



ACADEMIC  
PRESS

Available at  
**WWW.MATHEMATICSWEB.ORG**  
POWERED BY SCIENCE @ DIRECT®

Journal of Approximation Theory 121 (2003) 54–60

JOURNAL OF  
**Approximation  
Theory**

<http://www.elsevier.com/locate/jat>

# Necessary and sufficient condition that the limit of Stieltjes transforms is a Stieltjes transform

Jeffrey S. Geronimo<sup>\*,1</sup> and Theodore P. Hill<sup>2</sup>

*Georgia Institute of Technology, School of Mathematics, Atlanta, GA 30332-0160, USA*

Received 20 March 2002; accepted in revised form 18 October 2002

Communicated by Walter Van Assche

---

## Abstract

The pointwise limit  $S$  of a sequence of Stieltjes transforms  $(S_n)$  of real Borel probability measures  $(P_n)$  is itself the Stieltjes transform of a Borel p.m.  $P$  if and only if  $iyS(iy) \rightarrow -1$  as  $y \rightarrow \infty$ , in which case  $P_n$  converges to  $P$  in distribution. Applications are given to several problems in mathematical physics.

© 2002 Elsevier Science (USA). All rights reserved.

**Keywords:** Borel probability measure; Weak convergence; Stieltjes transform; Grommer–Hamburger theorem

---

## 1. Main theorem

The main goal of this note is to clear up some confusion in the literature about the relationship between limits of Stieltjes transforms of probability distributions and weak convergence, by characterizing which limits are themselves Stieltjes transforms of probability measures (Theorem 1). As such, this result may be considered a gloss on a theorem of Grommer and Hamburger [17], and also as a direct analog of Lévy's classical continuity theorem, complementing those in [6,7]. Since no reference is known to the authors, a detailed proof is included for completeness.

Throughout this note,  $\mathbb{R}$  and  $\mathbb{C}$  denote the real and complex numbers, respectively; p.m. and s.p.m. denote Borel probability measures, and sub-probability (mass  $\leq 1$ )

---

\*Corresponding author.

E-mail addresses: [geronimo@math.gatech.edu](mailto:geronimo@math.gatech.edu) (J.S. Geronimo), [hill@math.gatech.edu](mailto:hill@math.gatech.edu) (T.P. Hill).

<sup>1</sup> Research partially supported by NSF Grant DMS-9970613.

<sup>2</sup> Research partially supported by NSF Grant DMS-9971146 and Göttingen Academy of Sciences Gauss Professorship (Fall 2000).

measures, respectively, on  $\mathbb{R}$ ; and s.p.m.'s  $(\mu_n)$  converge vaguely to a s.p.m.  $\mu$  [3, p. 80], if there exists a dense subset  $D$  of  $\mathbb{R}$  such that for all  $a, b \in D$ ,  $a < b$ ,  $\mu_n((a, b]) \rightarrow \mu((a, b])$ . (Thus if  $(\mu_n)$ ,  $\mu$  are p.m.'s, vague convergence is equivalent to convergence in distribution.)

**Definition 1.** The *Stieltjes transform*  $S_P$  of a p.m.  $P$  is the function  $S_P : \{\operatorname{Im}(z) > 0\} \rightarrow \mathbb{C}$  given by

$$S_P(z) = \int_{-\infty}^{\infty} \frac{1}{w - z} dP(w).$$

(Note: In some texts the Stieltjes transform is defined as the negative of the one given here, cf. [5, 16].)

A basic property of Stieltjes transforms, which has important applications in the theory of moments, orthogonal polynomials, and mathematical physics (cf. [1, 2]; [11, pp. 48, 59]; [13–15]), is that they are a representing class for finite measures; [4, Chapter 14] has an extensive table of Stieltjes transforms.

**Lemma 1.** For s.p.m.'s  $P$  and  $Q$ ,  $P = Q$  iff  $S_P = S_Q$ .

**Proof.** Follows immediately from the Stieltjes transform inversion formula [1, p. 125].  $\square$

Just as limits of characteristic functions of p.m.'s are in general not characteristic functions, and limits of Hardy–Littlewood functions or expected-extrema functions are not in general Hardy–Littlewood or expected-extrema functions [6], limits of Stieltjes transforms of p.m.'s are not always Stieltjes transforms of p.m.'s, as the next easy example shows.

**Example 1.** For  $n = 1, 2, \dots$ , let  $P_n = \delta_{(n)}$ , the Dirac point mass at  $n$ . Then  $S_{P_n}(z) = (n - z)^{-1}$  for all  $n$  and all  $z$  with  $\operatorname{Im}(z) > 0$ , so  $\lim_{n \rightarrow \infty} S_{P_n}(z) \equiv 0$ , which is clearly not the Stieltjes transform for any p.m.  $P$  (see Lemma 2).

On the other hand, just as with Lévy's theorem, the limit of Stieltjes transforms is itself a Stieltjes transform if and only if it satisfies one single universal limit condition.

**Theorem 1.** Suppose that  $(P_n)$  are real Borel probability measures with Stieltjes transforms  $(S_n)$ , respectively. If  $\lim_{n \rightarrow \infty} S_n(z) = S(z)$  for all  $z$  with  $\operatorname{Im}(z) > 0$ , then there exists a Borel probability measure  $P$  with Stieltjes transform  $S_P = S$  if and only if

$$\lim_{y \rightarrow \infty} iyS(iy) = -1 \tag{1}$$

in which case  $P_n \rightarrow P$  in distribution.

**Corollary 1.** *If  $P, (P_n)$  are real Borel p.m.'s with Stieltjes transforms  $S, (S_n)$ , respectively, then  $P_n \rightarrow P$  in distribution if and only if  $S_n \rightarrow S$  pointwise.*

**Proof.** If  $S_n \rightarrow S$ , then  $P_n \rightarrow P$  in distribution by Theorem 1. Conversely, suppose that  $P_n \rightarrow P$  in distribution. Since  $f_z(w) := (w - z)^{-1}$  is continuous and bounded in  $w$  for fixed  $z$  in  $\{\text{Im}(z) > 0\}$ , then  $\text{Im}(f_z)$  and  $\text{Re}(f_z)$  are also continuous and bounded, so by the basic equivalence of convergence in distribution of p.m.'s and convergence of integrals of bounded continuous functions [3, Theorem 4.4.2],  $\int \text{Im}(f_z) dP_n \rightarrow \int \text{Im}(f_z) dP$  and  $\int \text{Re}(f_z) dP_n \rightarrow \int \text{Re}(f_z) dP$ , so  $S_n(z) \rightarrow S(z)$ .  $\square$

To facilitate the proof of Theorem 1, two additional lemmas are useful, which are stated here for ease of reference.

**Lemma 2.** *Let  $S : \{\text{Im}(z) > 0\} \rightarrow \mathbb{C}$  be analytic. Then there exists a p.m.  $P$  with  $S_P(z) = S(z)$  for all  $z$  with  $\text{Im}(z) > 0$  if and only if (1) holds and*

$$\text{Im}(S(z)) > 0 \quad \text{for all } z \text{ with } \text{Im}(z) > 0. \quad (2)$$

**Proof.** By the classical Akhiezer–Krein theorem [1, p. 93],  $S = S_P$  for some finite positive Borel measure  $P$  if and only if:  $S$  is analytic in  $\{\text{Im}(z) > 0\}$ ;  $S$  satisfies (2); and

$$\sup_{y \geq 1} |yS(iy)| < \infty. \quad (3)$$

Suppose  $P$  is a p.m. with  $S = S_P$ . The Akhiezer–Krein theorem implies that (2) holds, and (1) follows immediately from the definition of  $S_P$ . Conversely, suppose that  $S$  is analytic and satisfies (1) and (2). Since  $yS(iy)$  is continuous in  $y$ , (1) easily implies (3), so by the Akhiezer–Krein theorem again, there is a finite positive Borel measure  $P$  with  $S_P = S$ . By the Dominated Convergence Theorem,  $\lim_{y \rightarrow \infty} [-iyS_P(iy)] = \text{mass}(P)$ , so (1) implies that  $P$  is a p.m.  $\square$

**Lemma 3.** *Let  $\mathcal{F}$  be a family of functions analytic in an open connected set  $D$ . If for each compact set  $K$  in  $D$  there is a constant  $M(K)$  such that*

$$|f(z)| \leq M(K) \quad \text{for all } f \in \mathcal{F} \text{ and } z \in K, \quad (4)$$

*then every sequence in  $\mathcal{F}$  has a subsequence that converges uniformly on compact subsets of  $D$  to a function analytic in  $D$ .*

**Proof** (Hille [8, Theorem 15.2.3]).  $\square$

**Proof of Theorem 1.** If  $\lim S_n = S = S_P$  for some p.m.  $P$ , then (1) follows by Lemma 2.

Conversely, suppose that  $S = \lim_n S_n$  satisfies (1). Let  $\mathcal{F} = \bigcup \{S_n\}$ , and for  $K \subset D : = \{\text{Im}(z) > 0\}$ , let  $d(K) = \inf\{\|y - z\| : y \in \mathbb{R}, z \in K\}$ , the smallest distance from  $K$  to the real line. Clearly  $0 < d(K) < \infty$  for all compact  $K \subset D$ , and  $M(K) = 1/d(K)$

satisfies (4), so Lemma 3 implies that  $S = \lim S_n$  is analytic in  $D$ . By Lemma 2,  $\text{Im}(S_n(z)) > 0$  for all  $z \in D$ , so  $\text{Im}(S(z)) \geq 0$  for all  $z \in D$ . Suppose, by way of contradiction to (2), that  $\text{Im}(S(z_0)) = 0$  for some  $z_0 \in D$ . Since  $S$  is analytic,  $\text{Im}(S)$  and  $\text{Re}(S)$  are harmonic on  $D$  [9, p. 590]. By the maximum principle [9, p. 760], a non-constant function which is harmonic in a simply connected bounded open set  $G$  has neither a maximum nor a minimum in  $G$ , so since  $\text{Im}(S(z)) \geq 0$  on  $G$  for every simply connected open bounded set  $G$  with  $z_0 \in G \subset D$ , it follows (taking  $G_t = \{z \in D : \|z\| < t\}$ , and letting  $t \rightarrow \infty$ ) that  $\text{Im}(S(z)) \equiv 0$  for all  $z \in D$ , which contradicts (1). Thus (2) holds, and since  $S$  is analytic and (1) holds by assumption, Lemma 2 implies there exists a real Borel p.m.  $P$  with  $S_P = S$ .

For the convergence in distribution conclusion, suppose that  $S_n = S_{P_n} \rightarrow S_P$  pointwise in  $D$  for p.m.'s  $(P_n)$ ,  $P$ . By the Helly selection theorem [3, Theorem 4.3.3], there exists a s.p.m.  $Q$  and a subsequence  $(P_{n_k})$  of  $(P_n)$  such that  $P_{n_k} \rightarrow Q$  vaguely. Fix  $z$  in  $D$ , and let  $f_z : \mathbb{R} \rightarrow \mathbb{C}$  be given by  $f_z(w) = (w - z)^{-1}$ . Since  $f_z$  is continuous in  $w$  and vanishes at infinity,  $\text{Re}(f_z)$  and  $\text{Im}(f_z)$  are continuous and vanish at infinity, so it follows by the equivalence of vague convergence of s.p.m.'s and convergence of integrals of continuous functions which vanish at infinity [3, Theorem 4.4.1] that  $S_{P_{n_k}}(z) \rightarrow S_Q(z)$  as  $k \rightarrow \infty$  for all  $z \in D$ . By hypothesis,  $S_{P_n} \rightarrow S_P$ , so  $S_P = S_Q$ , which by Lemma 1 implies that  $P = Q$ . Since every vaguely convergent subsequence of  $(P_n)$  thus converges to  $P$ , this implies [3, Theorem 4.3.4] that  $P_n$  converges vaguely to  $P$ , that is, since  $(P_n)$  and  $P$  are p.m.'s,  $P_n$  converges to  $P$  in distribution.  $\square$

A slight generalization of Theorem 1 will be needed for one of the examples below.

**Theorem 2.** *Suppose that  $(P_n)$  are real Borel probability measures with Stieltjes transforms  $(S_n)$ , respectively. Let  $K \subset \{\text{Im } z > 0\}$  be an infinite set with a limit point  $z_0$ ,  $\{\text{Im } z_0 > 0\}$ . If  $\lim S_n(z) = S(z)$  for all  $z \in K$ , then there exists a Borel probability measure  $P$  with Stieltjes transform  $S_P = S$  if and only if (1) holds, in which case  $P_n \rightarrow P$  in distribution.*

**Proof.** It is enough to show that the pointwise convergence on  $K$  uniquely fixes the limiting functions in Lemma 3. Suppose that  $f$  and  $g$  are two limiting functions obtained from different subsequences of the family  $\{S_n\}$  and set  $h = f - g$ . Then  $h$  is analytic in  $D$  and zero on  $K$ . Thus  $h$  is identically zero on  $D$  [12, Theorem 10.18] which shows that all the limit functions of  $S_n$  are again equal to  $S$ . The rest of the proof follows as in Theorem 1.  $\square$

## 2. Applications

Theorem 1 can also be proved using the Grommer–Hamburger Theorem (e.g. [17, p. 105]), which states (in the probability measure context) that if the pointwise limit of Stieltjes transforms of a sequence of probability measures  $(P_n)$  exists, then the limit is always the Stieltjes transform of a sub-probability  $Q$ , and  $P_n$  converges vaguely to  $Q$ . There has been some confusion in the literature concerning the

statement and applications of this theorem (e.g. [5,16]), where uniform convergence on compact subsets was assumed to be sufficient to imply weak convergence. However, the Stieltjes transforms in Example 1 converge uniformly (to the zero function) on compacts, yet the measures do not converge weakly. As seen in Theorems 1 and 2, however, uniform convergence on compacts is automatic, but does not guarantee weak convergence. Convergence in distribution occurs if and only if the limit condition (1) holds, and it is this observation that simplifies the arguments in many applications, as is seen in the following three examples involving weak convergence to uniform, Cauchy, and arcsin distributions, respectively.

**Example 2.** Let  $\{L_n(x) : n = 1, 2, \dots\}$  be the unique polynomials of degree  $n$ , and positive leading coefficients, which satisfy the orthonormality relations

$$\int_{-1}^1 L_n(x) L_m(x) \frac{dx}{2} = \delta_{n,m}$$

(where  $\delta_{n,m}$  is the Kronecker delta function  $\delta_{n,m} = 1$  if  $n = m$ , and  $= 0$  otherwise), and let  $v_n$  be the discrete probability measure with masses equal to one-half (to normalize the weights) the Christoffel numbers, or the weights in the Gauss quadrature formula, located at the zeroes of  $L_n(x)$ ,  $n = 1, 2, \dots$  as given in [15, p. 48]. Using [15, Theorem 3.5.4], it follows that the Stieltjes transforms  $(S_n)$  of  $(v_n)$  satisfy

$$S_n(z) \rightarrow S(z) := -\frac{1}{2} \ln((z+1)/(z-1)) \quad \text{for } z \notin [-1, 1],$$

so since  $\lim_{y \rightarrow \infty} iyS(iy) = -1$ , it follows from Theorem 1 that  $v_n$  converges weakly to the probability measure with Stieltjes transform  $S$ , which is easily seen to be normalized Lebesgue measure  $dx/2$  on  $[-1, 1]$ . (This result is well known, cf. [15], and may also be derived using moment methods and Weierstrass approximation.)

**Example 3** (Van Assche [16, Theorem 4.17]). Let  $w : \mathbb{R} \rightarrow \mathbb{R}$  be an even measurable (weight) function which is positive a.s., and which satisfies  $\lim_{|x| \rightarrow \infty} \frac{\log w(x)}{|x|^\alpha} = -1$  for some  $\alpha > 1$ ; let  $\{x_{j,n} : j = 1, \dots, n; n = 1, 2, \dots\}$  be the zeroes of the orthogonal polynomials  $\{p_n(x) : n = 1, 2, \dots\}$  constructed via the weight function  $w$ ; and let  $\{\xi_n^\alpha : n = 1, 2, \dots\}$  be the normalized sequence of purely atomic probability measures with mass  $\frac{\alpha-1}{\alpha} \left(\frac{2}{\lambda_\alpha}\right)^{1/\alpha} \frac{n^{-1+1/\alpha}}{1+x_{j,n}^2}$  at  $x_{j,n}$ ,  $j = 1, \dots, n$ , where  $\lambda_\alpha = \frac{2}{\sqrt{\pi}} \frac{\Gamma((\alpha+1)/2)}{\Gamma(\alpha/2)}$ . As in [16], it can be shown that the Stieltjes transforms  $(S_n)$  of  $(\xi_n^\alpha)$  converge pointwise to  $S(z) = -(z+i)^{-1}$  for  $\text{Im}(z) > 0$ . Since  $S$  satisfies (1), Theorem 1 implies that  $\xi_n^\alpha$  converges weakly to a Borel probability measure  $P$ , which is easily seen (from  $S$ ) to be the standard Cauchy distribution with density  $\frac{1}{\pi} \frac{1}{1+x^2}$ ,  $-\infty < x < \infty$ . (Thus the argument of uniform convergence on compact sets used in [16] is not necessary.)

**Example 4** (Goh and Wimp [5, Theorem 1]; for a standard argument, see [10]). The Tricomi–Carlitz polynomials satisfy the following recurrence formula:

$$(n+1)f_{n+1}(x) - (n+\alpha)xf_n(x) + f_{n-1}(x) = 0,$$

with initial conditions  $f_0(x) = 1$  and  $f_1(x) = \alpha x$ . These polynomials satisfy the orthogonality relations:

$$\int_{-\infty}^{\infty} f_m(x) f_n(x) d\psi^\alpha(x) = \frac{2e^\alpha}{(n+\alpha)n!} \delta_{m,n},$$

where  $\psi^\alpha$  is a discrete mass measure with masses

$$\frac{(k+\alpha)^{k-1} e^{-k}}{k!} \quad \text{at } x = \pm(k+\alpha)^{-1/2}, \quad k = 0, 1, 2, \dots$$

Goh and Wimp [5] compute the Stieltjes transforms associated with the scaled Tricomi–Carlitz polynomials  $f_n(\frac{z}{\sqrt{n}})$ ,

$$S_n(z) = \frac{1}{n^{\frac{3}{2}}} \frac{f'_n(\frac{z}{\sqrt{n}})}{f_n(\frac{z}{\sqrt{n}})} = \int_{-\infty}^{\infty} \frac{dv_n(x)}{x-z},$$

where  $v_n$  is the discrete uniform probability measure with mass  $\frac{1}{n}$  at each zero of  $f_n(\frac{z}{\sqrt{n}})$ , and conclude, using the saddle point method, that

$$S_n(z) \rightarrow S(z) = -\frac{1}{z} + \frac{2}{z^3} (z_\beta(z) + \log(1 - z_\beta(z))),$$

for  $z \in N_\varepsilon$ , where:  $N_\varepsilon = \{z : |z - \frac{i}{2}| \leq \varepsilon\}$ ;  $\varepsilon$  is sufficiently small; and  $z_\beta(z) = z \frac{z - \sqrt{z^2 - 4}}{2}$ . (Here the branch of the square root used is such that  $\sqrt{z^2 - 4}$  is analytic for  $z \in \mathbb{C} \setminus [-2, 2]$  and positive for  $z > 2$ .) The function  $S(z)$  is the Stieltjes transform of the probability measure (e.g. [5,10, p. 194])

$$\frac{dv(x)}{dx} = \begin{cases} \frac{1}{\pi} \left( \frac{4 \arcsin(\frac{|x|}{2})}{|x|^3} - \frac{\sqrt{4-x^2}}{x^2} \right), & -2 < x < 2 \\ \frac{2}{|x|^3}, & |x| \geq 2. \end{cases}$$

By Theorem 2 it follows that  $v_n \rightarrow v$  weakly. (The argument of uniform convergence on compacts used in [5] is not needed.)

## Acknowledgments

The authors thank Professors Barry Simon and Walter Van Assche for several useful conversations and suggestions.

## References

- [1] N. Akhiezer, *The Classical Moment Problem*, Hafner, New York, 1965.
- [2] N. Akhiezer, M. Krein, *Some Questions in the Theory of Moments*, Translation of Mathematical Monographs, Vol. 2, Amer. Math. Soc., Providence, RI, 1962.
- [3] K. Chung, *A Course in Probability Theory*, 2nd Edition, Academic Press, San Diego, 1968.

- [4] A. Erdelyi (Ed.), Tables of integral transforms, Bateman Manuscript Project, Vol. 2, McGraw-Hill, New York, 1954.
- [5] W. Goh, J. Wimp, The zero distribution of the Tricomi–Carlitz polynomials, *Comput. Math. Appl.* 33 (1997) 119–127.
- [6] T. Hill, U. Krengel, Lévy-like continuity theorems for convergence in distribution, in: *Proceedings of the Göttingen Academy of Sciences*, 2001, *Nachr. Akad. Wiss. Göttingen* 1 (2002) 1–21.
- [7] T. Hill, M. Spruill, On the relationship between convergence in distribution and convergence of expected extremes, *Proc. Amer. Math. Soc.* 121 (1994) 1235–1243.
- [8] E. Hille, *Analytic Function Theory*, Vol. II, Chelsea, New York, 1959.
- [9] E. Kreyszig, *Advanced Engineering Mathematics*, Wiley, New York, 1962.
- [10] A. Kuijlaars, W. Van Assche, The asymptotic zero distribution of orthogonal polynomials with varying recurrence coefficients, *J. Approx. Theory* 99 (1999) 167–197.
- [11] B.M. Levitan, I.S. Sargsjan, *Sturm–Liouville and Dirac Operators*, Kluwer Academic, Boston, 1991.
- [12] W. Rudin, *Real and Complex Analysis*, Tata McGraw-Hill, New Dehli, 1974.
- [13] J. Shohat, J. Tamarkin, The problem of moments, *Amer. Math. Soc. Surveys* 1 (1950). Erratum 128 (2000) 625–626.
- [14] B. Simon, The classical moment problem as a self-adjoint finite difference operator, *Adv. in Math.* 137 (1998) 82–203.
- [15] G. Szegő, *Orthogonal polynomials*, American Mathematical Society Colloquium Publications, 4th Edition, Vol. 23, Amer. Math. Soc., Providence, RI, 1975.
- [16] W. Van Assche, *Asymptotics for Orthogonal Polynomials*, in: *Lecture Notes in Mathematics*, Vol. 1265, Springer, New York, 1980.
- [17] A. Wintner, *Spektraltheorie der Unendlichen Matrizen*, Hirzel, Leipzig, 1929.