ARKIV FÖR MATEMATIK Band 7 nr 24

1.67 05 3 Communicated 8 March 1967 by HARALD CRAMÉR and OTTO FROSTMAN

Determining the absolutely continuous component of a probability distribution from its Fourier-Stieltjes transform

By H. ROBERT VAN DER VAART

1. Introduction

Most discussions of the inversion formula

$$\mu(]a,b[) + \tfrac{1}{2}\mu(\{a\}) + \tfrac{1}{2}\mu(\{b\}) = (2\pi)^{-1} \lim_{T\to\infty} \int_{-T}^{T} \varphi(t) \frac{e^{-itb} - e^{-ita}}{-it}\, dt, \qquad (1.1)$$

which recovers a probability measure μ from its characteristic function (Fourier–Stieltjes transform)

$$\varphi(t) = \int_R e^{ity}\mu(dy), \qquad (1.2)$$

contain a remark to the effect that if $\varphi \in L_1(R)$ then μ is absolutely continuous and has a (continuous) density function f given by

$$f(x) = (2\pi)^{-1}\int_R \varphi(t)e^{-itx}\,dt \qquad (1.3)$$

(e.g., see Lukacs [13], p. 40, Th. 3.2.2).

For the case that $\varphi \notin L_1(R)$ there seems, however, to be some confusion. Lévy [12], pp. 167–168 recommends the use of the *Cauchy principal value* in equation (1.3) in case μ is *absolutely continuous*. The same recommendation is made by Kendall and Stuart [10], p. 94 for the case where the distribution function "is continuous everywhere and *has* a *density function*", and by Richter [16], p. 329 for the case where μ *has* a *differentiable density function*. Dugué [6], p. 24 quotes essentially a theorem of Jordan (cf. Goldberg [8], Th. 5C) to point out that if the Cauchy principal value of the integral in eq. (1.3) exists and if also $f(x+0)$ and $f(x-0)$ exist (note that these three conditions are satisfied if the *density* f of μ is of *bounded variation* in a neighborhood of x) then equation (1.3) is valid in the sense that

$$\frac{f(x+0)+f(x-0}{2} = \frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T}\varphi(t)e^{-itx}\,dt. \qquad (1.4)$$

At the same time, Robinson [17], p. 30 states that equation (1.3) holds as printed if the random variable in question *has* a *density function*. On the other hand Lukacs

[14], p. 177, citing Goldberg [8], p. 14, Th. 6C, recommends the use of $(C, 1)$-summation in equation (1.3) when φ is *not absolutely integrable* and μ is *absolutely continuous*.

The present paper wants to dissolve some of this confusion, and also to furnish criteria by which to select a suitable interpretation, if any, of equation (1.3). In many applications of the method of characteristic functions we first obtain what we *know* to be the *characteristic function* of some k-variate distribution, then we ask questions about that distribution: whether it is absolutely continuous, and if so what its density is, and so on. For such problems, knowledge of φ has to suffice. Accordingly, our criteria will be *in terms of* φ only, as opposed to the above quotations which impose conditions on the probability measure μ.

The fundamental result is: the integral of equation (1.3) *exists* (for Lebesgue-almost all x) *in the* $(C, 1)$*-sense whenever* φ is the characteristic function of a probability measure μ (*regardless* whether μ is absolutely continuous or purely discontinuous or singular continuous), and the resulting function of x gives the *density* of the *absolutely continuous component* of μ. Some corollaries and examples are also discussed. This paper corrects and strengthens the results listed in: van der Vaart [20].

2. Preliminaries

For easy reference we are listing some definitions, notations, and results to be used in the sequel.

2.1. Lebesgue decomposition. Let $\mathcal{B}_k = \mathcal{B}(R^k)$ denote the σ-algebra of Borel sets in R^k. Let λ denote Lebesgue measure on $\mathcal{B}_k$ and ψ a σ-finite signed measure. Then

$$\psi = \psi_a + \psi_s, \tag{2.1.1}$$

where uniquely determined ψ_a is absolutely continuous with respect to λ and ψ_s is singular with respect to λ (e.g., see Hewitt and Stromberg [9], sec. 19.42). In case $k=1$ let μ be a probability measure defined on $\mathcal{B}_1$. Then one can even say that

$$\mu = \mu_a + \mu_{sc} + \mu_{sd}, \tag{2.1.2}$$

where the μ_i are uniquely determined, μ_{sd} is purely discontinuous, μ_{sc} is singular continuous (re λ), and μ_a is absolutely continuous (re λ) (e.g., see Hewitt and Stromberg [9], sec. 19.61). In the sequel, the terms 'absolutely continuous' and 'singular' are always understood as relative to λ.

2.2. Derivatives of set functions. Let ψ be a finite signed measure on $\mathcal{B}_k$. The derivative of ψ with respect to Lebesgue measure λ at the point p is defined as

$$\dot{\psi}(p) = \frac{d\psi}{d\lambda}(p) \overset{\text{def}}{=} \lim_{\lambda(C)\to 0} \frac{\psi(C)}{\lambda(C)} \quad \text{when this limit exists,} \tag{2.2.1}$$

where C denotes any closed convex set containing p (and satisfying certain additional conditions: for details see Doob [5], p. 291). Hereafter C_ϱ will stand for a closed ball, centered at a point appearing from context, with radius ϱ. It is known that the derivative (2.2.1) exists λ-almost everywhere. It is also known that $\dot{\psi}_s$ (see eq. (2.1.1)) and $\dot{\mu}_{sc}$ and $\dot{\mu}_{sd}$ (see eq. (2.1.2)) are zero where they exist. Hence λ-almost everywhere $\dot{\psi}$ exists and equals $\dot{\psi}_a$, $\dot{\mu}$ exists and equals $\dot{\mu}_a$ (see Dunford and Schwartz [7], sec. III. 12.6).

2.3. Notation for integrals. The (Lebesgue-Stieltjes) integral of f relative to the (possibly signed) measure ψ will be denoted by $\int f(x)\psi(dx)$.

If $\psi=\lambda$, the Lebesgue measure, we will sometimes write dx for $\lambda(dx)$.

If g is a monotone non-decreasing point function then $\int \dots g(dx)$ will denote integration relative to the (positive) measure determined by the point function g (via the interval function $g(x_2)-g(x_1)$).

2.4. A special case of integration by parts. Let $r(z)$ depend on $|z|$ only, $r(z)=r^*(|z|)$, and let r^* be non-increasing. Let $z=0$ be the center of the closed balls C_ϱ. Finally let θ be a finite (positive) measure. Put $\theta(C_q)=\nu(q)$, a non-decreasing point function $(q\in R^1)$. Then

$$\int_{C_\varrho} r(z)\theta(dz)=\int_0^\varrho r^*(q)\nu(dq)=r^*(\varrho)\nu(\varrho)-\int_0^\varrho \nu(q)r^*(dq)=r^*(\varrho)\theta(C_\varrho)-\int_0^\varrho \theta(C_q)r^*(dq), \tag{2.4.1}$$

provided $\theta(C_0)=\theta(\{0\})=0$. Note that the last term in the last member of this equation is positive.

2.5. Total variation. The total variation $|\psi|$ of a signed measure ψ is a set function defined by

$$|\psi|(A)=\psi^+(A)+\psi^-(A) \quad \text{for all} \quad A\in\mathcal{B}_k, \tag{2.5.1}$$

where ψ^+ and ψ^- are the (positive) measures ensuing from the Jordan decomposition (e.g., see Royden [18], sec. 11.4, or Dunford and Schwartz [7], sec. III. 4.11). From the proof given in Dunford and Schwartz [7], sec. III. 12.6 (or a bit more explicitly in Rudin [19], sec. 8.6) it follows that

$$\frac{d|\psi|}{d\lambda}(p)=\left|\frac{d\psi}{d\lambda}(p)\right|=|\dot\psi(p)|, \tag{2.5.2}$$

wherever $\dot\psi(p)$ exists, i.e., λ-almost everywhere.

Now let $\psi=\mu-\alpha\lambda$ (α a real number, μ a probability measure on R^k, λ Lebesgue measure on R^k), and define

$$\theta=|\psi|=|\mu-\alpha\lambda|. \tag{2.5.3}$$

Then by (2.5.2)

$$\dot\theta(p)=|\dot\mu(p)-\alpha| \quad \text{for } \lambda\text{-almost all } p.$$

By the same argument that is commonly used in the discussion of Lebesgue points (e.g., see Doob [5], p. 291; or Dunford and Schwartz [7], III. 12.8; or Alexits [1], § 4.4.1) it follows that if

$$\psi_p(A)\overset{\text{def}}{=}\mu(A)-\dot\mu(p)\lambda(A) \quad \text{for all} \quad A\in\mathcal{B}_k \tag{2.5.4}$$

and

$$\theta_p\overset{\text{def}}{=}|\psi_p|$$

then

$$\dot\theta_p(p)=\frac{d\theta_p}{d\lambda}(p)=\frac{d|\psi_p|}{d\lambda}(p)=0 \quad \text{for } \lambda\text{-almost all } p.$$

This means that to λ-almost all p and to any $\varepsilon > 0$ a real number $\delta_\varepsilon(p)$ can be assigned such that

$$\theta_p(C_\varrho) = |\psi_p|(C_\varrho) < \varepsilon \cdot \lambda(C_\varrho) \quad \text{for} \quad \varrho < \delta_\varepsilon(p), \tag{2.5.5}$$

where the C_ϱ are closed balls with radius ϱ and center p.

We will also use the inequality

$$\left|\int_A f(x)\psi(dx)\right| \leqslant \int_A |f(x)| \cdot |\psi|(dx) \leqslant M|\psi|(A), \tag{2.5.6}$$

where $|f(x)| \leqslant M$ for $|\psi|$-almost all $x \in A$ (e.g., see Dunford and Schwartz [7], III. 2.20).

2.6. A lemma on integrals with kernels. The following lemma combines various methods used in this area; see a.o. Dunford and Schwartz [7], III. 12.10 and III. 12.11, and Doob [5], who also gives a short history of the problem (*l.c.*, p. 505).

Lemma. *Let λ be Lebesgue measure and μ a probability measure, both defined on $\mathcal{B}_k$. For all $x \in R^k$ let $K_T(x, .)$ be a summable function defined on R^k such that*

(i) $\lim_{T\to\infty} \int_{C_1} K_T(x,y)\lambda(dy) = 1$ (*C_1 being a closed ball with radius* 1 *and center* $y = x$),

and

(ii) $|K_T(x,y)| \leqslant r_T(|x-y|) = \begin{cases} \alpha T^{\varkappa} & \text{for} \quad |x-y| \leqslant T^{-1}, \\ \alpha T^{-l_1}|x-y|^{-l_2} & \text{for} \quad |x-y| > T^{-1}, \end{cases}$

where $\varkappa > 0$, $l_1 > 0$, $l_2 > 0$, $l_2 - l_1 = \varkappa$, and either $l_2 - k \leqslant 0$ or $0 < l_2 - k \leqslant l_1$. Then at λ-almost all points p, namely at all points p which satisfy condition (2.5.5):

$$\lim_{T\to\infty} \int_{R^k} K_T(p,y)\mu(dy) = \dot{\mu}(p) \tag{2.6.1}$$

Remark. Remember that $\dot{\mu}(p) = \dot{\mu}_a(p)$, see section 2.2. The rather special choice of majorant r_T simplifies the technical execution of the proof, and is sufficient for our purpose.

Proof. Because of condition (i) our claim is equivalent to

$$\lim_{T\to\infty}\left\{\int_{R^k} K_T(p,y)\mu(dy) - \int_{C_1} K_T(p,y)\dot{\mu}(p)\lambda(dy)\right\} = 0.$$

Since the subsequent argument will center around the fixed value p of y we will use the coordinate $z = y - p$ rather than y; accordingly $K_T(p, y)$ becomes $K_T(p, p+z)$ and $\mu(dy)$ becomes $\mu(p+dz)$. The *center of* the *balls* C_ϱ is $z = 0$. Now the expression whose limit we will investigate is the sum of three integrals:

$$I_1 = \int_{R^k \setminus C_1} K_T(p,p+z)\mu(p+dz);$$

$$I_2 = \int_{C_1 \setminus C_\delta} K_T(p, p+z)\, \psi_p(p+dz)$$

(for ψ_p see eq. (2.5.4)); and

$$I_3 = \int_{C_\delta} K_T(p, p+z)\, \psi_p(p+dz),$$

where δ is some number <1 and $<\delta_\varepsilon(p)$ (see eq. 2.5.5)). We will prove that for any $\eta > 0$ we can choose $T_\eta > \delta^{-1}$ large enough that $|I_1 + I_2 + I_3| < \eta$ for $T > T_\eta$.

Applying (2.5.6) and condition (ii) of the lemma we find

$$|I_1| \leqslant \alpha T^{-l_1} < \varepsilon\alpha \quad \text{for} \quad T^{l_1} > \varepsilon^{-1},$$

and also $\qquad |I_2| \leqslant \alpha T^{-l_1}\delta^{-l_2} |\psi_p|(C_1) < \varepsilon\alpha |\psi_p|(C_1) \quad \text{for} \quad T^{l_1} > \varepsilon^{-1}\delta^{-l_2}.$

Finally, applying successively (2.5.6), condition (ii), (2.4.1) (note that $\theta_p(C_0) = 0$ if p satisfies (2.5.5)), (2.5.5), the fact that $\lambda(C_q) = \beta q^k$, and the restriction $T > \delta^{-1}$, we find:

$$\begin{aligned}
|I_3| &\leqslant \int_{C_\delta} |K_T(p, p+z)| \cdot |\psi_p|(p+dz) \leqslant \int_{C_\delta} r_T(|z|)\, \theta_p(p+dz) = \alpha\theta_p(C_\delta) T^{-l_1}\delta^{-l_2} \\
&\quad - \int_0^\delta \theta_p(C_q)\, r_T(dq) \leqslant \varepsilon\alpha\lambda(C_\delta) T^{-l_1}\delta^{-l_2} - \varepsilon \int_0^\delta \lambda(C_q)\, r_T(dq) = \varepsilon\alpha\beta T^{-l_1}\delta^{k-l_2} \\
&\quad + \varepsilon\alpha\beta T^{-l_1} \int_{1/T}^\delta l_2 q^{k-l_2-1}\, dq \\
&\quad < \begin{cases} \varepsilon\alpha\beta l_2/(l_2 - k) & \text{if} \quad l_2 - k > 0 \\ \varepsilon\alpha\beta \quad \text{for} \quad T > \tau_0 \quad (\text{where } \tau_0^{-l_1}(1 + l_2 \log(\delta\tau_0)) < 1) & \text{if} \quad l_2 - k = 0 \\ \varepsilon\alpha\beta k/(k - l_2) & \text{if} \quad l_2 - k < 0. \end{cases}
\end{aligned}$$

Thus we find that $|I_1 + I_2 + I_3| < \varepsilon\alpha[1 + |\psi_p|(C_1) + \beta g(k, l_2)]$ for T large enough, where the coefficient of $\varepsilon\alpha$ is a finite number. This completes the proof.

Remark. If μ were a *signed* measure of *finite total variation*, the proof would remain virtually unaltered. Therefore the lemma is *valid* if μ is such a signed measure.

2.7. Examples of kernels K_T. We shall give a few examples of kernels of the special form $K_T(x, y) = Q_T(x - y)$, where $Q_T(z) = Q_T(-z) = (2\pi)^{-k} \int_{R^k} \gamma(t/T)\, e^{-itz}\, dt$, the function γ being a 'convergence factor' (see Bochner [3], definition 2.1.1). All listed kernels K_T satisfy conditions (i) and (ii) of our lemma, unless the opposite is mentioned explicitly. First a few instances with *dimension* $k = 1$, with $\varkappa = 1 = l_1$, $l_2 = 2$, and with α as listed below.

Non-periodic Fejér kernel (corresponding to $(C, 1)$-summation):

$$\gamma(\tau) = (1 - |\tau|) \text{ for } |\tau| \leqslant 1, \ 0 \text{ for } |\tau| > 1; \quad Q_T(0) = (2\pi)^{-1} T;$$

$$\begin{aligned}
Q_T(z) = (2\pi)^{-1} \int_{-T}^{T} \left(1 - \frac{|t|}{T}\right) e^{-itz}\, dt &= (2\pi T)^{-1} \int_0^T du \int_{-u}^{u} e^{-itz}\, dt \\
&= (1 - \cos Tz)(\pi T z^2)^{-1}; \ \alpha = \pi^{-1}.
\end{aligned}$$

Gaussian kernel:

$$\gamma(\tau) = e^{-\frac{1}{2}\tau^2}; \quad Q_T(0) = (2\pi)^{-\frac{1}{2}} T;$$

$$Q_T(z) = (2\pi)^{-\frac{1}{2}} T e^{-\frac{1}{2}T^2 z^2} \leqslant \sqrt{\frac{2}{\pi}} \frac{1}{T z^2} \text{ (for } z > T^{-1}); \quad \alpha = \sqrt{\frac{2}{\pi}}.$$

Kernel corresponding to Cesàro–Riesz summability (see Bochner [3], sec. 2.5): $\gamma(\tau) = 1 - \tau^2$ for $|\tau| \leqslant 1$, 0 for $|\tau| > 1$;

$$Q_T(0) = \frac{4}{3} \frac{T}{2\pi};$$

$$Q_T(z) = (2\pi)^{-1} \int_{-T}^{T} \left(1 - \frac{t^2}{T^2}\right) e^{-itz} dt = (2\pi T^2)^{-1} \int_0^T 2u\, du \int_{-u}^{u} e^{-itz} dt$$

$$= \frac{2}{\pi} T^{-2} z^{-3} (-Tz \cos Tz + \sin Tz);$$

$$|Q_T(z)| \leqslant \frac{4}{\pi} T^{-1} z^{-2} \text{ for } z \geqslant T^{-1}; \quad \alpha = \frac{4}{\pi}.$$

More general Cesàro–Riesz kernels have $\gamma(\tau) = (1 - \tau^2)^m$ for $|\tau| \leqslant 1$, 0 for $|\tau| > 1$;

$$Q_T(0) = \frac{T}{2\pi} B(m+1, \tfrac{1}{2}); \; Q_T(z) = \frac{T}{2\sqrt{\pi}} \Gamma(m+1) J_{m+\frac{1}{2}}(Tz) \left(\frac{Tz}{2}\right)^{-m-\frac{1}{2}};$$

these kernels furnish examples of l_1 and l_2 values different from 1 and 2 (see Bochner [3], sec. 2.5 and p. 169).

Other kernels (see Cramér [4], p. 192):

$$\gamma(\tau) = e^{-|\tau|}; \quad Q_T(0) = \frac{T}{\pi}; \quad Q_T(z) = \frac{T}{\pi} \frac{1}{1 + T^2 z^2} < \frac{1}{\pi} T^{-1} z^{-2}; \quad \alpha = \frac{1}{\pi}.$$

$$\gamma(\tau) = \frac{1}{1 + \tau^2}; \quad Q_T(0) = \frac{T}{2}; \quad Q_T(z) = \frac{T}{2} e^{-T|z|} < T^{-1} z^{-2}; \quad \alpha = 1.$$

Now three instances with *dimension* $k > 1$; here $t'z = \sum_i t_i z_i$; $|t|^2 = t't$.

Gaussian kernel: $\gamma(\tau) = e^{-\frac{1}{2}|\tau|^2}$;

$$Q_T(z) = (2\pi)^{-k/2} T^k e^{-\frac{1}{2}T^2|z|^2} \leqslant \left(\frac{2}{\pi}\right)^{k/2} \frac{k!}{T^k |z|^{2k}}; \quad \alpha = \left(\frac{2}{\pi}\right)^{k/2} k!$$

Other kernels: $\gamma(\tau) = \prod_j e^{-|\tau_j|}$;

$$Q_T(z) = \prod_j \frac{T}{\pi} \frac{1}{1 + T^2 z_j^2} \leqslant \frac{1}{\pi^k} \prod_j \frac{1}{T z_j^2};$$

and

$$\gamma(\tau) = \prod_j \frac{1}{1 + \tau_j^2}; \quad Q_T(z) = \prod_j \tfrac{1}{2} T e^{-T|z_j|} \leqslant \prod_j \frac{1}{T z_j^2}.$$

The Gaussian kernel is a function of $|z|$ only; hence the lemma of section 2.6 can be applied to it; the other two kernels are product kernels and the lemma is of no avail: the product $z_1^2 z_2^2 \dots z_k^2$ cannot be majorated efficiently by a function of $(z_1^2+z_2^2+\dots+z_k^2)$.

3. Results

We will now apply the results listed in Chapter 2 to the problems introduced in Chapter 1.

3.1. **Theorem 1.** *Let φ be the characteristic function of any probability measure $\mu=\mu_a+\mu_s$ defined on $\mathcal{B}_1$. Let $\gamma(\tau)$ be a convergence factor which, according to*

$$Q_T(z)=(2\pi)^{-1}\int_R \gamma\left(\frac{t}{T}\right)e^{-itz}\,dt,$$

corresponds to a kernel $K_T(x,y)=Q_T(x-y)=Q_T(y-x)$ satisfying the conditions of the lemma in section 2.6. Then the following limit exists λ-almost everywhere, and we have, wherever it exists:

$$\lim_{T\to\infty}(2\pi)^{-1}\int_R \varphi(t)\gamma(t/T)e^{-itp}\,dt=\dot{\mu}_a(p). \tag{3.1.1}$$

Remark. All functions γ discussed in section 2.7 for dimension $k=1$ were shown to satisfy the conditions of this theorem. Part of the convergence factor definition is that $\gamma\in L_1(R)$ (see Bochner [3], section 2.1).

Proof. Using the fact that φ is the characteristic function of some probability measure μ, we find for the integral in eq. (3.1.1):

$$\begin{aligned}(2\pi)^{-1}\int_R \gamma\left(\frac{t}{T}\right)e^{-itp}\,dt\int_R e^{ity}\mu(dy)&=(2\pi)^{-1}\int_R \mu(dy)\int_R \gamma(t/T)e^{-it(p-y)}\,dt\\&=\int_R K_T(p,y)\mu(dy),\end{aligned} \tag{3.1.2}$$

where the first equality sign follows from Fubini's theorem since $\gamma\in L_1(R)$ and $\int\mu(dy)=1$. Application of the lemma of section 2.6 yields the desired result immediately.

Remark. The application of an inversion formula of the type of eq. (3.1.1) just destroys (λ-almost everywhere) any contribution which a possible non-absolutely continuous component of μ may have made towards φ. Example: let μ be defined by $\mu(\{a\})=1$, then $\varphi(t)=e^{ita}$; now choose $\gamma(\tau)=e^{-|\tau|}$; then the integral in (3.1.1) equals

$$(2\pi)^{-1}\int_R e^{-it(p-a)}e^{-|\tau|/T}\,dt=\frac{T}{1+T^2(p-a)^2}\to 0$$

with $T\to\infty$ for *all* p, *except* for $p=a$ (where no finite limit exists).

Corollary 1. *Let φ be the characteristic function of any probability measure $\mu = \mu_a + \mu_s$, defined on $\mathcal{B}_1$. Define*

$$f_T(p) = (2\pi)^{-1}\int_{-T}^{T}\left(1-\frac{|t|}{T}\right)\varphi(t)e^{-itp}\,dt = (2\pi T)^{-1}\int_0^T du\int_{-u}^{u}\varphi(t)e^{-itp}\,dt$$
$$= (2\pi T)^{-1}\int_0^T du\int_0^T \varphi(u-v)e^{-i(u-v)p}\,dv. \quad (3.1.3)$$

Then:

(i) $f_T(p) \geq 0$ *for all* p, *all* $T>0$;
(ii) $\int f_T(p)dp = 1$ *for all* $T>0$;
(iii) $\lim_{T\to\infty} f_T(p)$ *exists for λ-almost all p, and equals $\dot{\mu}_a(p)$.*

Proof. Ad(i) Follows from last member of eq. (3.1.3). Cramér showed that "$\varphi(0)=1$ and $f_T(p)\geq 0$ for all p, T" is necessary and sufficient for a bounded and continuous complex function φ to be a characteristic function (cf. Lukacs [13], Th. 4.2.3).

Ad(ii) By (3.1.2) and sec. 2.7:

$$\int_R f_T(p)dp = \int_R dp\int_R Q_T(p-y)\mu(dy) = \int_R \mu(dy)\int_R Q_T(p-y)dp = \int_R \mu(dy) = 1,$$

since
$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1-\cos T(p-y)}{T(p-y)^2}dp = \int_R Q_T(p-y)dp = 1.$$

Ad(iii) Immediate from Theorem 1: take $\gamma(\tau) = 1-|\tau|$ for $|\tau|\leq 1$, 0 for $|\tau|>1$.

Remark. Theorem 1 and result (iii) of Corollary 1 are also valid if μ is not a probability measure, but any signed measure of finite total variation.

Remark. Corollary 1 remains valid for γ being any of the convergence factors discussed in section 2.7 (except that result (i) does not obviously hold for the Cesàro-Riesz kernels).

Remark. The discussion in section 3.1 has been restricted to the case of dimension $k=1$. The argument extends immediately to the Gaussian kernel for $k>1$ (being a radial function). However, the product kernels do need a separate study. For product measures μ it takes a trivial argument to show that product kernels do the job. For non-product measures a rather involved approach through conditional measures (sections) seems indicated. The complexity of the analogous situation in multiple Fourier series (e.g., see Bochner [2] and Mitchell [15]) shows that this problem falls outside the scope of the present paper.

3.2. Criterion for absolute continuity of μ. If the only thing known about a probability measure μ is its characteristic function φ, then one can find out if it is absolutely continuous and construct its density function in one operation according to:

Theorem 2. *Let φ be the characteristic function of some probability measure $\mu = \mu_a + \mu_s$ defined on $\mathcal{B}_1$. Let $f_T(p)$ be defined as in Corollary 1. Write $f(p)$ for $\lim_{T\to\infty} f_T(p)$. Then*

μ is absolutely continuous if and only if

$$\int_R f(p)\,dp = 1,$$

in which case f is the density function of μ.

Remark. The results of Corollary 1 warrant the application of Fatou's lemma, which gives $\int_R f(p)\,dp \leqslant 1$.

Proof. μ is absolutely continuous iff $1 = \mu_a(R) = \int_R \dot{\mu}_a(p)\,dp = \int_R f(p)\,dp$.

Example 1. It is indeed possible to prove directly that if

$$\int_R |\varphi(t)|\,dt < \infty \text{ then } f(p) = \lim_{T\to\infty} f_T(p) = (2\pi)^{-1}\int_R \varphi(t)e^{-itp}\,dt$$

and $\lim_{T\to\infty}\int_a^b f_T(p)\,dp = \int_a^b f(p)\,dp$ for all $a \in R$, $b \in R$, which by Theorem 2 shows that then φ is the characteristic function of an absolutely continuous probability measure.

Example 2. Consider the characteristic function e^{ita} and construct the corresponding $f_T(p) = (2\pi T)^{-1}\int_0^T du \int_{-u}^{u} e^{it(a-p)}dt$; then $\lim_{T\to\infty} f_T(p) = 0$, λ-almost everywhere; $\int_R 0\, dp \neq 1$. So this characteristic function corresponds to a non-absolutely continuous probability measure μ, and the density of the absolutely continuous component is 0, λ-almost everywhere: the abs. cont. component is absent.

Example 3. Let $\varphi(t) = \frac{1}{3}e^{-\frac{1}{2}t^2} + \frac{2}{3}$. We find that the density of the a. c. component is $\frac{1}{3}(2\pi)^{-\frac{1}{2}}e^{-\frac{1}{2}p^2}$, which sums to $\frac{1}{3}$, not to 1. So here the absolutely continuous component is not absent yet this measure is not absolutely continuous.

Example 4. We have proved that for any measure μ, or any characteristic function φ, a function f exists such that $\lim_{T\to\infty} f_T = f$ (λ-almost everywhere); further, if μ is positive and absolutely continuous then $\lim_{T\to\infty}\int f_T = \int f$. In view of the fact that $f_T(p) \geqslant 0$ for all p, Theorem $2H$ in Goldberg [8] is applicable (with $p = 1$) and yields Cramér's [4] (p. 192) necessary condition of type (g) for the special case of (positive) measures. The sufficiency of his condition follows trivially from our results by means of $\left|\int (f_T - f)\right| \leqslant \int |f_T - f|$.

3.3 The Cauchy principal value approach. If $\lim_{T\to\infty}\int_{-T}^{T} a(x)\,dx$ exists, then so does $\lim_{T\to\infty} T^{-1}\int_0^T du \int_{-u}^{u} a(x)\,dx$ and the two limits are equal.

Applying this observation to Corollary 1 and Theorem 2 proves:

Theorem 3. *Let φ be the characteristic function of some probability measure $\mu = \mu_a + \mu_s$ defined on $\mathcal{B}_1$. Define*

$$f^T(p) = (2\pi)^{-1} \int_{-T}^{T} \varphi(t) e^{-itp} dt.$$

(i) *If $\lim_{T\to\infty} f^T(p)$ exists λ-almost everywhere, then wherever it exists:*

$$\lim_{T\to\infty} f^T(p) = \dot{\mu}_a(p).$$

(ii) *Write $f(p)$ for $\lim_{T\to\infty} f^T(p)$; then μ is absolutely continuous if and only if*

$$\int_R f(p)\, dp = 1.$$

Example 1. The chi-square distributions with 1 or 2 degrees of freedom have characteristic functions $(1-2it)^{-n/2}$ $(n=1, 2)$, which do not belong to $L_1(R)$. Yet these distributions are absolutely continuous and their densities can be obtained by the Cauchy principal value approach.

Example 2. Purely discontinuous distributions or components cause their characteristic functions to have non-convergent f^T : if $\varphi(t) = e^{iat}$ then $f^T(p) = \pi^{-1} \sin T(a-p)/(a-p)$, which has no limit as $T \to \infty$ for any p.

3.4. Conclusions. We will use the following abbreviations:

BV = of bounded variation
AC = absolutely continuous
SC = singular continuous
SD = purely discontinuous
λ–a. = λ-almost

We will discuss only univariate distributions and distinguish two cases.

Case 1. φ is *known* to be a *characteristic function*, the Fourier–Stieltjes transform of a probability measure.

1a. $\varphi \in L_1(R)$; then $(2\pi)^{-1} \int_R \varphi(t) e^{-itp} dt$ exists for all $p \in R$ and constitutes the continuous density function of the AC probability measure determined by φ.

1b. $\varphi \notin L_1(R)$, but $f(p) = \lim_{T\to\infty} f^T(p) = (2\pi)^{-1} \lim_{T\to\infty} \int_{-T}^{T} \varphi(t) e^{-itp} dt$ exists λ–a. everywhere; this limit function then constitutes the density function of the AC component of the probability measure μ determined by φ. In this case μ cannot have a SD component, and we have not determined if the existence of a SC component is always compatible with the convergence of $f^T(p)$. Of course, μ is AC iff $\int_R f(p)\, dp = 1$. Jordan's theorem (Goldberg [8], Th. 5C) shows that whenever μ is AC and $\dot{\mu}$ is locally BV λ–a. everywhere, then the corresponding f^T has a limit

λ–a. everywhere. So if $\lim_{T\to\infty} f^T$ does not exist λ–a. everywhere, μ either must have a SD component or might have a SC component, or if μ is AC then $\dot{\mu}$ must be not locally BV on a set of positive Lebesgue measure.

Note that the confusion around the use of the *Cauchy principal value* method in evaluating integral (1.3) for a characteristic function φ has been dissolved to the effect that the *only justification* needed for its use *is* the λ-almost everywhere *existence* of $\lim_{T\to\infty} f^T(p)$; that is, the method gives a meaningful result whenever it is formally feasible. Note that $f^T(p)$ need not be non-negative, as opposed to $f_T(p)$. Also note that one need not immediately resort to $(C, 1)$-summation if φ is not absolutely integrable: one may first try his luck with the Cauchy principal value approach.

1c. $\operatorname{Lim}_{T\to\infty} f^T(p)$ does not exist (λ–a. everywhere). In this case, as in the two other cases,

$$f(p)=\lim_{T\to\infty} f_T(p)=(2\pi)^{-1}\lim_{T\to\infty} T^{-1}\int_0^T du\int_{-u}^{u}\varphi(t)e^{-itp}\,dt$$

does exist, λ–a. everywhere, and equals $\dot{\mu}_a$, the density of the AC component of μ; μ is AC iff $\int_R f(p)\,dp=1$. This $(C, 1)$-summation will *always identify* $\dot{\mu}_a$, but the computation is unduly cumbersome if any of the other two is available. Fortunately most of the work to be done in case (1*b*) is needed in case (1*c*) anyway.

Note that possibly another convergence factor γ may yield simpler computations for special cases (sections 2.7 and 3.1).

Case 2. φ is *not known* to be a *characteristic function*. There are two ways of finding out. As a preliminary, use the fact that all characteristic functions φ are everywhere continuous (even uniformly so), and also $\varphi(0)=1$, $|\varphi(t)|\leqslant 1$ everywhere. Then

2a. If in addition $\varphi\in L_1(R)$, use Theorem 2.1 in Letta [11], which amounts to the result that such φ is a characteristic function if also $\int\varphi(t)e^{-itp}\,dt\geqslant 0$ for all $p\in R$;

2b. if $\varphi\notin L_1(R)$, apply point (i) of Corollary 1.

ACKNOWLEDGEMENT

This work was partially supported by Public Health Service Grant GM-678 from the National Institute of General Medical Sciences to the Biomathematics Training Program, Institute of Statistics, Raleigh Section.

The author wants to thank Dr. T. Kawata for several useful comments he made during our correspondence.

Institute of Statistics (Biomathematics Program) and Department of Mathematics, North Carolina State University at Raleigh, North Carolina, U.S.A.

REFERENCES

1. Alexits, G., Convergence problems of orthogonal series. Pergamon, New York–London–Paris (1961).
2. Bochner, S., Summation of multiple Fourier series by spherical means. *Transactions of the American Mathematical Society 40*, 175–207 (1936).
3. Bochner, S., Harmonic analysis and the theory of probability. University of California Press, Berkeley–Los Angeles (1955).
4. Cramér, H., On the representation of a function by certain Fourier integrals. *Transactions of the American Mathematical Society 46*, 191–201 (1939).

5. DOOB, J. L., Relative limit theorems in analysis. *Journal d'analyse mathématique 8*, 289–306 (1960/61).
6. DUGUÉ, D., Traité de statistique théorique et appliquée, 1er fasc.: analyse aléatoire. Masson, Paris 1958.
7. DUNFORD, N., and SCHWARTZ, J. T., Linear operators, part 1: general theory. Interscience, New York–London (1958).
8. GOLDBERG, R. R., Fourier transforms. Cambridge University Press, London–New York 1961.
9. HEWITT, E., and STROMBERG, K., Real and abstract analysis. Springer, Berlin–Heidelberg–New York 1965.
10. KENDALL, M. G., and STUART, A., The advanced theory of statistics (three volume edition), vol. 1: distribution theory. Griffin, London 1958.
11. LETTA, G., Eine Bemerkung zur Kennzeichnung der charakteristischen Funktionen. Zeitschrift für Wahrscheinlichkeitstheorie *2*, 69–74 (1963).
12. LÉVY, P., Calcul des probabilités. Gauthier-Villars, Paris 1925.
13. LUKACS, E., Characteristic functions (Number 5 of Griffin's statistical monographs and courses). Griffin, London 1960.
14. LUKACS, E., Applications of characteristic functions in statistics. *Sankhyā, the Indian Journal of Statistics, Ser. A, 25*, 175–188 (1963).
15. MITCHELL, J., On the spherical summability of multiple orthogonal series. *Transactions of the American Mathematical Society 71*, 136–151 (1951).
16. RICHTER, H., Wahrscheinlichkeitstheorie, 2. Auflage. Springer, Berlin–Heidelberg–New York 1966.
17. ROBINSON, E. A., An introduction to infinitely many variates (Number 6 of Griffin's statistical monographs and courses). Griffin, London 1959.
18. ROYDEN, H. L., Real analysis. MacMillan, New York (1963).
19. RUDIN, W., Real and complex analysis. McGraw-Hill, New York–London (1966).
20. VAART, H. R. VAN DER, On the characteristic functions of absolutely continuous distribution functions. Abstract No. 36 in *Annals of Mathematical Statistics 33*, 824 (1962).

Tryckt den 29 december 1967

Uppsala 1967. Almqvist & Wiksells Boktryckeri AB