01 Dec 2020

A Brief on Characteristic Functions

Austin G. Vandegriffe

A Brief on Characteristic Functions

A Presentation for Harmonic Analysis
Missouri S&T : Rolla, MO

Presentation by

Austin G. Vandegriffe

2020

Contents

Notation

$\forall$	For all
$\forall_{\mathbb{P}}$	For $\mathbb{P}$-almost all, where $\mathbb{P}$ is a measure
$\exists$	There exists
$\iff$	If and only if
$\biguplus$	Disjoint union
$\rightharpoonup$	Weak convergence
λ_n	Lebegue measure on $\mathbb{R}^n$; the n is omitted if $n = 1$
$\hat{\mu}$	Characteristic function of a measure μ
$\to$	Complete convergence
∂B	The boundary of a set B
$\bar{B}$	The closer of a set B
B°	The interior of a set B
B^c	The complement of a set B
$\mathscr{B}(\Omega)$	σ-algebra generated by open sets of the implicit topology $\tau(\Omega)$
$\mathscr{B}_n$	σ-algebra generated by the usual topology on $\mathbb{R}^n$
$\mathcal{B}_r(\omega)$	A ball of radius r about an element ω of an implicit metric space (Ω, d)
$\mathscr{C}^k(X, Y)$	Space of continuous functions from X to Y with k continuous derivatives
$\mathscr{C}_b(X, Y)$	Space of bounded continuous functions from X to Y
$\mathscr{C}_c(X, Y)$	Space of continuous functions from X to Y with compact support
$\mathscr{C}_p(X, Y)$	Space of periodic continuous functions from X to Y
$\mathbb{C}$	The complex numbers
$\mathrm{diag}(\alpha_1, ..., \alpha_n)$	An $n \times n$ matrix with diagonal elements $\alpha_1, ..., \alpha_n$
F_μ	The distribution function of μ, that is, $F_\mu(x) = \mu((-\infty, x))$
iff	If and only if
$L^p(\Omega, \Omega'; \mu)$	Space of functions $f : \Omega \to \Omega'$ such f^p is μ-integrable
$\mathscr{P}(\mathcal{F})$	Set of probability measures on $\mathcal{F}$
$\mathscr{P}_n$	Set of probability measures on $\mathscr{B}_n$
$\mathbb{Q}$	The rational numbers
$\mathscr{R}(\mathcal{F})$	Set of Radon measures on the σ-algebra $\mathcal{F}$
$\mathscr{R}_n$	Set of Radon measures on $\mathscr{B}_n$
$\mathbb{R}$	The real numbers
$s.t.$	Such that
$U_\omega \in \tau$	An open set in τ containing ω
$\overline{z}$	Complex conjugate

Ackowledgement

I would like to thank Dr. Jason C. Murphy[1] for motivating my studies in harmonic analysis which ultimately enabled me to write this work. Dr. Murphy also guided me through some proofs and replaced my doubts with clarity; his input was of great value.

[1] Jason Murphy: https://scholar.google.com/citations?user=32q4x_cAAAAJ&hl=en

Characteristic Functions

Introduction

Characteristic functions (CFs) are often used in problems involving convergence in distribution, independence of random variables, infinitely divisible distributions, and stochastics [5]. The most famous use of characteristic functions is in the proof of the Central Limit Theorem, also known as the Fundamental Theorem of Statistics. Though less frequent, CFs have also been used in problems of nonparametric time series analysis [6] and in machine learning [7–9]. Moreover, CFs uniquely determine their distribution, much like the moment generating functions (MGFs), but the major difference is that CFs *always* exists, whereas MGFs can fail, e.g. the Cauchy distribution. This makes CFs more robust in general.

In the following, I will present an introduction and basic properties of the Fourier-Stieltjes transform, it's inverse and relation to the Radon-Nikodym derivative, then go on to prove the Lévy Continuity Theorem, and finally a short presentation of measure convolutions. Much of the following presentation will be for probability measures and their distribution functions; however, some results can be generalized to (un)signed finite measures [1, 5, 10]. One can find an overview of background knowledge in the Appendix.

1 Basic Properites of Characteristic Functions

Definition 1.1 (Characteristic Function / Fourier–Stieltjes Transform)**.** *Let* $\boldsymbol{X} : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^n, \mathscr{B}_n, \langle \bullet, \bullet \rangle)$ *and* $\mu = X_{\#}\mathbb{P}$*, the pushforward measure / distribution of* X*. The* characteristic function *or* Fourier–Stieltjes transform *of* μ *is*

$$\hat{\mu}(\boldsymbol{t}) = \int\limits_{\Omega} e^{i\langle \boldsymbol{t}, \boldsymbol{X}(\omega)\rangle}\ d\mathbb{P}(\omega) = \int\limits_{\mathbb{R}^n} e^{i\langle \boldsymbol{t}, \boldsymbol{x}\rangle}\ d\mu(\boldsymbol{x})$$

where equality comes from Theorem B.7.

In terms of the Fourier transform, if μ admits a Radon–Nikodym derivative f with respect to the n-Lebesgue measure λ_n (Theorem B.26), then

$$\hat{\mu}(\boldsymbol{t}) = \int\limits_{\mathbb{R}^n} e^{i\langle \boldsymbol{t}, \boldsymbol{x}\rangle}\ d\mu(\boldsymbol{x}) = \int\limits_{\mathbb{R}^n} e^{i\langle \boldsymbol{t}, \boldsymbol{x}\rangle} f(\boldsymbol{x})\ d\lambda_n(\boldsymbol{x}) = \hat{f}(\boldsymbol{t})$$

That is, the characteristic function is nothing but the Fourier transform of the density of μ if one exists.

Theorem 1.2 (Properties of the Characteristic Function)**.** *Let* μ *be as above, then*

i) $\hat{\mu}(0) = 1$

ii) $|\hat{\mu}(t)| \leqq 1$

iii) $\hat{\mu}(-t) = \overline{\hat{\mu}(t)}$

iv) $\hat{\mu}$ *is uniformly continuous on* $\mathbb{R}^n$

v) $(\widehat{\mu_1 + \mu_2}) = \hat{\mu}_1 + \hat{\mu}_2$

Proof.

i)

$$\hat{\mu}(0) = \int_{\mathbb{R}^n} e^{i\langle 0, \boldsymbol{x}\rangle} \ d\mu(\boldsymbol{x}) = \int_{\mathbb{R}^n} \ d\mu(\boldsymbol{x}) = \mu(\mathbb{R}^n) = 1$$

ii)

$$\begin{aligned}
|\hat{\mu}(\boldsymbol{t})| &= \left| \int_{\mathbb{R}^n} e^{i\langle 0, \boldsymbol{x}\rangle} \ d\mu(\boldsymbol{x}) \right| \\
&\leqq \int_{\mathbb{R}^n} |e^{i\langle 0, \boldsymbol{x}\rangle}| \ d\mu(\boldsymbol{x}) \\
&\leqq \int_{\mathbb{R}^n} 1 \ d\mu(\boldsymbol{x}) \\
&= \mu(\mathbb{R}^n) = 1
\end{aligned}$$

iii)

$$\begin{aligned}
\hat{\mu}(-\boldsymbol{t}) &= \int_{\mathbb{R}^n} e^{i\langle \boldsymbol{-t}, \boldsymbol{x}\rangle} \ d\mu(\boldsymbol{x}) \\
&= \int_{\mathbb{R}^n} e^{\boldsymbol{-}i\langle \boldsymbol{t}, \boldsymbol{x}\rangle} \ d\mu(\boldsymbol{x}) \\
&= \int_{\mathbb{R}^n} \overline{e^{i\langle t, x\rangle}} \ d\mu(x) \quad \backslash\backslash \text{ since } e^{ix} = \overline{e^{-ix}} \\
&= \overline{\int_{\mathbb{R}^n} e^{i\langle t, x\rangle} \ d\mu(x)} \\
&= \overline{\hat{\mu}(t)}
\end{aligned}$$

iv) Let $\delta > 0$

$$|\hat{\mu}(t+\delta) - \hat{\mu}(t)| = \left| \int_{\mathbb{R}^n} e^{i\langle t+\delta, x\rangle} - e^{i\langle t, x\rangle} \ d\mu(x) \right|$$

$$
\begin{aligned}
&= \left| \int\limits_{\mathbb{R}^n} e^{i\langle t,x\rangle + i\langle \delta,x\rangle} - e^{i\langle t,x\rangle}\ d\mu(x) \right| \\
&= \left| \int\limits_{\mathbb{R}^n} e^{i\langle t,x\rangle} \cdot e^{i\langle \delta,x\rangle} - e^{i\langle t,x\rangle}\ d\mu(x) \right| \\
&\leqq \int\limits_{\mathbb{R}^n} \left| e^{i\langle t,x\rangle} \cdot e^{i\langle \delta,x\rangle} - e^{i\langle t,x\rangle} \right|\ d\mu(x) \\
&= \int\limits_{\mathbb{R}^n} \left| e^{i\langle t,x\rangle} \right| \left| e^{i\langle \delta,x\rangle} - 1 \right|\ d\mu(x) \\
&= \int\limits_{\mathbb{R}^n} \left| e^{i\langle \delta,x\rangle} - 1 \right|\ d\mu(x) \\
&\leqq \sup_{x\in\mathbb{R}^n} \left| e^{i\langle \delta,x\rangle} - 1 \right| \xrightarrow{\delta\downarrow 0} 0
\end{aligned}
$$

v)

$$
\begin{aligned}
(\widehat{\mu_1 + \mu_2}) &= \int\limits_{\mathbb{R}^n} e^{i\langle t,x\rangle}\ d[\mu_1 + \mu_2](x) \\
&= \int\limits_{\mathbb{R}^n} e^{i\langle t,x\rangle}\ d\mu_1(x) + \int\limits_{\mathbb{R}^n} e^{i\langle t,x\rangle}\ d\mu_2(x) \\
&= \hat{\mu}_1 + \hat{\mu}_2
\end{aligned}
$$

□

Theorem 1.3. *Let* $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n), \lambda)$ *be as usual,* $\mu \in \mathscr{P}(\mathscr{B}(\mathbb{R}^n))$, *and* $f \in L^1(\mathbb{R}^n, \mathbb{C}; \lambda)$, *then*

$$
\int\limits_{\mathbb{R}^n} \hat{f}\ d\mu = \int\limits_{\mathbb{R}^n} f\hat{\mu}\ d\lambda_n
$$

Proof. This follows from Fubini's theorem

$$
\begin{aligned}
\int\limits_{\mathbb{R}^n} \hat{f}\ d\mu &= \int\limits_{\mathbb{R}^n} \left(\int\limits_{\mathbb{R}^n} e^{i\pi\langle \boldsymbol{t},\boldsymbol{x}\rangle} f(\boldsymbol{x})\ d\lambda_n(\boldsymbol{x}) \right)\ d\mu(\boldsymbol{t}) \\
&= \int\limits_{\mathbb{R}^n} \left(\int\limits_{\mathbb{R}^n} e^{i\pi\langle \boldsymbol{t},\boldsymbol{x}\rangle} f(\boldsymbol{x})\ d\mu(\boldsymbol{t}) \right)\ d\lambda_n(\boldsymbol{x}) \\
&= \int\limits_{\mathbb{R}^n} f(\boldsymbol{x}) \left(\int\limits_{\mathbb{R}^n} e^{i\pi\langle t,x\rangle}\ d\mu(\boldsymbol{t}) \right)\ d\lambda_n(\boldsymbol{x})
\end{aligned}
$$

$$= \int_{\mathbb{R}^n} [f \cdot \hat{\mu}](\boldsymbol{x}) \; d\lambda_n(\boldsymbol{x})$$

□

Theorem 1.4 (Uniqueness)**.** *If $\mu_1, \mu_2 \in \mathscr{P}_n$, then*

$$\mu_1 = \mu_2 \iff \hat{\mu}_1 = \hat{\mu}_2 \; .$$

Note: the equality of the characteristic functions is over the real numbers*.*

Proof.

" $\Longrightarrow$ ": This is obvious

$$\hat{\mu}_1(t) = \int_{\mathbb{R}^n} e^{i\langle t,x\rangle} \; d\mu_1(x) = \int_{\mathbb{R}^n} e^{i\langle t,x\rangle} \; d\mu_2(x) = \hat{\mu}_2(t)$$

" $\Longleftarrow$ ": Suppose $\hat{\mu}_1(t) = \hat{\mu}_2(t)$ for all t, then

$$\int_{\mathbb{R}^n} e^{i\langle t,x\rangle} d[\mu_1 - \mu_2](x) = \int_{\mathbb{R}^n} e^{i\langle t,x\rangle} d\bar{\mu}(x) = 0 \quad \forall t \in \mathbb{R}^n$$

then for all trigonometric polynomials $\gamma_N(\bullet) = \sum_{t_1,\ldots,t_n=-N}^{N} c_{(t_1,\ldots,t_n)} e^{i\langle t_{(1,\ldots,n)},\bullet\rangle}$ we have

$$\int_{\mathbb{R}^n} \gamma_N(x) d\bar{\mu}(x) = 0$$

and so for the uniform limit $\gamma(\bullet)$ in N, which are the continuous periodic functions (Theorem C.6), we have

$$\int_{\mathbb{R}^n} \gamma(x) d\bar{\mu}(x) = 0$$

Let $c(\bullet)$ be a continuous function which vanishes outside a fixed, bounded support S, choose m such that $S \subseteq (-m, m]^{\otimes n}$, and choose $\gamma_m(\bullet)$ to be a continuous periodic function with period $2m$ such that $\gamma_m(x) = c(x) \; \forall x \in (-m, m]^{\otimes n}$. Since $\bar{\mu}$ is the difference of of two monotone increasing functions ($F_{\bar{\mu}} = F_{\mu_1} - F_{\mu_2}$), it is a function of bounded variation; hence $\forall \epsilon > 0 \; \exists \; m$ s.t. $\bar{\mu}(x) < \epsilon$ for $|x| \geqq m\mathbf{1}$ and we have

$$0 = \int_{R^n} \gamma_m \; d\bar{\mu} \xrightarrow{m\uparrow\infty} \int_{R^n} c \; d\bar{\mu} = \int_S c \; d\bar{\mu} = 0$$

Since c was arbitary, we have, by Theorem B.20 and taking c to be indicator functions on sets in $\mathscr{B}_n$, that

$$\bar{\mu} = \mu_1 - \mu_2 = 0$$

and so $\mu_1 = \mu_2$. □

2 Inversion Formula

Theorem 2.1 (Inversion)**.** *Let* $(\mathbb{R}^n, \mathscr{B}_n, \lambda)$ *be as usual;* $\mu \in \mathscr{P}_n$*;* $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ $(\boldsymbol{a} < \boldsymbol{b})$ *and let* $(\boldsymbol{a}, \boldsymbol{b}) = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a} < \boldsymbol{x} < \boldsymbol{b}\}$*, then*

$$\mu((\boldsymbol{a}, \boldsymbol{b})) + 2^{-n}\mu(\{\boldsymbol{a}, \boldsymbol{b}\}) = (2\pi)^{-n} \lim_{\boldsymbol{k}\uparrow\infty} \int\limits_{(-\boldsymbol{k},\boldsymbol{k})} \left[\prod_{j=1}^{n} \left(\frac{e^{-it_j a_j} - e^{-it_j b_j}}{it_j}\right)\right] \hat{\mu}(\boldsymbol{t})\ d\lambda_n(\boldsymbol{t})$$

there $\boldsymbol{k} = [k_1, ..., k_n]^T$ *and* $(-\boldsymbol{k}, \boldsymbol{k}) = \{\boldsymbol{x} \in \mathbb{R}^n : -\boldsymbol{k} \leqq \boldsymbol{x} \leqq \boldsymbol{k}\}$*.*

Proof. We have that

$$\begin{aligned}
\Phi(\mathbf{k}) &= \int\limits_{(-\mathbf{k},\mathbf{k})} \left[\prod_{j=1}^{n} \left(\frac{e^{-it_j a_j} - e^{-it_j b_j}}{it_j}\right)\right] \hat{\mu}(\mathbf{t})\ d\lambda_n(\mathbf{t}) \\
&= \int\limits_{(-\mathbf{k},\mathbf{k})} \left[\prod_{j=1}^{n} \left(\frac{e^{-it_j a_j} - e^{-it_j b_j}}{it_j}\right)\right] \left(\int\limits_{\mathbb{R}^n} e^{i\langle \mathbf{t},\mathbf{z}\rangle}\ d\mu(\mathbf{z})\right)\ d\lambda_n(\mathbf{t}) \\
&= \int\limits_{(-\mathbf{k},\mathbf{k})} \left(\int\limits_{\mathbb{R}^n} \left[\prod_{j=1}^{n} \left(\frac{e^{-it_j a_j} - e^{-it_j b_j}}{it_j}\right) e^{i\langle \mathbf{t},\mathbf{z}\rangle}\right]\ d\mu(\mathbf{z})\right)\ d\lambda_n(\mathbf{t}) \\
&= \int\limits_{(-\mathbf{k},\mathbf{k})} \left(\int\limits_{\mathbb{R}^n} \prod_{j=1}^{n} \left(\frac{e^{it_j(z_j - a_j)} - e^{it_j(z_j - b_j)}}{it_j}\right)\ d\mu(\mathbf{z})\right)\ d\lambda_n(\mathbf{t}) \\
&\overset{Fub.}{=} \int\limits_{\mathbb{R}^n} \left(\int\limits_{(-\mathbf{k},\mathbf{k})} \prod_{j=1}^{n} \left(\frac{e^{it_j(z_j - a_j)} - e^{it_j(z_j - b_j)}}{it_j}\right)\ d\lambda_n(\mathbf{t})\right)\ d\mu(\mathbf{z})
\end{aligned}$$

Where Fubini's theorem comes from

$$\begin{aligned}
|\mu_{\mathbf{k}}([\mathbf{a}, \mathbf{b}])| &= \left|(2\pi)^{-n} \int\limits_{(-\mathbf{k},\mathbf{k})} \left[\prod_{j=1}^{n} \left(\frac{e^{-it_j a_j} - e^{-it_j b_j}}{it_j}\right)\right] \hat{\mu}(\mathbf{t})\ d\lambda_n(\mathbf{t})\right| \\
&\leqq (2\pi)^{-n} \int\limits_{(-\mathbf{k},\mathbf{k})} \left|\left[\prod_{j=1}^{n} \left(\frac{e^{-it_j a_j} - e^{-it_j b_j}}{it_j}\right)\right] \hat{\mu}(\mathbf{t})\right|\ d\lambda_n(\mathbf{t}) \\
&= (2\pi)^{-n} \int\limits_{(-\mathbf{k},\mathbf{k})} \left|\left[\prod_{j=1}^{n} \left(\frac{e^{-it_j a_j} - e^{-it_j b_j}}{it_j}\right)\right] \left(\int\limits_{\mathbb{R}^n} e^{i\langle \mathbf{t},\mathbf{z}\rangle}\ d\mu(\mathbf{z})\right)\right|\ d\lambda_n(\mathbf{t})
\end{aligned}$$

$$= (2\pi)^{-n} \int\limits_{(-\mathbf{k},\mathbf{k})} \left| \left(\int\limits_{\mathbb{R}^n} \left[\prod_{j=1}^{n} \left(\frac{e^{-it_j a_j} - e^{-it_j b_j}}{it_j} \right) \right] e^{i\langle \mathbf{t},\mathbf{z} \rangle} \; d\mu(\mathbf{z}) \right) \right| \; d\lambda_n(\mathbf{t})$$

$$= (2\pi)^{-n} \int\limits_{(-\mathbf{k},\mathbf{k})} \left| \left(\int\limits_{\mathbb{R}^n} \left[\prod_{j=1}^{n} \left(\int\limits_{a_j}^{b_j} e^{-it_j u} \; d\lambda_1(u) \right) \right] e^{i\langle \mathbf{t},\mathbf{z} \rangle} \; d\mu(\mathbf{z}) \right) \right| \; d\lambda_n(\mathbf{t})$$

$$\leqq (2\pi)^{-n} \int\limits_{(-\mathbf{k},\mathbf{k})} \left(\int\limits_{\mathbb{R}^n} \left[\prod_{j=1}^{n} \left(\int\limits_{a_j}^{b_j} \left| e^{-it_j u} \right| \; d\lambda_1(u) \right) \right] \left| e^{i\langle \mathbf{t},\mathbf{z} \rangle} \right| \; d\mu(\mathbf{z}) \right) \; d\lambda_n(\mathbf{t})$$

$$= (2\pi)^{-n} \int\limits_{(-\mathbf{k},\mathbf{k})} \left(\int\limits_{\mathbb{R}^n} \lambda_n([\mathbf{a},\mathbf{b}]) \; d\mu(\mathbf{z}) \right) \; d\lambda_n(\mathbf{t})$$

$$= (2\pi)^{-n} \lambda_n((-\mathbf{k},\mathbf{k})) \lambda_n([\mathbf{a},\mathbf{b}]) \mu(\mathbb{R}^n)$$

$$= (2\pi)^{-n} \lambda_n((-\mathbf{k},\mathbf{k})) \lambda_n([\mathbf{a},\mathbf{b}]) < \infty \quad \forall k \in \mathbb{R}^{+n}$$

Continuing from the application of Fubini's theorem we get

$$\Phi(\mathbf{k}) = \lim_{\mathbf{k}\uparrow\infty} \int\limits_{\mathbb{R}^n} \prod_{j=1}^{n} \left[\int\limits_{-k_j}^{k_j} \frac{e^{it_j(z_j - a_j)} - e^{it_j(z_j - b_j)}}{it_j} \; d\lambda_1(t_j) \right] \; d\mu(\mathbf{z})$$

$$= \lim_{\mathbf{k}\uparrow\infty} \int\limits_{\mathbb{R}^n} \prod_{j=1}^{n} \left[\int\limits_{-k_j}^{k_j} \frac{e^{it_j(z_j - a_j)}}{it_j} \; dt_j - \int\limits_{-k_j}^{k_j} \frac{e^{it_j(z_j - b_j)}}{it_j} \; d\lambda_1(t_j) \right] \; d\mu(\mathbf{z})$$

$$= \lim_{\mathbf{k}\uparrow\infty} \int\limits_{\mathbb{R}^n} \prod_{j=1}^{n} \left[\int\limits_{-k_j}^{k_j} \frac{\cos(t_j(z_j - a_j) + i\sin(t_j(z_j - a_j)}{it_j} \; d\lambda_1(t_j) \right.$$

$$\left. - \int\limits_{-k_j}^{k_j} \frac{\cos(t_j(z_j - b_j) + i\sin(t_j(z_j - b_j)}{it_j} \; d\lambda_1(t_j) \right] \; d\mu(\mathbf{z})$$

Now, the trick here is to use the Cauchy Principal Value C.9 of $\dfrac{\cos(t(z-c))}{it}$, that is

$$PV \int\limits_{-k}^{k} \frac{\cos(t(z-c)}{it} \; d\lambda_1(t) = \lim_{\epsilon\downarrow 0} \left[\int\limits_{-k}^{0-\epsilon} \frac{\cos(t(z-c))}{it} \; d\lambda_1(t) + \int\limits_{0+\epsilon}^{k} \frac{\cos(t(z-c))}{it} \; d\lambda_1(t) \right] = 0$$

by antisymmetry. So we obtain

$$\mu(B) = \lim_{\mathbf{k}\uparrow\infty} \int\limits_{\mathbb{R}^n} \prod_{j=1}^{n} \left[\int\limits_{-k_j}^{k_j} \frac{\cos(t_j(z_j - a_j) + i\sin(t_j(z_j - a_j)}{it_j} \; d\lambda_1(t_j) \right.$$

$$
- \int_{-k_j}^{k_j} \frac{\cos(t_j(z_j - b_j) + i\sin(t_j(z_j - b_j)}{it_j} \, d\lambda_1(t_j) \Bigg] \, d\mu(\mathbf{z})
$$

$$
= \lim_{\mathbf{k}\uparrow\infty} \int_{\mathbb{R}^n} \prod_{j=1}^{n} \left[\int_{-k_j}^{k_j} \frac{\sin(t_j(z_j - a_j)}{t_j} \, d\lambda_1(t_j) - \int_{-k_j}^{k_j} \frac{\sin(t_j(z_j - b_j)}{t_j} \, d\lambda_1(t_j) \right] \, d\mu(\mathbf{z})
$$

$$
= \lim_{\mathbf{k}\uparrow\infty} \int_{\mathbb{R}^n} \prod_{j=1}^{n} \left[\int_{-k_j(z_j-a_j)}^{k_j(z_j-a_j)} \frac{\sin(u_j)}{u_j} \, d\lambda_1(u_j) - \int_{-k_j(z_j-b_j)}^{k_j(z_j-b_j)} \frac{\sin(v_j)}{v_j} \, d\lambda_1(v_j) \right] \, d\mu(\mathbf{z})
$$

where $u_j = t_j(z_j - a_j)$ and $v_j = t_j(z_j - b_j)$. Since $\dfrac{\sin(x)}{x}$ is uniformly continuous, and so, for all $\forall (k, c) \in \mathbb{R}^2$, the integral functions

$$
\psi^{\pm}_{(k,c)}(z) = \pm \int_{0}^{\pm k(z-c)} \frac{\sin(x)}{x} \, d\lambda_1(x)
$$

is uniformly continuous in k, and so

$$
\Psi_{(k,c)}(z) = \psi^{+}_{(k,c)}(z) + \psi^{-}_{(k,c)}(z)
$$

is uniformly continuous with limits

$$
\lim_{k\uparrow\infty} \Psi_{(k,c)}(z) = \begin{cases} 0 & : z = c \\ \frac{\pi}{2} & : o.w. \end{cases}
$$

This gives five cases for

$$
\Psi(\mathbf{z}; (\mathbf{a}, \mathbf{b})) = \lim_{\mathbf{k}\uparrow\infty} \prod_{j=1}^{n} \left[\int_{-k_j(z_j-a_j)}^{k_j(z_j-a_j)} \frac{\sin(u_j)}{u_j} \, d\lambda_1(u_j) - \int_{-k_j(z_j-b_j)}^{k_j(z_j-b_j)} \frac{\sin(v_j)}{v_j} \, d\lambda_1(v_j) \right]
$$

$$
= \lim_{\mathbf{k}\uparrow\infty} \prod_{j=1}^{n} \left[\Psi_{(k_j,a_j)}(z_j) + \Psi_{(k_j,b_j)}(z_j) \right]
$$

namely

$$
\Psi(\mathbf{z}; (\mathbf{a}, \mathbf{b})) = \begin{cases} 0 & : \mathbf{z} < \mathbf{a} \\ 2^{-n} & : \mathbf{z} = \mathbf{a} \\ 1 & : \mathbf{a} < \mathbf{z} < \mathbf{b} \\ 2^{-n} & : \mathbf{z} = \mathbf{b} \\ 0 & : \mathbf{b} < \mathbf{z} \end{cases}
$$

Finally, we have

$$\Phi(\mathbf{k}) \xrightarrow{\mathbf{k}\uparrow\infty} \int_{\mathbb{R}^n} \Psi(\mathbf{z};(\mathbf{a},\mathbf{b}))\ d\mu(\mathbf{z}) = \mu((\mathbf{a},\mathbf{b})) + 2^{-n}\mu(\{\mathbf{a},\mathbf{b}\})$$

□

Corollary 2.2. *If $\hat{\mu}$ is integrable, then μ is absolutely continuous with respect to the Lebesgue measure λ and its Radon-Nikodym derivative is given by*

$$D_\lambda\mu(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{-i\langle t,x\rangle}\hat{\mu}\ d\lambda_1(t)$$

Proof. We will first show that μ has not atoms. Let $M = \int_{\mathbb{R}^n} |\hat{\mu}(\mathbf{t})|\ d\lambda_n(\mathbf{t})$

$$\begin{aligned}
\mu((\mathbf{a},\mathbf{b})) + 2^{-n}\mu(\{\mathbf{a},\mathbf{b}\}) &= (2\pi)^{-n} \lim_{\mathbf{k}\uparrow\infty} \int_{(-\mathbf{k},\mathbf{k})} \left[\prod_{j=1}^{n}\left(\frac{e^{-it_ja_j} - e^{-it_jb_j}}{it_j}\right)\right]\hat{\mu}(\mathbf{t})\ d\lambda_n(\mathbf{t}) \\
&\leqq (2\pi)^{-n} \lim_{\mathbf{k}\uparrow\infty} \int_{(-\mathbf{k},\mathbf{k})} \left|\left[\prod_{j=1}^{n}\left(\frac{e^{-it_ja_j} - e^{-it_jb_j}}{it_j}\right)\right]\right| |\hat{\mu}(\mathbf{t})|\ d\lambda_n(\mathbf{t}) \\
&= (2\pi)^{-n} \lim_{\mathbf{k}\uparrow\infty} \int_{(-\mathbf{k},\mathbf{k})} \left|\left[\prod_{j=1}^{n}\left(\int_{a_j}^{b_j} e^{-it_ju}\ d\lambda_1(u)\right)\right]\right| |\hat{\mu}(\mathbf{t})|\ d\lambda_n(\mathbf{t}) \\
&\leqq (2\pi)^{-n} \lim_{\mathbf{k}\uparrow\infty} \int_{(-\mathbf{k},\mathbf{k})} \left[\prod_{j=1}^{n}\left(\int_{a_j}^{b_j} \left|e^{-it_ju}\right|\ d\lambda_1(u)\right)\right] |\hat{\mu}(\mathbf{t})|\ d\lambda_n(\mathbf{t}) \\
&= (2\pi)^{-n} \lim_{\mathbf{k}\uparrow\infty} \int_{(-\mathbf{k},\mathbf{k})} \lambda_n([\mathbf{a},\mathbf{b}])\,|\hat{\mu}(\mathbf{t})|\ d\lambda_n(\mathbf{t}) \\
&= (2\pi)^{-n}\lambda_n([\mathbf{a},\mathbf{b}]) \int_{\mathbb{R}^n} |\hat{\mu}(\mathbf{t})|\ d\lambda_n(\mathbf{t}) \\
&= (2\pi)^{-n}\lambda_n([\mathbf{a},\mathbf{b}])M < \infty
\end{aligned}$$

Now, focus on an arbitrary point $\mathbf{x} \in \mathbb{R}^n$. Let $\boldsymbol{\delta} > 0$

$$\begin{aligned}
\mu((\mathbf{x} - \frac{\boldsymbol{\delta}}{2}, \mathbf{x} + \frac{\boldsymbol{\delta}}{2})) + 2^{-n}\mu(\{\mathbf{x} - \frac{\boldsymbol{\delta}}{2}, \mathbf{x} + \frac{\boldsymbol{\delta}}{2}\}) &\leqq (2\pi)^{-n}\lambda_n([\mathbf{x} - \frac{\boldsymbol{\delta}}{2}, \mathbf{x} + \frac{\boldsymbol{\delta}}{2}])M \\
&= (2\pi)^{-n}\delta M \xrightarrow{\boldsymbol{\delta}\downarrow 0} 0
\end{aligned}$$

Hence, μ has no atoms. Now for the Radon-Nikodym derivative we use Theorems B.26 and B.27:

$$\begin{aligned}
D_\lambda \mu(\mathbf{x}) &= \lim_{\delta\downarrow 0} \frac{\mu(\overline{\mathcal{B}_\delta(x)})}{\lambda_n(\overline{\mathcal{B}_\delta(x)})} = \lim_{\delta\downarrow 0} \frac{\mu(\overline{\mathcal{B}_\delta(x)})}{\delta^n} \\
&= \lim_{\delta\downarrow 0} \left[(2\pi\delta)^{-n} \lim_{\mathbf{k}\uparrow\infty} \int\limits_{(-\mathbf{k},\mathbf{k})} \left[\prod_{j=1}^{n} \left(\frac{e^{-it_j(x_j-\frac{\delta}{2})} - e^{-it_j(x_j+\frac{\delta}{2})}}{it_j} \right) \right] \hat{\mu}(\mathbf{t})\ d\lambda_n(\mathbf{t}) \right] \\
&= \lim_{\delta\downarrow 0}\ (2\pi)^{-n} \int\limits_{\mathbb{R}^n} \left[\prod_{j=1}^{n} \left(\frac{e^{-it_j(x_j-\frac{\delta}{2})} - e^{-it_j(x_j+\frac{\delta}{2})}}{it_j\delta} \right) \right] \hat{\mu}(\mathbf{t})\ d\lambda_n(\mathbf{t}) \\
&\overset{D.C.}{=} (2\pi)^{-n} \int\limits_{\mathbb{R}^n} \left[\lim_{\delta\downarrow 0} \prod_{j=1}^{n} \left(\frac{e^{-it_j(x_j-\frac{\delta}{2})} - e^{-it_j(x_j+\frac{\delta}{2})}}{it_j\delta} \right) \right] \hat{\mu}(\mathbf{t})\ d\lambda_n(\mathbf{t}) \\
&= (2\pi)^{-n} \int\limits_{\mathbb{R}^n} e^{-i\langle \mathbf{t},\mathbf{x}\rangle} \hat{\mu}(\mathbf{t})\ d\lambda_n(\mathbf{t})
\end{aligned}$$

□

3 Convergence & Continuity of Characteristic Functions

We want to further relate measures and their characteristic functions. In the following, we develop convergence of characteristic functions from convergence in measures, and vice versa plus properties of said convergences.

The main result of this section is

Theorem 3.1 (Lévy's Continuity Theorem)**.**

$$\mu_n \xrightarrow{n\uparrow\infty} \mu \iff \hat{\mu}_n \to f$$

where f is continuous at the origin. Moreover, $\hat{\mu} = f$.

Proof. We need a couple lemmas:

Lemma 3.2. *Let $(\mu_k)_{k\geqq 1} \subset \mathscr{R}_n$ and $\mu \in \mathscr{R}_n$. If $\mu_k \xrightarrow{k\uparrow\infty} \mu$, then $\hat{\mu}_k(t) \xrightarrow{k\uparrow\infty} \hat{\mu}(t)$ uniformly for t belonging to a compact subset of $\mathbb{R}^n$.*

Proof. We must show that $(\hat{\mu}_k)_{k\geqq 1}$ is a uniformly equicontinuous family of functions (see Definition C.1): Let $\epsilon > 0$. By Corollary B.15, the sequence $(\mu_k)_{k\geqq 1}$ is tight; hence, we can find a compact interval $K = \overline{\mathcal{B}_r(0)}$ such that

$$\mu_k(K^c) \leqq \epsilon\ \forall k$$

By the differentiability of $e^{i\langle t,x\rangle}$ $\exists\delta > 0$ such that for $\xi_1, \xi_2 \in \mathbb{R}^n$ $(|\xi_1 - \xi_2| < \delta)$

$$\begin{aligned}|e^{i\langle\xi_1,x\rangle} - e^{i\langle\xi_2,x\rangle}| &= |e^{i\langle\xi_1,x\rangle}||1 - e^{i\langle\xi_2-\xi_1,x\rangle}| \\ &\leqq |1 - e^{i\langle\xi_2-\xi_1,x\rangle}| \\ &< \epsilon\end{aligned}$$

for all $x \in K$. Hence,

$$\begin{aligned}|\hat{\mu}_k(\xi_1) - \hat{\mu}_k(\xi_2)| &\leqq \int\limits_K |1 - e^{i\langle\xi_1-\xi_2,x\rangle}|\ d\mu_k(x) + \int\limits_{K^c} |1 - e^{i\langle\xi_1-\xi_2,x\rangle}|\ d\mu_k(x) \\ &\leqq [\epsilon\mu_n(K)] + \left[\ \int\limits_{K^c} |1| + |e^{i\langle\xi_1-\xi_2,x\rangle}|\ d\mu_k(x)\right] \\ &< \epsilon\mu_n(\mathbb{R}^n) + 2\mu_k(K^c) \\ &\leqq \epsilon + 2\epsilon = 3\epsilon \perp\!\!\!\perp n\end{aligned}$$

recalling that $\mu_k(K^c) \leqq \epsilon$. This give uniform equicontinuity on K; hence,

$$\begin{aligned}|\hat{\mu}_n(\xi_1) - \hat{\mu}(\xi_1)| &\leqq \underbrace{|\hat{\mu}_n(\xi_1) - \hat{\mu}_n(\xi_2)|}_{\text{equicontinuity}} + \underbrace{|\hat{\mu}_n(\xi_2) - \hat{\mu}(\xi_2)|}_{\text{convergence}} + \underbrace{|\hat{\mu}(\xi_1) - \hat{\mu}(\xi_2)|}_{\text{Continuity: Thm. 1.2}} \\ &\leqq 3\epsilon \perp\!\!\!\perp n\end{aligned}$$

which gives uniform convergence on a compact set. $\square$

Lemma 3.3. *Let* $\mu \in \mathscr{R}_n \cap \mathscr{P}_n$*, then* $\forall k > 0$

$$\mu\left(\left\{\boldsymbol{x} \in \mathbb{R}^n : |\boldsymbol{x}| > \frac{2}{k}\right\}\right) \leqq k^{-n} \int\limits_{\mathcal{B}_k(0)} \hat{\mu}(0) - \hat{\mu}(\boldsymbol{t})\ d\lambda_n(\boldsymbol{t})$$

where $|\boldsymbol{x}| > \frac{2}{k} \implies |x_j| > \frac{2}{k}$ *(vector absolute value).*

Proof. With $\hat{\mu}(0) = 1$ and $\mu(\mathbb{R}^n) = 1$, we have immediately that

$$
\begin{aligned}
k^{-n} \int\limits_{\mathcal{B}_k(0)} \hat{\mu}(0) - \hat{\mu}(\boldsymbol{t}) \ d\lambda_n(\boldsymbol{t}) &= k^{-n} \int\limits_{\mathcal{B}_k(0)} \left[1 - \int\limits_{\mathbb{R}^n} e^{i\langle \boldsymbol{t}, \boldsymbol{x} \rangle} \ d\mu(\boldsymbol{x}) \right] \ d\lambda_n(\boldsymbol{t}) \\
&= k^{-n} \int\limits_{\mathcal{B}_k(0)} \left[\int\limits_{\mathbb{R}^n} \cancel{\mu(\mathbb{R}^n)}\, 1 - e^{i\langle \boldsymbol{t}, \boldsymbol{x} \rangle} \ d\mu(\boldsymbol{x}) \right] \ d\lambda_n(\boldsymbol{t}) \\
&\overset{Fub.}{=} k^{-n} \int\limits_{\mathbb{R}^n} \left[\int\limits_{\mathcal{B}_k(0)} 1 - e^{i\langle \boldsymbol{t}, \boldsymbol{x} \rangle} \ d\lambda_n(\boldsymbol{t}) \right] \ d\mu(\boldsymbol{x}) \\
&= k^{-n} \int\limits_{\mathbb{R}^n} \left[\int\limits_{\mathcal{B}_k(0)} 1 - \left[\prod_{j=1}^{n} e^{it_j x_j} \right] \ d\lambda_n(\boldsymbol{t}) \right] \ d\mu(\boldsymbol{x}) \\
&= k^{-n} \int\limits_{\mathbb{R}^n} \left[\lambda_n(\mathcal{B}_k(0)) - \int\limits_{\mathcal{B}_k(0)} \left[\prod_{j=1}^{n} e^{it_j x_j} \right] \ d\lambda_n(\boldsymbol{t}) \right] \ d\mu(\boldsymbol{x}) \\
& \qquad (\star)
\end{aligned}
$$

Noting the independence of the Lebesgue measure: $\int\limits_{\mathbb{R}^n} \prod\limits_{j \in [n]} f(t_j) \ d\lambda_n(\boldsymbol{t}) = \prod\limits_{j \in [n]} \int\limits_{\mathbb{R}^n} f(t_j) \ d\lambda_1(t_j)$ and that $\mathcal{B}_k(0) \subset [-k\mathbf{1}, k\mathbf{1}]$ we continue to obtain

$$
\begin{aligned}
(\star) &\geqq k^{-n} \int\limits_{\mathbb{R}^n} \left[(2k)^n - \left[\prod_{j=1}^{n} \int_{-k}^{k} e^{it_j x_j} \ d\lambda_1(t_j) \right] \right] \ d\mu(\boldsymbol{x}) \\
&= \int\limits_{\mathbb{R}^n} \left[2^n - k^{-n} \left[\prod_{j=1}^{n} \int_{-k}^{k} \cos(t_j x_j) + i \sin(t_j x_j) \ d\lambda_1(t_j) \right] \right] \ d\mu(\boldsymbol{x}) \qquad (\star\star)
\end{aligned}
$$

Now by antisymmetry of sine, we have that

$$
\int_{-k}^{k} \cos(t_j x_j) + i \sin(t_j x_j) \ d\lambda_1(t_j) = \int_{-k}^{k} \cos(t_j x_j) \ d\lambda_1(t_j) + i \int_{-k}^{k} \sin(t_j x_j) \ d\lambda_1(t_j) = \int_{-k}^{k} \cos(t_j x_j) \ d\lambda_1(t_j)
$$

And so, noting that $1 - \frac{\sin(kx)}{kx} \geqq 0$, we have

$$
(\star\star) = \int\limits_{\mathbb{R}^n} \left[2^n - k^{-n} \left[\prod_{j=1}^{n} \int_{-k}^{k} \cos(t_j x_j) \ d\lambda_1(t_j) \right] \right] \ d\mu(\boldsymbol{x})
$$

$$
\begin{aligned}
&= \int\limits_{\mathbb{R}^n} \left[2^n - k^{-n} \left[\prod_{j=1}^{n} \frac{2\sin(kx_j)}{x_j} \right] \right] \, d\mu(\boldsymbol{x}) \\
&= 2^n \int\limits_{\mathbb{R}^n} 1 - \left[\prod_{j=1}^{n} \frac{\sin(kx_j)}{kx_j} \right] \, d\mu(\boldsymbol{x}) \\
&\geqq 2^n \int\limits_{\{\boldsymbol{x}:|\boldsymbol{x}|>\frac{2}{k}\}} 1 - \left[\prod_{j=1}^{n} \frac{\sin(kx_j)}{kx_j} \right] \, d\mu(\boldsymbol{x}) \\
&\geqq 2^n \int\limits_{\{\boldsymbol{x}:|\boldsymbol{x}|>\frac{2}{k}\}} 1 - \left[\prod_{j=1}^{n} \frac{1}{k|x_j|} \right] \, d\mu(\boldsymbol{x}) \\
&= \int\limits_{\{\boldsymbol{x}:|\boldsymbol{x}|>\frac{2}{k}\}} 2^n \left(1 - \left[\prod_{j=1}^{n} \frac{1}{k|x_j|} \right] \right) \, d\mu(\boldsymbol{x}) \\
&= \int\limits_{\{\boldsymbol{x}:|\boldsymbol{x}|>\frac{2}{k}\}} 2^n \left(1 - 2^{-n} \right) \, d\mu(\boldsymbol{x}) \qquad (\overset{\star}{\star\star}) \\
&\geqq \int\limits_{\{\boldsymbol{x}:|\boldsymbol{x}|>\frac{2}{k}\}} 1 \, d\mu(\boldsymbol{x}) \\
&= \mu\left(\left\{ \boldsymbol{x} \in \mathbb{R}^n : |\boldsymbol{x}| > \frac{2}{k} \right\} \right)
\end{aligned}
$$

where $(\overset{\star}{\star\star})$ is because $\left(1 - \left[\prod_{j=1}^{n} \frac{1}{k|x_j|}\right]\right)$ is minimized when $\frac{1}{k|x_j|}$ is maximized at $x_i = \frac{2}{k}$. □

Lemma 3.4. *Let $(\mu_k)_{k \geqq 1} \subset \mathscr{R}_n \cap \mathscr{P}_n$. If $\hat{\mu}_k \xrightarrow{k\uparrow\infty} \hat{\mu}$ is continuous at the origin, then $(\mu_k)_{k \geqq 1}$ tight.*

Proof. From the previous lemma, we have, for all $k > 0$

$$
\mu\left(\left\{ \boldsymbol{x} \in \mathbb{R}^n : |\boldsymbol{x}| > \frac{2}{k} \right\} \right) \leqq k^{-n} \int\limits_{\mathcal{B}_k(0)} \hat{\mu}(0) - \hat{\mu}(\boldsymbol{t}) \, d\lambda_n(\boldsymbol{t})
$$

And since $\hat{\mu}_k \xrightarrow{k\uparrow\infty} \hat{\mu}$: $\forall \epsilon > 0 \; \exists \; k_* > 0$ s.t.

$$
k_*^{-n} \int\limits_{\mathcal{B}_{k_*}(0)} \hat{\mu}(0) - \hat{\mu}(\boldsymbol{t}) \, d\lambda_n(\boldsymbol{t}) < \epsilon
$$

and by dominated convergence $\exists\ J$ s.t.

$$k_*^{-n} \int_{\mathcal{B}_{k_*}(0)} \hat{\mu}_j(0) - \hat{\mu}_j(\boldsymbol{t})\ d\lambda_n(\boldsymbol{t}) < \epsilon \ \ \forall j > J$$

which implies

$$\mu_j\left(\left\{\boldsymbol{x} \in \mathbb{R}^n : |\boldsymbol{x}| > \frac{2}{k_*}\right\}\right) < \epsilon$$

From Lemma 3.3, $\exists\ k_\ell$ for each $(\mu_\ell)_{\ell \leqq J}$ s.t.

$$\mu_\ell\left(\left\{\boldsymbol{x} \in \mathbb{R}^n : |\boldsymbol{x}| > \frac{2}{k_\ell}\right\}\right) < \epsilon$$

So if we take $\tilde{k} = \min(\{k_*, k_1, ..., k_J\})$, we have that

$$\mu_j\left(\left\{\boldsymbol{x} \in \mathbb{R}^n : |\boldsymbol{x}| > \frac{2}{\tilde{k}}\right\}\right) < \epsilon\ \forall j \geqq 1$$

hence (μ_k) is tight. □

Lemma 3.5. *Let $(\mu_k) \subset \mathscr{R} \cap \mathscr{P}_n$. If $\hat{\mu}_k \xrightarrow{k\uparrow\infty} f$, where f is continuous at the origin, then $\exists\ \mu$ s.t. $\mu_k \xrightarrow{k\uparrow\infty} \mu$ and $\hat{\mu} = f$.*

Proof. From Lemma 3.4 and Theorem B.16, there exists $(\mu_{k_j})_{j\geqq 1} \subset (\mu_k)_{k\geqq 1}$ which converges completely to some μ. Now, by Lemma 3.2, we have that $\hat{\mu}_{k_j} \xrightarrow{j\uparrow\infty} \hat{\mu}$, and so $f = \hat{\mu}$. From the uniqueness of the characteristic function, the limit μ is independent of the sequence $(\mu_{k_j})_{j\geqq 1}$, and thus, by Corollary B.19, $\mu_k \xrightarrow{k\uparrow\infty} \mu$ □

Proof of Theorem 3.1.
"$\Longrightarrow$": By weak convergence we have

$$\hat{\mu}_n(\boldsymbol{t}) = \int_{R^n} e^{i\langle \boldsymbol{t}, \boldsymbol{x}\rangle} d\mu_n(\boldsymbol{x}) \xrightarrow{n\uparrow\infty} \int_{R^n} e^{i\langle \boldsymbol{t}, \boldsymbol{x}\rangle} d\mu(\boldsymbol{x}) = \hat{\mu}(\boldsymbol{t})$$

"$\Longleftarrow$": The result follows from Lemma 3.5.. □

4 Convolution of Measures

For all $B \in \mathscr{B}_n$ and $x \in \mathbb{R}^n$, let

$$B - u = \{b - u : b \in B\}$$

For any $\mu_1, \mu_2 \in \mathscr{P}_n$ and $B \in \mathscr{B}_n$, the function

$$z \mapsto \mu_i(B - z)$$

is a bounded function and measureable.

Theorem 4.1 (Convolution of Measures). *Let $\mu_1, \mu_2 \in \mathscr{P}_n \cap \mathscr{R}_n$, then the convolution*

$$\mu(\bullet) = [\mu_1 * \mu_2](\bullet) = \int\limits_{\mathbb{R}^n \times \mathbb{R}^n} \mathbb{1}_\bullet(x+y)\ d[\mu_1 \otimes \mu_2](x,y) = \int\limits_{x+y\in\bullet} d[\mu_1 \otimes \mu_2](x,y)$$

is a finite measure, and

$$\mu(\bullet) = \int\limits_{\mathbb{R}^n} \mu_2(x - \bullet)\ d\mu_1(x) = \int\limits_{\mathbb{R}^n} \mu_1(\bullet - y)\ d\mu_2(y)$$

Proof. Clearly

$$\begin{aligned}
\mu(B) &= \int\limits_{\mathbb{R}^n \times \mathbb{R}^n} \mathbb{1}_B(x+y)\ d[\mu_1 \otimes \mu_2](x,y) \\
&\leqq \int\limits_{\mathbb{R}^n \times \mathbb{R}^n} d[\mu_1 \otimes \mu_2](x,y) \\
&= [\mu_1 \otimes \mu_2](\mathbb{R}^n \times \mathbb{R}^n) = \mu_1(\mathbb{R}^n)\mu_2(\mathbb{R}^n) < \infty\ \forall B \in \mathscr{B}_n
\end{aligned}$$

Now let $(A_n)_{n\geqq 1} \subset \mathscr{B}_n$ be a disjoint sequence, then, noting that $\mathbb{1}_{A_i \uplus A_j} = \mathbb{1}_{A_i} + \mathbb{1}_{A_j}\ \forall i \neq j$,

$$\begin{aligned}
\mu\Big(\biguplus_{n\geqq 1} A_n\Big) &= \int\limits_{\mathbb{R}^n \times \mathbb{R}^n} \mathbb{1}_{\uplus A_n}(x+y)\ d[\mu_1 \otimes \mu_2](x,y) \\
&= \int\limits_{\mathbb{R}^n \times \mathbb{R}^n} \left[\sum_{n\geqq 1} \mathbb{1}_{A_n}(x+y)\right] d[\mu_1 \otimes \mu_2](x,y) \\
&\overset{DC}{=} \sum_{n\geqq 1} \left[\int\limits_{\mathbb{R}^n \times \mathbb{R}^n} \mathbb{1}_{A_n}(x+y)\ d[\mu_1 \otimes \mu_2](x,y)\right] \\
&= \sum_{n\geqq 1} \mu(A_n)
\end{aligned}$$

where dominated convergence comes from the disjointness of A_n, that is, given an (x,y), $x+y$ only belongs to at most one A_n, hence $\sum\limits_{n\geqq 1} \mathbb{1}_{A_n}(x+y) \leqq 1$, which is integrable. Finally, let $B \in \mathscr{B}_n$, then

$$\begin{aligned}
\mu(B) &= \int\limits_{x+y\in B} d[\mu_1 \otimes \mu_2](x,y) \\
&= \int\limits_{x\in B-y} d[\mu_1 \otimes \mu_2](x,y) \\
&= \int\limits_{\mathbb{R}^n\times\mathbb{R}^n} \mathbb{1}_{B-y}(x)\ d[\mu_1 \otimes \mu_2](x,y) \\
&= \int\limits_{\mathbb{R}^n} \mu_1(B-y)\ d\mu_2(y)
\end{aligned}$$

and similarly we obtain $\mu(B) = \int\limits_{\mathbb{R}^n} \mu_2(x-B)\ d\mu_1(x)$. □

Theorem 4.2 (Convolution Theorem)**.** *Let $\mu_1, \mu_2 \in \mathscr{R}_n$, and $\mu = \mu_1 * \mu_2$, then $\hat{\mu} = \hat{\mu}_1 \cdot \hat{\mu}_2$.*

Proof. With $d\mu(u) = \mathbb{1}_u(x+y)d[\mu_1 \otimes \mu_2](x,y)$, we have

$$\begin{aligned}
\hat{\mu}(t) &= \int\limits_{\mathbb{R}^n} e^{i\langle t,u\rangle} d\mu(u) \\
&= \int\limits_{\mathbb{R}^n\times\mathbb{R}^n} e^{i\langle t,x+y\rangle} d[\mu_1 \otimes \mu_2](x,y) \\
&= \int\limits_{\mathbb{R}^n\times\mathbb{R}^n} e^{i\langle t,x\rangle} e^{i\langle t,y\rangle} d[\mu_1 \otimes \mu_2](x,y) \\
&= \left[\int\limits_{\mathbb{R}^n} e^{i\langle t,x\rangle} d\mu_1(x)\right] \left[\int\limits_{\mathbb{R}^n} e^{i\langle t,y\rangle} d\mu_2(y)\right] \\
&= [\hat{\mu}_1 \cdot \hat{\mu}_2](t)
\end{aligned}$$

□

Now, using the uniqueness of the characteristic functions, we immediately have

Corollary 4.3. *The product of the characteristic functions of two Radon measures is the characteristic function of a Radon measure.*

Consequences of the above corollary is that the product of the characteristic functions of two *probability measures* is the characteristic function of a probability measure.

Preliminaries*

*This section has not been developed in detail as the focus of this text is characteristic functions. The details are left as a future work.

A Topology

Definition A.1 (Topology, Open/Closed Sets)**.** *Let Ω be a non-empty set, then a collection $\tau \subseteq 2^{\Omega}$ is a* topology *if*

i. $\emptyset, \Omega \in \tau$

ii. $T_1, T_2 \in \tau \implies T_1 \cap T_2 \in \tau$

iii. $(T_i)_{i \in \mathcal{I} \subset \mathbb{R}} \subset \tau \implies \bigcup_{i \in \mathcal{I}} T_i \in \tau$

Sets in τ are called open, *if $T^c \in \tau$ then T is closed, and if T is both open and closed it is* clopen.

Definition A.2 (Set Closer)**.** *Let (Ω, τ) be a topological space. The* closer *of a set $B \subset \Omega$ is the smallest closed set $\bar{B}$ such that $B \subseteq \bar{B}$.*

Definition A.3 (Interior Points)**.** *Let (Ω, τ) be a topological space and $B \subset \Omega$. A point $b \in B$ is an* interior point *if there exists an open set containing b completely contained in B.*

Definition A.4 (Boundary of a Set)**.** *Let (Ω, τ) be a topological space. The* boundary *of a set $B \subset \Omega$ is its closer less its interior, that is*

$$\partial B = \bar{B} \setminus B^{\circ}$$

B Measure Theory

B.1 Basic Measure Theory

Definition B.1 (σ-Algebra)**.** *Let Ω be a non-empty set. A collection $\mathcal{F} \subset 2^{\Omega}$ is a σ-algebra if*

i. $\Omega \in \mathcal{F}$

ii. $A \in \mathcal{F} \implies A^c \in \mathcal{F}$

iii. $(A_i)_{i \geqq 1} \subset \mathcal{F} \implies \bigcup_{i \geqq 1} A_i \in \mathcal{F}$

Definition B.2 (Measurable Maps / Random Variables)**.** *A map $X : (\Omega, \mathcal{F}) \to (\Omega', \mathcal{F}')$ is called* measurable *if*

$$X^{-1}(\mathcal{F}') \subset \mathcal{F}$$

Definition B.3 ((Probability) Measure). *Let* $(\Omega, \mathcal{F})$ *be a measureable space, the set function* $\mu : \mathcal{F} \to \mathbb{R}_0^+ \cup \{\infty\}$ *is a measure if*

i. $\mu(\emptyset) = 0$

ii. $\mu(\biguplus A_i) = \sum_{i \geqq 1} \mu(A_i)$

The measure μ *is called a* probability measure *if* $\mu(\Omega) = 1$. *The space of all probability measures on* $(\Omega, \mathcal{F})$ *is denoted* $\mathscr{P}(\mathcal{F})$*; we will reserve* $\mathscr{P}_n$ *to be* $\mathscr{P}(\mathscr{B}(\mathbb{R}^n))$.

Definition B.4 (Inner Regularity, Borel Measures, & Radon Measures). *A* σ*-finite measure* μ *on a topological* [1] *measure space* $(\Omega, \mathcal{F}, \tau)$ *is*

i) Inner regular / tight *if*

$$\mu(B) = \sup(\mu(K) : K \subseteq B \textit{ is compact}) \quad \forall B \in \mathcal{F}$$

ii) a Borel measure *if*

$$\forall \omega \in \Omega \ \exists \ U_\omega \in \tau \textit{ s.t. } \mu(U_\omega) < \infty$$

iii) a Radon measure *if* μ *is Borel and inner regular*

Further, the set of Radon measures on $\mathcal{F}$ *will be denoted* $\mathscr{R}(\mathcal{F})$*; we will reserve* $\mathscr{R}_n$ *to be* $\mathscr{R}(\mathscr{B}(\mathbb{R}^n))$.

Corollary B.5 (Approximation by Compact Sets). *If* $\mu \in \mathscr{R}(\mathcal{F})$*, then*

$$\forall(\epsilon > 0, B \in \mathcal{F}) \ \exists K \in \mathcal{F} \ (\textit{compact}) \textit{ s.t. } \mu(B \setminus K) < \epsilon$$

Proof. Immediate from inner regularity. □

Definition B.6 (Image Measure / Distribution). *Let* $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\Omega', \mathcal{F}')$ *be a random variable, then we can endow* $(\Omega', \mathcal{F}')$ *with the* distribution*, often called the* image measure, $\mu = X_\# \mathbb{P} = \mathbb{P} \circ X^{-1}$

Theorem B.7 (Image Measure Integration). *Let* $X : (\Omega, \mathcal{F}, \mu) \to (\Omega', \mathcal{F}')$ *and* $\mu' = X_\# \mu$. *Assume* $f \in L^1(\Omega', \overline{\mathbb{R}}; \mu')$*, then* $f \circ X \in L^1(\Omega, \overline{\mathbb{R}}; \mu)$ *and*

$$\int_\Omega f \circ X \ d\mu = \int_{\Omega'} f \ d\mu \circ X^{-1} = \int_{\Omega'} f \ d\mu'$$

[1] the topology will usually be implicit

Proof. Use the standard mechanism[2]. □

Definition B.8 (Distribution Function). *If* $(\mathbb{R}^n, \mathscr{B}_n, \mu)$ *is a finite measure space, then the* distribution function *of* μ *is given by*

$$\begin{aligned} F_\mu : \mathbb{R}^n &\to [0, \mu(\mathbb{R}^n)] \\ x &\mapsto \mu((-\infty, x)) \end{aligned}$$

Theorem B.9 (Properties of the Distribution Function). *A distribution function,* F_μ*, has the following properties*

i) $x_1 \leqq x_2 \implies F_\mu(x_1) \leqq F_\mu(x_2)$ *(nondecreasing)*

ii) $\lim\limits_{x \downarrow \text{-}\infty} F_\mu(x) = 0$

iii) $\lim\limits_{x \uparrow \infty} F_\mu(x) = \mu(\mathbb{R}^n)$

iv) $\lim\limits_{\delta \downarrow 0} F_\mu(x - \delta) = F_\mu(x)$ *(left continuous)*

Note: Sometimes distribution functions are defined on right inclusive intervals (i.e. $x \mapsto \mu((-\infty, x]))$*, in which case property (iv) would become right continuous* $\lim\limits_{\delta \downarrow 0} F_\mu(x + \delta) = F_\mu(x)$

Proof.

i)
$$\begin{aligned} F_\mu(x_2) = \mu((\text{-}\infty, x_2)) &= \mu((\text{-}\infty, x_1) \uplus [x_1, x_2)) \\ &= \mu((\text{-}\infty, x_1)) + \mu([x_1, x_2)) \\ &\geqq \mu((\text{-}\infty, x_1)) = F_\mu(x_1) \end{aligned}$$

ii) Let $x_n \xrightarrow{n\uparrow\infty} \text{-}\infty$. By monotone convergence

$$0 = \mu(\emptyset) = \mu(\lim_{n\uparrow\infty} \bigcap(-\infty, x_n)) \overset{M.C.}{=} \lim_{n\uparrow\infty} \mu(\bigcap(-\infty, x_n))$$

iii) Let $x_n \xrightarrow{n\uparrow\infty} \infty$.

$$\mu(\mathbb{R}^n) = \mu(\lim_{n\uparrow\infty} \bigcup(-\infty, x_n)) \overset{M.C.}{=} \lim_{n\uparrow\infty} \mu(\bigcup(-\infty, x_n))$$

iv) $\mu((\text{-}\infty, x)) = \mu\left(\lim\limits_{\delta\downarrow 0}(\text{-}\infty, x - \delta)\right) \overset{M.C.}{=} \lim\limits_{\delta\downarrow 0} \mu\left((\text{-}\infty, x - \delta)\right)$

□

[2] Show it works for indicator functions, then step functions, then positive integrable functions, then general integrable functions.

Definition B.10 (Continuity Sets and (Dis)Continuity Points). *Let $(\Omega, \mathcal{F}, \mu)$ be a probability space. A set $C \in \mathcal{F}$ is a* continuity set *of μ if*

$$\mu(\partial C) = 0$$

Further, if $\mathcal{F} = \mathscr{B}(\mathbb{R}^n)$, then a point $x \in \mathbb{R}^n$ is a continuity point *of μ if*

$$\lim_{\delta \downarrow 0} \mu((-\infty, x - \delta]) = \mu((-\infty, x])$$

and a discontinuity point *if*

$$\lim_{\delta \downarrow 0} \mu((-\infty, x + \delta]) = \mu((-\infty, x]) \neq \lim_{\delta \downarrow 0} \mu((-\infty, x - \delta))$$

Definition B.11 (Uniform Boundedness). *A collection $\mathcal{B}$ of measures on $\mathcal{F}$ is uniformly bounded if $\exists M > 0$ such that*

$$\mu(\Omega) \leqq M \ \forall \mu \in \mathcal{B}$$

Definition B.12 (Tightness of Measures). *A collection $\mathcal{T} \subset \mathscr{P}(\mathcal{F})$ is tight if*

$$\forall \epsilon > 0 \ \exists K_\epsilon \in \mathcal{F}(compact) \ s.t. \ \mu(\Omega \setminus K_\epsilon) < \epsilon \ \forall \mu \in \mathcal{T}$$

B.2 Convergence in Measure & Its Consequences

Definition B.13 (Weak Convergence in Measure / Convergence in Distribution). *A sequence of measures $(\mu_n)_{n \geqq 1}$ converges* weakly *to μ, denoted $\mu_n \xrightarrow{w} \mu$, if*

$$\mu_n(C) \xrightarrow{n \uparrow \infty} \mu(C) \quad \forall C \in \mathcal{F} \ (bounded) \ s.t. \ \mu(\partial C) = 0$$

That is, μ_n converges to μ on the bounded continuity sets of μ.

Definition B.14 (Complete Convergence in Measure). *A sequence of measures $(\mu_n)_{n \geqq 1}$ converges* completely *to μ, denoted $\mu_n \xrightarrow{w} \mu$, if it converges weakly and $\mu_n(\Omega) \xrightarrow{n \uparrow \infty} \mu(\Omega)$.*

Corollary B.15. *If $\mu_n \xrightarrow{n \uparrow \infty} \mu$, then $(\mu_k)_{k \geqq 1}$ is tight.*

Proof. See Corollary 2.6.2 pg. 33 of [1]. □

Theorem B.16 (Helly's First Theorem). *If a sequence of measures $(\mu_n)_{n \geqq 1}$ is uniformly bounded, then $\exists (\mu_{n_k})_{k \geqq 1}$ which converges weakly to some measure μ. Further, if $(\mu_n)_{n \geqq 1}$ is a sequence of measures which is uniformly bounded and tight $\exists (\mu_{n_k})_{k \geqq 1}$ which converges completely to some measure μ.*

Proof. See Theorem 2.6.1 pg. 29 of [1]. □

Theorem B.17 (Helly's Second Theorem)**.** *Let $(\mu_k)_{k\geqq 1} \subset \mathscr{R}_n$. If (μ_k) is uniformly bounded and converges weakly to μ and if $f \in \mathscr{C}_b^0(\mathbb{R}^n, \mathbb{C})$, then*

$$\int_B f \; d\mu_k \xrightarrow{k\uparrow\infty} \int_B f \; d\mu$$

for any bounded continuity set B of μ.

Proof. See Theorem 2.6.4 pg. 34 in [1]. □

Theorem B.18 (Helly-Bray / Narrow Convergence)**.** *A sequence of measures $(\mu_n)_{n\geqq 1}$ converges weakly to μ iff $\forall\ \psi \in \mathscr{C}_b^0(\Omega, \mathbb{R})$*

$$\int \psi \; d\mu_n \xrightarrow{n\uparrow\infty} \int \psi \; d\mu$$

Proof. See Theorem 2.6.5 pg. 35 in [1]. □

Theorem B.19.

i) If $(\mu_k)_{k\geqq 1}$ is uniformly bounded and if $\mu_{k_j} \xrightarrow{j\uparrow\infty} \mu$ for any weakly convergent subsequence $(\mu_{k_j})_{j\geqq 1} \subset (\mu_k)_{k\geqq 1}$, then $\mu_k \xrightarrow{k\uparrow\infty} \mu$.

ii) If $(\mu_k)_{k\geqq 1}$ is uniformly bounded and tight and if $\mu_{k_j} \xrightarrow{j\uparrow\infty} \mu$ for any weakly convergent subsequence $(\mu_{k_j})_{j\geqq 1} \subset (\mu_k)_{k\geqq 1}$, then $\mu_k \xrightarrow{k\uparrow\infty} \mu$.

Proof. For (i) we will do a proof by contradiction: We have that for all weakly convergent subsequences $(\mu_{k_j})_{j\geqq 1}$ that $\mu_{k_j} \xrightarrow{j\uparrow\infty} \mu$ for some μ; now suppose that $\neg(\mu_k \xrightarrow{k\uparrow\infty} \mu)$, then $\exists$ a bounded μ-continuity set $B \in \mathcal{F}$ and some $\epsilon > 0$ s.t. $\forall K > 0\ \exists\ k_j > K$ s.t.

$$|\mu_{k_j}(B) - \mu(B)| > \epsilon \tag{$\star$}$$

Now, let's focus on these $(\mu_{k_j})_{j\geqq 1}$; since (μ_k) is uniformly bounded, the subsequence (μ_{k_j}) too is uniformly bounded; hence, by Theorem B.16, we can extract a weakly convergent subsequence $(\mu_{k_{j_i}})_{i\geqq 1} \subset (\mu_{k_j})_{j\geqq 1}$. By hypothesis, all weakly convergent subsequences converge to μ, but this contradicts ($\star$). (ii) follows by a similar argument. □

Theorem B.20 (Equivalence of Measures)**.** *Let* $(\Omega, \mathcal{F})$ *be an arbitrary measurable space and* $\mu_i \in \mathscr{P}(\mathcal{F})$ $(i = 1, 2)$*, then the following are equivalent:*

i) $\mu_1 = \mu_2$

ii) $\displaystyle\int \psi \; d\mu_1 = \int \psi \; d\mu_2 \;\; \forall \psi \in \mathscr{C}_b^0(\Omega)$

iii) $\displaystyle\int \psi \; d\mu_1 = \int \psi \; d\mu_2 \;\; \forall \psi \in \mathscr{C}_{b,p}^0(\Omega)$

iv) $\displaystyle\int \psi \; d\mu_1 = \int \psi \; d\mu_2 \;\; \forall \psi \in \mathscr{C}_{b,c}^0(\Omega)$

Proof. See Theorem 2.2.1, pg. 13 in [1]. □

Example B.21. *Let* $\mu_k \sim \mathcal{N}\left(\boldsymbol{x}, \frac{1}{k}\boldsymbol{I}_n\right)$ *be an* n*-dimensional Gaussian distribution where* $\boldsymbol{I}_n$ *is the* $n \times n$ *identity matrix. Then*

$$\int\limits_{\mathbb{R}^n} \psi \; d\mu_k = \left(\frac{k}{2\pi}\right)^{\frac{n}{2}} \int\limits_{\mathbb{R}^n} \psi(\boldsymbol{z}) \; e^{-\frac{k}{2}||\boldsymbol{z}-\boldsymbol{x}||_2^2} \; d\lambda_n(\boldsymbol{z})$$

and by taking $\tilde{\boldsymbol{z}} = \sqrt{k}(\boldsymbol{z} - \boldsymbol{x})$ *we obtain*

$$\frac{1}{(2\pi)^{\frac{n}{2}}} \int\limits_{\mathbb{R}^n} \psi\left(\boldsymbol{x} + \frac{1}{\sqrt{k}}\tilde{\boldsymbol{z}}\right) \; e^{-\frac{1}{2}||\tilde{\boldsymbol{z}}||_2^2} \; d\lambda_n(\tilde{\boldsymbol{z}}) \xrightarrow[D.C.]{k\uparrow\infty} \frac{1}{(2\pi)^{\frac{n}{2}}} \int\limits_{\mathbb{R}^n} \psi(\boldsymbol{x}) \; e^{-\frac{1}{2}||\tilde{\boldsymbol{z}}||_2^2} \; d\lambda_n(\tilde{\boldsymbol{z}}) = \phi(\boldsymbol{x})$$

where the dominating function is $||\psi||_\infty$*; hence*

$$\int\limits_{\mathbb{R}^n} \psi \; d\mu_k \xrightarrow{k\uparrow\infty} \psi(\boldsymbol{x}) = \int\limits_{\mathbb{R}^n} \psi \; d\delta_{\boldsymbol{x}}$$

Therefore, $\mu_k \xrightarrow{k\uparrow\infty} \delta_{\boldsymbol{x}}$.

B.3 Derivatives of Measures

Definition B.22 (Lower and Upper Measure Derivatives)**.** *Let* $\mu, \nu \in \mathscr{R}_n$*. The lower and upper derivatives of* ν *w.r.t.* ν *are*

$$\underline{D}_\mu \nu(x) = \begin{cases} \liminf\limits_{\delta\downarrow 0} \dfrac{\nu(\overline{\mathcal{B}_\delta(x)})}{\mu(\overline{\mathcal{B}_\delta(x)})} & : \mu(\overline{\mathcal{B}_\delta(x)}) > 0 \;\forall \delta > 0 \\ \infty & : \exists \delta > 0 \; s.t. \; \mu(\overline{\mathcal{B}_\delta(x)}) = 0 \end{cases}$$

and

$$\overline{D}_\mu\nu(x) = \begin{cases} \limsup\limits_{\delta\downarrow 0} \dfrac{\nu(\overline{\mathcal{B}_\delta(x)})}{\mu(\overline{\mathcal{B}_\delta(x)})} & : \mu(\overline{\mathcal{B}_\delta(x)}) > 0 \ \forall\delta > 0 \\ \infty & : \exists\delta > 0 \ s.t. \ \mu(\overline{\mathcal{B}_\delta(x)}) = 0 \end{cases}$$

Note: $0 \leqq \underline{D}_\mu\nu(x) \leqq \overline{D}_\mu\nu(x) \ \forall x$

Lemma B.23. *Let $\mu, \nu \in \mathscr{R}_n$, then $\forall B \in \mathscr{B}(\mathbb{R}^n)$ and $\alpha \in (0, \infty)$*

i) $B \subseteq \{x : \underline{D}_\mu\nu(x) \leqq \alpha\} \implies \nu(B) \leqq \alpha\mu(B)$

ii) $B \subseteq \{x : \overline{D}_\mu\nu(x) \geqq \alpha\} \implies \alpha\mu(B) \leqq \nu(B)$

Proof. Utilizes a corollary of the Besicovitch's Covering Theorem: See Lemma 6.5 pg. 7 in [2]. □

Theorem B.24. *Let $\mu, \nu \in \mathscr{R}_n$, then*

$$\mu(\{z \in \mathbb{R}^n : \overline{D}_\mu(z) = \infty\} \cup \{z \in \mathbb{R}^n : \underline{D}_\mu\nu(z) < \overline{D}_\mu\nu(z) < \infty\}) = 0$$

That is, $\underline{D}_\mu\nu(x) = \overline{D}_\mu\nu(x) < \infty \ \forall_\mu x \in \mathbb{R}^m$.

Proof. For a fixed k,

$$\{x \in \mathcal{B}_k(0) : \overline{D}_\mu(v) = \infty\} \subseteq \{x \in \mathcal{B}_k(0) : \overline{D}_\mu(v) \geqq \alpha\} \ \forall\alpha > 0$$

Hence, by Lemma B.23,

$$\begin{aligned} \mu(\{x \in \mathcal{B}_k(0) : \overline{D}_\mu(v) = \infty\}) &\leqq \nu(\{x \in \mathcal{B}_k(0) : \overline{D}_\mu(v) \geqq \alpha\}) \\ &\leqq \frac{1}{\alpha}\nu(\mathcal{B}_k(0)) \xrightarrow{\alpha\uparrow\infty} 0 \end{aligned}$$

Now, take $0 < \alpha < \beta$ and $A_{\alpha\beta} = \{x \in \mathcal{B}_k(0) : \underline{D}_\mu(x) < \alpha < \beta < \overline{D}_\mu(x) < \infty\}$. Then, again by Lemma B.23, we have

$$\nu(A_{\alpha\beta}) \leqq \alpha\mu(A_{\alpha\beta}) \ \text{ and } \ \beta\mu(A_{\alpha\beta}) \leqq \nu(A_{\alpha\beta})$$

hence, $\alpha\mu(A_{\alpha\beta}) = \beta\mu(A_{\alpha\beta})$. But since $\alpha \neq \beta$, we have $\mu(A_{\alpha\beta}) = 0$. Observing that

$$\{ \ x \in \mathcal{B}_k(0) : \underline{D}_\mu\nu(u) < \overline{D}_\mu\nu(u) < \infty\} = \bigcup_{\substack{\alpha,\beta\in\mathbb{Q}^+ \\ \alpha<\beta}} A_{\alpha,\beta}$$

we have

$$\mu(\{ \ x \in \mathcal{B}_k(0) : \underline{D}_\mu\nu(u) < \overline{D}_\mu\nu(u) < \infty\}) \leqq \sum_{\substack{\alpha,\beta\in\mathbb{Q}^+ \\ \alpha<\beta}} \mu(A_{\alpha,\beta}) = 0$$

□

With the above theorem, taking

$$N = \{x \in \mathbb{R}^n : [\![\ \overline{D}_\mu(x) = \infty \ ]\!] \ \vee \ [\![\ \underline{D}_\mu\nu(x) < \overline{D}_\mu\nu(x) < \infty \ ]\!]\}$$

we define

$$D_\mu\nu(x) = \begin{cases} \overline{D}_\mu\nu(u) & : x \notin N \\ \infty & : x \in N \end{cases}$$

Theorem B.25 ($D_\mu\nu$-Measurability)**.** *Let $\mu, \nu \in \mathscr{R}_n$, then $D_\mu\nu(u)$ is measurable.*

Proof. First we show that $x \mapsto \mu(\overline{\mathcal{B}_r(x)})$ is upper semicontinuous: Let $x_n \to x$, then

$$\forall(\epsilon > 0, r > 0) \ \exists \ N \in \mathbb{N} \ s.t. \ \overline{\mathcal{B}_r(x_n)} \subseteq \overline{\mathcal{B}_{r+\epsilon}(x)}$$

By monotonicity of μ

$$\mu(\overline{\mathcal{B}_r(x_n)}) \leqq \mu(\overline{\mathcal{B}_{r+\epsilon}(x)})$$

and so

$$\limsup_{n\uparrow\infty} \mu(\overline{\mathcal{B}_r(x_n)}) \leqq \mu(\overline{\mathcal{B}_{r+\epsilon}(x)})$$

By taking $\epsilon \downarrow 0$, and again by monotonicity,

$$\limsup_{n\uparrow\infty} \mu(\overline{\mathcal{B}_r(x_n)}) \leqq \mu(\overline{\mathcal{B}_r(x)})$$

In a similar fashion we can show that $x \mapsto \nu(\overline{\mathcal{B}_r(x)})$ is upper semicontinuous. Now take N as above, and for $x \in \mathbb{R}^n \setminus N$ we have $\mu(\overline{\mathcal{B}_r(x)}) > 0 \ \forall r > 0$, and so

$$D_\mu\nu(x) = \lim_{r\downarrow 0} \frac{\nu(\overline{\mathcal{B}_r(x)})}{\mu(\overline{\mathcal{B}_r(x)})} < \infty$$

Since N is a μ-measurable null set where $D_\mu\nu = \infty$, we have that $D_\mu\nu$ is μ-measurable. □

Theorem B.26 (Radon–Nikodym on $\mathbb{R}^n$)**.** *Let $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ be a measurable space and $\mu, \nu \in \mathscr{P}(\mathcal{F})$ such that $\mathcal{N}_\mu \subseteq \mathcal{N}_\nu$ $(\nu \ll \mu)$, then*

$$\nu(B) = \int\limits_B D_\mu\nu \ d\mu \ \ \forall B \in \mathscr{B}(\mathbb{R}^n)$$

Proof. See Theorem 2.26 pg. 10 in [2]. □

Theorem B.27 (Lebesgue(-Besicovitch) Differentiation)**.** *Let μ be a Radon measure on $\mathbb{R}^n$. For $f \in L^1_{loc}(\mathbb{R}^n;\mu)$*

$$\lim_{\delta \downarrow 0} \frac{1}{\mu(\overline{\mathcal{B}_\delta(x)})} \int\limits_{\overline{\mathcal{B}_\delta(x)}} f \ d\mu = f(x) \ \ \forall_\mu x \in \mathbb{R}^n$$

Proof. It is sufficient to prove the theorem for $f \geqq 0$ and $f \in L^1(\mathbb{R}^n;\mu)$. We can define a positive linear functional

$$\Lambda\psi = \int\limits_{\mathbb{R}^n} \psi f \ d\mu \ \forall \psi \in \mathscr{C}^0_c(\mathbb{R}^n)$$

By the Riesz representation theorem, there exists a Radon measure ν satisfying

$$\int\limits_{\mathbb{R}^n} \psi \ d\nu = \int\limits_{\mathbb{R}^n} \psi f \ d\mu \ \forall \psi \in \mathscr{C}^0_c(\mathbb{R}^n)$$

Now, by Theorem B.20, we have that

$$\nu(B) = \int\limits_B f \ d\mu \ \ \forall B \in \mathscr{B}(\mathbb{R}^n)$$

Now, by Theorem B.24,

$$\lim_{r \to 0} \frac{\nu(\overline{\mathcal{B}_r(x)})}{\mu(\overline{\mathcal{B}_r(x)})} = \lim_{r \to 0} \frac{1}{\mu(\overline{\mathcal{B}_r(x)})} \int\limits_{\overline{\mathcal{B}_r(x)}} f \ d\mu = D_\mu \nu(x) \ \ \forall_\mu x \in \mathbb{R}^n \qquad (\star)$$

and, by Theorem B.26,

$$\nu(B) = \int\limits_B f \ d\mu = \int\limits_B D_\mu \nu(x) \ d\mu \ \ \forall B \in \mathscr{B}(\mathbb{R}^n)$$

Hence

$$f = D_\mu \nu \ \ \forall_\mu x \in \mathbb{R}^n \qquad (\star\star)$$

With $(\star)$ and $(\star\star)$ we have the result. □

C Analysis

Definition C.1 (Equicontinuity)**.** *A collection of functions $(f_i)i \in \mathcal{I}$ on Ω is equicontinuous if*

$$\forall \epsilon > 0 \ \exists \ \delta \ s.t. \ |f_i(\omega) - f_i(\tilde{\omega})| < \epsilon \ \forall (i \in \mathcal{I} \ ; \ \omega, \tilde{\omega} \in \Omega \ s.t. \ |\omega - \tilde{\omega}| < \delta)$$

Definition C.2 (Algebra of Functions). *A collection of* $\mathbb{K}$*-valued functions on* Ω*, denoted* $\mathcal{A}$*, is an* algebra *if for all* $f, g \in \mathcal{A}$ *and* $c \in \mathbb{K}$

i) $f + g \in \mathcal{A}$

ii) $fg \in \mathcal{A}$

iii) $cf \in \mathcal{A}$

Definition C.3 (Function Set Separating Points). *Let* $\mathcal{A}$ *be a set of functions on* Ω*.* $\mathcal{A}$ *is said to* separate points of Ω *if*

$$\forall\ x, y \in \Omega\ (x \neq y)\ \exists\ f \in \mathcal{A}\ s.t.\ f(x) \neq f(y)$$

Theorem C.4 (Complex Form of the Stone-Weierstrass). *Let* Ω *be a compact metric space. If a subalgebra* $\mathcal{A} \subseteq \mathscr{C}^0(\Omega, \mathbb{C})$*, containing the constant functions on* Ω*, separates points of* Ω*, then* $\mathcal{A}$ *is dense in the Banach space* $\mathscr{C}^0(\Omega, \mathbb{C})$*.*

Proof. See 7.3.2, pg. 139 in [3]. □

Definition C.5 (Trigonometric Polynomials). *A* trigonometric polynomial *is a complex valued function of the form*

$$x \mapsto \sum_{n=-N}^{N} c_n e^{i\pi n x}$$

and in the multivariate case

$$\boldsymbol{x} \mapsto \sum_{t_1=-N}^{N} \cdots \sum_{t_n=-N}^{N} c_{(t_1,\ldots,t_n)} e^{i\langle \boldsymbol{t}_{(1,\ldots,n)}, \boldsymbol{x} \rangle}$$

where $\boldsymbol{t}_{(1,\ldots,n)} = [t_1, \ldots, t_n]^T$*.*

Corollary C.6. *Let* f *be a function defined on* $\mathbb{R}^n$ *which is periodic with period* T_j *with respect to the* j^{th} *variable* $(j = 1, \ldots, n)$*. Then* f *is the uniform limit in* N *of a sequence of trigonometric polynomials*

$$\begin{aligned} f_N(\boldsymbol{x}) &= \sum_{t_1=-N}^{N} \cdots \sum_{t_n=-N}^{N} c_{(t_1,\ldots,t_n)} \exp\left(2\pi i \sum_{j=1}^{n} \frac{t_j x_j}{T_j}\right) \\ &= \sum_{t_1=-N}^{N} \cdots \sum_{t_n=-N}^{N} c_{(t_1,\ldots,t_n)} e^{i\langle \tilde{\boldsymbol{t}}_{(1,\ldots,n)}, \boldsymbol{x} \rangle} \end{aligned}$$

where $\tilde{\boldsymbol{t}}_{(1,\ldots,n)} = 2\pi \cdot \boldsymbol{T}^{-1} \boldsymbol{t}_{(1,\ldots,n)}$ *and* $\boldsymbol{T} = \mathrm{diag}(T_1, \ldots T_n)$*.*

Proof. See 7.4.1 pg. 139 in [3]. □

Definition C.7 (Positive Linear Functional & Sublinear)**.** $\ell \in \Omega^*$ *is* positive linear *if*

$$\forall\ \omega \geqq 0 \implies \ell(\omega) \geqq 0$$

and is sublinear *if* $\forall \omega_1, \omega_2 \in \Omega$ *and* $\alpha \in \mathbb{R}_0^+$

i. $\ell(\omega_1 + \omega_2) \leqq \ell(\omega_1) + \ell(\omega_2)$

ii. $\ell(\alpha\omega_1) = \alpha\ell(\omega_1)$

Theorem C.8 (Riesz Representation)**.** *Let* (Θ, d) *be a metric space, then* $\forall$ ***positive*** $\ell \in \mathscr{C}_b^0(\Theta)^*$ $\exists$ *inner regular* $\mu \in \mathscr{P}(\mathscr{B}(\tau_d(\Theta)))$ *s.t.*

$$\ell(f) = \int\limits_{\Theta} f\ d\mu\ \ \forall\ f \in \mathscr{C}_b^0(\Theta)$$

Proof. See Theorem 7.3, pg. 22 in [4]. □

Definition C.9 (Cauchy Principal Value)**.** *Let* $f : \to \mathbb{R}$ *be a function such that* $\forall \epsilon > 0\ \ f \in L^1(\mathbb{R} \setminus (\text{-}\epsilon, \epsilon); \lambda)$*, then the* Cauchy principal value *of f is, if it exists, is*

$$PV(f) := \lim_{\epsilon \downarrow 0} \int\limits_{\mathbb{R}\setminus(\text{-}\epsilon,\epsilon)} f\ d\lambda$$

Example C.10. *Let* f *be a symmetric integrable function, then*

$$PV\left(\frac{f(x)}{x}\right) = 0$$

by antisymmetry of $\frac{1}{x}$.

References

[1] Cuppens, R. (1975). Decomposition of Multivariate Probabilities. Academic Press. DOI: 10.1016/C2013-0-07308-2

[2] K.-S. Chou, "Differentiation of Measures and Functions," in Real Analysis II. The Chinese University of Hong Kong, 2017.

[3] J. Dieudonne, Foundations of Modern Analysis. Academic Press, 1969

[4] O. van Gaans, Probability measures on metric spaces. Universiteit Leiden.

[5] K.-I. Sato, Levy Processes and Infinitely Divisible Distributions. Cambridge Press, 1999.

[6] W. Hardle, H. Lutkepohl, R. Chen, W. Hardle, and H. Lutkepohl, "A Review of Nonparametric Time Series Analysis," International Statistical Review / Revue Internationale de Statistique, vol. 65, no. 1, p. 49, apr 1997.

[7] A. F. Ansari, J. Scarlett, and H. Soh, "A Characteristic Function Approach to Deep Implicit Generative Modeling,"

[8] M.-D. C. de-la Vega, R. Pino-Mej´ıas, and E.-L. Silva-Ram´ırez, "Clustering multivariate empiric characteristic functions for multi-class svm classification," International Journal of Electrical and Computer Engineering, vol. 6, no. 4, pp. 380 – 384, 2012.

[9] L. A. Waller, B. W. Turnbull, and J. M. Hardin, "Obtaining Distribution Functions by Numerical Inversion of Characteristic Functions with Applications," The American Statistician, vol. 49, no. 4, p. 346, nov 1995.

[10] E. Lukacs, Characteristic Functions, 2nd ed. Charles Griffin & Company, 1970.