locally integrable so that $T_{\frac{1}{\|x\|}}$ exists. Now recall (or check!) that

$$\widehat{\frac{1}{\|x\|} e^{-2\pi a \|x\|}}(k) = \frac{1}{\pi} \frac{1}{\|k\|^2 + a^2}.$$

Let T_n be the distribution corresponding to the function $\frac{1}{\|x\|} e^{-\frac{2\pi}{n}\|x\|}$. We have $T_n \to T_{\frac{1}{\|x\|}}$ (by monotone convergence), and $\hat{T}_n$ is the distribution that corresponds to $\frac{1}{\pi}\frac{1}{\|k\|^2 + \frac{1}{n^2}}$. Then $\hat{T}_n \to T_{\frac{1}{\pi\|x\|^2}}$ (monotone convergence again). We conclude that, in $d=3$, "the Fourier transform of $\frac{1}{\|x\|}$ is $\frac{1}{\pi\|k\|^2}$, in the sense of distributions".

If $T \in S'(\mathbb{R}^d)$, let $x^\alpha T$ denote the distribution $(x^\alpha T)(\phi) = T(x^\alpha \phi)$. It is well-defined since $x^\alpha \phi \in S(\mathbb{R}^d)$.

Proposition 6.4

(a) $\widehat{\frac{\partial^\alpha}{\partial x^\alpha} T} = (2\pi i k)^\alpha \hat{T}$

(b) $\frac{\partial^\alpha}{\partial k^\alpha} \hat{T} = (-2\pi i)^{|\alpha|} \widehat{x^\alpha T}$, $|\alpha| = |\alpha_1| + \dots + |\alpha_d|$.

Proof: exercise. □

Proposition 6.5

(a) The map $T \to \hat{T}$ is a bijection from S' to S'.

(b) The inverse map is $T \mapsto T^\vee$, where $T^\vee(\phi) = T(\phi^\vee)$, and $\phi^\vee(x) = \int_{\mathbb{R}^d} e^{2\pi i k x} \phi(k)\, dk$.

Proof: We start with (b). $\hat{T}^\vee(\phi) = \hat{T}(\phi^\vee) = T(\widehat{\phi^\vee}) = T(\phi)$, and $\widehat{T^\vee}(\phi) = T^\vee(\hat{\phi}) = T(\hat{\phi}^\vee) = T(\phi)$. Then $\hat{T}^\vee = \widehat{T^\vee} = T$. Next, we check that $\hat{T}$ is one-to-one. This is equivalent to $\hat{T}(\phi) = 0\ \forall \phi \Rightarrow T = 0$. If $\hat{T}(\phi) = 0$ $\forall \phi$, then $T(\phi) = 0\ \forall \phi$ since the Fourier transform is a bijection between S and S. (The latter follows from Proposition 4.2.)
The map $T \mapsto \hat{T}$ is onto: $\forall T \in S'$, we have $T = \widehat{T^\vee}$ by (b), so $T \in \mathrm{ran}(\mathcal{F})$. □

6.4. The Poisson equation

We consider the equation in $\mathbb{R}^3$: $-\Delta u = f$, where $f \in S(\mathbb{R}^3)$. This equation is not linear. In Physics, f is the density of electric charge and u is the electric potential; or f is the density of mass and u is the negative of the gravitational potential.

We first search for a solution using formal calculations. The Fourier transform of the equation is

$$4\pi^2 \|k\|^2 \hat{u}(k) = \hat{f}(k).$$

Then
$$\hat{u}(k) = \frac{1}{4\pi^2\|k\|^2}\,\hat{f}(k).$$

The inverse Fourier transform of $\frac{1}{4\pi^2\|k\|^2}$ is $\frac{1}{4\pi\|x\|}$ as we have seen in the previous section. This suggests that

$$u(x) = \left(\frac{1}{4\pi\|x\|} * f\right)(x) = \int_{\mathbb{R}^d} \frac{1}{4\pi\|x-y\|} f(y)\,dy.$$

The function u above is well-defined, and we now check that it is indeed a solution to the Poisson equation. It follows from the expression $u(x) = \int \frac{1}{4\pi\|y\|} f(x-y)$ that u is C^∞. We could prove that $-\Delta u = f$ using analysis, but we can proceed elegantly with distributions.

Formally, we have for all $f \in S(\mathbb{R}^3)$ that

$$f(x) = -\Delta u(x) = \int \left(-\Delta \frac{1}{4\pi\|x-y\|}\right) f(y)\,dy = \int \left(-\Delta \frac{1}{4\pi\|y\|}\right) f(x-y)\,dy.$$

This suggests that $-\Delta \frac{1}{4\pi\|x\|} = \delta_0$. This is a useful relation, let us now prove it.

Lemma 6.6

Let T be the distribution corresponding to $\frac{1}{4\pi\|x\|}$. Then

$$-\Delta T = \delta_0.$$

The function $\frac{1}{4\pi\|x\|}$ is then called the "Green function" of the Poisson equation.

Proof: By Proposition 6.5, two distributions are equal iff their Fourier transforms are equal. We check that $\widehat{-\Delta T} = \hat{\delta}_0$. We have $\hat{\delta}_0 = T_1$, i.e. the distribution corresponding to the function 1. Using Proposition 6.4 (a), $\widehat{-\Delta T} = 4\pi\|k\|^2 \hat{T}$, and $\hat{T}$ is the distribution corresponding to $\widehat{\frac{1}{4\pi\|x\|}}(k) = \frac{1}{4\pi^2\|k\|^2}$. Then $\widehat{-\Delta T} = T_1$ indeed. □

Given $\phi \in S(\mathbb{R}^3)$, let ϕ_x denote the Schwartz function $\phi_x(y) = \phi(x-y)$. Clearly, the map $\phi \mapsto \phi_x$ is a bijection $S \to S$. The function u above can be written as

$$u(x) = T(f_x)$$

Then

$$-\Delta u(x) = -\Delta_x T(f_x) \underset{\substack{\uparrow \\ \text{continuity} \\ \text{of distributions}}}{=} -T(\Delta_x f_x) \underset{\substack{\uparrow \\ f_x(y) = f(x-y)}}{=} -T(\Delta f_x)$$

$$\underset{\substack{\uparrow \\ \text{def. of distr.} \\ \text{derivative}}}{=} -\Delta T(f_x) \underset{\substack{\uparrow \\ \text{Lemma 6.6}}}{=} \delta_0(f_x) = f(x).$$

This concludes the chapter on distributions. It may be worth pointing out that convolutions of distributions are not defined in general. Nor are products of distributions.

7. APPLICATIONS TO PROBABILITY THEORY

Let us recall the basics of probability theory. One is given a measure space $(\mathcal{X}, \Sigma, \mathbb{P})$ where $\mathbb{P}$ is a probability measure (i.e. it is a measure, and $\mathbb{P}(\mathcal{X}) = 1$). The notation and the language differ from the analyst's and it may be a source of confusion. See Folland's dictionary between analysts' language and probabilists' dialect (p. 314 in his Real Analysis book).

A <u>random variable</u> X is a measurable function $\mathcal{X} \to \mathbb{R}$.
Its <u>distribution function</u> is

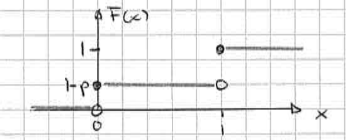
$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(\{\omega \in \mathcal{X} : X(\omega) \leq x\}).$$

Any distribution function satisfies the following properties:

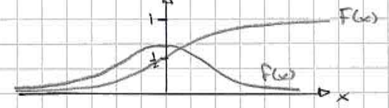
- it is increasing and $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x\to\infty} F(x) = 1$;
- it is right-continuous.

The random variable X is called <u>continuous</u> if its distribution function is continuous. If $F(x)$ is differentiable, then X has a <u>density</u> $f = F'$, such that $\mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(y)\,dy$.

<u>Examples</u>: (a) Bernoulli random variable, where 0 occurs with probability $1-p$ and 1 occurs with probability p. This is a discrete random variable and its distribution function is:



(b) Gaussian or normal random variable, denoted $\mathcal{N}(0,1)$. It has a density, $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.



An important characterisation of a random variable is its <u>mean</u> (or <u>average</u>) and its <u>variance</u>. If X has density f, we define

- the mean: $\mu = \int_{-\infty}^{\infty} x f(x)\,dx$ (the integral may not exist)
- the variance: $\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$ (σ^2 may be infinite)

Since random variables do not always have a density, we need a more general definition. If g is a continuous function on $\mathbb{R}$, we define its Riemann-Stieltjes integral with respect to F by

$$\int_a^b g(x)\,dF(x) = \lim_{n\to\infty} \sum_{i=1}^{n} g\left(a + i\tfrac{b-a}{n}\right)\left[F\left(a+(i+1)\tfrac{b-a}{n}\right) - F\left(a+i\tfrac{b-a}{n}\right)\right].$$

(If $F(x) = x$, we get the usual Riemann integral.) One can then take $a \to -\infty$ and $b \to +\infty$ if needed, provided these limits exist. Choosing $g(x) = x$, one gets the mean μ of X. If μ exists, one can choose $g(x) = (x-\mu)^2$ and one gets the variance σ^2 of X.

Remark: There are other ways to define the mean and variance of a general random variable. One can show for instance that to any distribution function F there corresponds a measure μ such that $\mathbb{P}(X \leq x) = F(x) = \mu((-\infty, x])$. Then one can use Lebesgue integration.

These notions generalise to the case of several random variables. Notion of joint distribution functions, e.g. $F(x_1, x_2) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2)$. The random variables are independent if their joint distribution function factorises.

7.1 Central limit theorem

The name was introduced by Pólya in 1920, because it plays a "zentralle Rolle" in probability theory. It states that a large sum of independent random variables has normal fluctuations around its mean. It was first understood by de Moivre in

the case of Bernoulli random variables. It was generalised by Gauss. The derivation using Fourier theory is due to Laplace.

We need to define the notion of convergence of a sequence of random variables.

Definition: Let (Y_n) be a sequence of random variables with distribution functions F_n. We say that (Y_n) converges in distribution to Y if for every point of continuity of F, we have

$$\lim_{n\to\infty} F_n(x) = F(x).$$

Here, F is the distribution function of Y.

We can now state the main theorem of this section.

Theorem 7.1 (Central limit theorem)
Let $X_1, X_2, \ldots$ be independent, identically distributed random variables with mean μ and variance σ^2 (we assume these exist). Then

$$\frac{X_1 + \ldots + X_n - \mu n}{\sigma \sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1).$$

In order to prove this theorem, we make the connection with Fourier theory.