Consider the error in using an N-term sum with coeffs $c_n = (f_n, f)$

$$E_N = f - \sum_{n=1}^{N} c_n f_n$$

Write out $\|E_N\|^2$, expand as much as possible, and finally use $\|E_N\|^2 \geq 0$:

including the products of sums, and use formula for c_n, then simplify.

BONUS: Use your above simplest expression for $\|E_N\|^2$ to show, for any coeffs a_n:

$$\| f - \sum_{n=1}^{N} a_n f_n \|^2 - \|E_N\|^2 = \sum_{n=1}^{N} (a_n - c_n)^2 \geq 0$$

so for any $\{a_n\}$ the error cannot do better than $\{c_n\}$, these c_n are optimal.

Consider the error in using an N-term sum with coeffs $c_n = (f_n, f)$

$$E_N = f - \sum_{n=1}^{N} c_n f_n$$

including the products of sums, and use formula for c_n, then simplify.

Write out $\|E_N\|^2$, expand as much as possible, and finally use $\|E_N\|^2 \geqslant 0$:

$$\|E_N\|^2 = (E_N, E_N) = \left(f - \sum c_n f_n\,,\ f - \sum c_n f_n\right)$$

by linearity $(a, b+c) = (a,b) + (a,c)$

$$= (f,f) - 2\left(f, \sum c_n f_n\right) + \left(\sum c_n f_n\,, \sum c_n f_n\right)$$

$$= \|f\|^2 - 2\sum c_n \underbrace{(f, f_n)}_{= c_n} + \sum_n \sum_m c_n c_m \underbrace{(f_n, f_m)}_{\begin{cases} 1 & \text{if } m=n \\ 0 & \text{otherwise} \end{cases}}$$

$$= \|f\|^2 - 2\sum c_n^2 + \sum c_n^2$$

$$= \|f\|^2 - \sum_{n=1}^{N} c_n^2$$

$\geqslant 0$ since $\|\cdot\|^2$ anything is non-negative! so $\sum_{n=1}^{N} c_n^2 \leqslant \|f\|^2$ Bessel's Inequality QED.

BONUS: Use your above simplest expression for $\|E_N\|^2$ to show, for any coeffs a_n:

$$\left\| f - \sum_{n=1}^{N} a_n f_n \right\|^2 - \|E_N\|^2 = \sum_{n=1}^{N} (a_n - c_n)^2 \geqslant 0$$

so for any $\{a_n\}$ the error cannot do better than $\{c_n\}$, these c_n are optimal.

See book p. 212