A strong Tauberian theorem for characteristic functions

R. Riedi† and P. Gonçalves‡
†Dept of Mathematics, EIA-Fribourg, Switzerland rudolf.riedi@hefr.ch
‡INRIA, ENS Lyon paulo.goncalves@inria.fr
March 9, 2015

Abstract
Using wavelet analysis we show that if the characteristic function of a random variable $X$ can be approximated at 0 by some polynomial of even degree $2p$ then the moment of order $2p$ of $X$ exists. This strengthens a Tauberian-type result by Ramachandran and implies that the characteristic function is actually $2p$ times differentiable at 0. This fact also provides the theoretical basis for a wavelet based non-parametric estimator of the tail index of a distribution.

Introduction
When studying the existence of moments $\beta_q = \mathbb{E}[X^q]$ of a random variable $X$ it is essential to understand the behavior of its characteristic function $\phi(u) = \mathbb{E}[\exp(iuX)]$ at 0, in particular its degree of differentiability, and more precisely its Hölder regularity.
A function $g$ is called Hölder-regular of degree $r \geq 0$ at $u_0$, denoted $g \in \mathcal{H}^r(u_0)$, if there exists a polynomial $P$ of degree at most $\lfloor r \rfloor := \max \{n \in \mathbb{N}_0 : n \leq r\}$ such that
\[ |g(u) - P(u)| \leq C \cdot |u - u_0|^r \] (1)
If $g$ is actually $\lfloor r \rfloor$ times differentiable, then $P$ must be equal to the Taylor polynomial and we call $g$ Taylor regular of degree $r$. Functions which obey (1) but are not $\lfloor r \rfloor$ times differentiable are abundant. An example is $g(u) = 1 + u + u^2 + u^{3.1} \sin(1/u)$ which is only once differentiable at $u = 0$ but Hölder regular for any $r \leq 3.1$.
However, differentiability of a characteristic function at 0 is intimately related to the existence of moments: for any positive integer $p$ we have $\phi^{(2p)}(0) = (-1)^p \mathbb{E}[X^{2p}]$ whenever one of the two is defined. To make this point even stronger, Ramachandran shows: if $\phi(u)$ is symmetric and Taylor regular of degree $r = 2p + \delta$ ($0 < \delta < 2$) at 0 then all moments $\beta_q$ of degree $q < r$ are finite (see [6, Thm. 3]).
Thereby, the assumption of $\phi$ being symmetric is not required (see, e.g., Lukacs [5, Thm. 2.2.2.]). Indeed, the symmetric function $\Re \phi(u)$ is the characteristic function of the random variable $Y$ which is equal to $X$ with probability 1/2, and equal to $-X$ with probability 1/2. Since $|Y| = |X|$ a.s. we have $\mathbb{E}|Y|^q = \mathbb{E}|X|^q = \beta_q$ and it suffices to consider $\Re \phi$.

The main result
We are able to strengthen Ramachandran’s result considerably via the following:
**Theorem 1** Let \( r > 0 \) be positive and \( p \) integer such that \( 2p \leq r < 2p + 2. \)
If \( \text{Re} \phi(u) \in \mathcal{H}^r(0) \) then \( \phi(u) \) is \( 2p \)-times differentiable at \( 0. \)

**Proof**

For \( p = 0 \) there is nothing to show. For \( p \geq 1 \) denote by \( \psi_p \) the function with real valued, non-negative Fourier transform

\[
\Psi_p(\nu) = \nu^{2p} \exp \left(-\nu^2/4\right).
\]  
(2)

Up to a constant factor, \( \psi_p(u) \) is the \( 2p \)-th derivative of \( \exp(-u^2) \).
Now consider the wavelet coefficients of \( \text{Re} \phi \) of scale \( s \) at \( t \) which are given by

\[
W_p(s, t) := \int \frac{1}{s} \psi_p \left( \frac{u-t}{s} \right) \text{Re} \phi(u) du.
\]  
(3)

The Fourier transform of \( \frac{1}{s} \psi_p \left( \frac{u}{s} \right) \) being \( \Psi_p(sv) \), Parseval’s identity yields

\[
W_p(s, 0) = \text{Re} \int \frac{1}{s} \psi_p \left( \frac{u}{s} \right) \phi(u) du = \text{Re} \int \Psi_p(sv)dF_X(\nu) = \mathbb{E}[\Psi_p(sX)].
\]  
(4)

Monotone convergence gives

\[
\sup_{s>0} \frac{W_p(s, 0)}{s^{2p}} = \sup_{s>0} \mathbb{E}[X^{2p} \exp(-s^2X^2/4)] = \mathbb{E}[X^{2p}] \in \mathbb{R}_0^+ \cup \{\infty\}.
\]

To control the decay of \( W_p(s, 0) \) at 0 note that \( \text{Re} \phi \) is Hölder regular of degree \( 2p \) which can be seen by absorbing higher order terms into \( C|u|^{2p} \) from (1) for \( |u| \leq 1 \) and by using \( |\phi(u)| \leq 1 \) for \( |u| > 1. \) Also, \( \text{Re} \phi \) is even. Thus, there is an even polynomial \( Q \) of degree at most \( 2p-2 \) and a constant \( C_1 \) such that \( |\text{Re} \phi(u) - Q(u)| \leq C_1|u|^{2p} \). By repeated integration by parts, we find that \( \int Q(u)\psi_p(u/s)du = 0 \) for any \( s > 0. \) From this, we find

\[
|W_p(s, 0)| = \left| \int \frac{1}{s} \psi_p \left( \frac{u}{s} \right) (\text{Re} \phi(u) - Q(u)) du \right| \leq C_1 \int \frac{1}{s} |\psi_p \left( \frac{u}{s} \right)| |u|^{2p} du
\[
= C_1 s^{2p} \int |\psi_p(v)||v|^{2p} dv = C_2 s^{2p}.
\]  
(5)

Note that \( C_2 < \infty \) due to the exponential tail of \( \psi_p \). The claim follows. \( \diamond \)

Theorem 1 strengthens Ramachandran [6, Thm. 3] as follows.

**Theorem 2** Let \( r > 0. \) If \( \text{Re} \phi \in \mathcal{H}^r(0) \) then \( \beta_q < \infty \) for \( 0 \leq q < r. \)

**Remark 1:** The stable laws with \( \phi(u) = \exp(-|u|^r) \) \((0 < r < 2)\) demonstrate that \( \beta_r \) may not be finite in general. However, if \( r \) is an even integer, we always have \( \beta_r < \infty. \)

**Proof**

In the light of Theorem 1 we may simply quote Ramachandran [6, Thm. 3] to complete the proof. However, just a little more work provides an alternative proof
to this classical result. In the notation of (1) and (2) we have \( \int P(u)\psi_{p+1}(u/s)du = 0 \) \((p \geq 0)\). Similar as in (5) we get

\[
E[\Psi_{p+1}(sX)] \leq C \int \frac{1}{s} |\psi_{p+1}(\frac{x}{s})| |u|^r du = C_3 |s|^r. \tag{6}
\]

Using \( \Psi_{p+1}(x) \geq e^{-1/4} > 1/2 \) for \( 1 \leq |x| \leq 2 \) \((p \geq 0)\) we find

\[
\int_{2^k \leq |x| \leq 2^{k+1}} dF_X(x) = E[|1_{|2^{-k}|X|<1}|] \leq 2E[\Psi_{p+1}(2^{-k}X)] \leq 2C_3 \cdot 2^{-kr}.
\]

It follows immediately that the moments \( \beta_q \) are finite for \( 0 \leq q < r \). \hfill \Box

**Classical results**

Theorem 1 is quite well known in the case \( r = 2 \):

**Theorem 3** If \( |\phi(u) - 1| \leq C \cdot u^2 \), then the variance of the law is finite.

**Proof**

Note that \( |\text{Re} \phi(u) - 1| \leq C \cdot u^2 \). But \( 1 - \text{Re} \phi(u) = E[2\sin^2(uX/2)] \). Since \( X^2 = \lim_{u \to 0} \sin^2(uX)^2 \) a.s. it suffices then to apply Fatou’s lemma. \hfill \Box

**Alternative choice of wavelet:**

Our choice of wavelet in the proofs of Theorem 1 and Theorem 2 (see (2)) is by far not the only possible. Here is one interesting alternative. Motivated by the classical argument of Theorem 3 we may replace \( \Psi_p \) by \( \sin^{2p} \) which leads to a somewhat more elementary computation akin to the proofs of Ramachandran and Lukacs.

Indeed, using \( \sin^{2p} = (2i)^{-2p}(\exp(isx) - \exp(-isx))^{2p} \) we obtain (cpre. (4))

\[
E[\sin^{2p}(sX)] = (-1/4)^p \sum_{k=0}^{2p} \binom{2p}{k} (-1)^k \phi(s(2p-2k)) =: \Delta^s_p \phi(0).
\]

Here, we introduced an order \( 2p \) difference operator \( \Delta^s_p \) centered at \( 0 \) which is in essence the \( 2p \)-fold convolution of the Haar wavelet. Being non-negative real \( \Delta^s_p \phi(0) \) equals \( \Delta^s_p \text{Re} \phi(0) \). Also, \( \Delta^s_p \) possesses the appropriate number of vanishing moments, meaning that \( \Delta^s_p Q = 0 \) for \( Q \) from (5), since \( Q \) is of degree at most \( 2p - 2 \) and \( \sum_{k=0}^{2p} \binom{2p}{k} (-1)^k \cdot k^m = 0 \) for any \( 0 \leq m \leq 2p - 2 \). Setting \( C_1 \) as in (5) we find

\[
E[\sin^{2p}(sX)] = |\Delta^s_p(\text{Re} \phi - Q)| \leq 4^{-p} \sum_{k=0}^{2p} \binom{2p}{k} C_1 |2p - 2k|^{2p} |s|^{2p}.
\]

Using Fatou as in Theorem 3 establishes finiteness of \( \beta_{2p} \) and, therefore, Theorem 1. To establish Theorem 2 in this setup, write \( E[\sin^{2p+2}(sX)] = |\Delta^s_{p+1}(\text{Re} \phi - P)| \leq C_4 |s|^r \) and use \( \sin \theta \geq \sqrt{3}/2 \) for \( \pi/3 \leq \theta \leq 2\pi/3 \) to establish finiteness of \( \beta_q \) for \( 0 \leq q < r \) as in Theorem 2 above. \hfill \Box

**Remark 2:** Lukacs [5] uses essentially the same operator \( \Delta^s_p \phi(t) \), centered at an arbitrary \( t \); he also uses the vanishing moments. But writing \( \Delta^s_p \phi(t) = \Delta^s_p P(t) + \Delta^s_p(\phi - P)(t) \) his estimates contain a term \( s^{2p} \) in addition to \( s^r \). Lukacs assumes
Lemma 5 At this point, the wavelet-maxima-property of characteristic functions proves useful:

\[ |\Delta_m^s \phi(0)| \leq C|s|^{2m}(s \to 0^+) \Rightarrow \phi^{(2m)}(u) \text{ exists.} \]

\[ |\Delta_n^s \phi(0)| \leq C|s|^r (s \to 0^+) \Rightarrow \beta_q < \infty (0 \leq q < r). \]

In the special case \( n = 1 \) we are lead to the following notion: A function \( g \) is said to lie in the Zygmund class \( \dot{C}^1(u_0) \) if \( g \) satisfies

\[ |g(u_0 + s) - 2g(u_0) + g(u_0 - s)| \leq C|s|^r \quad (7) \]

with \( r = 1 \). The condition (7) is equivalent to \( g \in \mathcal{H}_r(u_0) \) for \( 0 < r < 1 \), but the inclusion \( \mathcal{H}_1(u_0) \subset \dot{C}^1(u_0) \) is strict. However, since \( 1 - \text{Re} \phi(2s) = 2 \Delta_1^s \phi(0) \) the real part of any characteristic function in Zygmund class \( \dot{C}^1(0) \) lies actually in \( \mathcal{H}_1(0) \).

We study the derivatives of characteristic functions in terms of the following spaces: we set \( \dot{C}^r(u_0) = \mathcal{H}_r(u_0) \) for \( 0 < r < 1 \); for \( r > 1 \) we denote by \( \dot{C}^r(u_0) \) the space of functions who’s derivatives exist around \( u_0 \) up to order \( m \) \((m < r \leq m + 1) \) and who’s \( m \)-th derivative lies in \( \dot{C}^{r-m}(u_0) \).

For general functions, control on the regularity of derivatives is not obvious (see the example of a function \( g \in \mathcal{H}^{3,1}(0) \) from the introduction). In order to gain such control we make use of the so-called two-microlocalization: with reference to (3) the function \( \text{Re} \phi(u) \) is said to belong to \( C^{r,r'}(u_0) \) if

\[ |W_{p+2}(s,t)| \leq Cs^{r}
\left(1 + \frac{|t-u_0|}{s}\right)^{-r'} . \quad (8) \]

Here, the wavelet \( \psi_{p+2} \) needs to be sufficiently regular, i.e., \( 2p + 4 > r + 2 \).

We will need the following results

\[ \mathcal{H}^1(u_0) \subset C^{r-r}(u_0) \quad (r > 0) \]

\[ C^{r,r'}(u_0) \subset \mathcal{H}^r(u_0) \quad (r' > -r, r > 0 \text{ non-integer}) \quad (9) \]

The first two inclusions are found in [3, Prop. 1] as well as [4, Prop. 1.3]. For the third inclusion one proves the somewhat more general fact that \( g \in C^{r,0}(u_0) \) \((1 \leq r < 2) \) implies (7) by applying the mean value theorem twice in [3, p. 292].

At this point, the wavelet-maxima-property of characteristic functions proves useful:

**Corollary 4** Let \( r \) be positive real and \( m \) and \( n \) positive integers. Then

\[ |\Delta_m^s \phi(0)| \leq C|s|^{2m}(s \to 0^+) \Rightarrow \phi^{(2m)}(u) \text{ exists.} \]

\[ |\Delta_n^s \phi(0)| \leq C|s|^r (s \to 0^+) \Rightarrow \beta_q < \infty (0 \leq q < r). \]

Further regularity results
The alternative wavelets mentioned above provide regularity conditions that do not require an approximating polynomial of the characteristic function:

**Lemma 5** The wavelet coefficients \((3)\) of \( \phi \) and \( \text{Re} \phi \) are both maximal at 0:

\[ |W_p^{\text{Re}\phi}(s,t)| \leq |W_p^{\phi}(s,t)| \leq W_p(s,0) = \mathbb{E}[\Psi_p(sX)]. \]
Proof
This follows easily from Parseval’s identity, i.e.,
\[
W_p^\phi(s,t) = \int \frac{1}{s} \psi_p(\frac{u-t}{s})\phi(u)du = \int \Psi_p(s\nu)\exp(-it\nu)dF_X(\nu),
\]
combined with \(\Psi_p \geq 0\), and (4).

Note that \(\Re \phi \in \mathcal{H}'(0)\) implies \(|W_{p+2}(s,0)| \leq Cs^r (s > 0)\) due to (9) and (8) with \(t = u_0\).
On the other hand, this bound implies with Lemma 5 that \(\phi \in C^{r,0}(u_0)\) for any \(u_0\).
Notably, the same constant \(C\) can be used in the wavelet bound (8) for all \(u_0\). This fact is indicated by writing \(2C\) in (8) and analogously for other regularity spaces.

Now, (9) implies

\[\text{Corollary 6} \quad \text{For non-integer } r > 0 \text{ we have} \]
\[\Re \phi \in \mathcal{H}'(0) \iff |W_{p+2}(s,0)| \leq Cs^r (s > 0) \iff \phi \in \mathcal{H}'(\mathbb{R}) \iff \Re \phi \in \mathcal{H}'(\mathbb{R}).\]

Due to Theorem 1 and classical results, \(\Re \phi \in \mathcal{H}'(0)\) implies that the derivatives of \(\phi\) exist and are uniformly continuous up to order \(m\) (\(m < r \leq m + 1\)). It is known that [2, 4]
\[\Re \phi(u) \in C^{r',r}(u_0) \Rightarrow \frac{\partial^m}{\partial u^m} \Re \phi(u) \in C^{r-m,r'}(u_0)\]
This precise control of the regularity of the derivatives via two-microlocalization can be easily verified and sharpened for even order derivatives of characteristic functions:

\[\text{Lemma 7} \quad \text{Assume } \beta_{2m} < \infty \text{ and } p > m. \text{ Then (analogous for } \Re \phi)\]
\[W_{p-m}(s,t) = (-1)^m s^{-2m} \cdot W_p^\phi(s,t).\]

Proof
Consider the probability distribution \(dF_m(x) = \frac{1}{\beta_{2m}} x^{2m} dF_X(x)\) with characteristic function
\[\phi_m(u) = \frac{1}{\beta_{2m}} \int e^{iux} x^{2m} dF_X(x) = \frac{1}{\beta_{2m}} \mathbb{E}[e^{iux} X^{2m}] = (-1)^m \frac{1}{\beta_{2m}} \phi^{(2m)}(u).\]
A direct computation using Parseval’s identity (10) leads to the claimed relation.

Remark 3: The Cauchy distribution with \(\phi(u) = \exp(-|u|)\) lies in \(\mathcal{H}^1(0)\) but is not differentiable. A condition for \(\phi^{(2p)}\) being differentiable at 0 and, thus, everywhere is provided in [2, Prop. 1] and [4, Prop 1.5].
An alternative approach to using wavelets is a Littlewood-Paley decomposition of $\text{Re}\phi$ into functions $\Delta_j\text{Re}\phi$. To allow recycling of notation, consider
\[ h(\nu) = \begin{cases} 
 h(\nu/2) - h(\nu) & \text{if } 0 \leq h \leq 1, \\
 1 & \text{if } |\nu| \leq 1/2 \\
 0 & \text{if } |\nu| \geq 1. 
\end{cases} \]
Then, by definition, the Fourier transform of $\Delta_j\text{Re}\phi$ equals $\Psi_0(2^{-j}\nu)$ times the Fourier transform of $\text{Re}\phi$.

In the notation of (3) we get $\Delta_j\text{Re}\phi(u) = W_0(2^{-j}, u)$. Since $\Psi_0$ is positive, Lemma 5 holds also in this context; with (4), (6) and $\Psi_0 \leq K\Psi_p$ for some $K > 0$ we see that $\text{Re}\phi \in \mathcal{H}^r(0)$ implies
\[ |\Delta_j\text{Re}\phi(u)| \leq \Delta_j\text{Re}\phi(0) = E[\Psi_0(2^{-j}X)] \leq KC\nu 2^{-jr} \]
These estimates can be used in a manner similar to the arguments provided earlier.

**Statistical estimation**

As an application we mention a wavelet-based non-parametric estimator of the *tail-exponent*
\[ \lambda := \sup\{q > 0 : \text{the moment } \beta_q \text{ is finite}\} \]
of the distribution of $X$. As is well known $\beta_q < \infty (q > 0)$ implies $\text{Re}\phi \in \mathcal{H}^q(0)$ (see [5, Thm. 2.2.1.]). With Theorem 2 it follows that $\lambda = \sup\{r > 0 : \text{Re}\phi \in \mathcal{H}^r(0)\}$.

With Corollary 6 we conclude

**Corollary 9** Provided $2n > \lambda + 2$, the tail exponent $\lambda$ satisfies
\[ \lambda = \sup\{r > 0 : E[\Psi_n(sX)] \cdot s^{-r} \text{ is bounded for } s \to 0^+\} \]

In conclusion, the estimation of $\lambda$ can be based on estimating $E[\Psi_n(sX)]$ for various values of $s > 0$ followed by a least square linear regression in log-log towards fitting $E[\Psi_n(sX)] \simeq s^\lambda$. Various values of $n$ should be employed to ensure $2n > \lambda$. The exact setup and performance of this estimation procedure is detailed in [1].

**Acknowledgements:** The first author is grateful to Yves Meyer and to an anonymous reviewer for useful and inspiring suggestions and to the GDR d’Analyse multifractale for travel support.

**References**


