General Mathematics Vol. 10, No. 3–4 (2002), 23–32

Fourier transform of the probability measures

Romeo Vomişescu

Dedicated to Professor D. D. Stancu on his 75th birthday.

Abstract

In this paper we make the connection between Fourier transform of a probability measure and the characteristic function in the $\mathbb{R}^2$ space; also we establish some properties.

2000 Mathematical Subject Classification: 60B99

1. Let Ω and Ω' be any sets, let $\mathcal{K}$ and $\mathcal{K}'$ to be two σ-algebras on Ω and Ω' respectively and the measurable spaces $(\Omega, \mathcal{K})$, $(\Omega', \mathcal{K}')$.

A function $f : \Omega \to \Omega'$ is said to be $(\mathcal{K}, \mathcal{K}')$ - measurable if the borelian field $f^{-1}(\mathcal{K}') \subset \mathcal{K}$. Let $f : (\Omega, \mathcal{K}) \to (\Omega', \mathcal{K}')$ be a measurable function and let $\mu : \mathcal{K} \to [0, \infty)$ be a measure on $\mathcal{K}$. Then the function of set $\mu \circ f^{-1}$

defined on $\mathcal{K}'$ by the rule $(\mu \circ f^{-1})(A) = \mu(F^{-1}(A)), A \in \mathcal{K}'$ is a measure on $\mathcal{K}'$. The measure $\mu \circ f^{-1}$ is called, the image of the measure μ by f.

The triple $(\Omega, \mathcal{K}, \mu)$ where $(\Omega, \mathcal{K})$ is a measurable space, and μ is a measure on $\mathcal{K}$, is called the space with measure. If $\mu(\Omega) = 1$, then μ is called the probability measure.

Let $\varphi : (\Omega, \mathcal{K}) \to (\Omega', \mathcal{K}')$ be a measurable function. Then a measurable function $f : (\Omega, \mathcal{K}) \to X(\mathbb{R} \vee \mathbb{C})$ is $\mu \circ \varphi^{-1}$ - integrable, if and only if $f \circ \varphi$ is μ-integrable. In this case the following relation holds

$$\int (f \circ \varphi) d\mu = \int f \cdot d\mu \circ \varphi^{-1} \tag{1}$$

and is called the transport formula.

2. We note with $(\Omega, \mathcal{K}, P)$ a probability field and let $(X, \mathcal{X})$ be a measurable space where $\mathcal{X}$ is a borelian field on X. A measurable function $f : (\Omega, \mathcal{K}, P) \to (X, \mathcal{X})$ is called a random variable. If the function f is a random variable, then the image of P by f we will note with $P \circ f^{-1}$ and will be called, distribution of f. In this case the distribution of f is the probability on $\mathcal{X}$ defined by $P \circ f^{-1}(A) = P(f^{-1}(A)),\ A \in \mathcal{X}$. This events $f^{-1}(A)$ we also denote by $\{f \in A\}$. If F is a probability on $\mathbb{R}^k$, then we will say that F has the density ρ if $F \prec m_k$ (m_k is the Lebesgue measure on $\mathbb{R}^k$) and ρ is a version of the Radon-Nikodým derivative dF/dm.

(For λ, μ-measures, $\lambda \prec \mu$ denote that λ is absolutely continuous with respect to μ, i.e. $\mu(A) = 0$ implies that $\lambda(A) = 0$)

If the function $f : (\Omega, \mathcal{K}, P) \to \mathbb{R}^k$ is a random variable, we will say that

f has the density ρ if the distribution $P \circ f^{-1}$ has the density ρ. Hence a function $\rho : \mathbb{R}^k \to \overline{\mathbb{R}}$ is the density of the random variable f if:

i) ρ is measurable and $\rho \geq 0$

ii) $P(f \in A) = \int\limits_A \rho(x) dm_k(x),\ A \in \mathcal{B}_{\mathbb{R}^k}$, where $\mathcal{B}$ is a borelian field.

For a random variable $f : (\Omega, \mathcal{K}, P) \to \mathbb{R}^k$ and for a measurable function $\varphi : \mathbb{R}^k \to \mathbb{C}$, the transport formula can be written as

$$\int\limits_{\Omega} \varphi \circ f dP = \int\limits_{\mathbb{R}^k} \varphi(x) dP \circ f^{-1}(x) \tag{2}$$

In particular, if f has the density ρ, then

$$\int\limits_{\Omega} \varphi \circ f dP = \int\limits_{\mathbb{R}^k} \varphi(x) \rho(x) dx \tag{3}$$

Let $\zeta = (\xi, \eta)$ be a random vector whose components are the random variables ξ and η. If so, the function F defined by the relation

$$F(z) = F(x, y) = P(\xi \leq x; \eta \leq y),\ \text{for every } (z) = (x, y) \in \mathbb{R}^2 \tag{4}$$

is called the distribution function of the random vector ζ, where $P(\xi \leq x, \eta \leq y)$ is the probability that an aleatory point $\xi \in (-\infty, x]$, $\eta \in (-\infty, y]$.

The function F has analogous properties with the distribution function from the unidimensional case:

$$0 \leq F(x, y) \leq 1,\ \lim_{x,y \to -\infty} F(x, y) = 0,\ \lim_{x,y \to \infty} F(x, y) = 1.$$

The monotony condition of the function F will be characterized by the following inequalities :

$$F(x + h, y) - F(x, y) \geq 0,\ F(x, y + k) - F(x, y) \geq 0$$

$$F(x+h, y+k) - F(x+h, y) \geq F(x, y+K) - F(x, y)$$

where h and k represent two positive increases. Let $\mathcal{V}$ be boolean algebra of all B-intervals of the form

$$\Delta = [a, b) \times [c, d),\ a, b, c, d \in \mathbb{R}$$

and let $\mu : \mathcal{V} \to [0, \infty]$ be a measure on $\mathcal{V}$ so that $\mu(\Delta) < \infty$.

We know (see [3]) that there exists a monotone nondecreasing and left-continuous function F on $\mathbb{R}^2$, so that for every $a, b, c, d \in \mathbb{R}$ we have

$$(5)\ \ \mu([a, b) \times [c, d)) = F(b, d) - F(a, d) - F(b, c) + F(a, c) = P(\zeta \in \Delta)$$

the reciprocal being also valid.

If F_1 and F_2 are monotone non-decreasing and left-continuous functions on $\mathbb{R}^2$, so that

$$\mu([a, b) \times [c, d)) = F_1(b, d) - F_1(a, d) - F_1(b, c) + F_1(a, c) =$$

$$= F_2(b, d) - F_2(a, d) - F_2(b, c) + F_2(a, c),\ \forall a, b, c, d \in \mathbb{R},$$

then there exists a hyperbolic constant.

$\Psi(x, y) = \varphi(x) + \psi(y)$ so that $F_2(x, y) = F_1(x, y) + \Psi$. If μ is a measure on $\mathcal{V}$ with $\mu(\mathbb{R}^2) = \alpha < \infty$, then a monotone non-decreasing and left-continuous function F on $\mathbb{R}^2$, can be found, having the properties $\lim\limits_{x,y \to -\infty} F(x, y) = 0$, $\lim\limits_{x,y \to \infty} F(x, y) = \alpha$ and (4) holds.

The function F so defined, is unique. If $\alpha = 1$, then the function F is called distribution function (probability).

3. Let $\zeta = (\xi, \eta)$ be a random vector. Then, one defines for each measure μ on $\mathbb{R}^2$, Fourier transform or otherwise characteristical function of the probability measure

$$\widehat{\mu}(t) = \int e^{i<t,z>} d\mu(z),\ t \in \mathbb{R}^2 \tag{6}$$

where $t = (u, v),\ z = (x, y) \in \mathbb{R}^2$. This function is called the distribution of μ. We have

$$\widehat{\mu}(t) = \int\limits_{\mathbb{R}^2} e^{i<t,z>} dF(z) \tag{7}$$

where $F(z)$ has the expression (4).

If the random variable $f : (\Omega, \mathcal{K}, P) \to \mathbb{R}^2$ and $\mu = P \circ f^{-1}$ is distribution of f, then the characteristical function $\widehat{\mu}$ is

$$\widehat{\mu}(t) = \int e^{i<t,z>} d\mu(z) = \int\limits_{\Omega} e^{i<t,f>} dP = M \cdot e^{i<t,f>} \tag{8}$$

where M is the mean value. In this case we say that $\widehat{\mu}$ is the characteristical function of the random variable f. If ρ is the density in the point (x, y) of a mass equal with the unit distributed in plane x, y, then

$$\widehat{\mu}(u, v) = \int\limits_{\mathbb{R}^2} e^{i(ux+vy)} d\mu(x, y) = \int\limits_{\mathbb{R}^2} e^{i(ux+vy)} \rho(x, y) dxdy \tag{9}$$

Theorem 1. *For each measure μ on $\mathbb{R}^2$ we have:*

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

ii) $\widehat{\mu}(-t) = \overline{\widehat{\mu}(t)}$

iii) for every $a_1, a_2, ..., a_n \in \mathbb{C}$ and $t_1, t_2, ..., t_n \in \mathbb{R}^2$ we have

$$\sum_{j,k=1}^{n} a_j \cdot \overline{a_k} \cdot \widehat{\mu}(t_j - t_k) \geq 0$$

iv) $\widehat{\mu}$ is a uniformly continuous function.

Proof. i) This follows from (8)

ii) $\widehat{\mu}(-t) = \int\limits_{\mathbb{R}^2} \overline{e^{i<t,z>}} d\mu(z) = \overline{\int\limits_{\mathbb{R}^2} e^{i<t,z>} d\mu(z)} = \overline{\widehat{\mu}(t)}$

iii) $\sum\limits_{j,k} a_j \overline{a_k} \widehat{\mu}(t_j - t_k) = \int\limits_{\mathbb{R}^2} \sum\limits_{j,k} a_j \cdot \overline{a_k} e^{i<t_j - t_k, z>} d\mu(z) =$

$$= \int\limits_{\mathbb{R}^2} |\sum_j a_j \cdot e^{i<t,z>}|^2 \cdot d\mu(z) \geq 0$$

iv) for every $\nu = (h,k) \in \mathbb{R}^2_+, |\widehat{\mu}(u+h, v+k) - \widehat{\mu}(u,v)| \leq \int\limits_{\mathbb{R}^2} |e^{i<\nu,z>} - 1| d\mu(z)$.

We have also

$$\sup_{(h,k)\in\mathbb{R}^2_+} |\widehat{\mu}(u+h, v+k) - \widehat{\mu}(u,v)| \leq \int\limits_{\mathbb{R}^2} |e^{i<\nu,z>} - 1| d\mu(z)$$

where $|\nu| = (h^2 + k^2)^{\frac{1}{2}}$.

The integrand is bounded and it tends to zero for $\nu \to 0$. Then according to Lebesgue's dominated convergence theorem, we have

$$\lim_{|\nu|\to 0} \sup_{(u,v)\in\mathbb{R}^2} |\widehat{\mu}(u+h, v+k) - \widehat{\mu}(u,v)| = 0$$

and so $\widehat{\mu}$ is uniformly continuous.

Theorem 2. *Let μ be a measure on $\mathbb{R}^k$ and*

$$\int |x_j| d\mu(x) < \infty,\ j = \overline{1,k}, x = (x_1, ..., x_k), t = (t_1, ..., t_k).$$

Then $\widehat{\mu}$ is partial derivative with respect to t_j and we have

$$\frac{\partial\widehat{\mu}}{\partial t_j}(t) = i \int x_j \cdot e^{i<t,x>} d\mu(x) \tag{10}$$

The partial derivatives $\dfrac{\partial\widehat{\mu}}{\partial t_j}(t)$ are uniformly continuous.

Proof. Let e_j be the vectors of an orthonormal base. Then we have

$$\frac{\widehat{\mu}(t+he_j)-\widehat{\mu}(t)}{h} = \int e^{i<t,x>} \cdot \frac{e^{ihx_j}-1}{h} d\mu(x)$$

and $\left| e^{i<t,x>} \cdot \frac{e^{ihx_j}-1}{h} \right| \leq |x_j|$.

For $h \to 0$ and using the Lebesgue's dominated convergence theorem, we obtain (10). The second part of the theorem follows from Theorem (1).

Observation 1. *It is easy to show that, if μ is a measure on $\mathbb{R}^k$ and*

$$\int |x^n| d\mu(x) < \infty, x^n = x_1^{n_1} ... x_k^{n_k}, n_i \geq 0, |n| = n_1 + ... + n_k,$$

then the $\frac{\partial^{|n|}}{\partial t_1^{n_1} ... \partial t_k^{n_k}} (\widehat{\mu}(t))$ exists and

$$\frac{\partial^{|n|}}{\partial t_1^{n_1} ... \partial t_k^{n_k}} \widehat{\mu}(t) = i^{|n|} \cdot \int x_1^{n_1} ... x_k^{n_k} \cdot e^{i<t,x>} \cdot d\mu(x) \tag{11}$$

where i-imaginary unit.

Theorem 3. *Let μ be a probability measure on $\mathbb{R}^2$, so that $\int |x| d\mu(x,y) < \infty$, $\int |y| d\mu(x,y) < \infty$. Then*

a) If $\int x d\mu(x,y) = 0$, $\int y d\mu(x,y) = 0$ we have

$$\lim_{n\to\infty} \left[\widehat{\mu}\left(\frac{u}{n}, \frac{v}{n}\right)\right]^n = 1 \tag{12}$$

b) If in addition,

$$\int x^2 d\mu(x,y) = 1, \int y^2 d\mu(x,y) = 1 \text{ and } \int xy \cdot d\mu(x,y) = 0$$

we have

$$\lim_{n\to\infty} \left[\widehat{\mu}\left(\frac{u}{n}, \frac{v}{n}\right)\right]^n = e^{-\frac{1}{2}(u^2+v^2)}$$

Proof. a) From hypothesis and from the Theorem (2) follows that $\widehat{\mu}(u,v)$ is differentiable and it's partial derivatives are continuous. Since

$$\frac{\partial\widehat{\mu}(u,v)}{\partial u}=i\int x\cdot e^{i(ux+vy)}d\mu(x,y),\ \frac{\partial\widehat{\mu}(u,v)}{\partial v}=i\int ye^{i(ux+vy)}d\mu(x,y)$$

follows that

$$\frac{\partial\widehat{\mu}}{\partial u}(0,0)=i\int xd\mu(x,y)=0,\ \frac{\partial\widehat{\mu}}{\partial v}(0,0)=i\int yd\mu(x,y)=0$$

Applying the formula of Mac-Laurin, we obtain

$$\widehat{\mu}(t)=\widehat{\mu}(u,v)=\widehat{\mu}(0,0)+u\frac{\partial\widehat{\mu}}{\partial u}(\theta u,\theta v)+v\frac{\partial\widehat{\mu}}{\partial v}(\theta u,\theta v)=$$

$$=1+u\cdot\alpha(t)+v\cdot\beta(t)$$

for $|u|\leq 1,\ |v|\leq 1$ where $\alpha(t),\beta(t)$ are continuous functions in (0,0), and $\alpha(0)=0$, $\beta(0)=0$, $0<\theta<1$. Then,

$$\widehat{\mu}\left(\frac{t}{n}\right)=\widehat{\mu}\left(\frac{u}{n},\frac{v}{n}\right)=1+\frac{u}{n}\alpha\left(\frac{t}{n}\right)+\frac{v}{n}\beta\left(\frac{t}{n}\right)=1+\gamma_n(t)$$

where $\gamma_n(t)=\frac{u}{n}\alpha\left(\frac{t}{n}\right)+\frac{v}{n}\beta\left(\frac{t}{n}\right)$ with $\lim\limits_{n\to\infty}n\gamma_n(t)=0$, for every $(u,v)\in\mathbb{R}^2$. Then for every (u,v), we have

$$\lim_{n\to\infty}\left[\widehat{\mu}\left(\frac{t}{n}\right)\right]^n=\lim_{n\to\infty}[1+\gamma_n(t)]^n=\lim_{n\to\infty}\left[(1+\gamma_n(t))^{\dfrac{1}{\gamma_n(t)}}\right]^{n\cdot\gamma_n(t)}=1$$

b) From the Theorem (2) follows that $\widehat{\mu}$ is twice differentiable with the partial derivatives of the second continuous order and

$$\frac{\partial\widehat{\mu}}{\partial u}(0,0)=0,\ \frac{\partial\widehat{\mu}}{\partial v}(0,0)=0,\ \frac{\partial^2\widehat{\mu}}{\partial u^2}(0,0)=-1,\ \frac{\partial^2\widehat{\mu}}{\partial v^2}(0,0)=-1,\ \frac{\partial^2\widehat{\mu}}{\partial u\partial v}(0,0)=0.$$

Applying again the formula of Mac-Laurin for $|u| \leq 1$, $|v| \leq 1$, $0 < \theta < 1$ we have

$$\widehat{\mu}(t) = \widehat{\mu}(u,v) = \widehat{\mu}(0,0) + u\frac{\partial\widehat{\mu}}{\partial u}(0,0) + v\frac{\partial\widehat{\mu}}{\partial v}(0,0)+$$

$$+\frac{1}{2}\left[u^2\frac{\partial^2\widehat{\mu}}{\partial u^2}(\theta u,\theta v) + 2uv\frac{\partial^2\widehat{\mu}}{\partial u\partial v}(\theta u,\theta v) + v^2\cdot\frac{\partial^2\widehat{\mu}}{\partial v^2}(\theta u,\theta v)\right] \text{ or}$$

$$\widehat{\mu}(t) = \widehat{\mu}(u,v) = 1 + \frac{1}{2}u^2\theta_1(t) + uv\theta_2(t) + \frac{1}{2}v^2\theta_3(t) \text{ where}$$

$$\theta_1(0,0) = -1, \theta_3(0,0) = -1, \theta_2(0,0) = 0. \text{ Then,}$$

$$\widehat{\mu}\left(\frac{t}{\sqrt{n}}\right) = \widehat{\mu}\left(\frac{u}{\sqrt{n}},\frac{v}{\sqrt{n}}\right) = 1 + \frac{1}{2}\cdot\frac{u^2}{n}\theta_1\left(\frac{t}{\sqrt{n}}\right) + \frac{uv}{n}\theta_2\left(\frac{t}{\sqrt{n}}\right) +$$

$$+\frac{1}{2}\cdot\frac{v^2}{n}\theta_3\left(\frac{t}{\sqrt{n}}\right) = 1 + \tau_n(t), \text{ where}$$

$$\tau_n(t) = \frac{1}{2}\frac{u^2}{n}\theta_1\left(\frac{t}{\sqrt{n}}\right) + \frac{uv}{n}\theta_2\left(\frac{t}{\sqrt{n}}\right) + \frac{1}{2}\frac{v^2}{n}\theta_3\left(\frac{t}{\sqrt{n}}\right)$$

and

$$\lim_{n\to\infty} n\tau_n(t) = -\frac{1}{2}(u^2+v^2), \ \forall(u,v)$$

.

Then

$$\lim_{n\to\infty}\left[\widehat{\mu}\left(\frac{t}{\sqrt{n}}\right)\right]^n = e^{-\frac{1}{2}(u^2+v^2)}.$$

References

[1] E. Kamke, *Integral Lebega-Stieltiesa*, Fizmatgiz - Moscova (1959) (in Russian).

[2] A.N. Şiriav, *Veroiatnosti*, Moscova, 1989, (in Russian).

[3] R. Vomişescu, *B-Stieltjes integrals with probability interpretations*, Acta Technica Napocensis, 1994.

Department of Mathematics
University "Lucian Blaga" of Sibiu,
Str. Dr. I. Ratiu Nr.5–7
550012 - Sibiu, Romania
E-mail: *depmath@ulbsibiu.ro*