§13.4 Implicit function theorem

Recap: Suppose $f: \mathbb{R}^2 \longrightarrow \mathbb{R}$. Then any subset $R = \{(x,y) \mid f(x,y) = 0\}$ gives a relation on $\mathbb{R}^2$, but it will only define a function $f: \mathbb{R} \longrightarrow \mathbb{R}$ if additionally we have: $\forall x \in p_1(R)$ $\exists$ a unique $y \in \mathbb{R}$ s.t $(x,y) \in R$. Here, $p_1: \mathbb{R}^2 \longrightarrow \mathbb{R}$ is projection onto the first factor. In this case, the set R consists of points of the form $(x, g(x))$ for some $g: \mathbb{R} \longrightarrow \mathbb{R}$.

The implicit function theorem answers the question of "when is R a function", but only locally. Ie it says that subject to certain conditions that must hold at (x_0, y_0) there is a neighbourhood U of (x_0, y_0) such that $R \cap U$ contains points of the form $(x, g(x))$ for some $g: \mathbb{R} \longrightarrow \mathbb{R}$.

In general, it is not about functions $\mathbb{R} \longrightarrow \mathbb{R}$ and subsets of $\mathbb{R}^2$, but subsets of $\mathbb{R}^{n+k}$ and functions

$$f: \mathbb{R}^{n+k} \longrightarrow \mathbb{R}^n$$ and sets written

$$f_r(\vec{x}_1, \ldots, \vec{x}_n, t_1, \ldots, t_k) = \vec{0}$$

for $r = 1, \ldots, n$, here f_r is a component function of f.

So we will use the notation $(\vec{x};\vec{t}) \in \mathbb{R}^{n+k}$ to denote points $(x_1,\ldots,x_n,t_1,\ldots,t_k)$ to streamline our proof.

Theorem: (Implicit function Theorem)
Let $f:\mathbb{R}^{n+k} \longrightarrow \mathbb{R}^n$ be defined on an open set $U \subseteq \mathbb{R}^n$. Suppose all the partials of f are continuous on U, and that $\exists$ a point $(\vec{x}_0;\vec{t}_0) \in U$ with $f(\vec{x}_0;\vec{t}_0)=0$ and $\det[D_j f_i(\vec{x}_0;\vec{t}_0)] \neq 0$ $(i,j=1,\ldots,n)$. Then there exists an open set $T_0 \subseteq \mathbb{R}^k$ containing t_0 and a unique function $g: T_0 \longrightarrow \mathbb{R}^n$ such that

(i) All partials of g are continuous on T_0,
(ii) $g(t_0) = x_0$,
(iii) $f(g(\vec{t});\vec{t}) = 0$ $\forall t \in T_0$.

Proof: We'll apply the inverse function theorem to a new function F in order to get our result.

Define F as follows. First,

$$F: S \longrightarrow \mathbb{R}^{n+k}$$

and we take the first n component functions of F to be the same as f, i.e. $F_m(\vec{x};\vec{t}) = f_m(\vec{x};t)$ for $1 \leq m \leq n$.

For the last k coordinate functions, we take the k^{th} coordinate function of F to be projection onto the k^{th} factor, i.e. $F_m(\vec{x};\vec{t}) = t_m$ for $1 \leq m \leq k$.

Now it is easy to check that

$$J_F(\vec{x};\vec{t}) = \det[D_j f_i(\vec{x};\vec{t})]$$

since the matrix defining $J_F(\vec{x};\vec{t})$ is a block matrix with a $k \times k$ identity block in the bottom right, and the matrix $[D_j f_i(\vec{x};\vec{t})]$ in the top left.

Therefore $J_F(\vec{x}_0;\vec{t}_0) \neq 0$, and note also that $F(\vec{x}_0;\vec{t}_0) = (\vec{0},\vec{t}_0)$.

Thus, by the inverse function theorem $\exists$ sets X, Y with $(\vec{x}_0;\vec{t}_0) \in X$ and $(\vec{0},\vec{t}_0) \in Y$ and a function $G: Y \longrightarrow X$ that serves as an inverse. Both F and G have continuous partials on X and Y respectively, and $F(X) = Y$, $G(Y) = X$.

The function G consists of component functions itself. Recall $G: \underset{\cup}{\mathbb{R}^{n+k}} \longrightarrow \underset{\cup}{\mathbb{R}^{n+k}}$, so write $G = (\vec{v},\vec{w})$ where
$\vec{y}$ $\vec{x}$

$v: Y \longrightarrow \mathbb{R}^n$ and $w: Y \longrightarrow \mathbb{R}^k$, such that $(v(\vec{x};\vec{t}), w(\vec{x};\vec{t})) \in X$ for all $(\vec{x};\vec{t}) \in Y$. Then since G is an inverse of F,

$$G(F(\vec{x};\vec{t})) = (\vec{x},\vec{t})$$

becomes

$$v(F(\vec{x};t)) = \vec{x} \quad \text{and} \quad w(F(\vec{x},\vec{t})) = \text{~~t~~}$$

But now observe that every point $(\vec{x},\vec{t})$ in Y can be written uniquely as $F(\vec{x}';\vec{t}')$ and that when we write $(\vec{x};\vec{t}) = F(\vec{x}';\vec{t}')$ we must actually have $\vec{t} = \vec{t}'$ by definition of F.

Thus $w(\vec{x};\vec{t}) = w(F(\vec{x}';\vec{t})) = \vec{t}$

and $v(\vec{x};\vec{t}) = v(F(\vec{x}';\vec{t})) = \vec{x}'$.

So $G: Y \longrightarrow X$ is the following function: Given $(\vec{x};\vec{t}) \in Y$, $G(\vec{x};\vec{t}) = (\vec{x}';\vec{t})$ where $\vec{x}'$ is the point in $\mathbb{R}^n$ such that $(\vec{x};\vec{t}) = F(\vec{x}';\vec{t})$. This means that

$$F(v(\vec{x};\vec{t});\vec{t}) = (\vec{x};\vec{t}) \text{ for every } (\vec{x},\vec{t}) \in Y.$$

So we set

$$T_0 = \{\vec{t} \mid \vec{t} \in \mathbb{R}^k \text{ and } (\vec{0};\vec{t}) \in Y\}$$

and define $g: T_0 \longrightarrow \mathbb{R}^n$ by $g(\vec{t}) = v(\vec{0};\vec{t})$.

Then T_0 is open in $\mathbb{R}^k \subseteq \mathbb{R}^{n+k}$, and the partials of g are continuous since the components of g are simply a subset of the components of G, and the partials of G are continuous on Y (and $T_0 \subset Y$).

Also $g(t_0) = v(\vec{0};t_0) = \vec{x}_0$, since $F(\vec{x}_0;t_0) = (\vec{0},\vec{t}_0)$.

Last, the equation $F(v(\vec{x};\vec{t});t) = (\vec{x},\vec{t})$ $\forall (\vec{x};\vec{t}) \in Y$ becomes

$$f(v(\vec{x};\vec{t});\vec{t}) = \vec{x}$$

by restricting to the first n components.

Setting $\vec{x}=0$, $\forall t \in T_0$ we get

$$f(g(\vec{t}); \vec{t}) = \vec{0}.$$

This proves all claims in the theorem, except for uniqueness of $g(\vec{t})$. This follows from f being one-to-one.

Example: Can the equation $(x^2+y^2+2z^2)^{1/2} = \cos(z)$ be solved uniquely for y in terms of x and z near $(0,1,0)$? For z in terms of x and y?

We apply the previous notation, where $\mathbb{R}^{n+k}$ comprises elements of the form $(\vec{x}; \vec{t})$: Here $n=1$, $k=2$; and $\vec{x} = y$, $\vec{t} = (x,z)$ (solving for y in terms of x, z).

Then f is

$$f(y; x, z) = \sqrt{x^2+y^2+2z^2} - \cos(z)$$

and observe that f has continuous partials everywhere except at $(0,0,0)$.

Now at $(x,y,z) = (0,1,0)$ (ie $(y;x,z) = (1;0,0)$)

$f(1;0,0) = \sqrt{0+1^2+0} - \cos 0 = 0$, and

$$\frac{\partial f}{\partial y} = \frac{1}{2\sqrt{x^2+y^2+2z^2}} \cdot 2y, \quad \text{so in our case}$$

$\det [D_j f_i(x_0; b_0)]$ is a 1×1 matrix:

$$\frac{\partial f}{\partial y}(1;0,0) = \frac{1}{1} \neq 0, \text{ so the implicit function}$$

theorem applies and y can be written as $g(x,z)$ there.

Example: Can you solve

$$x^2 - y^2 - u^3 + v^2 + 4 = 0$$
$$2xy + y^2 - 2u^2 + 3v^4 + 8 = 0$$

for u and v in terms of x and y in a nbhd of the solution $(x, y, u, v) = (2, -1, 2, 1)$?

Set $F(x, y, u, v) = (x^2 - y^2 - u^3 + v^2 + 4,\ 2xy + y^2 - 2u^2 + 3v^2 + 8)$, so that $F: \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ and $F(2, -1, 2, 1) = (0, 0)$.

Here, $\vec{x} = (u, v)$ and $\vec{t} = (x, y)$, so whether or not we can apply the IFT depends on $f(u, v; x, y) = F(x, y, u, v)$ and the det:

$$\det\left[D_j f_i(\vec{x}_0; \vec{t}_0)\right] \quad (j = 1, 2) = \det\begin{pmatrix} -3u^2 & 2v \\ -4u & 12v^3 \end{pmatrix}\Bigg|_{(2,-1,2,1)} = \begin{pmatrix} -12 & 2 \\ -8 & 12 \end{pmatrix} = -128$$

Since the det $\neq 0$, and since all partials are continuous on $\mathbb{R}^4$, there's a $g: \underset{T_0}{\underset{\cup}{\mathbb{R}^2}} \longrightarrow \mathbb{R}^2$ with

$t_0 = (2, -1, 2, 1) \in T_0$ such that $f(g(\vec{t}); \vec{t}) = 0 \ \forall t \in T_0$,

i.e. we've solved for u, v in terms of x, y as required.

MATH 3472

Fourier Series (Chapter 11)

From now on, our functions f will be single-variable, defined on an arbitrary subinterval $I \subseteq \mathbb{R}$. The interval I may be bounded/unbounded, open/closed, half-open.

Recall: $L^2(I)$ denotes the set of (possibly complex valued) functions f that are measurable on I and such that $|f|^2 \in L(I)$. Writing $f \in L(I)$ means that f is Lebesgue integrable on I.

If $f, g \in L^2(I)$ then their inner product is

$$(f, g) = \int_I f(x)\,\overline{g(x)}\,dx,$$ this always exists.

We write $\|f\|$ for

$$(f, f)^{1/2} = \left(\int_I f(x)\,\overline{f(x)}\,dx \right)^{1/2},$$

which again always exists. Thus we may define:

Definition 11.1: Let $S = \{\varphi_0, \varphi_1, \varphi_2, \dots\}$ and suppose $\varphi_i \in L^2(I)$ $\forall i$. If $(\varphi_m, \varphi_n) = 0$ whenever $n \neq m$, then S is said to be an orthogonal system on I. If also $(\varphi_n, \varphi_n) = 1$ $\forall i$ then S is orthonormal on I.

Remark: If $S=\{\varphi_0, \varphi_1, \dots\}$ is orthogonal but not orthonormal then $S' = \{\varphi_0/\|\varphi_0\|, \varphi_1/\|\varphi_1\| \dots\}$ is orthonormal.

Example: $S = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{\cos(x)}{\sqrt{\pi}}, \frac{\sin(2x)}{\sqrt{\pi}}, \dots \right\}$

ie. $\varphi_0 = \frac{1}{\sqrt{2\pi}}$, $\varphi_{2n-1} = \frac{\cos(nx)}{\sqrt{\pi}}$, $\varphi_{2n} = \frac{\sin(nx)}{\sqrt{\pi}}$

is an orthonormal system on $I=[0, 2\pi]$. The proof is simply to do the integrals:

$$\int_0^{2\pi} \left(\frac{1}{\sqrt{2\pi}}\right)^2 dx = \int_0^{2\pi} \frac{\cos^2 nx}{\pi} dx = \int_0^{2\pi} \frac{\sin^2 nx}{\pi} dx = 1$$

and

$$0 = \int_0^{2\pi} \frac{1}{\sqrt{2\pi}} \cos nx \, dx = \int_0^{2\pi} \frac{1}{\sqrt{2\pi}} \sin nx \, dx = \int_0^{2\pi} \frac{1}{\pi} \cos nx \sin mx \, dx$$

and

$$\int_0^{2\pi} \cos nx \cos kx \, dx = \int_0^{2\pi} \sin nx \sin kx \, dx = 0 \quad \forall k \neq n.$$

An easier way: consider the function $e^{inx} = \cos(nx) + i\sin(nx)$

Then note that $\int_0^{2\pi} \frac{1}{2\pi} e^{inx} e^{-imx} dx = \delta_{nm}$, since if $n \neq m$:

$$\frac{1}{2\pi} \int_0^{2\pi} e^{inx} e^{-imx} dx = \frac{1}{i(n-m)} e^{i(n-m)x} \Big|_0^{2\pi} = 0 \cdot (n \neq m)$$

Taking real and imaginary parts of this gives the desired relations.

Correspondingly, an orthonormal system on $[0, 2\pi]$ is also

$$\varphi_n(x) = \frac{e^{inx}}{\sqrt{2\pi}} = \frac{\cos nx + i \sin nx}{\sqrt{2\pi}}$$

if we allow complex numbers.

Remark: We can actually show the systems above are orthonormal on any interval of length 2π, by shifting and using periodicity of the functions.

The problem we'll address:

Let $f \in L^2(I)$ be given, and suppose that S is an orthonormal system ($S = \{\varphi_0, \varphi_1, \dots\}$). Let $t_n(x)$ denote the partial sum

$$t_n(x) = \sum_{k=0}^{n} b_k \varphi_k(x), \text{ where } b_k \in \mathbb{C}.$$

We want to choose $b_k \in \mathbb{C}$ so that the "error" $\|f - t_n(x)\|$ is as small as possible, in particular so that $\lim_{n\to\infty} \|f - t_n(x)\| = 0$. In this case our function $f(x)$ will hopefully be approximated by $\{t_n(x)\}$ in some sense, in particular we'd like something like

$$f(x_0) = \lim_{n\to\infty} t_n(x_0) \quad \forall\, x_0 \in \text{domain}(f).$$

(This is asking for quite a lot!)

In order to show how to choose b_k's in the previous formula, consider the case where the f we wish to approximate has an 'easy' form:

$$f(x) = \sum_{k=0}^{n}{}' c_k \varphi_k(x) \quad , \text{ some } c_k \in \mathbb{C}.$$

Then obviously $t_n(x) = f(x)$ will make $\|t_n(x) - f(x)\|$ as small as possible, but without knowing this choice ahead of time how can we construct $t_k(x)$ (ie choose c_k) so that $t_n(x) = f(x)$? Well for $0 \leq k \leq n$, we have

$$(f, \varphi_m) = \left(\sum_{k=0}^{n}{}' c_k \varphi_k, \varphi_m \right) = \sum_{k=0}^{n}{}' c_k (\varphi_k, \varphi_m) = c_m$$

by properties of inner product

since $(\varphi_k, \varphi_m) = 0$ if $k \neq m$ and $(\varphi_k, \varphi_m) = 1$ if $k = m$.

In other words, choosing $c_m = (f, \varphi_m)$ is forced on us by the special case $f(x) = \sum_{k=0}^{n}{}' c_k \varphi_k(x)$. Does it work in general?

<u>Theorem</u>: Let $\{\varphi_0, \varphi_1, \ldots\}$ be orthonormal on I, assume $f \in L^2(I)$. Define sequences of functions $\{s_n\}$ and $\{t_n\}$ by

$$s_n(x) = \sum_{k=0}^{n}{}' c_k \varphi_k(x) \quad \text{and} \quad t_n(x) = \sum_{k=0}^{n}{}' b_k \varphi_k(x),$$

where $c_k = (f, \varphi_k)$ and b_k are arbitrary. Then $s_n(x)$ are better, in the sense that

$$\|f - s_n\| \leq \|f - t_n\|$$

with equality iff $b_k = c_k$ $\forall k$.

Proof: We'll arrive at $\|f - s_n\| \leq \|f - t_n\|$ from the equation

$$\|f - t_n\|^2 = \|f\|^2 - \sum_{k=0}^{n} |c_k|^2 + \sum_{k=0}^{n} |b_k - c_k|^2.$$

This equation implies $\|f - s_n\| \leq \|f - t_n\|$ since it shows that $\|f - t_n\|$ is the smallest when $b_k = c_k$ for all k (since $|b_k - c_k|$ appears on the RHS). So we focus on proving this equation.

Write

$$\|f - t_n\|^2 = (f - t_n, f - t_n) = (f, f) - (f, t_n) - (t_n, f) + (t_n, t_n).$$

Now compute, using the properties of inner products:

$$(t_n, t_n) = \left(\sum_{k=0}^{n} b_k \varphi_k, \sum_{k=0}^{n} b_k \varphi_k\right)$$

$$= \sum_{k=0}^{n} \sum_{m=0}^{n} b_k \bar{b}_m (\varphi_k, \varphi_m) = \sum_{k=0}^{n} |b_k|^2 \quad \text{(orthonormality)}$$

and

$$(f, t_n) = \left(f, \sum_{k=0}^{n} b_k \varphi_k\right) = \sum_{k=0}^{n} \bar{b}_k (f, \varphi_k) = \sum_{k=0}^{n} \bar{b}_k c_k$$

$$(t_n, f) = \left(\sum_{k=0}^{n} b_k \varphi_k, f\right) = \sum_{k=0}^{n} b_k (\varphi_k, f) = \sum_{k=0}^{n} b_k \bar{c}_k, \text{ since}$$

$\overline{(\varphi_k, f)} = (f, \varphi_k)$. Plug in these quantities, and get:

$$\|f - t_n\|^2 = \|f\|^2 - \sum_{k=0}^{n} \bar{b}_k c_k - \sum_{k=0}^{n} \bar{c}_k b_k + \sum_{k=0}^{n} |b_k|^2$$

$$= \|f\|^2 - \sum_{k=0}^{n} |c_k|^2 + \sum_{k=0}^{n} (b_k - c_k)(\overline{b_k} - \overline{c_k})$$

$$= \|f\|^2 - \sum_{k=0}^{n} |c_k|^2 + \sum_{k=0}^{n} |b_k - c_k|^2, \text{ so the theorem follows.}$$

§11.4 MATH 3472

Definition: Given $f \in L^2(I)$, and $S = \{\varphi_0, \varphi_1, \dots\}$ an orthonormal system, we write

$$f(x) \sim \sum_{k=0}^{\infty} c_k \varphi_k(x)$$

when the coefficients c_k are given by

$$c_k = \int_I f(x)\,\overline{\varphi_k(x)}\,dx = (f, \varphi_k).$$

This is called the Fourier series of f on I relative to S, the coefficients c_k are the Fourier coefficients.

Remark: We use the symbol "$\sim$" above when relating a function $f(x)$ to its Fourier series, because we are not making any claims about the convergence of $\sum_{k=0}^{\infty} c_k \varphi_k$ at this time.

Remark: If $I = [0, 2\pi]$ and S is the set

$$\varphi_0 = \frac{1}{\sqrt{2\pi}}, \quad \varphi_{2n-1} = \frac{\cos(nx)}{\sqrt{\pi}}, \quad \varphi_{2n} = \frac{\sin(nx)}{\sqrt{\pi}}$$

Then

$$f \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos(nx) + b_n \sin(nx)\right)$$

is how the Fourier series is commonly written. It is sometimes called simply "the Fourier series of f".

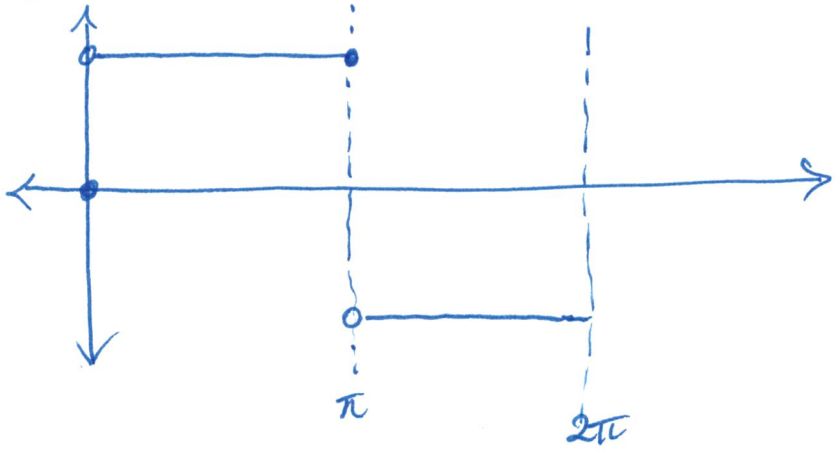
Example: Suppose

$$H(x) = \begin{cases} 0 & \text{if } x<0 \\ \frac{1}{2} & \text{if } x=0 \\ 1 & \text{if } x>0 \end{cases}$$

and consider

$$f(x) = 2\left(H\left(\frac{x}{2\pi}\right) - H\left(\frac{x}{2\pi} - 1\right)\right) - 1$$

whose graph over $[0, 2\pi]$ is



Now to compute the a_n's and b_n's, note that $f(x) = f(-x)$ so the function is odd on its domain $[0, 2\pi]$.

Thus

$$a_n = \frac{1}{2\pi}\int_0^{2\pi} f(x)\cos(nx)\,dx \quad \text{for } n \geq 0$$

$= 0$, since $\cos(nx)$ is even on $[0, 2\pi]$,

ie

$$\int_0^{2\pi} f(x)\cos(nx)\,dx = \int_0^{\pi} f(x)\cos(nx)\,dx + \int_{\pi}^{2\pi} f(x)\cos(nx)\,dx$$

$$= \int_0^{\pi} f(x)\cos(nx)\,dx - \int_0^{\pi} f(x)\cos(nx)\,dx$$

(using $f(\pi + x) = -f(x)$ for $x \in [0, \pi)$)

$= 0$.

On the other hand

$$b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(nx)\,dx$$

$$= \frac{2}{\pi}\int_0^{2\pi} f(x)\sin(nx)\,dx \quad \left(\text{since } \sin(nx) \text{ is odd on } [0, 2\pi]\right)$$

$$= \frac{2}{\pi}\left[-\cos(nx)\cdot\frac{1}{n}\right]_0^{\pi}$$

$$= \frac{2}{n\pi}\left((-1)^n - 1\right), \text{ therefore we get}$$

$$f(x) \sim \frac{4}{\pi}\sum_{k=0}^{\infty}{}' \frac{1}{2k+1}\sin((2k+1)x).$$

---

§11.5 Properties of Fourier Coefficients

The purpose of the inequalities that follow will be to aid in the investigation of convergence properties of Fourier series.

Theorem 11.5: Suppose $\{\varphi_0, \varphi_1, \dots\}$ are orthonormal on I, and assume $f \in L^2(I)$, and suppose that

$$f(x) \sim \sum_{n=0}^{\infty}{}' c_n \varphi_n(x).$$

Then

a) The series $\sum |c_n|^2$ converges and satisfies

$$\sum_{n=0}^{\infty}{}' |c_n|^2 \le \|f\|^2 \quad \text{(Bessel's inequality)}$$

b) $\sum_{n=0}^{\infty} |c_n|^2 = \|f^*\|^2$ (Parseval's formula)

holds iff $\lim_{n\to\infty} \|f - s_n\| = 0$, where $s_n(x) = \sum_{k=0}^{n} c_k \varphi_k(x)$.

Proof: We already established that when

$$t_n(x) = \sum_{k=0}^{n} b_k \varphi_k(x), \quad \text{where } b_k \in \mathbb{C} \text{ are arbitrary,}$$

then

$$\|f - t_n\|^2 = \|f\|^2 - \sum_{k=0}^{n} |c_k|^2 + \sum_{k=0}^{n} |b_k - c_k|^2$$

where $c_k = (f, \varphi_k)$. In particular if we take $b_k = c_k$ then we get

$$\|f\|^2 = \sum_{k=0}^{n} |c_k|^2 + \|f - t_n\|^2 \quad \text{for all } n.$$

This establishes $\sum_{n=0}^{\infty} |c_n|^2 \le \|f\|^2$, since $\|f - t_n\|^2$ is always positive. In fact, from

$$\|f\|^2 = \sum_{k=0}^{n} |c_k|^2 + \|f - s_n\|^2$$

we also see that

$$\|f\|^2 = \sum_{k=0}^{\infty} |c_k|^2 \quad \text{iff} \quad \lim_{n\to\infty} \|f - s_n\|^2 = 0.$$