\mathbb{R}^n} \frac{1}{(2\pi)^{\frac{n}{2}}} \int\limits_{\mathbb{R}^n} \varphi(x) e^{-i y \cdot x} \, d x \, \psi(y) \, d y$$

$$= \int\limits_{\mathbb{R}^n} \varphi(x) \frac{1}{(2\pi)^{\frac{n}{2}}} \int\limits_{\mathbb{R}^n} e^{-i x \cdot y} \psi(y) \, d y \, d x \qquad \text{As } \int_{\mathbb{R}^{2n}} \left|\varphi(x) e^{-i y \cdot x} \psi(y)\right| d x \, d y < \infty, \text{ we can use Fubini's theorem.}$$

$$= \int\limits_{\mathbb{R}^n} \varphi(y) \widehat{\psi}(y) \, d y$$

$$\square$$

Proposition. *The Fourier transform $\widehat{\,\cdot\,}$ is a continuous bijective linear map from the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ to itself. $\widecheck{\widehat{\varphi}} = \varphi$, where $\widecheck{\psi}(x) := \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \psi(k) e^{i k \cdot x} \, d x$, read "unhat", "bird", "seagull", or whatever you like.*

Proof.

$$\varphi \in \mathcal{S}(\mathbb{R}^n) \Rightarrow x^\alpha \partial_x^\beta \in \mathcal{S}(\mathbb{R}^n)$$

$$\left(\widehat{x^\alpha \partial_x^\beta \varphi}\right)(\boldsymbol{k}) = \mathrm{i}^{|\alpha|} \partial_k^\alpha \left(\widehat{\partial_x^\beta \varphi}\right)(\boldsymbol{k}) \qquad \text{from the second equality in (5), applied } |\boldsymbol{\alpha}| \text{ times.}$$

$$= \mathrm{i}^{|\alpha|} \partial_k^\alpha \left(\mathrm{i}^{|\beta|} \boldsymbol{k}^\beta \widehat{\varphi}\right)(\boldsymbol{k})$$

$$= \mathrm{i}^{|\alpha|+|\beta|} \partial_k^\alpha \left(\boldsymbol{k}^\beta \widehat{\varphi}\right)(\boldsymbol{k})$$

"I should have done it the other way around. [...] Let's try it the other way around." We want:

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

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

$$\left(\partial_x^\alpha \widehat{(\boldsymbol{k}^\beta \varphi)}\right)(\boldsymbol{k}) = \mathrm{i}^{|\alpha|+|\beta|} \boldsymbol{k}^\alpha \partial_k^\beta \widehat{\varphi}(\boldsymbol{k}) \qquad \text{by applying the first part of (5) } |\boldsymbol{\alpha}| \text{ times and the second part } |\boldsymbol{\beta}| \text{ times.}$$

$$\forall \boldsymbol{\alpha}, \boldsymbol{\beta}, \; \left| \boldsymbol{k}^\alpha \partial_k^\beta \widehat{\varphi}(\boldsymbol{k}) \right| \le \frac{1}{(2\pi)^{\frac{n}{2}}} \left\| \partial_x^\beta (x^\beta \varphi) \right\|_1 < \infty \qquad \text{from (2).}$$

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

$$\breve{\widehat{\varphi}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \widehat{\varphi}(\boldsymbol{k}) \mathrm{e}^{\mathrm{i}\,\boldsymbol{k}\cdot\boldsymbol{x}} \, \mathrm{d}\boldsymbol{k}$$

$$= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \varphi(\boldsymbol{y}) \mathrm{e}^{-\mathrm{i}\,\boldsymbol{k}\cdot\boldsymbol{y}} \, \mathrm{d}\boldsymbol{y} \, \mathrm{e}^{\mathrm{i}\,\boldsymbol{k}\cdot\boldsymbol{x}} \, \mathrm{d}\boldsymbol{k}$$

We don't interchange the integrals here because that would lead to ugly calculations. "If pou paint the walls before you ſtart building, it's not a good idea." At this point, somebody suggeſts replacing $\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \mathrm{e}^{\mathrm{i}\,\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{y})} \, \mathrm{d}\boldsymbol{k}$ by $\delta(\boldsymbol{x}-\boldsymbol{y})$. "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."

$$\breve{\widehat{\varphi}} = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \widehat{\varphi}(\boldsymbol{k}) \mathrm{e}^{\mathrm{i}\,\boldsymbol{k}\cdot\boldsymbol{x}} \underbrace{\lim_{\epsilon \downarrow 0} \mathrm{e}^{-\epsilon \frac{|\boldsymbol{k}|^2}{2}}}_{\substack{1 \text{ written in} \\ \text{some other way}}} \, \mathrm{d}\boldsymbol{k}$$

$$= \lim_{\epsilon \downarrow 0} \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \widehat{\varphi}(\boldsymbol{k}) \mathrm{e}^{\mathrm{i}\,\boldsymbol{k}\cdot\boldsymbol{x}} \mathrm{e}^{-\epsilon \frac{|\boldsymbol{k}|^2}{2}} \qquad \begin{array}{l}\text{"Don't worry." This is actually a one-liner using} \\ \text{Lebesgue's dominated convergence theorem.}\end{array}$$

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

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

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

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.

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

We therefore get:

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

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(\boldsymbol{y}) \frac{e^{-\frac{1}{2\epsilon}|\boldsymbol{x}-\boldsymbol{y}|^2}}{\epsilon^{\frac{n}{2}}} \, \mathrm{d}\boldsymbol{y}$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(\boldsymbol{y}\sqrt{\epsilon} + \boldsymbol{x}) \frac{e^{-\frac{1}{2\epsilon}|\boldsymbol{y}\sqrt{\epsilon}|^2}}{\epsilon^{\frac{n}{2}}} \, \mathrm{d}(\boldsymbol{y}\sqrt{\epsilon}) \qquad \text{“Epsilons everywhere!”}$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \varphi(\boldsymbol{x} + \boldsymbol{y}\sqrt{\epsilon}) e^{-\frac{1}{2}|\boldsymbol{y}|^2} \, \mathrm{d}\boldsymbol{y} \qquad \text{“Now we can paint.”}$$

$$\overset{\epsilon\downarrow 0}{\to} \frac{1}{(2\pi)^{\frac{n}{2}}} \varphi(\boldsymbol{x}) \int_{\mathbb{R}^n} e^{-\frac{1}{2}|\boldsymbol{y}|^2} \, \mathrm{d}\boldsymbol{y} = \varphi(\boldsymbol{x}).$$

$\square$

3 Eigenfunctions of the Fourier transform

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

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

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

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

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 $H \in \mathrm{End}(\mathcal{S}(\mathbb{R}^n))$ 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 $n = 1$.

$$\left.\begin{aligned}(\widehat{x^2\varphi})(k) &= -\frac{\mathrm{d}^2}{\mathrm{d}k^2}\widehat{\varphi}(k) \\[1em] \left(\widehat{\frac{\mathrm{d}^2}{\mathrm{d}x^2}\varphi}\right)(k) &= -k^2\widehat{\varphi}(k)\end{aligned}\right\} \qquad \text{From the lemma your life depends on.}$$

Subtracting the second equality from the first one above:

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

where $\mathcal{F} = \widehat{\cdot}$ because things are getting a bit too big for the hat. We smell a harmonic potential. "Don't get your fingers too close, they sometimes bite."
  Let $H := \frac{1}{2}\left(-\frac{\mathrm{d}^2}{\mathrm{d}x^2} + x^2 - 1\right)$. 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

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

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

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.

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:

$$\langle \varphi, P\psi \rangle = \int_{\mathbb{R}} \overline{\varphi}(x) \frac{d}{dx}\psi(x) \, dx$$

$$= [\overline{\varphi}\psi]_{-\infty}^{\infty} - \int_{\mathbb{R}} \frac{d}{dx}\varphi(x)\psi(x) \, dx$$

$$= -\langle P\varphi, \psi \rangle, \qquad \qquad [\overline{\varphi}\psi]_{-\infty}^{\infty} = 0 \text{ as } \varphi \in \mathcal{S}(\mathbb{R}), \psi \in \mathcal{S}(\mathbb{R}), \text{ and therefore } \overline{\varphi}\psi \in \mathcal{S}(\mathbb{R}).$$

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

If we let $h := \hbar$, the conditions verified by P and Q are the conditions for infinite matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ 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.*

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

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

$$\left(e^{itP}f\right)(x) = f(x+t) \qquad \text{as } \left(e^{it\left(\frac{1}{i}\frac{d}{dx}\right)}\right)f(x) \overset{\text{formally}}{=} \sum_{l=0}^{\infty} \frac{t^l}{l!}\frac{d^l}{dx^l}f(x), \text{ the Taylor series of } f(x+t) \text{ at } x.$$

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

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

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

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

4 The finite Heisenberg group

The professor whistles "Singin' in the Rain" while waiting for the lecture to start.

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

"You write commutators to see the level of how much it doesn't commute (*sic*)."

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

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

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

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

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

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

"Out of this trivial stuff comes something deep."

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

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

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 $\mathbb{R}^n \times \widehat{\mathbb{R}^n}$ is as the cotangent bundle $\mathbb{R}^n \times (\mathbb{R}^n)^*$. "If you would have the energy to learn ancient Greek, and you talked about the real housewives of Atlanta, it would be a waste." It is not sufficient to have a fancy language, you actually have to say something.

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

<div align="right">SUPPLEMENTAL</div>

4.1 Ἡ φανταχτερή γλῶσσα τῶν ἀστικῶν μαθηματικῶν[2]

Assume the space G of generalised coordinates of a physical system is a Lie group, i.e., a group which is also a differentiable manifold, and in which composition and inversion are smooth maps. Then if $q : \mathbb{R} \to G$ describes the evolution of the generalised coordinates of the system with time, the generalised velocity at time t, $\dot{q}(t)$, lies in the tangent space $T_{q(t)} G$ — recall that the tangent space of g at x is defined as the set of derivatives[3] $\dot{\gamma}(0)$ at 0 of smooth curves $\gamma : \mathbb{R} \to G$ with $\gamma(0) = x$.

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

$$(q(t), \dot{q}(t)) \in T G := \bigcup_{x \in G} \{x\} \times T_x G.$$

---

[2] The fancy language of bourgeois mathematicians.

[3] 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 γ in the set of curves through x with $\gamma(0) = x$ under the relation $\gamma_1 \sim \gamma_2 \Leftrightarrow (f \circ \gamma_1)'(0) = (f \circ \gamma_2)'(0)$, where f is a differentiable chart of a neighbourhood of x onto $\mathbb{R}^n$. It can then be shown that this is independent of f. Quite often though, G is naturally a submanifold of $\mathbb{R}^n$ and this trick is unneeded.

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

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

$$\mathrm{d}\,\varphi_q(\boldsymbol{q})\dot{\boldsymbol{\gamma}}(0) = \mathrm{d}\,\varphi_q(\boldsymbol{\gamma}(0))\dot{\boldsymbol{\gamma}}(0) = \left(\varphi_q \circ \boldsymbol{\gamma}\right)'(0) = \dot{\boldsymbol{\gamma}}_0(0),$$

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

$$\mathrm{T}\,G = \bigcup_{\boldsymbol{x} \in G}\{\boldsymbol{x}\} \times \mathrm{T}_{\boldsymbol{x}}\,G \cong \bigcup_{\boldsymbol{x} \in G}\{\boldsymbol{x}\} \times \mathrm{T}_{\mathbb{1}}\,G = G \times \mathrm{T}_{\mathbb{1}}\,G = G \times \mathfrak{g},$$

where the isomorphism is canonical.

Recall from Hamiltonian mechanics that for a Lagrangian $\mathcal{L} : \mathrm{T}\,G \times \mathbb{R} \to \mathbb{R}, (\boldsymbol{q}, \boldsymbol{v}, t) \mapsto \mathcal{L}(\boldsymbol{q}, \boldsymbol{v}, t)$, 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* $\boldsymbol{p} : \boldsymbol{v} \mapsto \sum_i v_i p_i$ is therefore a linear form on the space $\mathfrak{g} \ni \boldsymbol{v}$. Using $\boldsymbol{q}(t)$ as the evolution of the system as above, $\dot{\boldsymbol{q}}(t) \in \mathrm{T}_{\boldsymbol{q}(t)}\,G$, and so $\boldsymbol{p}(t) \in \left(\mathrm{T}_{\boldsymbol{q}(t)}\,G\right)^* =: \mathrm{T}^*_{\boldsymbol{q}(t)}\,G$, the *cotangent space* of G at $\boldsymbol{q}$, and the tuples $(\boldsymbol{q}, \boldsymbol{p})$ lie in the *cotangent bundle* $\mathrm{T}^*\,G$,

$$\mathrm{T}^*\,G := \bigcup_{\boldsymbol{q} \in G}\{\boldsymbol{q}\} \times \mathrm{T}^*_{\boldsymbol{q}}\,G.$$

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

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

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, $\mathcal{H} : \mathrm{T}^*\,G \times \mathbb{R} = G \times \mathfrak{g}^* \times \mathbb{R} \to \mathbb{R}$.

The Pontryagin dual[4] $\widehat{G}$ is the set of homomorphisms from G to $\mathrm{U}(1)$. 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?*

<div align="right">SUPPLEMENTAL</div>

4.2   Examples from Raisa Galimova's exercise class

Let G be a group acting on a measure space[5] (X, Σ, μ) by the left action $f : G \to X^X, g \mapsto f_g$, 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

$$\mathrm{L}^2(X) := \left\{\varphi \in X^{\mathbb{C}} \,\middle|\, \|\varphi\|_2 < \infty\right\}$$

where

$$\|\varphi\|_2 := \langle \varphi, \varphi \rangle$$

$$\langle \psi, \varphi \rangle := \int_X \overline{\psi}\varphi \, \mathrm{d}\mu.$$

Note that if X is finite, and μ is the counting measure,

$$\langle \psi, \varphi \rangle = \int_X \overline{\psi} \varphi \, d\mu = \sum_{x \in X} \overline{\psi}(x) \varphi(x)$$

$$\forall \varphi \in \mathbb{C}^X \quad \|\varphi\|_2 = \int_X |\varphi|^2 \, d\mu = \sum_{x \in X} |\varphi|^2 < \infty$$

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

$$\langle \psi, \varphi \rangle = \int_X \overline{\psi} \varphi \, d\mu = \int_{\mathbb{R}^n} \overline{\psi}(x) \varphi(x) \, dx$$

$$\|\varphi\|_2 = \int_X |\varphi|^2 \, d\mu = \int_{\mathbb{R}^n} |\varphi(x)|^2 \, dx$$

and again we recover the definition we gave for the inner product on $\mathcal{S}(\mathbb{R}^n) \subsetneq L^2(\mathbb{R}^n)$. Now consider the representation $\rho : G \to \mathrm{GL}\left(L^2(X)\right), g \mapsto \rho_g$ of G on $L^2(X)$ defined by $\left(\rho_g(\varphi)\right)(x) = \varphi\left(f_g(x)\right)$.

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

$$\widehat{\varphi}(j) = \frac{1}{2\pi} \int_0^{2\pi} \varphi(\theta) e^{-ij\theta} \, d\theta = \int_{S^1} \varphi e^{-ij\cdot} \, d\mu = \langle e^{ij\cdot}, \varphi \rangle \quad \text{where } e^{ij\cdot} := \theta \mapsto e^{ij\theta}, \quad e^{ij\cdot} \in \widehat{S^1} \cong \mathbb{Z}$$

by taking μ normed, i.e., $\mu(S^1) = 1$. In other words, $\widehat{\cdot}(j) = \langle e^{ij\cdot}, \cdot \rangle$ is the projection of φ onto the subspace $V_j := \mathrm{span}\{e^{ij\cdot}\}$ of $L^2(S^1)$. Note that the $e^{ij\cdot}$ are exactly the elements of $\widehat{S^1}$. Moreover, V_j is ρ-invariant, as ρ acts by translation on $\varphi \in L^2(S^1)$, namely $(\rho_\alpha(\varphi))(\theta) = \varphi(\alpha + \theta)$, and hence

$$\forall \lambda e^{ij\cdot} \in V_j \quad \left(\rho_\alpha\left(\lambda e^{ij\cdot}\right)\right) = \lambda e^{ij(\alpha + \cdot)} = \lambda e^{ij\alpha} e^{ij\cdot} \in V_j.$$

As it is one-dimensional, V_j is irreducible. We also have, with $\langle e^{ij\cdot}, e^{ik\cdot} \rangle = \widehat{e^{ik\cdot}}(j) = \delta_{jk}$, $V_j \perp V_k$ for $k \neq j$. By Parseval's identity, $\sum_{j \in \mathbb{Z}} \widehat{\varphi}(j) e^{ij\cdot}$ converges absolutely to φ for all $\varphi \in L^2(S^1)$, so

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

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

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

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 $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot} := \boldsymbol{x} \mapsto \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}$ are exactly the elements of $\widehat{\mathbb{R}^n} \cong \mathbb{R}^n$, and that the $V_{\boldsymbol{k}} := \operatorname{span}\{\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot}\}$ are ρ-invariant, where ρ acts by translation and are orthogonal. The projector onto $V_{\boldsymbol{k}}$ is $\widehat{\cdot}(\boldsymbol{k})$. As for $\varphi \in \mathrm{L}^2(\mathbb{R}^n)$,

$$\widecheck{\widehat{\varphi}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \widehat{\varphi}(\boldsymbol{k}) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \, \mathrm{d}\boldsymbol{k} \varphi(\boldsymbol{x}) = \int_{\mathbb{R}^n} \widehat{\varphi} \mathrm{e}^{\mathrm{i}\boldsymbol{x}\cdot} \, \mathrm{d}\hat{\mu} = \varphi(\boldsymbol{x})$$

for almost all x — $\hat{\mu}$ is the dual measure [6] —, the representation ρ of $\mathbb{R}$ on $\mathrm{L}^2(\mathbb{R})$ by translation decomposes in irreducible representations as

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

Once again, it makes sense to think of the projectors as $\widehat{\cdot}(\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot})$, 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 $\boldsymbol{k} \mapsto \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}\cdot}$ 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 $\mathrm{e}^{-\frac{1}{2}|\boldsymbol{x}|^2}$ is no longer self-dual, i.e., the same as its Fourier transform, but $\mathrm{e}^{-\pi|\boldsymbol{x}|^2}$ is instead.

This also gives us an isomorphism between the momentum space $\mathrm{T}_0^*\,\mathbb{R}^n = (\mathbb{R}^n)^*$ and the Pontryagin dual $\widehat{\mathbb{R}^n}$, namely $\boldsymbol{\lambda} \mapsto \mathrm{e}^{2\pi\mathrm{i}\lambda}$, with $\widehat{\mathbb{R}^n} \ni \mathrm{e}^{2\pi\mathrm{i}\lambda} : \boldsymbol{k} \mapsto \mathrm{e}^{2\pi\mathrm{i}\lambda\boldsymbol{k}}$, 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 the tangent space. Thus the phase space should indeed be $\mathrm{T}^*\,\mathbb{R}^n$, and not $\mathrm{T}\,\mathbb{R}^n$.

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

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

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

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

Lecture 10, 2013-03-21.
There are official notes for the material covered in this lecture.

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

We switched from an additive to a multiplicative notation. "Algebraists... you have to be careful. They're sensitive."

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

---

[6] 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 μ in such a way that $\mu = \widehat{\mu}$, but if we had defined a non-unitary version, e.g. $\widehat{\varphi}(\boldsymbol{k}) := \int_{\mathbb{R}^n} \varphi(\boldsymbol{x}) \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \, \mathrm{d}\boldsymbol{x}$, we would need to define the inverse Fourier transform accordingly as $\widecheck{\psi}(\boldsymbol{x}) := \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \psi(\boldsymbol{k}) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \, \mathrm{d}\boldsymbol{x}$, i.e., for $\mu = \mathcal{L}^n$, $\widehat{\mu} = \frac{1}{(2\pi)^n} \mathcal{L}^n$.

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."

Recall:

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

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

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

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

So we have this very nice guy here: $\frac{d^2}{dx^2} + x^2$, and having gone to Elementärschule, 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 $\left[\frac{d}{dx}, x\right] = \mathbb{1}$. 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 is 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;

$$\langle \psi, A\varphi \rangle = \frac{1}{\sqrt{2}} \int_{\mathbb{R}} \overline{\psi}(x)\left(\frac{d}{dx} + x\right)\varphi(x) \, dx$$

$$= \frac{1}{\sqrt{2}} \int_{\mathbb{R}} \overline{\psi}(x)\frac{d}{dx}\varphi(x) \, dx + \frac{1}{\sqrt{2}} \int_{\mathbb{R}} \overline{(x\psi(x))}\varphi(x) \, dx$$

$$= \int_{\mathbb{R}} \overline{\left(\frac{1}{\sqrt{2}}\left(-\frac{d}{dx} + x\right)\psi(x)\right)}\varphi(x) \, dx$$

$$= \langle A^\dagger\psi, \varphi \rangle$$

Write

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

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

and similarly for k instead of x.

Observe:

$$\mathcal{F}(A_x\varphi)(k) = \mathrm{i}A_k\widehat{\varphi}(k)$$

$$\mathcal{F}\left(A_x^\dagger\varphi\right)(k) = \mathrm{i}A_k^\dagger\widehat{\varphi}(k)$$

$$A_x\mathrm{e}^{-\frac{1}{2}x^2} = \left(\frac{\mathrm{d}}{\mathrm{d}\,x} + x\right)\mathrm{e}^{-\frac{1}{2}x^2} = 0.$$

define

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

so that $\langle h_0, h_0\rangle = 1$.

We now can perform the following induction:

$$Ah_0 = 0$$

$$\widehat{h_0} = h_0$$

$$\widehat{A_x^\dagger h_0}(k) = (-\mathrm{i})A_k^\dagger h_0(k)$$

$$\mathcal{F}\left(\left(A_x^\dagger\right)^n h_0\right)(k) = (-\mathrm{i})A_k^\dagger\,\mathcal{F}\left(\left(A^\dagger\right)^{n-1}h_0\right)(k) = (-\mathrm{i})^n\left(A_k^\dagger\right)^n h_0(k)$$

"I'll do some coherent states. I like coherent states." Next week we shall prove:

$$\left[A, A^\dagger\right] = \mathbb{1}$$

$$\left[A, \left(A^\dagger\right)^n\right] = n\left(A^\dagger\right)^{n-1}$$

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