ON THE REPRESENTATION OF A FUNCTION AS A HELLINGER INTEGRAL

RICHARD H. STARK

We derive in this note a necessary and sufficient condition that a nondecreasing, continuous function \( h \) of a single variable \( x \) be representable as a Hellinger integral of the form \( \int_0^x \frac{(df)^2}{dg} \). This condition was first proved by Hellinger in his dissertation \([1]\).\(^1\) Other proofs have been given by Hahn \([2]\) and Hobson \([3]\), who transform to Lebesgue integrals and make use of Lebesgue theory. Hellinger's proof and the less complicated proof given here have a certain simplicity in that they avoid reliance on these notions and even remain entirely within the range of monotone functions.

We consider nondecreasing functions of a real variable \( x \) on the interval \( 0 \leq x \leq 1 \) (henceforth denoted as \([0, 1]\)). For such a function \( f(x) \) and a closed interval \( \Delta \) with end points \( x_1 \) and \( x_2 \) \((#1 ^#2)\), we define a new function \( f_\Delta(x) \) to denote the length of the interval on the \( x \)-axis determined by the interval on the \( x \)-axis common to \( \Delta \) and \((0, x)\). More precisely, denoting

\[
\begin{align*}
  f(x \pm 0) &= \lim_{h \to 0} f(x \pm h) \quad \text{if } 0 < x < 1, \\
  f(0 - 0) &= f(0); \quad f(1 + 0) = f(1),
\end{align*}
\]

we define

\[
 f_\Delta(x) = \begin{cases}
  0 & \text{if } 0 \leq x < x_1, \\
  f(x + 0) - f(x_1 - 0) & \text{if } x_1 \leq x \leq x_2, \\
  f(x_2 + 0) & \text{if } x_2 < x \leq 1.
\end{cases}
\]

Received by the editors January 30, 1948.

\(^1\) Numbers in brackets refer to the references cited at the end of the paper.
The interval \([f(x_1 - 0), f(x_2 + 0)]\) on the \(f\)-axis will be denoted as \(f(\Delta)\). We retain the notation \(\Delta f\) for the proper difference \(f(x_2) - f(x_1)\).

Given sets \(\{E_i\} \subset [0, 1]\) for which the symbols \(f_{E_i}(x)\) and \(f(E_i)\) have been defined, we define

\[
\begin{align*}
& f_{x_{E_i}}(x) = \sum f_{E_i}(x) \quad \text{if } E_i \cap E_j = \emptyset \text{ when } i \neq j, \\
& f(\sum E_i) = \sum f(E_i) \quad \text{if } E_1 \subset E_2.
\end{align*}
\]

The sets which we shall consider will be constructed from sets of intervals in a manner such that (1) and (2) will assign to each such set \(E\) a function \(f_E(x)\) called the measure function of \(E\) on the \(f\)-axis and a set \(f(E)\) of points on the \(f\)-axis. Our \(f_E(1)\) is just the measure with respect to \(f(x)\) of the set \(E\) as defined in [4; p. 277]. It is immediate that if \(f_{E_D}(x)\) denotes the measure on the \(f_E(x)\) axis of the set \(D\) on the \(x\)-axis, then

\[
f_{E_D}(x) = f_{E\cdot D}(x).
\]

We now consider functions \(f(x), g(x), \) and \(h(x)\) which are nondecreasing and continuous in \([0, 1]\) and satisfy the inequality

\[
(\Delta f)^2 \leq \Delta g \Delta h \quad \text{on every subinterval } \Delta \text{ of } [0, 1].
\]

Let \(M\) be an arbitrary set on the \(x\)-axis. For arbitrary \(\epsilon > 0\), the set \(g(M)\) on the \(g\)-axis may be enclosed in a set \(I_\epsilon\) of countably many mutually disjoint intervals \(d_i\) such that the measures on the \(g\)-axis of the sets \(g(M)\) and \(I_\epsilon\) differ by less than \(\epsilon\); that is, the inverse function \(x(g)\) defines a corresponding set \(x(I_\epsilon)\) of the \(x\)-axis which we denote for brevity as \(X_\epsilon\) such that

\[
0 \leq g_{X_\epsilon}(1) - g_M(1) \leq \epsilon.
\]

Now for \(\Delta\) an arbitrary interval of the \(x\)-axis and \(\Delta_i\) the intersection of \(\Delta\) with \(x(d_i)\), we have:

\[
\Delta f_M \leq \Delta f_{x_\epsilon} = \sum_{i=1}^\infty \Delta_i f.
\]

But \((\Delta f)^2 \leq \Delta_i g \Delta_i h\) on every \(\Delta_i\) so that from the Cauchy-Schwarz inequality:

\[
\left( \sum_{i=1}^\infty \Delta_i f \right)^2 \leq \sum_{i=1}^\infty \Delta_i g \sum_{i=1}^\infty \Delta_i h \leq \Delta g_{X_\epsilon} \Delta h.
\]
With use of (5) and (6), this gives

\[(\Delta f_M)^2 \leq \Delta g_M \Delta h.\]

Replacing (4) by (8) and carrying out an analogous argument for a set \(N\) and the \(h\)-axis, we have

\[(\Delta f_{M,N})^2 \leq \Delta g_M \Delta h_N \quad \text{if (4) holds.}\]

**Theorem.** To a given pair of nondecreasing, continuous functions \(g(x)\) and \(h(x)\) for which \(g(0) = 0 = h(0)\) there corresponds an \(f(x)\) such that

\[h(x) = \int_0^x \left( \frac{df}{dg} \right)^2 \frac{dg}{1/2} \]

if and only if for every set \(E\) of the \(x\)-axis such that \(g_E(1) = g(1)\), we have \(h_E(1) = h(1)\).

For the proof, we use the following elementary properties of Hellinger integrals \([1]\) which hold for arbitrary functions \(g(x)\) and \(h(x)\) that are nondecreasing and continuous in \([0, 1]\):

a. Existence of \(t(x) = \int_0^x (du)^2/dg\) implies \(t(x) = \int_0^x (dgdt)^{1/2} / dg\).

b. \(f(x) = \int_0^x (dgdh)^{1/2}\) exists and \((\Delta f)^2 \leq \Delta g \Delta h\).

c. The inequality \((\Delta f)^2 \leq \Delta g \Delta h\) implies that \(s(x) = \int_0^x (df)^2 / dg\) exists and \((\Delta f)^2 / \Delta g \leq \Delta s \leq \Delta h\) if \(\Delta g \neq 0\).

It follows from (10a) that the desired representation of \(h(x)\) exists only if it is given by \(\int_0^x (du)^2 / dg\) where \(f(x) = \int_0^x (dgdh)^{1/2}\). By (10b), \((\Delta f)^2 \leq \Delta g \Delta h\). Let \(E\) be any set such that \(g_E(1) = g(1)\). Then from (8), we have on replacing \(M\) by the complement \(\overline{E}\) of \(E\) with respect to \([0, 1]\) and taking for \(\Delta\) the interval \([0, 1]\) that

\[(f_{\overline{E}}(1))^2 = g_{\overline{E}}(1) h(1) = 0\]

and consequently \(f_{\overline{E}}(1) = f(1)\). It follows from application of (11) to (9) that

\[(\Delta f)^2 \leq \Delta g \Delta h_{\overline{E}}.\]

Property (10c) with (12) gives that \(\Delta s \leq \Delta h_{\overline{E}}\) so that the function \(a(x) = h_{\overline{E}}(x) - s(x)\) is nondecreasing and if \(\Delta g \neq 0\)

\[0 \leq \Delta a = \Delta h_{\overline{E}} - \Delta s \leq \Delta h - (\Delta f)^2 / \Delta g\]

\[= ((\Delta g \Delta h)^{1/2} - \Delta f)((\Delta h / \Delta g)^{1/2} + \Delta f / \Delta g).\]

Consequently
We next choose a sequence of divisions $D_n$ of the interval (0, 1) into finitely many nonoverlapping intervals (that is, no two intervals of $D_n$ have more than an end point in common) such that each $D_n$ is formed from $D_{n-1}$ by addition of finitely many division points and

$$\sum_{D_n} ((\Delta g \Delta h)^{1/2} - \Delta f) \leq 4^{-n}. \tag{14}$$

That such a choice is possible is implied by (10a).

For each division $D_n$, we distinguish two types of intervals:

(a) The set $G_n$ of intervals such that $\Delta g \geq 4^{-n} \Delta h$;
(b) The set $S_n$ of intervals such that $\Delta g < 4^{-n} \Delta h$.

Then from (13), (14), (15), we have

$$a_{G_n}(1) = \sum_{G_n} \Delta a \leq 2^{-n+1}, \tag{16}$$

$$g_{S_n}(1) = \sum_{S_n} \Delta g \leq 4^{-n} \Delta h(1). \tag{17}$$

We define $R_n = \bigsqcup_{m=n}^\infty G_m$ and $R = \lim_{n \to \infty} R_n$. It is immediate that

$$a_R(1) = \lim_{n \to \infty} a_{R_n}(1) \leq \lim_{n \to \infty} a_{G_n}(1) = 0 \quad \text{(see (16))}, \tag{18}$$

and since $R \subset \sum_{m=n}^\infty S_m$,

$$g_R(1) \leq \lim_{n \to \infty} \sum_{m=n}^\infty g_{S_m}(1) = 0 \quad \text{(see (17))}. \tag{19}$$

Thus $R$ is a set of the type $E$ assumed in (13). Consequently, $0 \leq \Delta a \leq \Delta h_R$, and it follows from (9) and (3) that $a_R(1) \leq h_R \cdot \overline{\bar{a}}(1) = 0$. Hence

$$a_R(1) = 0. \tag{19}$$

The function $a(x)$ is nondecreasing with $a(0) = 0$ and by (18) and (19) has $a(1) = 0$. Thus $a(x) \equiv 0$, that is,

$$s(x) = h_B(x). \tag{20}$$

Therefore,

$$s(x) = \int_0^x \left( \frac{d}{dg} \int_0^x (dg \Delta h)^{1/2} \right)^2 \frac{dg}{dg}$$

is the measure function on the $h$-axis of a set $R$ of the $x$-axis. This
integral is actually equal to \( h(x) \) if and only if \( h_R(1) = h(1) \). The condition that \( g_E(1) = g(1) \) implies \( h_E(1) = h(1) \) is surely necessary to provide that \( h(x) = s(x) \), for \( g_E(1) = g(1) \) implies \( s_E(1) = s(1) \). On the other hand, if \( g_E(1) = g(1) \) does imply \( h_E(1) = h(1) \), the condition that \( h_R(1) = h(1) \) is fulfilled and \( h(x) = s(x) \). This completes the proof.

REFERENCES