PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 126, Number 2, February 1998, Pages 553–560
S 0002-9939(98)04168-9

A REPRESENTATION THEOREM FOR SCHAUDER BASES IN HILBERT SPACE

STEPHANE JAFFARD AND ROBERT M. YOUNG

(Communicated by Dale Alspach)

ABSTRACT. A sequence of vectors $\{f_1, f_2, f_3, \dots\}$ in a separable Hilbert space H is said to be a Schauder basis for H if every element $f \in H$ has a unique norm-convergent expansion

$$f = \sum c_n f_n.$$

If, in addition, there exist positive constants A and B such that

$$A \sum |c_n|^2 \leq \left\| \sum c_n f_n \right\|^2 \leq B \sum |c_n|^2,$$

then we call $\{f_1, f_2, f_3, \dots\}$ a Riesz basis. In the first half of this paper, we show that every Schauder basis for H can be obtained from an orthonormal basis by means of a (possibly unbounded) one-to-one positive self adjoint operator. In the second half, we use this result to extend and clarify a remarkable theorem due to Duffin and Eachus characterizing the class of Riesz bases in Hilbert space.

1. The main theorem

A sequence of vectors $\{f_1, f_2, f_3, \dots\}$ in a separable Hilbert space H is said to be a Schauder basis for H if every element $f \in H$ has a unique norm-convergent expansion

$$f = \sum c_n f_n.$$

If, in addition, there exist positive constants A and B such that

$$A \sum |c_n|^2 \leq \left\| \sum c_n f_n \right\|^2 \leq B \sum |c_n|^2,$$

then we call $\{f_1, f_2, f_3, \dots\}$ a Riesz basis. Riesz bases have been studied intensively ever since Paley and Wiener first recognized the possibility of nonharmonic Fourier expansions

$$f = \sum c_n e^{i\lambda_n t}$$

for functions f in $L^2(-\pi, \pi)$ (see, for example, [2] and [5] and the references therein).

In the first half of this paper, we show that every Schauder basis for H can be obtained from some orthonormal basis by means of a (possibly unbounded) one-to-one positive self adjoint operator. In the second half, we use this result to extend

Received by the editors April 17, 1996 and, in revised form, August 22, 1996.
1991 *Mathematics Subject Classification.* Primary 46B15; Secondary 47A55.
Key words and phrases. Schauder basis, Riesz basis.

and clarify a remarkable theorem due to Duffin and Eachus [1] characterizing the class of Riesz bases in Hilbert space.

Theorem 1. *Let H be a separable Hilbert space and $\{f_n\}$ a Schauder basis for H. Then there exists an orthonormal basis $\{e_n\}$ and a one-to-one positive self adjoint transformation T of the Hilbert space such that $T\colon e_n \to f_n$ for $n = 1, 2, 3, \ldots$.*

Proof. Define a linear transformation S on H by setting

$$S(x) = \sum_{n=1}^{\infty} (x, f_n) f_n$$

for those points x in H for which the series converges *weakly.* If $\{f_n\}$ forms a conditional basis for the Hilbert space, then S may not be bounded and may not even be defined for all elements of the Hilbert space. Let us use the symbols $D(T)$ and $R(T)$ to designate the domain and range, respectively, of a linear transformation T.

We are going to show that S is self adjoint, positive, and one-to-one. In order to define the adjoint transformation S^*, we must first show that $D(S)$ is everywhere dense in H. Let $\{g_n\}$ be the system in H biorthogonal to $\{f_n\}$, so that $(f_n, g_m) = \delta_{nm}$. Then $S\colon g_n \to f_n$ $(n = 1, 2, 3, \ldots)$ and therefore every finite linear combination $c_1 g_1 + \cdots + c_n g_n$ also belongs to $D(S)$. Since the system biorthogonal to a Schauder basis is itself a Schauder basis (see [5], p. 29), the set of all such combinations forms a dense subspace of H, and the adjoint of S can be defined.

It follows at once from the definition that S is symmetric,

$$(Sx, y) = (x, Sy)$$

for every x and y in $D(S)$, and hence that $D(S^*) \supseteq D(S)$. Now the domain of S^* consists, by definition, of all y for which (Sx, y) is continuous for x in $D(S)$. Since $D(S)$ is dense, there is a uniquely determined point y^* in H such that $(Sx, y) = (x, y^*)$ for every x in $D(S)$. Accordingly,

$$\sum_{n=1}^{\infty} (x, f_n)(f_n, y) = (x, y^*)$$

whenever $x \in D(S)$ and $y \in D(S^*)$. It is crucial to show that the series on the left converges for *all* x in the Hilbert space. For this purpose, take $x = g_n$ in the expansion above, so that $(f_n, y) = (g_n, y^*)$ for $n = 1, 2, 3, \ldots$. Since $\{f_n\}$ is a Schauder basis, the biorthogonal expansion $y^* = \sum (y^*, g_n) f_n$ is valid in norm, and hence weakly. Thus, the series $\sum (y^*, g_n)(f_n, x)$ is convergent for every x in H and y in $D(S^*)$ and hence for these values of x and y,

$$\sum_{n=1}^{\infty} (y, f_n)(f_n, x)$$

is also convergent. This shows that $\sum (y, f_n) f_n$ converges weakly whenever $y \in D(S^*)$, and hence that $D(S^*) \subseteq D(S)$. Thus, $D(S) = D(S^*)$ and S is self adjoint.

To see that S is one-to-one, suppose that $S(x) = 0$, so that

$$\sum_{n=1}^{\infty} (x, f_n) f_n = 0 \quad \text{(weakly)}.$$

Taking the inner product of both sides with g_n $(n = 1, 2, 3, \dots)$, we see that $(x, f_n) = 0$ for $n = 1, 2, 3, \dots$, and, since the f_n are complete, x must be zero. Thus, S is one-to-one. That S is positive follows at once since

$$(Sx, x) = \sum_{n=1}^{\infty} |(x, f_n)|^2 \geq 0.$$

Now, the null space of an operator is the orthogonal complement of the range of its adjoint, and since S is one-to-one and self adjoint, its range is dense in H. Therefore, S^{-1} is also one-to-one, positive, and self adjoint. It follows that there exists a unique one-to-one, positive, self adjoint transformation A such that

$$A^2 = S^{-1}$$

(see [3], p. 265). It is to be shown that the system $\{e_n\}$ defined by

$$e_n = A(f_n) \qquad (n = 1, 2, 3, \dots)$$

forms an orthonormal basis for H.

That the e_n are well-defined is clear, since

$$f_n \in R(S) = D(S^{-1}) \subset D(A).$$

Orthonormality follows at once from

$$\begin{aligned}(e_n, e_m) &= (Af_n, Af_m) = (f_n, A^2 f_m)\\ &= (f_n, S^{-1} f_m) = (f_n, g_m) = \delta_{nm}.\end{aligned}$$

To establish completeness, let x be a fixed but arbitrary element in $D(S^{-1})$. Then $S^{-1}(x) \in R(S^{-1}) = D(S)$, and hence we can write

$$\begin{aligned}x = S(S^{-1}x) &= \sum_1^{\infty} (S^{-1}x, f_n) f_n\\ &= \sum_1^{\infty} (Ax, Af_n) f_n = \sum_1^{\infty} (Ax, e_n) f_n\end{aligned}$$

where each of the series is weakly convergent to x. If x_n denotes the nth partial sum of the last series, then

$$Ax_n = \sum_1^{n} (Ax, e_m) e_m \to P(Ax),$$

where P denotes the orthogonal projection onto the subspace generated by the e_n $(n = 1, 2, 3, \dots)$. Since every self adjoint transformation is closed, the relations

$$\begin{aligned}x_n &\to x \text{ weakly},\\ Ax_n &\to P(Ax)\end{aligned}$$

imply that $P(Ax) = Ax$. Thus, P reduces to the identity transformation on a dense subspace of H and hence on all of H. This shows that the system $\{e_n\}$ is complete and so forms an orthonormal basis for H. The required transformation T is obtained by taking $T = A^{-1}$. This completes the proof of the theorem. □

2. On a classic characterization of Riesz bases

Let us now turn our attention to the class of Riesz bases in a separable Hilbert space H. If $\{f_n\}$ is such a basis, then for *any* orthonormal basis $\{e_n\}$, the mapping

$$T\colon e_n \to f_n \qquad (n = 1, 2, 3, \dots)$$

can be extended to a bounded invertible operator on all of H. Riesz bases constitute the largest and most tractible class of bases known. It is extremely difficult to exhibit at least one bounded basis for a Hilbert space that is not equivalent to an orthonormal basis. An example in $L^2(-\pi, \pi)$ was discovered by Babenko (see [4], p. 351).

The simplest and perhaps the most obvious way of ensuring that a bounded linear transformation T is invertible is by showing that it is "close" to the identity transformation I in the sense that $\|I - T\| < 1$. This is a stringent requirement to place on a linear operator and one might well suppose that the class of Riesz bases is very small. A remarkable theorem due to Duffin and Eachus [1] shows, surprisingly, that just the opposite is true: *Every Riesz basis for H is obtained in essentially this way.*

Theorem 2. *Let H be a separable Hilbert space and $\{f_n\}$ a Riesz basis for H. Then there exists an orthonormal basis $\{e_n\}$ and a positive number ρ such that the mapping*

$$T\colon e_n \to f_n \qquad (n = 1, 2, 3, \dots)$$

can be extended to a bounded invertible operator on all of H such that

$$\|I - \rho T\| < 1.$$

Let us briefly review the construction of the orthonormal basis $\{e_n\}$.

> Start with any orthonormal basis $\{\varphi_n\}$ for H and let V be the unique bounded invertible operator on H that maps $\varphi_n \to f_n$. Factorize V by writing $V = PU$, where P is a positive operator and U is a unitary operator. Then $e_n = U(\varphi_n), n = 1, 2, 3, \dots,$ is the desired basis.

We are going to show that the basis $\{e_n\}$ so constructed is independent of the choice of $\{\varphi_n\}$ and that it is, in fact, the same basis guaranteed by Theorem 1.

We have

$$V(x) = \sum (x, \varphi_n) f_n$$

and

$$V^*(x) = \sum (x, f_n) \varphi_n,$$

and hence

$$\begin{aligned} VV^*(x) &= \sum_n \sum_m (x, f_m)(\varphi_m, \varphi_n) f_n \\ &= \sum_n (x, f_n) f_n = S(x), \end{aligned}$$

where S is the same operator used in the proof of Theorem 1. We shall continue to use the symbol A to designate the positive square root of S^{-1}. Since $\{f_n\}$ is a Riesz basis, both A and S are bounded invertible operators defined on all of H.

Using the canonical factorization $V = PU$, we find $V^* = U^*P^* = U^{-1}P$, and hence

$$S = VV^* = P^2.$$

Thus,

$$P^{-1} = S^{-1/2} = A$$

and

$$U = P^{-1}V = AV.$$

Since $V\colon \varphi_n \to f_n$, it follows that

$$U(\varphi_n) = A(f_n) \qquad (n = 1, 2, 3, \ldots).$$

Thus the two algorithms of Theorems 1 and 2 generate the same orthonormal basis.

The next theorem shows that this distinguished basis is "closest" to $\{f_n\}$ in a certain precise sense.

Theorem 3. *Let $\{f_n\}$ be a Riesz basis for a separable Hilbert space H and let $\{e_n\}$ be the distinguished orthonormal basis guaranteed by Theorems 1 and 2. If T denotes the operator on H for which*

$$T\colon e_n \to f_n \qquad (n = 1, 2, 3, \ldots),$$

then there exists a positive number ρ with the following property: If $\{h_n\}$ is any other orthonormal basis for H and G the operator for which

$$G\colon h_n \to f_n \qquad (n = 1, 2, 3, \ldots),$$

then

$$\|I - \lambda G\| \geq \|I - \rho T\|$$

for any positive real number λ.

Proof. Since $\{f_n\}$ is a Riesz basis, we know from Theorem 1 that the mapping

$$S(x) = \sum_{n=1}^{\infty} (x, f_n) f_n$$

defines a positive self adjoint invertible operator on all of H and that T is the positive square root of S. If $\alpha = 1/\|S^{-1}\|$ and $\beta = \|S\|$, then

$$\alpha I \leq S \leq \beta I.$$

Here, as usual, the order relation $A \leq B$ between symmetric operators A and B means that $(Ax, x) \leq (Bx, x)$ for all elements x of H. We define

$$\rho = \frac{2}{\sqrt{\alpha} + \sqrt{\beta}}.$$

Select an arbitrary orthonormal basis $\{h_n\}$ for H and let G be the operator for which $G(h_n) = f_n$ $(n = 1, 2, 3, \ldots)$. Then for any x in H, $x = \sum (x, h_n) h_n$ and hence

$$G(x) = \sum_{n=1}^{\infty} (x, h_n) f_n$$

and

$$G^*(x) = \sum_{n=1}^{\infty} (x, f_n) h_n,$$

so that

$$GG^* = \sum_n \sum_m (x, f_m)(h_m, h_n) f_n = \sum_n (x, f_n) f_n = S(x).$$

We can now calculate $\|G\|$ and $\|G^{-1}\|$. We have

$$\begin{aligned} \|G\|^2 = \|G^*\|^2 &= \sup_{\|x\|=1} (G^*x, G^*x) \\ &= \sup_{\|x\|=1} (GG^*x, x) = \sup_{\|x\|=1} (Sx, x) = \|S\|, \end{aligned}$$

and hence $\|G\| = \sqrt{\beta}$. A similar calculation applied to G^{-1} shows that $\|G^{-1}\| = 1/\sqrt{\alpha}$.

Let ε be an arbitrary positive number and select unit vectors x_1 and x_2 such that

$$\sqrt{\beta} - \varepsilon \le \|G(x_1)\| \le \sqrt{\beta}$$

and

$$\sqrt{\alpha} \le \|G(x_2)\| \le \sqrt{\alpha} + \varepsilon.$$

For any positive number λ, set $a = \|(I - \lambda G)(x_1)\|$ and $b = \|(I - \lambda G)(x_2)\|$. Then

$$a \ge |\|x_1\| - \|\lambda G(x_1)\|| \ge |1 - \lambda\sqrt{\beta}| - \lambda\varepsilon,$$

and

$$b \ge |\|x_2\| - \|\lambda G(x_2)\|| \ge |1 - \lambda\sqrt{\alpha}| - \lambda\varepsilon,$$

and hence

$$\|I - \lambda G\| \ge \max\{a, b\} \ge \max\{|1 - \lambda\sqrt{\alpha}|, |1 - \lambda\sqrt{\beta}|\} - 2\lambda\varepsilon.$$

By inspecting the graphs of $y = |1 - \lambda\sqrt{\alpha}|$ and $|1 - \lambda\sqrt{\beta}|$, we see that the smallest possible value of the second maximum occurs when

$$1 - \lambda\sqrt{\alpha} = \lambda\sqrt{\beta} - 1,$$

in other words, when

$$\lambda = \frac{2}{\sqrt{\alpha} + \sqrt{\beta}} = \rho.$$

Accordingly,

$$\|I - \lambda G\| \ge \|I - \rho G\| \ge \frac{\sqrt{\beta} - \sqrt{\alpha}}{\sqrt{\beta} + \sqrt{\alpha}} - 2\lambda\varepsilon,$$

and, since ε was arbitrary,

$$\|I - \lambda G\| \ge \frac{\sqrt{\beta} - \sqrt{\alpha}}{\sqrt{\beta} + \sqrt{\alpha}}$$

for any positive λ.

It remains only to show that this lower bound is in fact attained when $G = T$. Since

$$\alpha I \le S \le \beta I$$

and $T = S^{1/2}$, it follows that

$$\sqrt{\alpha} I \leq T \leq \sqrt{\beta} I.$$

Therefore,

$$-\left(\frac{\sqrt{\beta}-\sqrt{\alpha}}{\sqrt{\beta}+\sqrt{\alpha}}\right) I \leq I - \rho T \leq \left(\frac{\sqrt{\beta}-\sqrt{\alpha}}{\sqrt{\beta}+\sqrt{\alpha}}\right) I,$$

which in turn implies that

$$\|I - \rho T\| \leq \frac{\sqrt{\beta}-\sqrt{\alpha}}{\sqrt{\beta}+\sqrt{\alpha}}$$

(see [4], p. 262). This shows that $\|I - \rho T\| = (\sqrt{\beta}-\sqrt{\alpha})/(\sqrt{\beta}+\sqrt{\alpha})$ and the proof is complete. □

Remark. It is worth pointing out that an orthonormal basis $\{e_n\}$ with this minimum norm property is not unique. To see this, take $H = R^4$ and let $\{e_1, e_2, e_3, e_4\}$ be the standard orthonormal basis for H. Choose

$$f_1 = e_1/2, \quad f_2 = e_2, \quad f_3 = e_3, \quad f_4 = 3e_4/2.$$

Then the mapping $T\colon e_n \to f_n$ extends to a positive self adjoint operator on H. The minimum value of $\|I - \lambda T\|$ occurs when $\lambda = 1$ and

$$\|I - T\| = 1/2.$$

Consider now the orthonormal basis $\{h_1, h_2, h_3, h_4\}$ defined by

$$\begin{aligned} h_1 &= e_1, \\ h_2 &= (\cos\theta)e_2 + (\sin\theta)e_3, \\ h_3 &= (-\sin\theta)e_2 + (\cos\theta)e_3, \\ h_4 &= e_4 \end{aligned}$$

where θ is a small positive number. If $G\colon h_n \to f_n$, then the matrix of $I - G$ is

$$\begin{pmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 1-\cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & 1-\cos\theta & 0 \\ 0 & 0 & 0 & 1/2 \end{pmatrix}$$

and

$$\|I - G\| = 1/2$$

provided that θ is sufficiently small. This shows that $\{e_n\}$ is not the only orthonormal basis with the minimum norm property described in Theorem 3.

References

1. R. J. Duffin and J. J. Eachus, *Some notes on an expansion theorem of Paley and Wiener*, Bull. Amer. Math. Soc. **48** (1942), 850–855. MR **4:**97e
2. S. V. Khrushchev, N. K. Nikolskii and B. S. Pavlov, *Unconditional bases of exponentials and reproducing kernels*, in "*Complex Analysis and Spectral Theory*" (V. P. Havin and N. K. Nikolskii, editors), pp. 214–335, Lecture Notes in Mathematics, vol. 864, Springer-Verlag, Berlin/Heidelberg (1981). MR **84k:**46019
3. F. Riesz and B. Sz.-Nagy, *Functional Analysis*, Frederick Ungar Publ. Co., New York (1955). MR **17:**175i

4. I. Singer, *Bases in Banach spaces* I, Springer-Verlag, New York/Heidelberg (1970). MR **45:**7451
5. R. M. Young, *An Introduction to Nonharmonic Fourier Series*, Academic Press, New York (1980). MR **81m:**42027

Centre de Mathématiques et Leurs Applications, 61 Avenue du Président Wilson, 94235 Cachan cedex, France

Department of Mathematics, Oberlin College, Oberlin, Ohio 44074