A NOTE ON THE RIESZ REPRESENTATION THEOREM

DON H. TUCKER

1. Introduction. In 1909, F. Riesz [5] gave an integral representation for the bounded linear transformations $T$ from the space of real valued continuous functions on $[0, 1]$ into the real numbers, where the norm on the space is defined $\|f\| = \max_{0 \leq x \leq 1} |f(x)|$. An extensive bibliography for representation theorems is given in [1]. In 1961, S. E. James [2] generalized this result by considering continuous functions whose range of values was a subset of a Banach space $S$ and considered bounded linear transformations $T$ from this space into $S$. James’ result required that the transformation $T$ be such that there exist a functional $\overline{T}$ from the real valued continuous functions on $[0, 1]$ into the reals such that for each real valued continuous function $g$ on $[0, 1]$ and for each $h$ in $S$, $T[g(x)h] = \overline{T}[g] \cdot h$.

The purpose of this note is to extend James’ result in the following way: suppose $S_1$ is a linear normed space, $S_2$ is a Banach space, $C$ is the space of continuous functions from $[0, 1]$ into $S_1$ with norm defined $\|g\|_C = \int_0^1 \|g(x)\|_{S_1} dx$ and $B[S_1, S_2]$ is the space of continuous linear transformations from $S_1$ into $S_2$.

**Theorem 1.** If $T$ is a bounded linear transformation from $C$ into $S_2$, then there exists a function $A$ defined and of bounded variation on $[0, 1]$ with values in $B[S_1, S_2]$ such that, for each function $f$ in $C$, $T[f] = \int_0^1 dA(x) \cdot f(x)$.

2. Preliminary remarks. Continuity and bounded variation are considered as defined in the usual way with the appropriate norm used instead of absolute values. Since on the interval $[0, 1]$ the Heine-Borel theorem holds, each function in $C$ is bounded and uniformly continuous. Furthermore, if $f$ is in $C$, then $f_n(x) = \sum_{v=0}^{\infty} \left( \frac{\nu}{n} \right)^{x(1-x)^{n-v}} \cdot f(v/n)$ converges uniformly and hence in norm to $f$. The argument in [6, p. 152] with absolute values replaced by norms goes through.

The integral used here is of the type defined by MacNerney [4]. The appropriate change of norm for absolute value in the argument in [6, p. 31] gives the following form of the Helly-Bray theorem: if $\{K_n(x)\}_{n=0}^\infty$ is uniformly of bounded variation on $[0, 1]$ and $K_n(x) \rightarrow K(x)$ as $n \rightarrow \infty$, the values of $K_n$ being in $B[S_1, S_2]$, then if $f$ is in $C$.

Presented to the Society, January 22, 1962; received by the editors February 19, 1962.

354
\[
\lim_{n \to \infty} \int_0^1 dK_n(x) \cdot f(x) = \int_0^1 dK(x) \cdot f(x).
\]

3. Proof of Theorem 1.

**Lemma 1.** If \( G_n(x, t) = \sum_{v/n < x} \binom{n}{v} t^v (1-t)^{n-v} \) for \( 0 < x < 1, \ 0 \leq t \leq 1, \) and \( k_x(t) = 1 \) for \( 0 \leq t \leq x; k_x(t) = 0 \) for \( x < t \leq 1, \) then \( \int_0^1 |k_x(t) - G_n(x, t)| dt \to 0 \) as \( n \to \infty. \)

**Proof.**

\[
\sum_{v/n \approx x} \binom{n}{v} x^v (1-x)^{n-v} \leq \sum_{v/n \approx x} \binom{n}{v} x^v (1-x)^{n-v} \frac{(nx-v)^2}{n^2 \delta^2} \leq \frac{1}{n^2 \delta^2} \sum_{v=0}^{n} (nx-v)^2 \binom{n}{v} x^v (1-x)^{n-v} = \frac{nx(1-x)}{n^2 \delta^2} = \frac{x(1-x)}{n \delta^2} \leq \frac{1}{4 \delta^2 n}.
\]

Now consider \( G_n(x, t) = \sum_{v/n \approx x} \binom{n}{v} t^v (1-t)^{n-v} \) and take \( t > (x+\epsilon). \) Then

\[
\sum_{v/n \approx x} \binom{n}{v} \epsilon t^v (1-t)^{n-v} \leq \sum_{l-t/v \approx x} \binom{n}{v} \epsilon t^v (1-t)^{n-v} \leq \sum_{l-t/v \approx x} \binom{n}{v} \epsilon (1-t)^{n-v} \leq \frac{1}{4 \epsilon^2 n}.
\]

Now take \( \epsilon = n^{-1/4}; \) then \( G_n(x, t) < n^{-1/2}/4 \) for \( t > x + n^{-1/4}, \) and so \( G_n(x, t) \) converges uniformly to zero in every interval \( x < t_0 \leq t \leq 1. \) From symmetry (i.e., consider \( 1 - G_n(x, t) \)) \( G_n(x, t) \) converges uniformly to 1 in every interval \( 1 \leq t \leq t_0 < x. \) The result then follows. The basic thought of this lemma is well known in the theory of probability. See comment by Lorentz [3, p. 4].

We shall denote by \( C(R) \) the space of continuous real valued functions on \( [0, 1] \) with norm defined by \( \|f\|_{C(R)} = \int_0^1 |f(x)| dx. \) Suppose \( T \) is a bounded linear transformation from \( C \) into \( S_2. \)

**Lemma 2.** The transformation defined by \( B_g \cdot k = T[g(x) \cdot k] \) for \( g \) in \( C(R) \) and \( k \) in \( S_1 \) is, for fixed \( g, \) a bounded linear transformation from \( S_1 \) into \( S_2. \) Furthermore, \( B_g \) is a bounded linear transformation from \( C(R) \) into the Banach space \( B[S_1, S_2]. \) The latter statement holds whether we use the l.u.b. norm in \( C(R) \) or the norm defined above.
Proof. \( B\alpha + \beta B\beta \) = \( T[g(x) \cdot (\alpha k + \beta h)] = \alpha T[g(x) \cdot k] + \beta T[g(x) \cdot h] \)

and

\[
\| B \cdot k \|_s = \| T[g(x) \cdot k] \|_s \leq T \| [g(x) \cdot k] \|_c = T \int_0^1 |g(x) \cdot k| \, dx
\]

\[
= T \left[ \int_0^1 |g(x)| \, dx \cdot \| k \|_s \right] \leq \max |g| \cdot \| k \|_s.
\]

Hence, \( \| B \| \leq T \| g \|_c(R) \) whichever norm is used in \( C(R) \). Furthermore,

\[
(\alpha B\alpha + \beta B\beta) \cdot k = T[\alpha g(x) \cdot k] + T[\beta h(x) \cdot k] = T[\alpha g(x) \cdot k + \beta h(x) \cdot k]
\]

\[
= T[(\alpha g(x) + \beta h(x)) \cdot k] = B_{\alpha \alpha + \beta \beta} \cdot k.
\]

Hence \( B\alpha \) is a bounded linear transformation from \( C(R) \) into \( B[Si, S\beta] \). We shall hereafter refer to this transformation from \( C(R) \) into \( B[Si, S\beta] \) as \( 3 \).

Suppose \( f \) is in \( C \); then \( f_n(x) = \sum_{n=0}^{N} \left( \begin{array}{c} n \\ v \end{array} \right) x^v(1-x)^n \cdot v/n \) converges uniformly and in norm to \( f \), and therefore \( T[f_n] \) converges to \( T[f] \). Also,

\[
T[f_n(x)] = \sum_{n=0}^{N} T[\lambda_{n,v}(x) \cdot f\left( \frac{v}{n} \right)] = \sum_{n=0}^{N} B_{\lambda_{n,v}} \cdot f\left( \frac{v}{n} \right)
\]

where

\[
\lambda_{n,v}(x) = \left( \begin{array}{c} n \\ v \end{array} \right) x^v(1-x)^n.
\]

Hence we may write \( T[f_n] = \int_0^1 dK_n(x) \cdot f(x) \), where \( K_n(x) = \sum_{x=0}^{x=N} B_{\lambda_{n,v}} \), for \( 0 < x < 1 \); \( K_n(0) = N \), where \( N \) denotes the transformation which maps \( S_1 \) into the zero point of \( S_2 \), and \( K_n(1) = B_1 \). Hence, for \( 0 < x < 1 \), \( K_n(x) = 3 \left[ \sum_{x=n}^{x=N} \lambda_{n,v}(t) \right] \); \( \sum_{x=n}^{x=N} \lambda_{n,v}(t) = G_n(x, t) \) of Lemma 1; and for each \( x \) this sequence converges in norm to \( k_n(t) \) as \( n \to \infty \). Since 3 is a continuous transformation from \( C(R) \) into \( B[Si, S\beta] \) and \( B[Si, S\beta] \) is complete, \( K_n(x) \) converges for each \( x \).

\[
V_0K_n = \sum_{n=0}^{N} \| B_{\lambda_{n,v}} \| \leq \sum_{n=0}^{N} T \int_0^1 |\lambda_{n,v}(x)| \, dx = T \int_0^1 \lambda_{n,v}(x) \, dx
\]

\[
= T \int_0^1 \sum_{n=0}^{N} \lambda_{n,v}(x) \, dx = |T|
\]

since \( \lambda_{n,v}(x) \geq 0 \) for \( 0 \leq x \leq 1 \) and \( \sum_{n=0}^{N} \lambda_{n,v}(x) = 1 \). Therefore, \( \{K_n\} \)
are uniformly of bounded variation on \([0, 1]\) and for each \(x\) converge to some point \(K(x)\) in \(B[S_1, S_2]\), the function \(K\) being of total variation not more than \(|T|\) and then, by the Helly-Bray theorem, in §2, \(T[f] = \int_0^1 dK(x) \cdot f(x)\).

4. Some remarks on the space \(B[C, S_2]\). It is easily seen that for a given function \(K\) of bounded variation on \([0, 1]\) with values in \(B[S_1, S_2]\) the transformation \(T[f] = \int_0^1 dK(x) \cdot f(x)\) is a linear transformation from \(C\) into \(S_2\) which is continuous if the uniform norm is used above. (Let us assume that each \(K\) considered has been minimized in total variation by defining \(K(x) = \frac{1}{2}[K(x^-) + K(x^+)]\); \(0 < x < 1\). This will not affect the transformation \(T\) which it produces.) A natural question now would be, “For what functions \(K\) is the corresponding \(T\) continuous in the integral norm?” The answer is given by the following.

**Theorem 2.** In order that \(T[f] = \int_0^1 dK(x) \cdot f(x)\) should be continuous in the integral norm it is necessary and sufficient that \(K\) should satisfy a Lipschitz condition on \([0, 1]\). Furthermore, the norm of the transformation \(T\) is the g.l.b. of the Lipschitz constants for \(K\).

**Proof.** The sufficiency being easily seen only the necessity will be proved here.

First, suppose \(K\) is not continuous on \([0, 1]\). Since \(K\) is quasi-continuous and the total variation of \(K\) has been minimized, there exists a point \(p\); \(0 < p < 1\) (if \(p\) were 0 or 1 the argument need be only slightly changed) such that \(K(p^-) \neq K(p^+)\) and a sequence of intervals \([p_i, q_i]\) such that \(p_i \neq p\) and \(q_i \neq p\) and \(K\) is continuous at the points \(p_i\) and \(q_i\), \(i = 0, 1, \ldots\). Choose points \(k_i\) in \(S_1\) such that \(\|k_i\|_{S_1} = 1\) and

\[
\|K(p_i) - K(q_i)\|_{S_2} \leq \|K(p_i) - K(q_i)\| - \frac{1}{i}
\]

and define

\[
g_i(x) = \begin{cases} 0 & 0 \leq x \leq p_i, \quad q_i \leq x \leq 1 \\ k_i & \text{otherwise.} \end{cases}
\]

Then \(\|g_i(x)\| = \|k_i\| \cdot \|q_i - p_i\| \to 0\) as \(i \to \infty\) and furthermore \(\int_0^1 dK(x) \cdot g_i(x)\) exists. Now choose \(f_i \in C\) so that \(\|f_i - g_i\|_C < 1/i\) and hence
\[ \|f_i\|_c \to 0. \] Furthermore \[ \|f_0 dK \cdot [f_i - g_i]\|_s \to 0 \] but \[ \int_0^1 dK \cdot g_i = [K(q_0) - K(p_0)] \cdot k_i \] so that
\[
\left\| \int_0^1 dK \cdot g_i \right\|_s \geq \left\{ \|K(p_i) - K(q_i)\| - \frac{1}{i} \right\} \to \|K(p+) - K(p-)\| > 0,
\]
so that \[ \|f_0 dK \cdot f_i\|_s \to \|K(p+) - K(p-)\| > 0, \] but \[ \|f_i\|_c \to 0. \] Hence \( K \) is continuous.

Second, suppose \( K \) is not Lipschitz on \([0,1]\). Then there exists a sequence \([p_i, q_i]\) of subintervals of \([0,1]\), whose lengths converge to zero and such that \[ \|K(q_i) - K(p_i)\| > i(q_i - p_i). \] Define \( g_i(x) = 1/(q_i - p_i) \cdot 1/i \cdot k_i \) for \( p_i \leq x \leq q_i \) where \( k_i \) is a point in \( S_i \) of norm 1 for which \[ \|K(q_i) - K(p_i)\| > i(q_i - p_i) \] and \( g_i(x) = N_{S_i} \) elsewhere. \[ \|f_0 dK \cdot g_i\|_s > 1 \] and \[ \|g_i\| = 1/i. \] Approximate \( g_i \) with \( f_i \) in \( C \) as before and obtain a contradiction which establishes the first statement of the theorem. The final statement of the theorem then follows readily by a similar argument.

**Bibliography**


**University of Utah**