On the Riesz-Fischer theorem

John Horváth
University of Maryland, College Park

To Ákos Császár on the occasion of his eightieth birthday
with admiration and friendship

Abstract

Are the two forms in which the theorem of the title is usually stated equivalent? We first summarize the three Comptes Rendus notes in which Frédéric Riesz published his results concerning L^2, and then, in somewhat more detail, an article from 1910 which has been published only in Hungarian. Riesz deduces the two forms not from each other but both from the Fréchet-Riesz representation theorem. A theorem states that some of Riesz's results hold in the case of an abstract inner product space, and leads to maximal orthonormal systems which are not total. We conclude with a proof due to Ákos Császár which shows that a variant of Riesz's condition implies the Fischer form (i.e., completeness).

1. According to folklore the two forms of the theorem in the title are the following:

Fischer: The normed space $L^2([a,b])$ is complete.

Riesz: Let (φ_k) be an orthonormal sequence in $L^2([a,b])$. Given a sequence (c_k) of scalars such that $\sum c_k^2 < \infty$, there exists an f in $L^2([a,b])$ for which

$$c_k = \int f\varphi_k.$$

It is also believed that the two statements are equivalent. Since both are true in $L^2([a,b])$, equivalence is meant in the sense that each follows from the other in a simple way. The true testing ground for the equivalence is an abstract inner product space, a concept which was not available in 1907. However, before discussing the abstract case let us see what Riesz had to say.

2. In 1907 Frédéric Riesz published three notes in the Comptes Rendus [C.R.] of the Paris Académie des Sciences.

1. *Sur les systèmes orthogonaux de fonctions*, C.R. **144** (1907), 615-619, listed under [C2] in the collected works of F. Riesz (Œuvres complètes = Gesammelte Arbeiten, Akadémiai Kiadó, Budapest, 1960), pp. 378-381. In this note

he proves his theorem first in the case when $a = 0, b = 2\pi$ and the orthonormal system is that of the trigonometric functions. If $\sum a_k^2 + b_k^2 < \infty$, then the formally integrated series

$$\sum \frac{1}{k}(a_k \sin kx - b_k \cos kx)$$

converges to a function F of bounded variation. By a theorem of H. Lebesgue (to which Riesz gave much later a famous proof, cf.the article of Michael W. Botsko, Amer. Math. Monthly **110** (2003), 834-838), the function F has almost everywhere a derivative $f \in L^2([0, 2\pi])$ which satisfies the condition $a_k = \int f(x) \cos kx dx, b_k = \int f(x) \sin kx dx$. The passage to an arbitrary orthonormal system is performed with the help of a theorem concerning a system of linear equations with infinitely many unknowns.

2. *Sur les systèmes orthogonaux de fonctions et l'équation de Fredholm*, C.R. **144** (1907), 734-736 (Œuvres [C3], pp. 382-385). This note is dated 8 avril 1907, while in his next note Riesz refers to it as "note publiée le 2 avril". This discrepancy might be explained by the manner in which the Comptes Rendus was organized in those days. A person living in Paris took his manuscript to the Académie des Sciences on Monday afternoon, when the weekly meetings take place, and handed it to one of the academicians. The next day he went to the editor Gauthier-Villars on the quai des Grands-Augustins to proofread it, and the issue with the note appeared on Thursday. The Café Mahieu, which still existed a few years ago on the corner of the Boulevard Saint-Michel and the rue Soufflot but has since been replaced by a fast food outlet, subscribed to the Comptes Rendus. When Frederick Riesz, from whom I know all these details, was in Paris, he visited the Café Mahieu on Thursdays and asked the waiter to bring the latest issue of the Comptes Rendus with his café au lait.

However, in April 1907 Frederick Riesz was in Göttingen as we shall see below. So the manuscript and the proofsheets were sent by mail, which explains why the note appeared the week after it was presented to the Académie des Sciences.

The note contains three results. The first is a Parseval formula from which it follows in particular that if $c_k = \int f\varphi_k$, then $\sum c_k^2 = \int f^2$. The second is an extension of the theorem of the last note to functions of several variables, and the third solves the Fredholm integral equation

$$\varphi(x) + \int_a^b K(x, y)\varphi(y)dy = f(x)$$

under the only condition that K and f are square integrable.

3. *Sur une espèce de Géométrie analytique des systèmes de fonctions sommables*, C.R. **144** (1907), 1409-1411 (Œuvres [C4], pp. 386-388). Riesz starts this short note by saying that on February 24 of "this year" (i.e., 1907) he gave a lecture at the Mathematical Society of Göttingen on his research concerning systems of

"summable" (i.e., Lebesgue integrable) functions. He published his main results in two notes in the Comptes Rendus (March 18 and "April 2"), and in a little different note in the Göttinger Nachrichten. His intention was to return to the subject only in a detailed memoir to appear in the Mathematische Annalen. However, the two notes of Monsieur Fischer (May 13 and 27) made him change his project.

Then follows a page long obscure philosophizing according to which Riesz is interested in creating an analytic geometry (i.e., introducing coordinates) in function space, while "Monsieur Fischer develops, in a very elegant manner, the synthetic theory".

At the end of the note Riesz states two "immediate consequences" of his theorem. The first is truly spectacular. It states that if U is a continuous linear form (operation) on L^2, then there exists $k \in L^2$ such that $U(f) = \int fk$ for all $f \in L^2$. This celebrated theorem was published in the very same issue of the Comptes Rendus by Maurice Fréchet and is now called (or should be called) the *Fréchet-Riesz theorem.* One appreciates the coincidence considering that the Comptes Rendus appears each week.

The second consequence gives an answer to a problem posed by Erhard Schmidt. Let (φ_k) be a sequence of continuous functions which are the indefinite integrals of square integrable functions ψ_k. Then every continuous function can be represented as the sum of a uniformly convergent series whose terms are linear combinations of 1 and of the φ_k if and only if there is no non-zero square integrable function which is orthogonal to each ψ_k.

The German note referred to above (*Über orthogonale Funktionensysteme*, Nachrichten von der Königl. Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-physikalische Klasse, 1907, 116-122; Œuvres [C5], pp. 389-395) was presented to the Society by David Hilbert on March 9. It contains the material of the first two Comptes Rendus notes with a few more indications concerning the proofs. In a footnote Riesz says that an exhaustive presentation ("ausführliche Darstellung") will appear in the Mathematische Annalen.

The "ausführliche Darstellung" never appeared in the Mathematische Annalen. It is true that three years later a large article by Friedrich Riesz was published in that journal (*Untersuchungen über Systeme integrierbarer Functionen*, Math. Ann. **69** (1910), 449-497; Œuvres [C10], pp. 441-489) but it is not about L^2. In this classical work Riesz introduces for $1 < p < \infty$ the space ("function class") L^p, proves that its dual is L^q with $\frac{1}{p} + \frac{1}{q} = 1$, and considers as its main result the solution of the "moment problem": given a sequence (φ_k) in L^p, under what condition on the sequence of scalars (c_k) does there exist an $f \in L^q$ such that $\int f\varphi_k = c_k$ for all k? In 1912 Eduard Helly gave a simple proof of Riesz's theorem and on that occasion proved essentially what became to be called the "Hahn-Banach theorem" (cf. H. Hochstadt, *Eduard Helly, father of the Hahn-Banach theorem*, Math. Intelligencer **2** (1980), 123-125).

The detailed presentation of the results concerning L^2 of Riesz and Fischer appeared in an article written in Hungarian and published the same year as the Annalen paper (*Integrálható függvények sorozatai*, Matematikai és Physikai Lapok **19** (1910), 165-182, 228-243; Œuvres [C9], pp. 407-440). It is accessible in the Collected Works but it has not been translated, possibly because its title ("Sequences of integrable functions") so closely resembles that of the Annalen paper. It is my understanding that there are people out there whose command of Hungarian is less than perfect, therefore I will give a succinct sketch of the article.

After some inequalities concerning the norm in $L^2 = L^2([a, b])$, Riesz proves Fatou's theorem: If (f_k) is a sequence of positive integrable functions such that $\int f_k \leq G$ for all k, and if f_k converges almost everywhere to f, then f is integrable and $\int f \leq G$. Riesz says that in his celebrated thesis (Acta Math. **30** (1906), 335-400; see p. 375) Fatou assumed unnecessarily that the f_k are bounded. Riesz then gives two applications of Fatou's theorem:

1) If the functions $g_k \in L^2$ converge almost everywhere to g, and $\int g_k^2 \leq G^2$ for all k, then $g \in L^2$, $\int g^2 \leq G^2$ and for any $h \in L^2$ one has $\int gh = \lim_{k\to\infty} \int g_k h$.

2) Let $f \in L^2$ and $a = x_0 < x_1 < \cdots < x_m = b$ be a subdivision of $[a, b]$. In the interval $x_k \leq x < x_{k+1}$ let the value of $\varphi_m(x)$ be the average

$$\frac{1}{x_{k+1} - x_k} \int_{x_k}^{x_{k+1}} f.$$

Then the φ_m converge almost everywhere to f as $\max(x_{k+1} - x_k) \to 0$, and $\lim_{m\to\infty} \int |f - \varphi_m|^2 = 0$.

Next Riesz proves that the function F on $[a, b]$ is the indefinite integral of some $f \in L^2$, i.e., $F(x) - F(a) = \int_a^x f$, if and only if there exists $G > 0$ such that for all subdivisions of $[a, b]$ one has

$$\sum_{k=0}^{m-1} \frac{|F(x_{k+1} - F(x_k)|^2}{x_{k+1} - x_k} \leq G^2.$$

The necessity is obvious, while the sufficiency follows from the fact that F is of bounded variation and so its derivative f exists almost everywhere. That $f \in L^2$ follows from 1) above.

And then follows the *Fréchet-Riesz theorem*: If A is a continuous linear map from L^2 into the field of scalars, then there exists an $a \in L^2$ such that $A[f] = \int af$ for all $f \in L^2$. *Proof.* Set $f(x; \xi) = 1$ for $x \leq \xi$, $f(x; \xi) = 0$ for $x > \xi$, and $A(\xi) = A[f(\cdot; \xi)]$. By the preceding theorem $A' = a \in L^2$ and a has the required property.

It is from this theorem that Riesz obtains his form of the Riesz-Fischer theorem. He first observes that an orthonormal system (a_i) in L^2 is always countable. This was proved by Erhard Schmidt when the a_i are continuous, and in the general case by Riesz himself in his first Comptes Rendus note ever (C.R. **143** (1906), 738-741; Œuvres [C1], pp. 375-377). Then he states his result the following way:

If $(a_i(x))$ is an orthonormal system of functions, then the system of equations

$$\int_a^b a_i(x)\xi(x)dx = c_i \quad (i = 1, 2, 3, \ldots)$$

has a solution belonging to the class L^2 if and only if the sum $\sum c_i^2$ exists. If this condition is satisfied, then there is an essentially unique solution function which is orthogonal to all functions that are orthogonal to all the functions $a_i(x)$. For this function $\xi(x)$ we have

$$\int_a^b [\xi(x)]^2 dx - \sum_i c_i^2 = \lim_{n\to\infty} \int_a^b [\xi(x) - \sum_{i=1}^n c_i a_i(x)]^2 dx = 0. \tag{1}$$

Corollary. If $(a_i(x))$ is an orthonormal system and $f(x) \in L^2$ is orthogonal to all functions that are orthogonal to every $a_i(x)$, then

$$\int_a^b (f(x))^2 dx = \sum_i [\int_a^b a_i(x) f(x) dx]^2.$$

This holds for *all* functions $f(x)$ in L^2 if there is no nonzero function in L^2 which is orthogonal to every $a_i(x)$. Such a system $(a_i(x))$ used to be called "complete". To avoid confusion with the sense in which this word is used in the next (Fischer's) theorem, I call such an orthonormal system *maximal.*

Proof of Riesz's Theorem. The condition is necessary by Bessel's inequality

$$\sum [\int_a^b a_i f]^2 \leq \int_a^b f^2.$$

Conversely, define the linear form Ξ on L^2 by

$$\Xi[f] = \sum c_i \int_a^b a_i f.$$

Using the Cauchy-Schwarz inequality one sees that Ξ is bounded (i.e., continuous) because of the hypothesis and Bessel's inequality. By the Fréchet-Riesz theorem there exists $\xi \in L^2$ such that

$$\int_a^b f\xi = \sum c_i \int_a^b a_i f$$

for all $f \in L^2$. Setting $f = a_k$ we get $\int_a^b a_k \xi = c_k$, and if $\int_a^b a_i f = 0$ for all i, then $\int_a^b f\xi = 0$. Substituting ξ for f we obtain (1).

At the end of the article Riesz gives three proofs for Fischer's theorem: L^2 *is complete.* For the sake of brevity, let me use the "modern" notation for the norm $\|f\| = (\int f^2)^{\frac{1}{2}}$.

The *first proof* uses the Fréchet-Riesz theorem. Let (f_k) be a Cauchy sequence in L^2, i.e., given $\epsilon > 0$ there exists $N(\epsilon)$ such that $\|f_k - f_l\| \leq \epsilon$ for

$k,l \geq N(\epsilon)$. Then $|\,\|f_k\| - \|f_l\|\,| \leq \|f_k - f_l\|$, so $(\|f_k\|)$ is a Cauchy sequence of scalars, hence it converges and in particular there exists $M > 0$ such that $\|f_k\| \leq M$ for all k.

Next, let $g \in L^2$. Then

$$|\int f_k g - \int f_l g\,| = |\int (f_k - f_l)g\,| \leq \|f_k - f_l\|.\|g\| \leq \epsilon\|g\|$$

for $k,l \geq N(\epsilon)$. So $(\int f_k g)$ is a Cauchy sequence and its limit $F[g]$ satisfies $|F[g]| \leq M\|g\|$. Clearly $g \mapsto F[g]$ is linear, hence by the above mentioned theorem there exists $f \in L^2$ such that $F[g] = \int fg$ for all $g \in L^2$. Furthermore

$$|\int (f - f_k)g\,| = |F[g] - \int f_k g| \leq \epsilon\|g\|$$

for $k \geq N(\epsilon)$ and all $g \in L^2$. Setting $g = f - f_k$ we get $\|f - f_k\|^2 \leq \epsilon\|f - f_k\|$, i.e., $\|f - f_k\| \leq \epsilon$ and therefore the sequence (f_k) converges to f.

The *second proof* is substantially different from the first one. To start, Riesz introduces weak convergence and proves an "existence theorem", what we now call a compactness theorem: If an infinite subset of L^2 is such that the square integrals (i.e., the norms) of the functions in the set are all $\leq G$, then there is a weakly convergent sequence in the subset. Next he proves that if (g_k) converges weakly to g and $\|g_k\| \leq G$ for all k, then $\|g\| \leq G$.

Let now (f_k) be a Cauchy sequence in L^2. Then the $\|f_k\|$ are bounded, so there exists a subsequence (f_{n_k}) which converges weakly to some $f \in L^2$. But then $f_{n_k} - f_m$ converges weakly to $f - f_m$. By the preceding result $\|f - f_m\| \leq \limsup_{n_k \to \infty} \|f_{n_k} - f_m\|$, hence $\lim_{m\to\infty} \|f - f_m\| \leq \lim_{k\to\infty, m\to\infty} \|f_{n_k} - f_m\| = 0$.

As Riesz observes in a footnote, his *third proof* works for any exponent $p > 1$ and was found independently also by Hermann Weyl (*Über die Konvergenz von Reihen die nach Orthogonalfunktionen fortschreiten*, Math. Ann. **67** (1909), 225-245; Gesammelte Abhandlungen, Springer, 1968, Band I.). It is based on the "Riesz selection principle": If (f_k) is a Cauchy sequence in L^p, there exists a subsequence which converges almost everywhere (to a function in L^p). This is the proof for the completeness which can now be found in most textbooks (e.g., F. Riesz - B. Sz.-Nagy: Leçons d'Analyse Fonctionnelle, §28).

3. Let us now examine what is left of the preceding results in the abstract situation. We denote by H an inner product space (also called a pre-Hilbert space), i.e., a vector space over the field $\mathbf{R}$ of real numbers with a map $(f,g) \mapsto (f|g)$ from $H \times H$ into the set $\mathbf{R}$ which satisfies the following conditions: $(f|g) = (g|f), (\alpha f_1 + \beta f_2|g) = \alpha(f_1|g) + \beta(f_2|g)$ for $\alpha, \beta \in \mathbf{R}, f_1, f_2, g \in H, (f|f) \geq 0$ and $(f|f) = 0$ only if $f = 0$. The norm of $f \in H$ is $\|f\| = \sqrt{(f|f)}$. For the sake of simplicity I consider only the real case as I have done so far; the passage to the complex case when $(f|g) = \overline{(g|f)}$ is straightforward. We denote by $\widehat{H}$ the completion of H, which with the inner product extended by continuity is a Hilbert space.

The equivalence of Fischer's theorem and of the Fréchet-Riesz theorem subsits also in the abstract situation:

Theorem. The following statements are equivalent:
(a) H is complete, i.e., $H = \widehat{H}$.
(b) For any closed subspace M of H, different from H, there exists a a nonzero $g \in H$ such that $(f|g) = 0$ for all $f \in M$ (i.e., such that g is orthogonal to M.)
(c) For any closed hyperplane M in H, there exists a nonzero $h \in H$ which is orthogonal to M.
(d) For any continuous ("bounded") linear form $L : H \to \mathbf{R}$ there exists $g \in H$ such that $L(f) = (f|g)$ for all $f \in H$.

Proof. The implication (a) $\Rightarrow$ (b) is of course well known, its proof is based on the important

F. Riesz Lemma. Let K be a complete convex subset of H and $f \in H$. There exists a unique $f_o \in K$ such that $\|f - f_o\| \leq \|f - g\|$ for all $g \in K$.

Quick proof. Assume f not in K. Then $\min \|g - f\| = \delta > 0$. Choose $g_n \in K$ so that $\|g_n - f\| \to \delta$. By the parallelogram law (g_n) is a Cauchy sequence and converges to the required f_o.

Now let $M \neq H$ and h not in M. Since M is convex and complete (because H is), there exists $h_o \in M$ such that $\|h - h_o\| \leq \|h - g\|$ for all $g \in M$. Then $h - h_o$ is orthogonal to M. Indeed, if there exists $f \in M$ such that $(f|h - h_o) \neq 0$, we may assume that $(f|h - h_o) > 0$, and then

$$\|h - h_o - \epsilon f\|^2 = \|h - h_o\|^2 - 2\epsilon(h - h_o|f) + \epsilon^2\|f\|^2,$$

which is $< \|h - h_o\|^2$ for small $\epsilon > 0$. Since $h_o + \epsilon f \in M$ this is a contradiction.

(c) is a special case of (b).

(c) $\Rightarrow$ (d) is again well known. If $L = 0$, take $g = 0$. If $L \neq 0$, then $M = \mathrm{Ker} L$ is a hyperplane, and it is closed because L is continuous. Let $g_o \neq 0$ be orthogonal to M and set $g = \frac{L(g_o)}{\|g_o\|^2} g_o$. Then $L(g) = \|g\|^2 = (g|g)$. Let h be arbitrary in H and set $L(h) = \alpha$. Then $L(h - \frac{\alpha}{\|g\|^2} g) = \alpha - \frac{\alpha}{\|g\|^2} L(g) = \alpha - \alpha = 0$, so $h = f + \frac{\alpha}{\|g\|^2} g$, where $f \in M$. We have $(h|g) = (f|g) + \frac{\alpha}{\|g\|^2}(g|g) = \alpha$, so $L(h) = (h|g)$.

We could use the argument of Riesz's first proof to show (d) $\Rightarrow$ (a), but it is simpler to avail ourselves of the existence of the completion $\widehat{H}$ of H. Assume that $H \neq \widehat{H}$ and chose $\xi \in \widehat{H}, \xi$ not in H. Then $f \mapsto (f|\xi)$ is a continuous linear form on H, so by assumption there exists $g \in H$ such that $(f|\xi) = (f|g)$ for all $f \in H$. Therefore $(f|\xi - g) = 0$ for all $f \in \widehat{H}$, hence $\xi = g \in H$, which is a contradiction.

The implication (d) $\Rightarrow$ (a) is not surprising. It expresses the well-known general fact that the dual of a normed vector space is complete (a "Banach" space). Of interest to us is the equivalence (a) $\Leftrightarrow$ (b), which makes it possible to have maximal orthogonal systems (φ_ι) in a non-complete inner product space

H such that the linear combinations of the φ_ι are not dense in H, i.e., such that (φ_ι) is not *total* in H. Indeed, to prove that in a Hilbert space $\widehat{H}$ a maximal orthonormal system is total, one considers the closure M of the space spanned by the linear combinations of the φ_ι. If $M \neq \widehat{H}$, then by (b) there exists $g \neq 0$ orthogonal to M, in particular to all the φ_ι, which contradicts maximality.

Jacques Dixmier (Acta Sci. Math. Szeged **15** (1954), 29-30) proved that if the Hilbert dimension of $\widehat{H}$ is sufficiently elevated, then there might exist in H maximal orthonormal systems that are not total. For the commodity of the reader I will reproduce part of his argument in a simple case.

Consider the two Hilbert spaces $H_1 = l^2(\mathbf{R})$ and $H_2 = l^2(\mathbf{N})$. Denote by $(f_\rho)_{\rho\in\mathbf{R}}$ the unit vector basis of H_1 and by $(e_n)_{n\in\mathbf{N}}$ the unit vector basis of H_2. By a theorem of Erdős and Kaplansky (G. Köthe, Topological Vector Spaces I, §9, 5.(3)) the *algebraic* dimension of H_2 has the power of the continuum, so there exists an algebraically free (i.e., linearly independent)family $(g_\rho)_{\rho\in\mathbf{R}}$ in H_2. Let $K = H_1 \oplus H_2$ be the orthogonal direct sum of the two spaces, and let H be the subspace of K generated *algebraically* by the family $(f_\rho + g_\rho)_{\rho\in\mathbf{R}}$ of elements.

We claim that every orthonormal family $(a_\iota)_{\iota\in I}$ in H is countable. Indeed, for any $n \in \mathbf{N}$ we have

$$\sum_\iota (a_\iota|e_n)^2 = \|e_n\|^2 = 1,$$

so $(a_\iota|e_n) \neq 0$ for only countably many indices $\iota \in I$. Since there are only countably many vectors e_n, only countably many vectors are not orthogonal to H_2. Let now a_ι be orthogonal to H_2. It can be written in the form

$$a_\iota = c_1 f_{\rho_1} + \cdots + c_n f_{\rho_n} + c_1 g_{\rho_1} + \cdots c_n g_{\rho_n}.$$

Both a_ι and $\sum_{j=1}^n c_j f_{\rho_j}$ are orthogonal to H_2, so

$$\|\sum_{j=1}^n c_j g_{\rho_j}\|^2 = (a_\iota|\sum_{j=1}^n c_j g_{\rho_j}) = 0,$$

i.e., $\sum_{j=1}^n c_j g_{\rho_j} = 0$. But the g_ρ are linearly independent, therefore $c_j = 0\,(1 \leq n)$ and finally $a_\iota = 0$. Thus the orthonormal family $(a_\iota)_{\iota\in I}$ contains only countably many vectors.

The orthogonal projection of H into H_1 contains all the vectors f_ρ, therefore it is dense in H_1. The Hilbert dimension of H_1 has the power of the continuum, hence that of $\widehat{H}$ is at least as large. Therefore $(a_\iota)_{\iota\in I}$ cannot be total in $\widehat{H}$, and neither in H.

4. We return to the question at the beginning. Again H is an inner product space, and we consider the condition

(A) H is complete, i.e., a Hilbert space.

The question we ask is whether (A) is equivalent to:

(B) If $(\varphi_\iota)_{\iota\in I}$ is an orthonormal system in H, and (c_ι) a family of scalars such that $\sum_{\iota\in I} c_\iota^2 < \infty$, then there exists $f \in H$ such that $c_\iota = (f|\varphi_\iota)$.

It is very easy to deduce (B) from (A). It is unlikely that the systems $(a_\iota)_{\iota\in I}$ in Dixmier's example satisfy (B), but that would show that (B) does not imply (A).

In a letter dated April 13, 2000 Ákos Császár deduced (A) from the following stronger condition:

(B+) If $(\varphi_\iota)_{\iota\in I}$ is an orthonormal system in H, and $(c_\iota)_{\iota\in I}$ a family of scalars such that $\sum c_\iota^2 < \infty$, then $\sum c_\iota\varphi_\iota$ is summable to an element $f \in H$ such that $c_\iota = (f|\varphi_\iota) \quad (\iota \in I)$.

. In fact, however, this result has little to do with the problem. Császár's proof shows that (A) follows already from the following conditions in which Riesz's main point, namely that the c_ι are the Fourier coefficients $(f|\varphi_\iota)$ of some $f \in H$, does not figure explicitly:

(B$*$) If $(\varphi_\iota)_{\iota\in I}$ is an orthonormal system in H and $(c_\iota)_{\iota\in I}$ a family of scalars such that $\sum c_\iota^2 < \infty$, then $\sum_{\iota\in I} c_\iota\varphi_\iota$ is summable in H.

Császár's proof of (B$*$) $\Rightarrow$ (A). Let (f_n) be a Cauchy sequence in H. With the Gram-Schmidt procedure we construct an orthonormal sequence (φ_n) in H. We take $\varphi_1 = \frac{f_1}{\|f_1\|}$ if $f_1 \neq 0$ and an arbitrary φ_1 if $f_1 = 0$. In either case $f_1 = \|f_1\|\varphi_1$.

Assume now that we have constructed the elements $\varphi_1, \dots, \varphi_n$ such that $(\varphi_i|\varphi_k) = \delta_{ik}$ and that each f_k $(1 \le k \le n)$ belongs to the span of $\varphi_1, \dots, \varphi_n$. Set $g = f_{n+1} - \sum_{k=1}^n (f_{n+1}|\varphi_k)\varphi_k$. Then

$$(g|\varphi_l) = (f_{n+1}|\varphi_l) - \sum_{k=1}^{n} (f_{n+1}|\varphi_k)(\varphi_k|\varphi_l) = (f_{n+1}|\varphi_l) - (f_{n+1}|\varphi_l) = 0$$

for $1 \le l \le n$. If $g \neq 0$ we set $\varphi_{n+1} = \frac{g}{\|g\|}$, if $g = 0$ then for φ_{n+1} we take any vector such that $(\varphi_k|\varphi_{n+1}) = 0$ for $1 \le k \le n$. If no such φ_{n+1} exists, then H is spanned by $\varphi_1, \dots, \varphi_n$ and therefore complete.

We have now $f_n = \sum_{k=1}^n (f_n|\varphi_k)\varphi_k$ for all $n \in \mathbf{N}$. Given $\epsilon > 0$, there exists $N(\epsilon) \in \mathbf{N}$ such that $\|f_n - f_m\| \le \epsilon$ for $n \ge m \ge N(\epsilon)$. By the orthogonality of the φ_k we have

$$\|f_n - f_m\|^2 = \|\sum_{k=1}^{n} (f_n|\varphi_k)\varphi_k - \sum_{k=1}^{m} (f_m|\varphi_k)\varphi_k\|^2 =$$

$$= \sum_{k=1}^{m} [(f_n|\varphi_k) - (f_m|\varphi_k)]^2 + \sum_{k=m+1}^{n} (f_n|\varphi_k)^2 \le \epsilon^2 \qquad (2)$$

for $n \geq m \geq N(\epsilon)$. In particular we have

$$|(f_n|\varphi_k) - (f_m|\varphi_k)| \leq \epsilon$$

whenever $n \geq m \geq \max(k, N(\epsilon))$. Thus for each k the scalars $(f_n|\varphi_k)\,(n \geq k)$ form a Cauchy sequence, so $\lim_{n\to\infty}(f_n|\varphi_k) = c_k$ exists. We have

$$\|f_n\| \leq \|f_n - f_{N(\epsilon)}\| + \|f_{N(\epsilon)}\| \leq \|f_{N(\epsilon)}\| + \epsilon$$

so $\|f_n\| \leq K$ for some $K > 0$ and all n. Therefore

$$\sum_{k=1}^{n}(f_n|\varphi_k)^2 = \|\sum_{k=1}^{n}(f_n|\varphi_k)\varphi_k\|^2 = \|f_n\|^2 \leq K^2$$

for all n. Fixing m we have

$$\sum_{k=1}^{m}(f_n|\varphi_k)^2 \leq K^2$$

for $n \geq m$, and letting $n \to \infty$ we obtain

$$\sum_{k=1}^{m} c_k^2 \leq K^2.$$

Since this holds for any m, we obtain $\sum c_k^2 < \infty$. According to hypothesis (B$*$) the series $\sum c_k\varphi_k$ is summable to some $f \in H$.

We claim that the sequence (f_n) converges to f. According to (2) we have

$$\sum_{k=1}^{m}[(f_n|\varphi_k) - (f_m|\varphi_k)]^2 \leq \epsilon^2$$

for $n \geq m \geq N(\epsilon)$. Letting $n \to \infty$ this yields $\sum_{k=1}^{m}[c_k - (f_m|\varphi_k)]^2 \leq \epsilon^2$, or equivalently

$$\|\sum_{k=1}^{m} c_k\varphi_k - \sum_{k=1}^{m}(f_m|\varphi_k)\varphi_k\| = \|\sum_{k=1}^{m} c_k\varphi_k - f_m\| \leq \epsilon$$

for $m \geq N(\epsilon)$. By the definition of f:

$$\|f - \sum_{k=1}^{m} c_k\varphi_k\| \leq \epsilon \qquad \text{for } m \geq N_2(\epsilon),$$

so finally

$$\|f - f_m\| \leq \|f - \sum_{k=1}^{m} c_k\varphi_k\| + \|\sum_{k=1}^{m} c_k\varphi_k - f_m\| \leq 2\epsilon$$

for $m \geq \max(N(\epsilon), N_2(\epsilon))$ which proves the claim and with it the theorem.

Of course $c_k = (f|\varphi_k)$ but this is not part of the hypothesis but a conclusion.