3.3. **The Central Limit Theorem.** Let $f \in W_{1,1}$ and set $\hat{f} := f - m(f)$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \hat{f} \circ T^k(x) = 0 \quad m - \text{a.e.}$$

Let us set $\Psi_n := \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \hat{f} \circ T^k$. We can consider Ψ_n a random variable with distribution $F_n(t) := m(\{x \,:\, \Psi_n(x) \leq t\})$. It is well know that, for each continuous function g holds[6]

$$m(g(\Psi_n)) = \int_{\mathbb{R}} g(t) dF_n(t)$$

where the integral is a Riemann-Stieltjes integral. It is thus clear that if we can control the distribution F_n, we have a very sharp understanding of the probability to have small deviations (of order $\sqrt{n}$) from the limit. From the work in the previous section it follows that there exists $\delta > 0$ such that, for each $|\lambda| \leq \delta\sqrt{n}$,

$$\begin{aligned} \varphi_n(\lambda) := m(e^{i\lambda\Psi_n}) = m(\mathcal{L}^n_{\lambda/\sqrt{n}} 1) &= (1 - \sigma\lambda^2 n^{-1} + \mathcal{O}(\lambda^3 n^{-\frac{3}{2}}))^n \\ &= e^{-\sigma\lambda^2}(1 + \mathcal{O}(\lambda^3 n^{-\frac{1}{2}})). \end{aligned} \tag{3.15}$$

The above quantity is called *characteristic function* of the random variable and determine the distribution via the formula

$$F_n(b) - F_n(a) = \lim_{\Lambda\to\infty} \frac{1}{2\pi} \int_{-\Lambda}^{\Lambda} \frac{e^{-ia\lambda} - e^{-ibt}}{i\lambda} \varphi_n(\lambda) d\lambda,$$

as can be seen in any basic book of probability theory.[7]

Formula (3.15) means in particular that

$$\lim_{n\to\infty} m(e^{\lambda\Psi_n}) = e^{-\sigma\lambda^2} =: \varphi(\lambda).$$

What can we infer out of the above facts? First of all a simple computation shows that

$$g(t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-it\lambda} \varphi(\lambda) d\lambda = \frac{1}{2\sqrt{\pi\sigma}} e^{-\frac{t^2}{4\sigma}}$$

a random variable with such a density is called a Gaussian random variable with zero average and variance σ. Accordingly, formula (3.15) can be interpreted by saying that there exists a Gaussian random variable G such that

$$\frac{1}{n} \sum_{k=0}^{n-1} \hat{f} \circ T^k \sim \frac{1}{\sqrt{n}} G(1 + \mathcal{O}(n^{-\frac{1}{2}}))$$

---

[6] If $g \in \mathcal{C}^1_0$, then

$$\int_{\mathbb{R}} g dF_n = -\int_{\mathbb{R}} F_n(t) g'(t) dt = -\int_{\mathbb{R}} dt \int_{\mathbb{T}^1} dx \chi_{\{z \,:\, \Psi_n(z) \leq t\}}(x) g'(t).$$

Applying Fubini yields

$$\int_{\mathbb{R}} g dF_n = -\int_{\mathbb{T}^1} dx \int_{\mathbb{R}} dt \chi_{\{z \,:\, \Psi_n(z) \leq t\}}(x) g'(t) = -\int_{\mathbb{T}^1} dx \int_{\Psi_n(x)}^{\infty} g'(t) dt = \int_{\mathbb{T}^1} dx g(\Psi_n(x)).$$

[7] In the case when there exists a density, that is an L^1 function f_n such that $F_n(b) - F_n(a) = \int_a^b f_n(t) dt$, then the formula above becomes simply

$$f_n(t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-it\lambda} \varphi_n(\lambda) d\lambda,$$

and follows trivially by the inversion of the Fourier transform.

DRAFT -- DRAFT -- DRAFT -- DRAFT -- DRAFT --

in distribution. But what does this means concretely. Actual estimates are made difficult by the fact that the distribution under study no not necessarily have a density, thus we are Fourier transforming function that behave quite badly at infinity. To overcome such a problem we can smoothen the quantities involved.

Let $j \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}_+)$ such that $\int_{\mathbb{R}} j(t)dt = 1$ and $j(t) = 0$ for all $|t| > 1$, for each $\varepsilon > 0$ defined then $j_\varepsilon(t) := \varepsilon^{-1} j(\varepsilon^{-1} t)$ and

$$F_{n,\varepsilon}(t) := \int_{\mathbb{R}} j_\varepsilon(t-s) F_n(s) ds. \tag{3.16}$$

A simple computation shows that, for each $a, b \in \mathbb{R}$, holds

$$F_n(b+\varepsilon) - F_n(a-\varepsilon) \geq F_{n,\varepsilon}(b) - F_{n,\varepsilon}(a) \geq F_n(b-\varepsilon) - F_n(a+\varepsilon)$$

that is: if the measurements have a precision smaller than 2ε, then $F_{n,\varepsilon}$ is as good as F_n to describe the resulting statistics. On the other hand calling $\varphi_{n,\varepsilon}$ the characteristic function associated to $F_{n,\varepsilon}$, holds $\varphi_{n,\varepsilon}(\lambda) = \varphi_n(\lambda)\hat{j}(\varepsilon\lambda)$, where $\hat{j}$ is the Fourier transform of j. Since now $F_{n,\varepsilon}$ is a smooth random variable it has a density $f_{n,\varepsilon}$ and

$$f_{n,\varepsilon}(t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\lambda t} \varphi_n(\lambda) \hat{j}(\varepsilon\lambda) d\lambda$$

since j is smooth it follows that there exists $C > 0$ such that $|\hat{j}(\lambda)| \leq C(1+\lambda^2)^{-2}$. We can finally use formula (3.15) to obtain a quantitative estimate

$$\begin{aligned}
f_{n,\varepsilon}(t) &= \frac{1}{2\pi} \int_{-\varepsilon\sqrt{n}}^{\varepsilon\sqrt{n}} e^{-i\lambda t} \varphi_n(\lambda) \hat{j}(\varepsilon\lambda) d\lambda + \mathcal{O}(\varepsilon^{-5} n^{-\frac{3}{2}}) \\
&= \frac{1}{2\pi} \int_{-\varepsilon\sqrt{n}}^{\varepsilon\sqrt{n}} e^{-i\lambda t} \varphi(\lambda) \hat{j}(\varepsilon\lambda) d\lambda + \mathcal{O}(\varepsilon^{-5} n^{-\frac{3}{2}} + n^{-\frac{1}{2}}) \\
&= g(t) + \mathcal{O}(\varepsilon + \varepsilon^{-5} n^{-\frac{3}{2}} + n^{-\frac{1}{2}}) = g(t) + \mathcal{O}(n^{-\frac{1}{2}})
\end{aligned}$$

provided we choose $n^{-\frac{1}{2}} \geq \varepsilon \geq n^{-5}$. Which, as announced, means that, if the precision of the instrument is compatible with the statistics, the typical fluctuations in measurements are of order $\frac{1}{\sqrt{n}}$ and Gaussian. This is well known by sperimentalists who routinely assume that the result of a measurement is distributed according to a Gaussian.[8]

Dipartimento di Matematica, Università di Roma (Tor Vergata), Via della Ricerca Scientifica, 00133 Roma, Italy, liverani@mat.uniroma2.it

[8]Note however that our proof holds in a very special case that has little to do with a real experimental setting. To prove the analogous statement in for a realistic experiment is a completely different ball game.

DRAFT -- DRAFT -- DRAFT -- DRAFT -- DRAFT --