CAPÍTOL 2

TEORIA DE STURM-LIOUVILLE

2.1 Introducció

Abans d'entrar en aquesta teoria recordem algunes de les propietats importants dels vectors en un espai euclidià de dimensió finita, que seran força útils a l'hora de fer l'extensió necessària als espais de funcions, que com veurem és un espai vectorial de dimensió infinita.

Recordem, per exemple, que la dimensió n d'un espai vectorial és el nombre de vectors linealment independents que pertanyen a l'espai. Considerem un conjunt (arbitrari) de vectors $\{u_j\}$ amb $j = 1, \ldots, n$ unitaris i ortogonals dos a dos. Aleshores qualsevol vector $\vec{A}$ es podrà expressar com

$$\vec{A} = a_1 u_1 + a_2 u_2 + \ldots + a_n u_n$$

i la forma de calcular els coeficients a_j serà tenint en compte les propietats d'ortonormalitat (ortogonalitat i unitarietat) del conjunt de vectors. Multipliquem escalarment per u_k

$$(\vec{A}, u_k) = a_1(u_1, u_k) + a_2(u_2, u_k) + \ldots + a_n(u_n, u_k)$$

i de tots aquests productes escalars només n'hi ha un que sobreviu que és el que ens dóna a_k; així doncs

$$a_k = (\vec{A}, u_k).$$

Recordem també que el concepte de producte escalar ens porta al de mòdul o norma d'un vector

$$\|\vec{A}\| = \sqrt{(\vec{A}, \vec{A})} = \sqrt{a_1^2 + \ldots + a_n^2}$$

i a partir d'aquí al de distància entre dos elements de l'espai vectorial com la norma del vector diferència.

Les equacions diferencials a les quals s'arriba després d'haver aplicat el mètode de separació de variables es poden expressar de tal manera que ens recorden el procés de diagonalització d'una matriu. Prenem, per exemple, l'equació diferencial

$$\Phi'' = M\Phi.$$

En aquesta equació podem interpretar l'operador que deriva dues vegades com una matriu actuant sobre la funció (vector) Φ i el seu resultat és un escalar que multiplica a la pròpia funció (vector); aleshores, trobar els valors de M que permetin resoldre l'equació diferencial serà anàleg a trobar els valors propis d'una matriu.

2.2 Espais de funcions

Denotarem per $\mathcal{C}^n([a,b])$ l'espai de funcions d'una variable real, $a \le x \le b$, a valors complexos, $f(x) \in \mathbb{C}$ que admeten derivades fins a l'ordre n en tot l'interval, les quals són contínues.

Com que si sumem dues funcions d'aquestes el resultat també té la mateixa propietat, i si multipliquem una d'aquestes funcions per una constant, $\lambda \in \mathbb{C}$, el resultat també és de $\mathcal{C}^n([a,b])$, aquest espai de funcions té estructura d'espai vectorial.

Resulta obvi de la definició que si $n < m$, llavors $\mathcal{C}^m([a,b]) \subset \mathcal{C}^n([a,b])$. També és fàcil veure que aquests espais tenen dimensió infinita. Efectivament, la infinitat de monomis elementals 1, x, x^2, ..., x^l, ..., pertanyen a qualsevol espai $\mathcal{C}^n([a,b])$ i al mateix temps són linealment independents, és a dir, no hi ha cap combinació lineal $k_0+k_1x+k_2x^2+...+k_lx^l = 0$, excepte la que té tots els $k_i = 0$.

Donades dues funcions $f, g \in \mathcal{C}^n([a,b])$ i una funció real ω contínua en $[a,b]$ i positiva (excepte potser en un nombre finit de punts), podem definir el producte de Hilbert (amb *pes* ω):[1]

$$(f,g) \equiv \int_a^b f(x)^* g(x)\omega(x)dx \tag{2.1}$$

el qual té les propietats següents:

- És hermític: $(f,g) = (g,f)^*$.
- És distributiu respecte de la suma:
  $(f_1+f_2, g) = (f_1,g) + (f_2,g)$ i $(f, g_1+g_2) = (f,g_1) + (f,g_2)$,

[1] Fixem-nos que aquest producte és l'equivalent del producte escalar (intern) en un espai vectorial de dimensió finita i la funció ω té el paper de la mètrica.

- Mentre que per al producte per constants, $k \in \mathbb{C}$:
  $(kf, g) = k^*(f, g)$ i $(f, kg) = k(f, g)$

  Aquesta última propietat s'anomena *sesquilinealitat*.

- És *definit positiu*, és a dir, si $g \neq 0$, llavors $(g, g) > 0$.
  Efectivament,
  $(g, g) = \int_a^b g^*(x)g(x)\omega(x)dx = \int_a^b |g(x)|^2\omega(x)dx > 0$

- És *no degenerat*. Si $(f, g) = 0$, $\forall g$, llavors $f = 0$.
  És conseqüència de l'anterior. Veiem que si $(f, g) = 0$, $\forall g$, tindrem en particular que $(f, f) = 0$, que per la propietat anterior implica $f = 0$.

Amb aquest producte escalar definirem la **norma** d'una funció per:

$$\|f\| \equiv \sqrt{(f, f)} = \left(\int_a^b dx\, |f(x)|^2\omega(x)\right)^{1/2}.$$

Utilitzant ara el concepte de norma es pot definir la **distància** entre dues funcions f i g

$$d(f, g) = \|f - g\| = \left[\int_a^b dx|f(x) - g(x)|^2\omega(x)\right]^{1/2}.$$

Ambdues definicions satisfan les propietats habituals que defineixen una norma i una distància en un espai vectorial de dimensió finita.

A continuació veurem quina relació tenen aquestes definicions amb la resolució d'equacions diferencials de segon ordre amb condicions de contorn. De fet, aquest procés de resolució el veurem com una diagonalització d'operadors diferencials que han de satisfer, a més, unes certes condicions de contorn.

2.3 Operadors diferencials lineals de segon ordre

Donada una equació diferencial de segon ordre

$$\mathcal{L}[y] = a_0(x)y'' + a_1(x)y' + a_2(x)y = 0,$$

sent $y(x)$ una funció complexa de variable real i a_0, a_1, a_2 funcions reals de variable real, s'anomena **operador adjunt** l'operador $\mathcal{L}^\dagger$ que actua segons

$$\mathcal{L}^\dagger[y] = (a_0(x)y)'' - (a_1(x)y)' + a_2(x)y = a_0y'' + (2a_0' - a_1)y' + (a_0'' - a_1' + a_2)y$$

L'operador s'anomena **autoadjunt** si $\mathcal{L} = \mathcal{L}^\dagger$, és a dir

$$a_0(x)y'' + a_1(x)y' + a_2(x)y = a_0 y'' + (2a_0' - a_1)y' + (a_0'' - a_1' + a_2)y,$$

per tant, l'operador és autoadjunt si i només si $a_0' = a_1$. Aleshores l'equació diferencial es pot expressar de la següent manera

$$\mathcal{L}[y] = (a_0(x)y')' + a_2(x)y.$$

Aquesta expressió s'anomena **forma canònica** de l'operador diferencial.

Encara que un operador diferencial de segon ordre no sigui autoadjunt sempre n'hi ha un d'equivalent que sí que ho és[2]

$$\omega(x)\mathcal{L}[y] = \omega(x)a_0(x)y'' + \omega(x)a_1(x)y' + \omega(x)a_2(x)y = \frac{d}{dx}\left[p(x)\frac{d}{dx}\right] + q(x)w(x)y(x).$$

Podem determinar els $\{p, q, w\}$ en funció dels $\{a_0, a_1, a_2\}$.

$$p(x) = \exp\left[\int^x \frac{a_1(t)}{a_0(t)}dt\right] \qquad q(x) = a_2(x) \qquad w(x) = \frac{p(x)}{a_0(x)},$$

que són funcions reals perquè el conjunt de les a's del qual provenen també ho és.

Definim ara l'operador diferencial en la seva forma canònica

$$\mathcal{L} \equiv \frac{1}{w(x)}\frac{d}{dx}\left(p(x)\frac{d}{dx}\right) + q(x).$$

A partir d'ara suposarem que l'operador diferencial pren la forma canònica. És la funció $w(x)$ la que permet que tot operador es pugui expressar en la forma canònica. Ja veurem més endavant quin és el paper que té aquesta funció.

2.3.1 Condicions de contorn

Per resoldre una equació diferencial de segon ordre,

$$\frac{1}{w(x)}(p(x)y'(x))' + q(x)y(x) = 0,$$

[2] Això no passa amb operadors diferencials d'ordre superior.

calen dues condicions suplementàries. Denotarem per $\mathcal{E}$ el subespai de funcions de $\mathcal{C}^n([a,b])$ que satisfan les condicions de contorn del problema físic associat a l'equació diferencial. Una forma bastant habitual que poden presentar aquestes condicions suplementàries ve donada a través de combinacions lineals de $y(a), y'(a), y(b), y'(b)$; tenim doncs

$$B_0[y] = \alpha_{00}y(a) + \alpha_{01}y'(a) + \beta_{00}y(b) + \beta_{01}y'(b) = \gamma_0$$

$$B_1[y] = \alpha_{10}y(a) + \alpha_{11}y'(a) + \beta_{10}y(b) + \beta_{11}y'(b) = \gamma_1$$

amb el conjunt $\{\alpha, \beta\}$ fixat. Aquestes s'anomenen *condicions de contorn mixtes no homogènies.* Les corresponents condicions de contorn *homogènies* s'obtenen si $\gamma_0 = \gamma_1 = 0$,[3] que són les que considerarem a partir d'aquest moment.

Un operador diferencial autoadjunt és **hermític** si

$$(f, \mathcal{L}g) = (\mathcal{L}f, g) \qquad \forall f, g \in \mathcal{C}^n([a,b]).$$

El producte escalar és el relatiu a la funció pes $w(x)$, que fa que l'operador $\mathcal{L}$ sigui autoadjunt. Fixem-nos en la diferència entre aquests dos productes escalars

$$(f, \mathcal{L}g) - (\mathcal{L}f, g) =$$

$$\begin{aligned} &= \int_a^b f^*(x)\left[\frac{d}{dx}p(x)\frac{dg}{dx} + wq(x)g(x)\right] - \int_a^b g(x)\left[\frac{d}{dx}p(x)\frac{df^*}{dx} + wq(x)f^*(x)\right] = \\ &= f^*pg'\big]_a^b - \int_a^b f'^*pg'dx - pf'^*g\big]_a^b + \int_a^b g'pf'^*dx = [pf^*g' - pf'^*g]_a^b = \left[p\overline{W}(f^*, g)\right]_a^b \end{aligned} \tag{2.2}$$

[3]Es pot demostrar que l'anàlisi general es pot fer prenent condicions de contorn homogènies sobre la funció. Sigui $y(x)$ la solució de

$$a_0y'' + a_1y' + a_2y = h(x),$$

que satisfà condicions de contorn mixtes no homogènies, i sigui $z(x)$ *qualsevol* funció satisfent les mateixes condicions de contorn. Resulta que la funció $u(x) = y(x) - z(x)$ satisfà l'equació diferencial

$$\mathcal{L}[u] = \mathcal{L}[y] - \mathcal{L}[z] = h(x) - \mathcal{L}[z] = g(x),$$

amb condicions de contorn homogènies

$$B_0[u] = B_0[y] - B_0[z] = \gamma_0 - \gamma_0 = 0,$$

$$B_1[u] = B_1[y] - B_1[z] = \gamma_1 - \gamma_1 = 0.$$

Per tant, un problema que satisfà condicions de contorn mixtes no homogènies és equivalent a un problema homogeni redefinint el terme inhomogeni de l'equació diferencial.

On podem veure que *l'hermiticitat de l'operador diferencial depèn directament de les condicions de contorn* imposades en el problema. Examinem ara alguns tipus de condicions sota les quals podrem assegurar l'hermiticitat de l'operador diferencial:

1. $\boxed{p(a) = p(b) = 0}$ Aquesta és una condició lligada estrictament a l'*operador diferencial* que ens permet d'obtenir hermiticitat sense haver d'exigir condicions de contorn a les funcions del domini.

2. $\boxed{\alpha_1 f(a) + \alpha_2 f'(a) = 0 \text{ i } \beta_1 f(b) + \beta_2 f'(b) = 0}$. Aquestes condicions s'anomenen no mixtes perquè no barregen els dos extrems (però barregen funció i derivada). Són les condicions de contorn més generals que anul·len el Wronskià en un punt.

3. $\boxed{f(a) = f(b) \text{ i } f'(a) = f'(b)}$ si $p(a) = p(b)$. Aquestes són condicions de contorn periòdiques. Si $p(x)$ és periòdica en un interval, podem obtenir hermiticitat exigint condicions de contorn periòdiques sobre les funcions del domini i les seves derivades primeres.

Aquesta llista de possibilitats no és exhaustiva, però recull les més importants. Al llarg dels capítols següents veurem exemples dels tres casos, que representen bona part dels problemes interessants en Física.

2.4 Problema de Sturm-Liouville

En molts problemes amb equacions diferencials en derivades parcials, la separació de variables condueix a equacions d'autovalors del tipus

$$\mathcal{L}[y] + \lambda y = 0 \quad \text{amb} \quad \mathcal{L} = \frac{1}{w(x)} \frac{d}{dx} p(x) \frac{d}{dx} + q(x).$$

En general, haurem de determinar totes les solucions y_λ diferents de la trivial que satisfan l'equació diferencial, associades a tots els possibles valors del paràmetre λ, sota unes certes condicions de contorn. Aquest problema es coneix amb el nom de *problema de Sturm-Liouville*. Fixem-nos, de nou, en què aquest és un problema anàleg al de calcular valors i vectors propis d'una matriu.

Exemple: Recordem com l'equació de Laplace en coordenades esfèriques ens porta a l'equació de Legendre

$$(1 - x^2)y'' - 2xy' + \lambda y = 0$$

que, en la forma canònica, es pot expressar com

$$\frac{d}{dx}\left[(1 - x^2)\frac{dy}{dx}\right] + \lambda y = 0$$

d'on identifiquem

$$p(x) = (1 - x^2) \quad q(x) = 0 \quad w(x) = 1 \quad \lambda = l(l+1).$$

En realitat no hi ha cap garantia que y existeixi per a qualsevol valor de λ. Exigir que λ tingui associada una funció pròpia sovint restringeix els valors acceptables de λ a un *conjunt discret de valors*. Per exemple, en el cas de l'equació de Legendre, $\lambda = l(l+1)$, amb $\ell \in \mathbb{N}$.

Les equacions diferencials ordinàries que provenen de les equacions en derivades parcials de la física es poden expressar en la forma canònica com un problema de Sturm-Liouville amb les següents funcions i paràmetres:

EQUACIÓ	DOMINI	$p(x)$	$q(x)$	$w(x)$	λ
Legendre	$-1 \leq x \leq 1$	$1 - x^2$	0	1	$l(l+1)$
Legendre associada	$-1 \leq x \leq 1$	$1 - x^2$	$-\frac{m^2}{1-x^2}$	1	$l(l+1)$
Bessel	$0 \leq x < \infty$	x	x^2	$1/x$	$-n^2$
Laguerre	$0 \leq x < \infty$	xe^{-x}	0	e^{-x}	a
Hermite	$-\infty < x < \infty$	e^{-x^2}	0	e^{-x^2}	2α
O.H.S.	$0 \leq x \leq 2\pi$	1	0	1	ω^2

En aquestes equacions sovint apareixen dos termes que multipliquen $y(x)$. Si és una funció de x es posa en $q(x)$ mentre que λ representa en general la constant de separació de variables de l'equació diferencial en derivades parcials i sobre el seu valor no tenim de moment cap mena d'informació.

2.4.1 Problema de Sturm-Liouville regular

Sigui

$$\mathcal{L} = \frac{1}{w(x)} \frac{d}{dx} p(x) \frac{d}{dx} + q(x)$$

un operador diferencial amb les següents condicions:

- q, w reals i continus en $[a, b]$
- p real i derivable amb continuïtat en $[a, b]$
- $w > 0$ en $[a, b]$.

Aleshores, imposant condicions de contorn tals que l'operador sigui hermític, definim el següent problema d'autovalors

$$\mathcal{L}[u_i] + \lambda_i u_i = 0, u_i \neq 0$$

conegut com a *problema de Sturm-Liouville regular*. Aquí u_i són les funcions pròpies i λ_i els valors propis associats a la funció pròpia u_i.

Teorema: *Si $\mathcal{L}$ és hermític aleshores els seus valors propis són reals i les funcions pròpies associades a valors propis diferents són ortogonals.*

DEMOSTRACIÓ:

Sigui

$$\mathcal{L}[u_i] + \lambda_i u_i = 0 \quad \text{i} \quad \mathcal{L}[u_j] + \lambda_j u_j = 0.$$

Conjuguem la darrera expressió

$$\mathcal{L}[u_j^*] + \lambda_j^* u_j^* = 0$$

i multipliquem la primera per u_j^* i la segona per u_i. Restant tenim

$$u_j^*\mathcal{L}[u_i] - u_i\mathcal{L}[u_j^*] + (\lambda_i - \lambda_j^*)u_i u_j^* = 0.$$

Finalment, integrem entre a i b amb $w(x)$:

$$\int_a^b w u_j^* \mathcal{L}[u_i] dx - \int_a^b w u_i \mathcal{L}[u_j^*] dx + (\lambda_i - \lambda_j^*) \int_a^b w u_i u_j^* dx = 0$$

$$\underbrace{(u_j, \mathcal{L}u_i) - (\mathcal{L}u_j, u_i)}_{=0} + (\lambda_i - \lambda_j^*)(u_j, u_i) = 0$$

i arribem a

$$(\lambda_i - \lambda_j^*)(u_j, u_i) = 0.$$

A partir d'aquesta darrera expressió podem deduir:

1. Si $i = j$, com que $(u_i, u_i) = \|u_i\|^2 > 0$, perquè $u_i \neq 0$. Per tant resulta que $\lambda_i - \lambda_i^* = 0 \Rightarrow \boxed{\lambda_i \in R}$
2. Si $\lambda_i \neq \lambda_j \Rightarrow (u_i, u_j) = 0$

□

Quan $i \neq j$ però $\lambda_i = \lambda_j$ tenim el que es diu *degeneració*.[4] Llavors (u_i, u_j) no té per què ser zero. Però sempre és possible de fer-los ortogonals, amb el mètode de Gram-Schmidt, per exemple.

[4]El conjunt de solucions $u \in \mathcal{E}$ de $\mathcal{L}[u] + \lambda u = 0$ és un subespai vectorial d'$\mathcal{E}$, que denotarem per $\mathcal{E}_\lambda$. Si λ no és un valor propi aleshores $\mathcal{E}_\lambda = \{0\}$, mentre que si λ és un valor propi $\dim\mathcal{E}_\lambda \geq 1$. En el cas $\dim\mathcal{E}_\lambda = m > 1$, hi ha més d'una funció pròpia per al valor propi λ, i diem que el valor propi λ és degenerat.

2.5 Ortonormalització de Gram-Schmidt

El procediment d'ortonormalització de Gram-Schmidt permet, per exemple, de trobar un conjunt ortonormal de solucions per a un λ degenerat. Siguin u_1, u_2, ..., u_n un conjunt linealment independent de solucions amb valor propi λ. Aleshores, existeix un conjunt ortonormal de solucions v_1, v_2, ..., v_n.

Definirem un procés iteratiu per construir un conjunt de funcions $\{v_i, i = 1..n\}$ que satisfan les propietats d'ortonormalitat:

$$(v_i, v_j) = \delta_{ij} , \qquad i, j = 1, ..., n .$$

El primer pas és prendre

$$v_1 = \frac{u_1}{\|u_1\|} , \text{ amb } \|u_1\|^2 \equiv (u_1, u_1) .$$

Després, si tenim els k primers termes v_1, ..., v_k, el $k+1$-èsim es defineix com:

$$v_{k+1} \equiv \frac{w_{k+1}}{\|w_{k+1}\|} ,$$

amb

$$w_{k+1} \equiv u_{k+1} - \sum_1^k (v_k, u_{k+1}) v_k .$$

Exemple: Aplicació del mètode de Gram-Schmidt al conjunt dels polinomis, amb una funció pes $\omega(x) = e^{-x}$, en l'interval $[0, \infty)$. Es pot comprovar que aquests polinomis són solució de l'equació diferencial de Laguerre.

Aquest exemple ens serveix per mostrar el mètode de Gram-Schmidt sobre un conjunt arbitrari de funcions. Partint del conjunt dels polinomis linealment independents $\{x^n\}$, obtenim un altre conjunt que satisfà propietats d'ortonormalitat. Fixem-nos que els $\{x^n\}$ *no són solució de l'equació de Laguerre*. Són simplement un conjunt de funcions que hem emprat per obtenir un conjunt ortonormal en un *interval* amb una *funció pes* donats. Aquestes dues condicions són les que determinen que el resultat final siguin els *polinomis de Laguerre*, que sí són solució de l'equació de Laguerre, $xy'' + (1 - x)y' + ny = 0$.

Escollint altres intervals o altres funcions pes el conjunt $\{x^n\}$ porta a altres conjunts de polinomis *ortonormals*.

Polinomi	Interval	$w(x)$	Normalització [a]
Legendre (P_n)	$[-1,1]$	1	$\frac{2}{2n+1}$
Laguerre (L_n)	$[0,\infty)$	e^{-x}	1
Associat de Laguerre (L_n^k)	$[0,\infty)$	$x^k e^{-x}$	$\frac{(n+k)!}{n!}$
Hermite (H_n)	$(-\infty,\infty)$	e^{-x^2}	$2^n\sqrt{\pi}n!$

[a] Aquest tipus de normalització és el que es fa servir habitualment de tal manera que, per exemple,

$$< P_n|P_m >= \frac{2}{2n+1}\delta_{nm}$$

Veurem en els propers temes com obtenir aquests polinomis ortogonals a partir de les solucions de les respectives equacions diferencials.

2.6 Completesa

Considerarem un conjunt ortonormal de funcions pròpies d'un determinat problema de Sturm-Liouville. Habitualment és un conjunt infinit numerable, és a dir, és de la forma $\{u_j, j \in \mathbb{N}\}$. Podrem intentar d'escriure una funció contínua qualsevol[5] d'un nombre finit de funcions pròpies:

$$f(x) = \sum_{j=1}^{n} a_j u_j(x)\,, \qquad a_j \in \mathbb{C}. \tag{2.3}$$

Això en general no serà possible, però podem buscar la combinació de les n primeres funcions pròpies que més s'hi aproximi:

$$S_n(x) \equiv \sum_{j=1}^{n} a_j u_j(x).$$

En alguns punts ens aproparem més i en altres menys, fins i tot podem trobar que en algun punt la sèrie i la funció coincideixin.

[5] De fet tot el desenvolupament seguit al llarg del capítol es pot estendre a un conjunt de funcions que, sense ser contínues, sí són funcions de quadrat integrable. El fet de considerar funcions contínues obeeix a una raó de simplicitat; no obstant això, el problema que plantegen aquestes funcions contínues és que una successió de funcions contínues pot tenir com a límit una funció discontínua. Aleshores la successió no tindria límit dins el propi espai; es tractaria doncs d'una successió de Cauchy, però no convergent. El requisit "de quadrat integrable" és simplement la condició de norma finita, que garanteix que els productes escalars són acotats. Afegint al conjunt de funcions contínues de quadrat integrable tots els possibles límits de successions d'aquestes funcions aconseguim la compleció de l'espai de funcions contínues: l'espai de funcions de quadrat integrable.

Una mesura de la distància entre f i S_n la dóna la **desviació mitjana quadràtica**,

$$\|f - S_n\|^2 \equiv \int_a^b |f(x) - S_n(x)|^2 dx = (f - S_n, f - S_n). \tag{2.4}$$

Per minimitzar aquesta **distància** ens adonem que f és una funció de les n variables complexes $a_1, ..., a_n$ o bé de les $2n$ variables reals $\alpha_1, ..., \alpha_n, \beta_1, ..., \beta_n$, on hem posat $a_j = \alpha_j + \imath\beta_j$,

$$F(a_1, ..., a_n) = (f - S_n, f - S_n) = \left(f - \sum_{j=1}^{n} a_j u_j, f - \sum_{i=1}^{n} a_i u_i \right) . \tag{2.5}$$

que, desenvolupant els productes escalars, s'escriu

$$F(a_1, ..., a_n) = (f, f) - \sum_{j=1}^{n} a_j^*(u_j, f) - \sum_{j=1}^{n} a_j(f, u_j) + \sum_{j=1}^{n} |a_j|^2.$$

La condició de mínim dóna

$$\frac{\partial F}{\partial \alpha_j} = 0 \Rightarrow -(u_j, f) - (f, u_j) + 2\alpha_j = 0 \tag{2.6}$$

i

$$\frac{\partial F}{\partial \beta_j} = 0 \Rightarrow +\imath(u_j, f) - \imath(f, u_j) + 2\beta_j = 0\,, \tag{2.7}$$

que té per solució

$$\alpha_j = \Re(u_j, f)\,, \quad \text{i} \quad \beta_j = \Im(u_j, f)\,,$$

és a dir,

$$a_j = (u_j, f)\,, \quad j = 1, ..., n. \tag{2.8}$$

La matriu de les derivades segones de F és la matriu unitat $2n \times 2n$, definida positiva, per la qual cosa (2.8) dóna efectivament el mínim de F.

Hem provat, doncs, que el sumatori

$$S_n(x) \equiv \sum_{j=1}^{n} (u_j, f) u_j \tag{2.9}$$

és la millor aproximació a f que podem tenir amb una combinació lineal de les n funcions u_1, ..., u_n. Aquesta aproximació millora com més funcions u_j considerem. Efectivament, com que $S_{n+1} = S_n + (u_{n+1}, f)u_{n+1}$, resulta que

$$\|f - S_{n+1}\|^2 = \|f - S_n - (u_{n+1}, f)u_{n+1}\|^2 = \|f - S_n\|^2 - \|(u_{n+1}, f)\|^2 \leq \|f - S_n\|^2.$$

La millor aproximació llavors la donarà la sèrie:

$$S(x) \equiv \sum_{j=1}^{\infty}(u_j, f)u_j(x). \tag{2.10}$$

En principi $S(x)$ només dóna la millor aproximació i no és necessàriament cert que $f - S = 0$. Tanmateix, és fàcil comprovar que $(f - S_n, u_j) = 0$, $\forall j \leq n$ i, en conseqüència, $(f - S, u_j) = 0$, $\forall j \in \mathbb{N}$ i també $(f - S, S) = 0$. D'aquí es dedueix que

$$(f, f) = (f - S, f - S) + (S, S) \tag{2.11}$$

i tenim l'anomenada **desigualtat de Bessel**:

$$\|f\|^2 \geq \|S\|^2 = \sum_{j=1}^{\infty} |(u_j, f)|^2 . \tag{2.12}$$

Quan podrem dir doncs que la sèrie és vàlida com un desenvolupament de la funció f, en el sentit de la convergència en mitjana quadràtica? Doncs quan es compleixi la següent relació[6]

$$\|f\|^2 = \sum_{j=1}^{\infty} |(u_j, f)|^2 , \tag{2.13}$$

coneguda com a **igualtat de Parseval**. Així, en aquest cas, podrem escriure

$$f = S = \sum_{j=1}^{\infty}(u_j, f)u_j . \tag{2.14}$$

Si aquesta igualtat es compleix $\forall f$ direm que el conjunt de funcions $\{u_j(x), j \in \mathbb{N}\}$ és un conjunt ortonormal **complet**, perquè qualsevol funció f de l'espai considerat es pot expressar com una suma infinita (*sèrie*) de múltiples de les u_n. Ara bé, una sèrie comporta dues operacions: la *suma* i el *pas al límit*. La igualtat (2.14) vol dir que

$$\lim_{n\to\infty} \|f - S_n\| = 0 ,$$

[6] Tenim una forma alternativa de veure-ho: $\{S_n\}$ és una successió (de Cauchy). Si, a més, és convergent aleshores $S_n \to f$ (entenent el límit en mitjana quadràtica). Aleshores de (2.11) es dedueix que $\|f\|^2 = \|S\|^2$.

és a dir, **convergència en mitjana quadràtica**, la qual cosa no vol dir que hi hagi **convergència puntual**, que s'expressaria com:

$$\lim_{n\to\infty} S_n(x) = f(x), \quad \forall x \in [a,b] \,.$$