



TREBALL DE FI DE GRAU

ESTUDI D'UN MODEL VARIACIONAL PER A LA RESTAURACIÓ D'IMATGES SUBMARINES

Bartomeu Garau Verger

Grau de Matemàtiques

Escola Politècnica Superior

Any acadèmic 2022-23

ESTUDI D'UN MODEL VARIACIONAL PER A LA RESTAURACIÓ D'IMATGES SUBMARINES

Bartomeu Garau Verger

Treball de Fi de Grau

Escola Politècnica Superior

Universitat de les Illes Balears

Any acadèmic 2022-23

Paraules clau del treball: Imatges submarines, Anàlisi, Càcul de Variacions, Optimització

Tutora: Catalina Sbert Juan

Autoritz la Universitat a incloure aquest treball en el repositori institucional per consultar-lo en accés obert i difondre'l en línia, amb finalitats exclusivament acadèmiques i d'investigació

Autor/a	Tutor/a
Sí <input checked="" type="checkbox"/>	No <input type="checkbox"/>
Sí <input checked="" type="checkbox"/>	No <input type="checkbox"/>

A la meva família i amics,
per haver-me donat suport i ànims
al llarg dels darrers quatre anys.

SUMARI

Sumari	iii
Acrònims	v
Resum	vii
1 Introducció	1
2 Conceptes bàsics d'Anàlisi Funcional	3
2.1 Espais normats	3
2.1.1 L'espai dual	5
2.1.2 Espais reflexius	7
2.2 Espais de Hilbert	10
3 Els espais de Sobolev, $W^{1,p}$	15
3.1 Un tast de teoria de la mesura i integració	15
3.2 Els espais de Lebesgue, $L^p(\Omega)$	17
3.2.1 Definició i propietats bàsiques	17
3.2.2 El dual de $L^p(\Omega)$. Reflexivitat	20
3.2.3 El cas del producte cartesià	22
3.3 Els espais de Sobolev, $W^{1,p}(\Omega)$	23
3.3.1 Definició i propietats bàsiques	23
3.3.2 Desigualtats de Sobolev. L'espai $W_0^{1,p}(\Omega)$	25
3.4 Introducció a l'espai de funcions de Variació Fitada	27
4 El mètode directe del Càcul de Variacions	29
4.1 El problema de minimització	29
4.2 El mètode directe	30
4.2.1 Convexitat, coercivitat i semi-continuitat	30
4.2.2 El principi variacional. Teorema general d'existència	31
5 Optimització convexa i algorismes Primal-Dual	35
5.1 Nocions bàsiques i propietats	35
5.1.1 Subdiferenciabilitat	35
5.1.2 La conjugada de Legendre-Fenchel	38
5.1.3 L'operador proximal	40
5.2 Algorismes Primal-Dual	41
5.2.1 Dualitat convexa	41

5.2.2	L'algorisme Primal-Dual de Chambolle-Pock	42
6	Tractament d'imatges	43
6.1	Imatges com a objectes matemàtics. Problemes mal posats	43
6.2	<i>Dehazing</i> o reducció de boira	45
6.2.1	Causes de la boira i el <i>dark channel prior</i>	45
6.2.2	Model proposat	47
6.2.3	Model proposat, cas $p = 2$	48
6.2.4	Model proposat, cas $p = 1$	53
6.2.5	Comparativa: cas $p = 1$ vs. cas $p = 2$	55
7	Conclusions	57
A	Apèndix	59
A.1	Codi Python implementat	59
A.1.1	Paquets emprats	59
A.1.2	Funcions auxiliars	59
A.1.3	Funcions auxiliars pròpies del <i>dehazing</i>	62
A.1.4	Funcions de correcció de color	65
A.1.5	Algorismes de <i>denoising</i> o reducció de renou	66
A.1.6	Algorismes de <i>dehazing</i> o reducció de boira	67
	Bibliografia	71

ACRÒNIMS

qpt quasi per tot

RESUM

La restauració d'imatges submarines és fonamental per a l'estudi del fons marí. La baixa qualitat d'aquestes imatges dificulta el seu estudi, i per això és important tractar-les i millorar-ne la qualitat. Això però, no és una tasca senzilla. Per tal de poder fer-ho haurem d'emprar eines de diferents camps de les matemàtiques i eines informàtiques per poder implementar els algorismes proposats. Si plantejam la restauració d'imatges submarines com un problema matemàtic, aquest passa per diferents etapes: la demonstració d'existència i unicitat de solució, l'estudi d'algorismes per trobar la solució, i la revisió i comparació dels resultats obtinguts. Per poder dur a terme correctament cada una d'aquestes passes, estudiarem les propietats bàsiques dels espais de Sobolev $W^{1,p}(\Omega)$, introduirem el mètode directe del Càlcul de Variacions, i farem servir tècniques d'Optimització Convexa per trobar la solució explícita del problema.

INTRODUCCIÓ

La restauració d'imatges submarines és un tema realment atractiu que sol passar desaparcebute. Actualment, moltes de les tasques d'arqueologia marina, de l'estudi de l'ecosistema submarí o, fins i tot, de salvament marítim, depenen de l'estudi d'imatges submarines preses per bucejadors o robots que exploren el fons marí. Però sota l'aigua, a causa de les característiques del medi on ens trobam, la llum no és comporta igual que amb l'aire, fent que es distorioni i no es captin les imatges com són realment. A més, normalment l'aigua no és totalment cristal·lina a causa de la presència de partícules en suspensió com grans d'arena, trossos de posidònia... Aquests dos factors fan que es creï una capa similar a la boira. Per tant, si volem restaurar aquestes imatges submarines, podem tractar el problema de recuperar la imatge neta com un problema de *dehazing* o reducció de boira.



Figura 1.1: Exemples d'imatges submarines.

A l'àmbit de *dehazing* destaca el treball fet per He, Sun i Tang [1], però nosaltres ens centrarem amb el model proposat per Fang, Li i Zeng a [2]. Al seu article, proposen recuperar una versió neta i sense renou d'una imatge amb boira a partir de minimitzar el següent funcional:

$$E(\mathbf{g}, d) = \int_{\Omega} h|D\mathbf{g}| + \lambda \int_{\Omega} |Dd| + \frac{1}{2} \int_{\Omega} |\mathbf{g} - \mathbf{f} - \mathbf{d}|^2 + \frac{\gamma}{2} \int_{\Omega} |d - d_0|^2.$$

Però per minimitzar aquest funcional cal estudiar en profunditat les propietats de l'espai BV , l'espai de funcions de variació fitada. Nosaltres estudiarem una versió

1. INTRODUCCIÓ

d'aquest problema que ens permetrà fer l'estudi en el context dels espais de Sobolev:

$$E_p(\mathbf{g}, d) = \int_{\Omega} h|\nabla \mathbf{g}|^p + \lambda \int_{\Omega} |\nabla d|^p + \frac{1}{2} \int_{\Omega} |\mathbf{g} - \mathbf{f} - \mathbf{d}|^2 + \frac{\gamma}{2} \int_{\Omega} |d - d_0|^2$$

per $p > 1$. Així, podrem emprar teoremes basats en el mètode directe del Càlcul de Variacions per poder trobar el seu mínim a l'espai de Sobolev, i després dissenyar algorismes per trobar el mínim de $E_p(\mathbf{g}, d)$ pel cas $p = 2$. Al llarg d'aquest document s'estudiaran diferents branques de les matemàtiques, però a causa de l'extensió limitada d'aquest treball només es faran les demostracions que siguin més rellevants per poder assolir els objectius proposats i les que hagin despertat l'interès matemàtic de l'autor.

El capítol (2) és la base del treball. Hi introduirem els conceptes bàsics d'Anàlisi Funcional per a la correcta comprensió del treball. Després, en el capítol (3), seguirem ampliant conceptes teòrics de l'Anàlisi Funcional: desenvoluparem la teoria dels espais $L^p(\Omega)$ necessària per a poder explicar els espais de Sobolev $W^{1,p}(\Omega)$. Aquests espais de funcions desenvoluparan un paper molt important en aquest treball, ja que serà on viurà la solució del nostre problema.

Una vegada coneixem l'espai on viurà la nostra solució, necessitam veure que el problema es pot resoldre. Per això, al capítol (4), estudiarem el mètode directe del Càlcul de Variacions. En concret, desenvoluparem la teoria necessària per a poder demostrar un teorema que ens garantirà l'existència i unicitat de solució del problema de *dehazing*.

La següent passa serà trobar la solució explícita. Per això haurem de recórrer a l'Optimització Convexa, que serà treballada en el capítol (5). Ens centrarem en estudiar la relació entre el problema primal i el problema primal-dual, i com la solució d'un ens pot donar la solució de l'altre. En el nostre cas, implementarem l'algorisme Primal-Dual de Chambolle-Pock [3].

Per últim, arribam al problema esmentat a l'incici d'aquest capítol: la restauració d'imatges submarines. En el capítol (6), explicarem en detall el problema de *dehazing* i com aquest es relaciona amb el tractament d'imatges submarines. Amb tot el que haurem vist als capítols anteriors, podrem demostrar que el problema té una única solució, per posteriorment desenvolupar el corresponent algorisme de Chambolle-Pock, implementar-ho amb codi Python, i observar i comparar els resultats obtinguts.

CONCEPTES BÀSICS D'ANÀLISI FUNCIONAL

Els segles XIX i XX van ser els segles on es van fonamentar la matemàtica moderna. Al segle XIX es van establir les bases de l'Anàlisi Matemàtica que coneixem avui dia gràcies a Cauchy, Weierstrass, Bolzano... Però no és fins al segle XX que apareix el que avui en dia coneixem com Anàlisi Funcional. Aquest neix, entre d'altres, amb el treball de David Hilbert (1862 - 1943) sobre equacions integrals, que el portaren a desenvolupar la teoria espectral; i amb la tesi de Stefan Banach (1892 - 1945), considerada com la base de l'Anàlisi Funcional moderna, on s'axiomatitza la noció d'espai normat complet. D'aquesta noció i de les seves propietats, així com d'una part de la feina duta a terme per Hilbert, en xerrarem al llarg d'aquest capítol.

Cal dir que aquest capítol és un recull de resultats la majoria dels quals s'ha vist a classe. Les demostracions que s'ometen es poden trobar a [4] i [5].

2.1 Espais normats

En aquesta secció considerarem X un espai vectorial sobre \mathbb{R} . Abans de parlar d'espais normats, començarem definint què és una *norma*.

Definició 2.1.1. *Sigui X un espai vectorial. Una norma X és una aplicació $\|\cdot\| : X \rightarrow \mathbb{R}$ que verifica:*

- (i) $\|x\| \geq 0$ i $\|x\| = 0$ si i només si $x = 0$,
- (ii) $\|\lambda x\| = |\lambda| \|x\|$ per a tota $\lambda \in \mathbb{R}$ i tot $x \in X$,
- (iii) $\|x + y\| \leq \|x\| + \|y\|$ per a tots $x, y \in X$.

Quan tenim un espai vectorial X dotat d'una norma $\|\cdot\|$, anomenarem *espai normat* al parell $(X, \|\cdot\|)$. Si no hi ha confusió, el denotarem per X . Ara bé, a vegades en un mateix espai X hi podem trobar definides diverses normes distintes.

Proposició 2.1.1. *Dues normes $\|\cdot\|_a, \|\cdot\|_b$ en X són equivalents si i només si existeixen dues constants $C_1, C_2 > 0$ tals que per a tot $x \in X$*

$$C_1 \|x\|_a \leq \|x\|_b \leq C_2 \|x\|_a$$

Un dels elements més importants a tractar en espais normats són les successions i el seu caràcter. Podem definir una successió convergent en un espai normat de la següent forma:

Definició 2.1.2. *Sigui X un espai normat. Una successió $\{x_n\}$ d'elements de X és convergent a \hat{x} si donat $\epsilon > 0$ existeix un $n_0 \in \mathbb{N}$ tal que $\|x_n - \hat{x}\| < \epsilon$ per a tot $n \geq n_0$.*

Observació 1. *Aquesta noció de convergència també rep el nom de convergència forta o convergència en norma.*

Ara bé, no sempre podem trobar successions convergents, però de vegades, malgrat que no convergeixin, hi ha successions que “s'estabilitzen”, és a dir, que a partir de cert element, els termes de la successió no difereixen gaire. Això és el que es coneix com a successió de Cauchy. Formalment:

Definició 2.1.3. *Sigui X un espai normat. Una successió $\{x_n\}$ d'elements de X és de Cauchy si donat $\epsilon > 0$ existeix un $n_0 \in \mathbb{N}$ tal que $\|x_n - x_m\| < \epsilon$ per a tot $n, m \geq n_0$.*

Al llarg d'aquest treball, ens interessaran els espais normats complets. A la seva tesi, Banach els va anomenar *espais de tipus E*. Posteriorment, va canviar el nom a *espais de tipus B*. Avui en dia, els coneixem com espais de Banach, i els definim com:

Definició 2.1.4. *Direm que un espai normat $(X, \|\cdot\|)$ és un espai de Banach si és complet respecte la seva norma, és a dir, si tota successió de Cauchy és convergent.*

En espais normats, si $M \subset X$ és un subespai, també serà normat amb la norma $\|\cdot\|_M$. Ara bé, si X és complet, M no necessàriament és complet, només quan sigui tancat.

Proposició 2.1.2. *Sigui X un espai de Banach. Un subespai $M \subset X$ és de Banach respecte la norma induïda si i només si és tancat.*

Quan treballam amb espais que són de dimensió infinita, el concepte de numerabilitat esdevé gaire difús. No podem dir a priori si un espai de dimensió infinita és numerable o no, però en tal cas podem parlar de separabilitat.

Definició 2.1.5. *Direm que X és un espai separable si té algun subconjunt numerable dens.*

L'estructura d'espai vectorial que presenten els espais normats ens porten a estudiar les aplicacions que els relacionen, conegudes en aquest camp com *operadors*. Anem a definir-los formalment.

Definició 2.1.6. *Siguin $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ espais normats i definim l'operador A com l'aplicació $A : X \rightarrow Y$. Aleshores:*

(i) Direm que A és un operador lineal si es compleix la condició:

$$A(\alpha x + \beta y) = \alpha A(x) + \beta A(y), \quad \forall x, y \in X, \forall \alpha, \beta \in \mathbb{R}.$$

(ii) Direm que A és un operador fitat si existeix K constant real positiva tal que:

$$\|Ax\|_Y \leq K\|x\|_X, \quad \forall x \in X.$$

Si consideram el conjunt de totes les aplicacions lineals entre dos espais normats, aquest té estructura d'espai vectorial amb les operacions de suma i producte per un escalar de funcions. Dins aquest conjunt, és de gran importància el que està format per totes les aplicacions lineals i fitades.

Definició 2.1.7. L'espai vectorial format per tots els operadors lineals i fitats de X en Y , denotat per $\mathcal{L}(X, Y)$, és un espai normat amb la següent norma:

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|_Y}{\|x\|_X} = \sup_{\|x\| \leq 1} \|Ax\|_Y, \quad \forall A \in \mathcal{L}(X, Y).$$

Proposició 2.1.3. Si Y és un espai de Banach, l'espai $\mathcal{L}(X, Y)$ és un espai de Banach.

La següent proposició resumeix les principals propietats dels operadors lineals entre espais normats.

Proposició 2.1.4. Sigui $A : X \rightarrow Y$ un operador lineal entre els espais normats $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$. Les següents afirmacions són equivalents:

- (i) A és continu.
- (ii) A és continu en qualque $x_0 \in X$.
- (iii) A és continu en el 0.
- (iv) A és fitat.

Per acabar aquesta secció, donam un resultat referent al producte cartesià d'espais de Banach que ens serà útil de cara als següents capítols.

Proposició 2.1.5. Sigui X_j un espai de Banach per $j = 1, \dots, N$ amb la norma $\|\cdot\|_j$. Aleshores, el producte cartesià $X = \prod_{j=1}^N X_j$, on els elements són del tipus $x = (x_1, \dots, x_N)$ amb cada $x_j \in X_j$, és un espai de Banach respecte les normes equivalents

$$\|x\|_{(p)} = \left(\sum_{j=1}^N \|x_j\|_j^p \right)^{1/p} \quad \text{si } 1 \leq p < \infty, \quad \text{o} \quad \|x\|_{(\infty)} = \max_{1 \leq j \leq N} \|x_j\|_j$$

2.1.1 L'espai dual

Quan parlem d'operadors lineals entre espais, hi ha un tipus que és fonamental a l'Anàlisi Funcional: els *funcionals lineals*.

Definició 2.1.8. Direm que un operador és un funcional quan l'espai d'arribada és el cos dels nombres reals, \mathbb{R} . És a dir, donat un espai normat X , un funcional és un operador $f : X \rightarrow \mathbb{R}$.

Quan hem introduït els operadors, hem destacat el conjunt dels operadors lineals i fitats. En el cas que considerem el conjunt de funcionals lineals i fitats, podem parlar de l'*espai dual*.

2. CONCEPTES BÀSICS D'ANÀLISI FUNCIONAL

Definició 2.1.9. Sigui X un espai normat. Definim l'espai dual de X com $X' = \mathcal{L}(X, \mathbb{R})$, és a dir, com l'espai dels funcionals lineals i fitats de X .

Observació 2. Al ser l'espai dual $X' = \mathcal{L}(X, \mathbb{R})$, i per ser \mathbb{R} un espai de Banach, es té que X' sempre és un espai de Banach amb la norma

$$\|f\|_* = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|} = \sup_{\|x\| \leq 1} |f(x)|, \quad \forall f \in X'.$$

Observació 3. Si no hi ha dubte, denotarem la norma dual per $\|\cdot\|_*$. En el cas que s'emprin diversos espais i/o subespais, per evitar confusió, la denotarem per la norma de l'espai dual corresponent. Per exemple, si ens trobam a l'espai X , escriurem $\|\cdot\|_{X'}$

En espais normats de dimensió infinita, en general, perdem les nocions geomètriques d'ortogonalitat que tenim quan consideram \mathbb{R}^N , però podem definir les nocions d'ortogonalitat a partir de l'espai dual.

Definició 2.1.10. Siguin X un espai normat i $M \subset X$, $N \subset X'$ subespais vectorials. Definim el complement ortogonal de M com:

$$M^\perp = \left\{ f \in X' \mid f(x) = 0, \forall x \in M \right\}$$

i el complement ortogonal de N com:

$$N^\perp = \left\{ x \in X \mid f(x) = 0, \forall f \in N \right\}.$$

L'Anàlisi Funcional se sosté gràcies als seus quatre pilars fonamentals: els teoremes de Hahn-Banach, el de Banach-Steinhaus, el de l'aplicació oberta i el del graf tancat. El primer és de gran importància, ja que gràcies a ell es poden demostrar molts de resultats on es veu involucrat l'espai dual.

Teorema 2.1.1 (Hahn-Banach en espais normats). Sigui X un espai normat i sigui $M \subset X$ un subespai vectorial. Si $g \in M'$, aleshores existeix un $f \in X'$ tal que $f(x) = g(x)$ per a tot $x \in M$ i a més $\|f\|_{X'} = \|g\|_{M'}$.

És a dir, el Teorema de Hahn-Banach ens diu que tot funcional lineal i fitat de M es pot extreure a un funcional lineal i fitat de X . En el següent corol·lari recollirem algunes de les principals conseqüències d'aquest teorema.

Corol·lari 2.1.1.1. Sigui X un espai normat.

(i) Per a tot $x \in X$ existeix un $f \in X'$ tal que $\|f\|_* = 1$ i $f(x) = \|f\|_* \|x\|$. Això implica que per a tot $x \in X$:

$$\|x\| = \max_{\|f\|_* = 1} f(x) = \max_{f \neq 0} \frac{|f(x)|}{\|f\|_*}.$$

(ii) Si $M \subset X$ és un subespai vectorial i $x \in X \setminus \overline{M}$, aleshores existeix un $f \in M^\perp$ tal que $f(x) \neq 0$.

(iii) Com a conseqüència, es dedueix que per a tot $M \subset X$ es té que $\overline{M} = (M^\perp)^\perp$. En particular, un subespai M és dens si i només si $M^\perp = \{0\}$.

2.1.2 Espais reflexius

Sigui X un espai normat. Una vegada hem definit el seu espai dual segons la Definició 2.2.2, denotat per X' , podem considerar $X'' = (X')'$, $X''' = (X'')'$, ...

Definició 2.1.11. *L'espai X'' rep el nom d'espai bidual.*

El nostre objectiu ara és determinar quina relació existeix entre un espai i el seu bidual. Això ens duu a parlar de la injecció canònica:

Definició 2.1.12. *Sigui X un espai normat. La injecció canònica és l'aplicació $J: X \rightarrow X''$ donada per*

$$Jx(f) = f(x), \quad \forall x \in X, \forall f \in X'.$$

La injecció canònica és, per construcció, una aplicació injectiva i lineal. A més està ben definida, i a conseqüència del corol·lari 2.1.1.1 (i), és isomètrica. Tot això ho recopilam a la següent proposició.

Proposició 2.1.6. *La injecció canònica és injectiva, està ben definida, és lineal i és isomètrica, és a dir,*

$$\|Jx\|_{X''} = \|x\|$$

Aquesta proposició ens permet identificar X amb el subespai $J(X) \subset X''$. Ara bé, en els casos on aquesta aplicació sigui exhaustiva, estarem identificant X amb $J(X) = X''$. Quan tenim aquesta relació, estam parlant d'espais reflexius.

Definició 2.1.13. *Direm que un espai normat X és reflexiu si la injecció canònica és exhaustiva. En tal cas direm que X s'identifica amb X'' i ho denotarem per $X \equiv X''$.*

Recordem que l'espai dual sempre és un espai de Banach, i per tant l'espai bidual també ho és. Llavors com un espai reflexiu s'identifica amb el seu bidual, en concret serà un espai de Banach.

Observació 4. *Perquè un espai sigui reflexiu necessàriament la injecció canònica ha de ser un isomorfisme isomètric. Si existeix un isomorfisme isomètric entre un espai i el seu bidual que no és la injecció canònica, l'espai no necessàriament és reflexiu. Per exemple, el definit a [6].*

A la proposició (2.1.2) hem vist que perquè un subespai vectorial d'un espai de Banach sigui complet, aquest ha de ser tancat. Aquesta propietat és similar en el cas de la reflexivitat, ja que els subespais d'un espai reflexiu també seran reflexius quan siguin tancats.

Proposició 2.1.7. *Tot subespai tancat d'un espai reflexiu és reflexiu.*

Demostració. Sigui X un espai reflexiu i $M \subset X$ subespai tancat. Volem veure que l'aplicació

$$\begin{aligned} \hat{J}: M &\longrightarrow M'' \\ m &\longmapsto \hat{J}m: M' \longrightarrow \mathbb{R} \\ g &\longmapsto \hat{J}m(g) = g(m) \end{aligned}$$

2. CONCEPTES BÀSICS D'ANÀLISI FUNCIONAL

és exhaustiva. Considerem un $\hat{F} \in M''$ i definim F com $F(f) = \hat{F}(f|_M)$ per a tot $f \in X'$. Aquesta aplicació és lineal, ja que per a tot $f, g \in X', \alpha, \beta \in \mathbb{R}$:

$$\begin{aligned} F(\alpha f + \beta g) &= \hat{F}((\alpha f + \beta g)|_M) = \hat{F}(\alpha f|_M) + \hat{F}(\beta g|_M) = \alpha \hat{F}(f|_M) + \beta \hat{F}(g|_M) \\ &= \alpha F(f) + \beta F(g), \end{aligned}$$

i a més és contínua perquè, com de la definició de norma dual tenim que

$$\|f|_M\|_{M'} = \sup_{\substack{x \in M \\ x \neq 0}} \frac{|f|_M(x)|}{\|x\|_M} \leq \sup_{\substack{x \in X \\ x \neq 0}} \frac{|f(x)|}{\|x\|_X} = \|f\|_{X'},$$

llavors es té que

$$|F(f)| = |\hat{F}(f|_M)| \leq \|\hat{F}\|_{M''} \cdot \|f|_M\| \leq \|\hat{F}\|_{M''} \cdot \|f\|_{X'}.$$

Per tant $F \in X''$. Com X és reflexiu, sabem que existeix un $x \in X$ tal que $Jx \equiv F$, on J és la injecció canònica de X en X'' . Així

$$f(x) = Jx(f) = F(f) = \hat{F}(f|_M), \quad \forall f \in X'. \quad (2.1)$$

Suposem que $x \notin M$. Com M és tancat, pel corol·lari 2.1.1.1 (ii) existeix un $f \in M^\perp$ tal que $f(x) \neq 0$, però si $f \in M^\perp$, llavors $f|_M \equiv 0$. Això contradiu (2.1), per tant $x \in M$. Així, podem considerar $\hat{J}x$.

Ara sigui $g \in M'$, pel teorema de Hahn-Banach (teorema 2.1.1) existeix una $f \in X'$ tal que $f|_M = g$. Així

$$\hat{J}x(g) = g(x) = f|_M(x) = f(x) = \hat{F}(f|_M) = \hat{F}(g).$$

Al ser g arbitrari, tenim que $\hat{J}x = \hat{F}$ i per tant \hat{J} és exhaustiva. □

Compacitat en espais reflexius

A continuació introduirem noves nocions de convergència: la convergència feble i la convergència $*$ -feble. Aquests conceptes són útils per poder demostrar resultats referents a la compacitat en els espais reflexius.

Definició 2.1.14. Sigui X un espai normat. Direm que una successió $\{x_n\} \subset X$ convergeix feblement a $\hat{x} \in X$, i ho denotarem per $x_n \rightharpoonup \hat{x}$, si i només si $f(x_n) \rightarrow f(\hat{x})$ per a tot $f \in X'$.

Proposició 2.1.8. La convergència en norma implica la convergència feble. És a dir, si $x_n \rightharpoonup \hat{x}$ en X un espai normat, aleshores $x_n \rightharpoonup \hat{x}$.

Demostració. Sigui $f \in X'$ fixat però arbitrari. Sigui $\{x_n\}$ una successió en X tal que $x_n \rightharpoonup \hat{x}$, és a dir, donat $\epsilon > 0$ existeix un $n_0 \in \mathbb{N}$ tal que $\|x_n - \hat{x}\| < \epsilon / \|f\|_*$ per a tot $n \geq n_0$. Aleshores per a tot $n \geq n_0$:

$$|f(x_n) - f(\hat{x})| = |f(x_n - \hat{x})| \leq \|f\|_* \|x_n - \hat{x}\| < \epsilon$$

□

Definició 2.1.15. Sigui X un espai normat. Direm que una successió $\{f_n\} \subset X'$ convergeix $*$ -feblement a un funcional $f \in X'$, i ho denotarem per $f_n \xrightarrow{*} f$, si i només si es té que $f_n(x) \rightarrow f(x) \forall x \in X$.

També podem pensar en la convergència feble i $*$ -feble sobre X' . Pel que fa a la convergència feble, direm que $f_n \rightarrow f$ en X' si $F(f_n) \rightarrow F(f)$ per a tot $F \in X''$. Si pensam en la convergència $*$ -feble, considerant la injecció canònica, podem dir que $f_n \xrightarrow{*} f$ si $Jx(f_n) \rightarrow Jx(f)$ per a tot $x \in X$, o equivalentment, si $F(f_n) \rightarrow F(f)$ per a tot $F \in J(X)$.

Observació 5. Podem veure que, en general, la convergència feble en X' implica la convergència $*$ -feble; i en el cas dels espais reflexius, són equivalents, ja que $J(X) = X''$.

Als cursos bàsics d'Anàlisi Matemàtica, en treballar en espais de dimensió finita, podem trobar resultats meravellosos, com el Teorema de Bolzano-Weierstrass, que ens diu que tot successió fitada té una subsuccessió convergent. Aquest resultat no és vàlid en espais de dimensió infinita, però podem veure com amb les noves nocions de convergència, podem establir un resultat equivalent si ens conformam amb la convergència feble. Primer, anem a introduir el següent teorema, que ens serà d'utilitat per demostrar el resultat desitjat:

Teorema 2.1.2. Sigui X un espai normat separable i $\{f_n\} \subset X'$ una successió fitada. Llavors existeix una subsuccessió de $\{f_n\}$ que convergeix $*$ -feblement.

Demostració. Consultar secció 1.10 de [5]. □

Volem adequar aquest teorema al cas dels espais reflexius amb la convergència feble, però abans de fer-ho necessitem un resultat previ.

Teorema 2.1.3. Sigui X un espai normat tal que X' és separable. Llavors X és separable.

Demostració. Sigui $\{f_n\}$ un subconjunt dens i numerable de X' . Per definició de norma dual, existeix un $\{x_n\} \subset X$ tal que $\|x_n\|_X = 1$ per a tot $n \in \mathbb{N}$ i

$$|f_n(x_n)| \geq \frac{1}{2} \|f_n\|_*, \text{ o equivalentment } \|f_n\|_* \leq 2|f_n(x_n)| \quad (2.2)$$

Sigui D l'espai vectorial sobre \mathbb{Q} generat pels x_n . Podem escriure'l com $D = \cup_{n \in \mathbb{N}} D_n$, on

$$D_n = \left\{ \sum_{j=1}^n \alpha_j x_j \mid \alpha_j \in \mathbb{Q} \right\}$$

i té cardinal menor o igual a \mathbb{Q}^n , deduïnt així que D és numerable per ser unió numerable de conjunts numerables.

Volem veure que D és dens en X . Sigui L l'espai vectorial sobre \mathbb{R} generat pels x_n . Llavors D és dens en L ja que donat un $x \in L$ expressat com $x = \sum_{j=1}^n r_j x_j$ amb $r_j \in \mathbb{R}$, podem trobar una successió $\{\alpha_j^k\} \subset \mathbb{Q}$ tal que $\alpha_j^k \rightarrow r_j$ i $\sum_{j=1}^n \alpha_j^k x_j \rightarrow x$ quan $k \rightarrow \infty$ per tot $j = 1, \dots, n$. Per tant per veure que D és dens en X , veurem que L és dens en X .

Per veure que L és dens, pel corol·lari 2.1.1.1.(iii) basta veure que $L^\perp = \{0\}$. Sigui $f \in L^\perp$, tendrem que $f(x) = 0$ per a tota $x \in L$. Ara bé, com $\{f_n\}$ és dens en X' , $\forall \epsilon > 0 \exists f_n$ tal que $\|f - f_n\|_* \leq \epsilon$, usant (2.2) tenim que

$$\|f_n\|_* \leq 2|f_n(x_n)| = 2|(f_n - f)(x_n)| \leq 2\epsilon.$$

2. CONCEPTES BÀSICS D'ANÀLISI FUNCIONAL

D'aquí obtenim que

$$\|f\|_* \leq \|f_n\|_* + \|f - f_n\| \leq 3\epsilon.$$

Com la desigualtat és certa $\forall \epsilon > 0$, i f és arbitrari, tenim que $f = 0$ i per tant $L^\perp = \{0\}$, és a dir, L és dens en X . \square

Ara, gràcies als dos teoremes anteriors i a l'equivalència de la convergència feble en X' i la $*$ -feble en espais reflexius, podem demostrar el resultat desitjat.

Teorema 2.1.4. *Sigui X un espai normat reflexiu. Llavors tota successió fitada en X admet una subsuccessió feblement convergent.*

Demostració. Sigui $\{x_n\}$ una successió fitada en X . Denotem per M la clausura de l'espai generat pels $\{x_n\}$. Usant el mateix raonament que a la demostració del teorema anterior amb l'espai L , es té que el subespai M és separable. A més, M és un subespai tancat d'un espai reflexiu, i per la proposició 2.1.7, M és reflexiu, és a dir, $M \equiv M''$. Per últim, $M \equiv M'' = (M')'$ és separable i és l'espai dual de M' , pel teorema 2.1.3, M' també és separable.

Llavors, com $\{x_n\}$ és una successió fitada en M , la successió $\{Jx_n\}$ està fitada en M'' per ser J una isometria. Així, pel teorema 2.1.2, existirà $\{Jx_{n_k}\}$ una subsuccessió que convergeix $*$ -feblement a un $y \in M''$. Ara bé com M és reflexiu, $M \equiv M''$ via J i per tant existirà un únic $\hat{x} \in M$ tal que $J\hat{x} = y$. Llavors, com

$$Jx_{n_k} \xrightarrow{*} J\hat{x},$$

per definició de convergència $*$ -feble:

$$Jx_{n_k}(g) \rightarrow J\hat{x}(g) \quad \forall g \in M',$$

i per la definició de la injecció canònica:

$$g(x_{n_k}) \rightarrow g(\hat{x}) \quad g \in M',$$

que significa que $x_{n_k} \rightarrow \hat{x}$ en M . Ara bé, si prenem un $f \in X'$ qualsevol, podem considerar la restricció $f|_M \in M'$. Així $f(x_{n_k}) = f|_M(x_{n_k}) \rightarrow f|_M(\hat{x}) = f(\hat{x})$. Per tant, $\{x_{n_k}\}$ convergeix feblement a \hat{x} en X . \square

2.2 Espais de Hilbert

Definició 2.2.1. *Sigui H un espai vectorial. Un producte escalar $\langle x, y \rangle$ és una forma bilineal simètrica de $H \times H$ en \mathbb{R} tal que:*

- (i) $\langle x, x \rangle \geq 0$ per a tot $x \in H$.
- (ii) $\langle x, x \rangle = 0$ si i només si $x = 0$.

Direm que un espai vectorial és pre-Hilbert si està dotat d'un producte escalar.

Proposició 2.2.1. *Sigui H un espai vectorial pre-Hilbert. Llavors l'aplicació definida per $\|x\| = \sqrt{\langle x, x \rangle}$ és una norma. A més, verifica la desigualtat de Cauchy-Schwarz, és a dir, que per a tot $x, y \in H$, es verifica*

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

Una vegada hem definit el concepte de producte escalar, i hem vist que d'aquest podem generar una norma, podem parlar d'espais de Hilbert.

Definició 2.2.2. *Un espai pre-Hilbert complet respecte la norma definida pel producte escalar s'anomena espai de Hilbert.*

Els espais de Hilbert van néixer de la necessitat de poder extender les bones propietats de l'espai euclidià a problemes de dimensió infinita. Per això, en aquests espais podem parlar, per exemple, de relacions d'ortogonalitat.

Definició 2.2.3. *Sigui H un espai pre-Hilbert. Direm que $x, y \in H$ són ortogonals, i ho denotarem per $x \perp y$, si $\langle x, y \rangle = 0$. Si $G \subset H$ és un subconjunt, el conjunt ortogonal de G és $G^\perp = \{x \in H \mid \langle x, y \rangle = 0 \ \forall y \in G\}$.*

I amb les nocions d'ortogonalitat, podem parlar de la projecció ortogonal d'un element sobre conjunt convex.

Teorema 2.2.1 (Teorema de la Projecció). *Sigui $K \subset H$ un subconjunt convex tancat i no buit d'un espai de Hilbert. Llavors per a tot $x \in H$ existeix un únic $y_0 \in K$ tal que*

$$\|x - y_0\| = d(x, K) = \inf_{y \in K} \|x - y\|.$$

A més, aquest y_0 és l'únic element tal que $x - y_0 \perp K$.

Definició 2.2.4. *L'element y_0 donat pel teorema anterior s'anomena projecció d' x sobre K , i el denotarem per $y_0 = \Pi_K x_0$.*

Una altra de les principals propietats dels espais de Hilbert és la possibilitat de representar cada funcional de l'espai dual com a producte escalar d'elements de l'espai original, així com diu el següent teorema:

Teorema 2.2.2 (Representació de Riesz-Fréchet). *Sigui H un espai de Hilbert. Donat un $f \in H'$, $\exists! y \in H$ tal que $f(x) = \langle y, x \rangle \ \forall x \in H$. A més, $\|y\|_H = \|f\|_{H'}$.*

Demostració. Vegem primer que $\|y\|_H = \|f\|_{H'}$. Considerem l'aplicació

$$\begin{aligned} T : H &\longrightarrow H' \\ y &\longmapsto f : H \longrightarrow \mathbb{R} \\ x &\longmapsto \langle y, x \rangle, \end{aligned}$$

que és una aplicació lineal. Fixat y tenim que $f(x) = \langle y, x \rangle$ per a tot $x \in H$. Així, usant la definició de norma dual i la desigualtat de Cauchy-Schwarz:

$$\|f\|_{H'} = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|_H} = \sup_{x \neq 0} \frac{|\langle y, x \rangle|}{\|x\|_H} \leq \sup_{x \neq 0} \frac{\|y\|_H \|x\|_H}{\|x\|_H} = \|y\|_H.$$

Usant la definició de norma tenim que:

$$\|y\|_H^2 = \langle y, y \rangle = |\langle y, y \rangle| = |f(y)| \leq \|f\|_{H'} \|y\|_H \implies \|y\|_H \leq \|f\|_{H'}.$$

Per tant $\|y\|_H = \|f\|_{H'}$. Vegem ara que realment $\exists y \in H$ que verifica la propietat desitjada.

2. CONCEPTES BÀSICS D'ANÀLISI FUNCIONAL

Si $\ker(f) = H$, per $y \equiv 0$ tendrem que $f(x) = \langle 0, x \rangle = 0$ per a tota $x \in H$. Ara bé, si $M = \ker(f) \neq H$, aleshores $M^\perp \neq \{0\}$, i sabem que al manco existirà un $v \in M^\perp$ tal que $v \neq 0$ i $\|v\| = 1$ (si no fos unitari, podríem normalitzar-lo). Així, per a tot $x \in H$, podem construir $m = x - \frac{f(x)}{f(v)} v$ que verifica $m \in M$, ja que $f(m) = 0$. Aleshores, com $v \in M^\perp$ i $m \in M$:

$$\langle \alpha v, m \rangle = 0 \quad \forall \alpha \in \mathbb{R}.$$

En particular, si $\alpha = f(v)$:

$$0 = \left\langle f(v)v, x - \frac{f(x)}{f(v)}v \right\rangle = f(v)\langle v, x \rangle - f(x)\langle v, v \rangle.$$

Així, com v és unitari podem reescriure

$$f(x) = f(v)\langle v, x \rangle = \langle f(v)v, x \rangle.$$

Per tant, per $y = f(v)v$ podem escriure $f(x) = \langle y, x \rangle$. A més, aquest y serà únic, ja que si existissin dos $y_1, y_2 \in H$ tals que

$$f(x) = \langle y_1, x \rangle = \langle y_2, x \rangle \quad \forall x \in H,$$

tendríem que

$$\langle y_1 - y_2, x \rangle = 0 \quad \forall x \in H,$$

i aleshores $y_1 = y_2$. □

El teorema de representació de Riesz-Fréchet diu que tot funcional de l'espai dual té un representant a l'espai original. D'aquesta forma podem establir una relació entre H i H' . Amb el següent teorema, veurem que també podem identificar H' amb H'' , i d'aquí podem deduir la reflexivitat dels espais de Hilbert.

Teorema 2.2.3. *Sigui H un espai de Hilbert. Llavors H és reflexiu*

Demostració. Sigui H un espai de Hilbert. Pel teorema de Representació de Riesz-Fréchet (2.2.2), tenim que existeix una aplicació T_1 donada per:

$$\begin{aligned} T_1 : H &\longrightarrow H' \\ y &\longmapsto f_y : H \longrightarrow \mathbb{R} \\ x &\longmapsto \langle y, x \rangle, \end{aligned}$$

que és un isomorfisme isomètric, és a dir, $\|y\|_H = \|f_y\|_{H'}$. Ara bé, sabem que H' és un espai de Banach dotat de la norma de l'espai dual. Podem definir en H' el producte escalar com $\langle f_y, f_u \rangle := \langle y, u \rangle$, on y i u són els únics elements de H que verifiquen $f_y(x) = \langle y, x \rangle$ i $f_u(x) = \langle u, x \rangle$ per a tot $x \in H$ respectivament. Llavors pel teorema de Representació de Riesz-Fréchet (2.2.2) tenim que

$$\|f\|_{H'} = \|y\|_H = \sqrt{\langle y, y \rangle} = \sqrt{\langle f, f \rangle},$$

és a dir, que la norma de H' prové del producte escalar. Aleshores H' és un espai de Hilbert i tornant a utilitzar el teorema (2.2.2), podem establir l'isomorfisme

$$\begin{aligned} T_2 : H' &\longrightarrow H'' \\ f &\longmapsto F_f : H' \longrightarrow \mathbb{R} \\ g &\longmapsto \langle f, g \rangle \end{aligned}$$

amb $\|f\|_{H'} = \|F\|_{H''}$. Llavors la composició

$$\begin{aligned} T_2 \circ T_1 : H &\longrightarrow H'' \\ y &\longmapsto F_{f_y} : H' \longrightarrow \mathbb{R} \\ f_u &\longmapsto \langle f_y, f_u \rangle = \langle y, u \rangle = f_u(y) \end{aligned}$$

serà un isomorfisme on

$$\|F_{f_y}\|_{H''} = \|T_2(T_1(y))\|_{H''} = \|T_1(y)\|_{H'} = \|y\|_H$$

A més, notem que $(T_2 \circ T_1(y))(f) = F_{f_y}(f) = f(y) = J_y(f)$, on J denota la injecció canònica de H en H'' . Llavors la injecció canònica és un isomorfisme isomètric entre H i H'' i, per tant, H és un espai reflexiu. \square

ELS ESPAIS DE SOBOLEV, $W^{1,p}$

Els espais de Sobolev duen el nom del matemàtic soviètic Sergei Lvovich Sobolev (1908 - 1989). La seva feina al món de les matemàtiques es concentrà principalment en l'anàlisi funcional i les equacions en derivades parcials. En aquests espais, basats en els espais L^p , s'afebleixen les condicions de diferenciabilitat i es permet trobar el que es coneixen com les solucions febles de les equacions diferencials, o també demostrar l'existència d'alguns problemes del càlcul de variacions. Cal esmentar que malgrat que els espais portin el seu nom, matemàtics com Beppo Levi o L. Tonelli varen treballar anteriorment amb funcions d'aquest tipus, i la majoria de les propietats d'aquests espais foren desenvolupades per Sobolev, així com per J. W. Calkin i C.B. Morrey. En aquest capítol, definirem i estudiarem les propietats bàsiques dels espais L^p per després poder usar-los i construir els espais $W^{1,p}$.

3.1 Un tast de teoria de la mesura i integració

Abans d'entrar de ple en els espais $W^{1,p}$, necessitam recordar alguns conceptes bàsics de teoria de la mesura i alguns resultats d'integració que ens seran útils per desenvolupar els principals resultats dels espais L^p . Pel que fa a la teoria de la mesura i les funcions mesurables, recordem que:

Definició 3.1.1. *Un espai de mesura és un triplet $(\Omega, \mathcal{M}, \mu)$ tal que:*

- Ω és un conjunt.
- \mathcal{M} és una σ -àlgebra en Ω , és a dir, \mathcal{M} és una família de subconjunts de Ω tals que:
 - (a) $\emptyset \in \mathcal{M}$,
 - (b) Si $A \in \mathcal{M}$, llavors $A^C \in \mathcal{M}$,
 - (c) Si $A_n \in \mathcal{M} \forall n$, llavors $\cup_{n=1}^{\infty} A_n \in \mathcal{M}$.
- μ és una mesura, és a dir, és una aplicació $\mu: \mathcal{M} \rightarrow [0, \infty]$ que satisfà:

- (a) $\mu(\emptyset) = 0$,
- (b) $\mu(\cup_{n \geq 1} A_n) = \sum_{n \geq 1} \mu(A_n)$ per a tota família $\{A_n\} \subset \mathcal{M}$ de subconjunts disjunts dos a dos.

Definició 3.1.2. Direm que un subconjunt $E \in \mathcal{M}$ és de mesura nul·la si $\mu(E) = 0$. Direm que una propietat es satisfà quasi per tot (qpt) si es satisfà sobre tot Ω excepte en un conjunt de mesura nul·la.

Definició 3.1.3. Donats dos espais mesurables, (Ω, \mathcal{M}) i (Ψ, \mathcal{M}') , direm que una funció $f : \Omega \rightarrow \Psi$ és una funció mesurable si $f^{-1}(E) \in \mathcal{M}'$ per a tot $E \in \mathcal{M}$.

Definició 3.1.4. Sigui $(\Omega, \mathcal{M}, \mu)$ un espai de mesura. Denotarem per $L^1(\Omega)$ l'espai de les funcions de Ω en \mathbb{R} mesurables i integrables, és a dir:

$$L^1(\Omega) = \left\{ f : \Omega \rightarrow \mathbb{R} \mid f \text{és mesurable i } \int_{\Omega} |f| d\mu < \infty \right\}.$$

A més, direm que

$$\|f\|_{L^1} = \|f\|_1 = \int_{\Omega} |f| d\mu$$

En particular, en aquest treball estudiam principalment espais de funcions reals de variable real. En aquest cas, podem definir en concret què és una funció real mesurable.

Definició 3.1.5. Sigui $(\Omega, \mathcal{M}, \mu)$ un espai de mesura i $f : \Omega \rightarrow [-\infty, +\infty]$. Direm que f és una funció real mesurable si $f^{-1}(E) \in \mathcal{M}$ per a tot $E \in \mathcal{B}(\mathbb{R})$, on $\mathcal{B}(\mathbb{R})$ és la σ -àlgebra de Borel¹.

Un tipus de funcions que necessitarem per la secció de Càlcul de Variacions són les funcions de Carathéodory, definides a continuació.

Definició 3.1.6. Sigui $\Omega \subset \mathbb{R}^N$ obert i sigui $f : \Omega \times \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$. Llavors f és una funció de Carathéodory si:

- (i) $f(x, \cdot)$ és contínua qpt $x \in \Omega$,
- (ii) $f(\cdot, \xi)$ és mesurable per a tota $\xi \in \mathbb{R}^n$

Vegem ara alguns dels principals resultats d'integració de funcions L^1 que necessitarem per demostrar alguns dels resultats dels espais L^p .

Teorema 3.1.1 (Convergència monòtona, Beppe-Levi). Sigui $\{f_n\}$ una successió de funcions L^1 que satisfan:

- (i) $f_1 \leq f_2 \leq \dots \leq f_n \leq f_{n+1} \leq \dots$ qpt en Ω ,
- (ii) $\sup_n \int f_n < \infty$.

Llavors si $f_n \rightarrow f$ qpt en Ω , la funció f pertany a L^1 i $\|f_n - f\|_1 \rightarrow 0$.

Teorema 3.1.2 (Lema de Fatou). Sigui $\{f_n\}$ una successió de funcions mesurables no negatives sobre un conjunt measurable E . Aleshores:

$$\int_E \left(\liminf_n f_n \right) d\mu \leq \liminf_n \int_E f_n d\mu.$$

¹Recordem que la σ -àlgebra de Borel és la generada pels oberts de $[-\infty, +\infty]$

3.2 Els espais de Lebesgue, $L^p(\Omega)$

3.2.1 Definició i propietats bàsiques

Els espais de funcions $L^p(\Omega)$ estan vinculats a una mesura. Nosaltres considerarem la mesura de Lebesgue, μ . Ara, una vegada fixada la nostra mesura, podem definir els espais $L^p(\Omega)$. Primer ho farem per $1 \leq p < \infty$ i llavors per $p = \infty$.

Definició 3.2.1. Sigui $p \in \mathbb{R}$ amb $1 \leq p < \infty$. Definim els espais de Lebesgue L^p com

$$L^p(\Omega) = \left\{ f : \Omega \rightarrow \mathbb{R} \text{ mesurables } \left| \int_{\Omega} |f|^p < \infty \right. \right\}.$$

Amb la norma

$$\|f\|_{L^p} = \|f\|_p = \left(\int_{\Omega} |f|^p \right)^{1/p}.$$

Definició 3.2.2. Definim l'espai de Lebesgue L^∞ com

$$L^\infty(\Omega) = \left\{ f : \Omega \rightarrow \mathbb{R} \text{ mesurables } \left| \exists C \text{ tal que } |f(x)| \leq C \text{ qpt en } \Omega \right. \right\}.$$

Amb la norma

$$\|f\|_{L^\infty} = \|f\|_\infty = \inf_C \{ |f(x)| \leq C \text{ qpt en } \Omega \}.$$

Realment treballarem amb l'espai quocient $L^p(\Omega)/\sim$ per $1 \leq p \leq \infty$, on \sim denota la relació d'equivalència $f \sim g$ si, i només si, $f(x) = g(x)$ qpt $x \in \Omega$, ja que sinó $\|\cdot\|_p$ seria una seminorma i no una norma. Per simplicitat, ho denotarem per $L^p(\Omega)$.

D'ara ençà, per a $1 \leq p \leq \infty$, denotarem l'*exponent conjugat de p* com q , és a dir, el $q \in \mathbb{R}$ que satisfà

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Anem a enunciar un lema tècnic que ens servirà per demostrar la desigualtat de Hölder:

Lema 3.2.1 (Desigualtat de Young). Siguin $a, b \in \mathbb{R}$ tals que $a, b \geq 0$. Siguin p, q exponents conjugats tals que $p, q > 1$. Aleshores

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q$$

Demostració. Utilitzant la concavitat del logaritme, prenent $t = \frac{1}{p}$ i $1 - t = \frac{1}{q}$ tenim que

$$\begin{aligned} \ln \left(\frac{1}{p}a^p + \frac{1}{q}b^q \right) &= \ln(ta^p + (1-t)b^q) \geq t \ln a^p + (1-t) \ln b^q \\ &= \ln a + \ln b = \ln(ab) \end{aligned}$$

D'aquí se'n dedueix la desigualtat desitjada, per ser el logaritme una funció creixent. \square

Teorema 3.2.2 (Desigualtat de Hölder). Siguin $1 \leq p, q \leq \infty$ exponents conjugats. Siguin $f \in L^p(\Omega)$ i $g \in L^q(\Omega)$. Aleshores $fg \in L^1(\Omega)$ i es satisfà

$$\|fg\|_1 \leq \|f\|_p \|g\|_q$$

3. ELS ESPAIS DE SOBOLEV, $W^{1,p}$

Demostració. Si $f = 0$ o $g = 0$, el resultat és directe i a més es satisfà la igualtat. Suposem que $f, g \neq 0$. Considerem

$$a = \frac{|f(x)|}{\|f\|_p} \quad \text{i} \quad b = \frac{|g(x)|}{\|g\|_q}.$$

Clarament $a, b \geq 0$. Llavors, usant la desigualtat de Young (lema 3.2.1) amb a i b tenim que

$$ab \leq \frac{1}{p} a^p + \frac{1}{q} b^q \implies \frac{|f(x)|}{\|f\|_p} \cdot \frac{|g(x)|}{\|g\|_q} \leq \frac{1}{p} \frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{q} \frac{|g(x)|^q}{\|g\|_q^q}.$$

Integrant ambdós costats de la desigualtat sobre Ω :

$$\frac{\|fg\|_1}{\|f\|_p \|g\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1$$

D'on aïllant arribam a $\|fg\|_1 \leq \|f\|_p \|g\|_q$. \square

Cal esmentar que la desigualtat de Hölder també té una versió per sumes finites o infinites, i es demostra anàlogament al cas integral:

$$\sum_n |a_n b_n| \leq \left(\sum_n |a_n|^p \right)^{1/p} \left(\sum_n |b_n|^q \right)^{1/q}.$$

Gràcies a la desigualtat de Hölder, podem provar una altra desigualtat pròpia dels espais $L^p(\Omega)$, la desigualtat de Minkowski:

Teorema 3.2.3 (Desigualtat de Minkowski). *Siguin $1 \leq p \leq \infty$ i $f, g \in L^p(\Omega)$. Aleshores $f + g \in L^p(\Omega)$ i es satisfà:*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Demostració. Per començar, vegem que $f + g \in L^p(\Omega)$. Utilitzant que la funció $y = |x|^p$ és convexa, de la definició de convexitat, prenent $\lambda = \frac{1}{2}$, tenim que

$$\left| \frac{f(x)}{2} + \frac{g(x)}{2} \right|^p \leq \frac{|f(x)|^p}{2} + \frac{|g(x)|^p}{2} \implies \frac{1}{2^p} |f + g|^p \leq \frac{1}{2} (|f|^p + |g|^p).$$

Integrant ambdós costats, obtenim que

$$\frac{1}{2^p} \int_{\Omega} |f + g|^p \leq \frac{1}{2} \left(\int_{\Omega} |f|^p + \int_{\Omega} |g|^p \right) < \infty,$$

per ser $f, g \in L^p(\Omega)$. Llavors la seva suma també pertany a $L^p(\Omega)$. Anem ara a provar la desigualtat. Podem descomposar

$$|f + g|^p = |f + g| |f + g|^{p-1} \leq |f + g|^{p-1} (|f| + |g|).$$

Integrant ambdós costats de la desigualtat sobre Ω i usant la desigualtat de Hölder (teorema 3.2.2) tenim que

$$\|f + g\|_p^p \leq \int_{\Omega} |f + g|^{p-1} (|f| + |g|) \leq \left(\int_{\Omega} |f + g|^{(p-1)q} \right)^{1/q} (\|f\|_p + \|g\|_p),$$

on $p = q(p-1)$. Sigui $I = (\int_{\Omega} |f + g|^p)^{1/q}$, tenim dos casos diferents segons el valor de I :

(i) Si $I = 0$, $\int_{\Omega} |f + g|^p = 0$ i per tant es té la igualtat.

(ii) Si $0 < I < \infty$, aïllant es té que

$$\left(\int_{\Omega} |f + g|^p \right)^{1-1/q} \leq \left(\int_{\Omega} |f|^p \right)^{1/p} + \left(\int_{\Omega} |g|^p \right)^{1/p},$$

que és el resultat desitjat, ja que $1 - 1/q = 1/p$

□

A conseqüència de les desigualtats demostrades, tenim el següent resultat fonamental referent als espais $L^p(\Omega)$.

Corol·lari 3.2.3.1. *Sigui $1 \leq p \leq \infty$. Llavors $L^p(\Omega)$ és un espai vectorial, $\|\cdot\|_p$ és una norma i $(L^p, \|\cdot\|_p)$ és un espai normat.*

Fins ara hem vist que els espais $L^p(\Omega)$ són espais normats. En el següent teorema veurem que, a més, aquests espais són complets respecte a les normes $\|\cdot\|_p$ definides per $1 \leq p \leq \infty$.

Teorema 3.2.4 (Riesz-Fischer). *Sigui $1 \leq p \leq \infty$. Llavors $(L^p, \|\cdot\|_p)$ és un espai de Banach.*

Observació 6. *A partir d'ara ens centrarem a estudiar els espais $L^p(\Omega)$ per $1 < p < \infty$. Per això només demostrarem aquest resultat pel cas $1 \leq p < \infty$, encara que també existeix la demostració pel cas $p = \infty$. La demostració es pot trobar a la tercera secció del capítol 2 de [4].*

Demostració. Sigui $\{f_n\} \subset L^p$ una successió de Cauchy. Si prenem $\epsilon_i = \frac{1}{2^i}$, existeix un $n_i \in \mathbb{N}$ tal que $\|f_n - f_m\|_p < \frac{1}{2^i}$ per a tots $n, m \geq n_i$. Llavors podem extreure una subsuccessió $\{f_{n_i}\}$ tal que $\|f_{n_{i+1}} - f_{n_i}\|_p < \frac{1}{2^i}$ per a tot $i \geq 1$.

Definim ara $g_k = \sum_{i=1}^k |f_{n_{i+1}} - f_{n_i}|$. Com $|f_{n_{i+1}} - f_{n_i}| \in L^p$, llavors $g_k \in L^p$ i així:

$$\|g_k\|_p = \left\| \sum_{i=1}^k |f_{n_{i+1}} - f_{n_i}| \right\|_p \leq \sum_{i=1}^k \|f_{n_{i+1}} - f_{n_i}\|_p < \sum_{i=1}^k \frac{1}{2^i} < 1$$

Com $\{g_k\}$ és una successió creixent de funcions mesurables no negatives (per ser suma de termes positius) i fitades, el teorema de la convergència monòtona (3.1.1) ens assegura que convergeix puntualment qpt a un límit finit $g(x)$ amb $g \in L^p$, on $g = \lim_k \sum_{i=1}^k |f_{n_{i+1}} - f_{n_i}| = \sum_{i=1}^{\infty} |f_{n_{i+1}} - f_{n_i}|$.

Com g és finit, tenim que la sèrie $\sum_{i=1}^{\infty} (f_{n_{i+1}} - f_{n_i})$ és absolutament convergent puntualment, així com ho serà $f_{n_1} + \sum_{i=1}^{\infty} (f_{n_{i+1}} - f_{n_i})$. Si definim la funció

$$f(x) = f_{n_1}(x) + \sum_{i=1}^{\infty} (f_{n_{i+1}}(x) - f_{n_i}(x)) \quad \text{qpt},$$

al poder reescriure $f_{n_1} + \sum_{i=1}^{k-1} (f_{n_{i+1}} - f_{n_i}) = f_{n_k}$, tenim que

$$f(x) = \lim_{k \rightarrow \infty} f_{n_k}(x)$$

3. ELS ESPAIS DE SOBOLEV, $W^{1,p}$

Vegem que $f \in L^p$ i que $\|f_n - f\|_p \rightarrow 0$. Donat $\epsilon > 0$ existeix un $N \in \mathbb{N}$ tal que $\|f_n - f_m\|_p < \epsilon$ per a tot $n, m \geq N$. Si prenem $m \geq N$, usant el lema de Fatou (3.1.2):

$$\begin{aligned} \|f - f_m\|_p^p &= \int_{\Omega} |f - f_m|^p = \int_{\Omega} \left| \lim_{k \rightarrow \infty} (f_{n_k} - f_m) \right|^p = \int_{\Omega} \lim_{k \rightarrow \infty} |f_{n_k} - f_m|^p \\ &\leq \liminf_k \int_{\Omega} |f_{n_k} - f_m|^p = \liminf_k \|f_{n_k} - f_m\|_p^p \leq \epsilon^p \end{aligned}$$

per a $n_k \geq N$. Llavors $f - f_m \in L^p$, i com $f = (f - f_m) + f_m$, llavors $f \in L^p$. A més $\|f - f_m\|_p \leq \epsilon$ per a tot $m \geq N$. \square

Exemple 3.2.1. Ja hem vist que els espais $L^p(\Omega)$ per $1 \leq p \leq \infty$ són espais de Banach. Si consideram l'espai $L^2(\Omega)$ dotat del producte escalar

$$\langle f, g \rangle = \int_{\Omega} f(x)g(x)dx,$$

tenim que és un espai de Hilbert. És fàcil veure que la norma L^2 prové del producte escalar donat, ja que:

$$\|f\|_2 = \left(\int_{\Omega} |f(x)|^2 dx \right)^{1/2} = \sqrt{\langle f, f \rangle}$$

3.2.2 El dual de $L^p(\Omega)$. Reflexivitat

Així com als espais de Hilbert teníem una forma d'identificar un espai amb el seu dual a través del teorema de representació de Riesz-Fréchet, en els espais L^p tenim un teorema de representació similar.

Teorema 3.2.5 (Teorema de representació de Riesz). Sigui $1 < p < \infty$ i $\varphi \in (L^p)'$. Llavors existeix una única $g \in L^q$ tal que

$$\varphi(f) = \int_{\Omega} fg$$

per a tota $f \in L^p$. A més, $\|g\|_q = \|\varphi\|_{(L^p)'}$

Demostració. Considerem l'aplicació

$$\begin{aligned} T : L^q &\longrightarrow (L^p)' \\ g &\longmapsto \varphi_g : L^p \longrightarrow \mathbb{R} \\ f &\longmapsto \int_{\Omega} fg, \end{aligned}$$

Volem veure que T és un isomorfisme isomètric lineal. Vegem primer que és lineal. Siguin $a, b \in \mathbb{R}$ i $g_1, g_2 \in L^q$. Llavors, per a cada $f \in L^p$

$$\begin{aligned} T(ag_1 + bg_2)(f) &= \varphi_{ag_1 + bg_2}(f) = \int_{\Omega} f(ag_1 + bg_2) = a \int_{\Omega} fg_1 + b \int_{\Omega} fg_2 \\ &= a\varphi_{g_1}(f) + b\varphi_{g_2}(f) = aTg_1(f) + bTg_2(f) \end{aligned}$$

Vegem ara que la norma es conserva. Usant la desigualtat de Hölder (teorema 3.2.2), tenim que

$$|\varphi_g(f)| = \left| \int_{\Omega} fg \right| \leq \int_{\Omega} |f||g| \leq \|f\|_p \|g\|_q$$

Arrel d'aquesta desigualtat i de la definició de norma dual obtenim que $\|\varphi_g\|_{(L^p)'} \leq \|g\|_q$. Per veure l'altra desigualtat, considerem la funció $f = g|g|^{q-2}$ amb $g \in L^q$. Notem que $f \in L^p$, ja que

$$\int_{\Omega} |f|^p = \int_{\Omega} |g|^p |g|^{p(q-2)} = \int_{\Omega} |g|^p |g|^{q-p} = \int_{\Omega} |g|^q < \infty.$$

A més

$$\varphi_g(f) = \int_{\Omega} g|g|^{q-2}g = \int_{\Omega} g^2|g|^{q-2} = \int_{\Omega} |g|^q = \|g\|_q^q.$$

Per tant, usant la definició de norma dual

$$\|\varphi_g\|_{(L^p)'} = \sup_{\substack{u \in L^p \\ u \neq 0}} \frac{|\varphi_g(u)|}{\|u\|_p} \geq \frac{|\varphi_g(f)|}{\|f\|_p} = \frac{\|g\|_q^q}{(\|g\|_q^q)^{1/p}} = \|g\|_q^{q-(q/p)} = \|g\|_q.$$

Llavors $\|\varphi_g\|_{(L^p)'} = \|g\|_q$. Falta veure que T és bijectiva. Com T és una isometria, tenim que també és injectiva. Ara bé, per a poder comprovar l'exhausitivitat s'empren conceptes i resultats de teoria de la mesura com el teorema de Radon-Nikodym, que surten dels continguts d'aquest treball. \square

Per tant, en els espais L^p , podem identificar el seu espai dual amb L^q , on recordem que q és l'exponent conjugat de p . I així com en els espais de Hilbert el teorema de representació de Riesz-Fréchet (2.2.2) ens permetia provar la seva reflexivitat, en els espais $L^p(\Omega)$ passa el mateix amb el teorema de representació de Riesz (3.2.5).

Teorema 3.2.6. *Els espais L^p per $1 < p < \infty$ són reflexius.*

Demostració. Pel teorema de Representació de Riesz (3.2.5), per a tot $\varphi \in (L^p)'$ existeix una única $g \in L^q$ tal que $\varphi(f) = \varphi_g(f) = \int_{\Omega} fg$ per a tota $f \in L^p$. Així podem definir la injecció canònica com

$$\begin{aligned} J : L^p &\longrightarrow (L^p)'' \\ f &\longmapsto Jf : (L^p)' \longrightarrow \mathbb{R} \\ \varphi_g &\longmapsto Jf(\varphi_g) = \varphi_g(f) = \int_{\Omega} fg. \end{aligned}$$

Volem veure que J és exhaustiva, és a dir, que $\forall F \in (L^p)'' \exists f \in L^p$ tal que $Jf = F$. Això per definició vol dir que per a tota $\varphi_g \in (L^p)'$ s'ha de satisfet que $Jf(\varphi_g) = \varphi_g(f) = F(\varphi_g)$. Si prenem $F \in (L^p)''$, tenim que $F : (L^p)' \rightarrow \mathbb{R}$ i llavors pel teorema de Representació de Riesz (3.2.5), que permet identificar $(L^p)'$ amb L^q , podem pensar F com $F : L^q \rightarrow \mathbb{R}$ definida com $F(g) = F(\varphi_g)$. Aquesta aplicació per construcció és lineal i fitada, i per tant $F \in (L^q)'$. Novament pel teorema (3.2.5), existeix una única $f \in L^p$ tal que $F = \psi_f$. Llavors per a tota $g \in L^q$ tenim

$$F(\varphi_g) = F(g) = \psi_f(g) = \int_{\Omega} fg = \varphi_g(f) = Jf(\varphi_g).$$

Lo que prova la exhaustivitat de J . \square

3.2.3 El cas del producte cartesià

De cara a desenvolupar els espais de Sobolev, algunes propietats seran més senzilles de demostrar si consideram $W^{1,p}$ com a subespai tancat del producte cartesià d'espais L^p .

Definició 3.2.3. Definim el producte cartesià d'espais L^p com $L_N^p(\Omega) = \prod_{j=1}^N L^p(\Omega)$, format per elements del tipus $u = (u_1, \dots, u_N)$, on cada $u_j \in L^p(\Omega)$.

Proposició 3.2.1. L'espai L_N^p , per $1 \leq p \leq \infty$, és un espai de Banach dotat de les normes equivalents

$$\|u\|_{L_N^p} = \left(\sum_{j=1}^N \|u_j\|_p^p \right)^{1/p} \quad \text{si } 1 \leq p < \infty, \quad \text{o} \quad \|u\|_{L_N^\infty} = \max_{1 \leq j \leq N} \|u_j\|_\infty$$

Demostració. Conseqüència directa de la proposició 2.1.5. \square

Quan ens trobam amb L_N^p , també podem generalitzar resultats importants com el Teorema de Representació de Riesz (3.2.5). Vegem-ho.

Teorema 3.2.7. Sigui $1 < p < \infty$. Per a tot $\varphi \in (L_N^p)'$ existeix un únic $g \in L_N^q$ tal que per a tot $f \in L_N^p$ es té que

$$\varphi(f) = \sum_{j=1}^N \int_{\Omega} f_j g_j.$$

A més $\|\varphi\|_{(L_N^p)'} = \|g\|_{L_N^q}$ i per tant $(L_N^p)' \equiv L_N^q$.

Demostració. Sigui $w \in L^p$. Definim $w_{(j)} = (0, \dots, 0, w, 0, \dots, 0)$, és a dir, l'element de L_N^p on la component j -èssima val w i la resta valen 0. Definim $\varphi_j(w) = \varphi(w_{(j)})$. Així, $\varphi_j \in (L^p(\Omega))'$, i pel teorema 3.2.5 existeix un únic $g_j \in L^q$ tal que per a tot $w \in L^p$:

$$\varphi(w_{(j)}) = \varphi_j(w) = \int_{\Omega} w g_j.$$

Així, per qualsevol $f \in L_N^p$ podem escriure:

$$\varphi(f) = \varphi \left(\sum_{j=1}^N f_j w_{(j)} \right) = \sum_{j=1}^N \varphi(f_j w_{(j)}) = \sum_{j=1}^N \int_{\Omega} f_j g_j$$

Ara usant la desigualtat de Hölder dues vegades: primer per a funcions de L^p i després per sumes finites:

$$|\varphi(f)| \leq \sum_{j=1}^N \|f_j\|_p \|g_j\|_p \leq \|f\|_{L_N^p} \|g\|_{L_N^q}$$

Així $\|\varphi\|_{(L_N^p)'} \leq \|g\|_{L_N^q}$, on $g = (g_1, \dots, g_N)$. Per veure l'altra desigualtat, definim $f_j = g_j \cdot |g_j|^{q-2}$ per $j = 1, \dots, N$ i, anàlogament a la demostració del teorema 3.2.5, arribam a veure que $\|\varphi\|_{(L_N^p)'} \geq \|g\|_{L^q}$. \square

Ara que tenim una versió del Teorema de Representació de Riesz pel cas del producte cartesià, igual que en el cas dels espais L^p , aquest pot ser utilitzat per veure que el producte cartesià d'espais L^p també és reflexiu.

Teorema 3.2.8. Els espais L_N^p per $1 < p < \infty$ són reflexius.

3.3 Els espais de Sobolev, $W^{1,p}(\Omega)$

3.3.1 Definició i propietats bàsiques

Siguin u i v funcions derivables definides en $[a, b]$. Si integraren per parts, obtenim:

$$\int_a^b u(x)v'(x)dx = u(x)v(x) \Big|_a^b - \int_a^b v(x)u'(x)dx.$$

Si, a més, v s'anula a la frontera o és de suport compacte en $[a, b]$, tendrem que

$$\int_a^b u(x)v'(x)dx = - \int_a^b v(x)u'(x)dx.$$

Ara bé, si u no és derivable, què podem fer? No podem aplicar el teorema d'integració per parts tal com el coneixem, però en alguns casos es pot trobar alguna funció, diguem-li g , tal que

$$\int_a^b u(x)v'(x)dx = - \int_a^b v(x)g(x)dx.$$

Aquesta g es coneix com la derivada feble de u . Això es pot generalitzar per a funcions de diverses variables. Aquesta idea de derivada feble d'una funció ens permet construir els espais de Sobolev de la següent forma:

Definició 3.3.1. Sigui $\Omega \subset \mathbb{R}^N$ un obert, i sigui $p \in \mathbb{R}$ amb $1 \leq p \leq \infty$. Definim l'espai de Sobolev $W^{1,p}$ com

$$W^{1,p}(\Omega) = \left\{ u \in L^p(\Omega) \left| \begin{array}{l} \exists g_1, g_2, \dots, g_N \in L^p(\Omega) \quad \text{tals que} \\ \int_{\Omega} u \frac{\partial \varphi}{\partial x_i} = - \int_{\Omega} g_i \varphi \quad \forall \varphi \in C_c^{\infty}(\Omega), \forall i = 1, 2, \dots, N \end{array} \right. \right\}.$$

A més, per $u \in W^{1,p}(\Omega)$, definim $\frac{\partial u}{\partial x_i} = g_i$. Aquests espais són espais normats dotats de la norma

$$\|u\|_{W^{1,p}} = \|u\|_p + \sum_{i=1}^N \left\| \frac{\partial u}{\partial x_i} \right\|_p.$$

Observació 7. A vegades es sol emprar la norma equivalent

$$\|u\|_{W^{1,p}} = \left(\|u\|_p^p + \sum_{i=1}^N \left\| \frac{\partial u}{\partial x_i} \right\|_p^p \right)^{1/p}.$$

Observació 8. Recordem que $C_c^{\infty}(\Omega)$ és l'espai de funcions infinitament diferenciables amb suport compacte, i ve definit com

$$C_c^{\infty}(\Omega) = \left\{ f \in C^{\infty}(\Omega) \mid \text{supp}(f) = \overline{\{x \in \Omega \mid f(x) \neq 0\}} \text{ és compacte} \right\}.$$

Notem que els espais de Sobolev estan definits a partir de funcions de L^p . Això significa que aquestes funcions ja satisfan algunes de les bones propietats d'aquests espais, com per exemple les desigualtats de Hölder i Minkowski. Gràcies a aquest fet, les demostracions d'aquesta secció se simplifiquen bastant. Primerament, vegem que els espais de Sobolev són espais normats complets.

3. ELS ESPAIS DE SOBOLEV, $W^{1,p}$

Proposició 3.3.1. *L'espai de Sobolev $W^{1,p}$ és un espai de Banach per tot $1 \leq p \leq \infty$*

Demostració. Sigui $\{u_n\} \subset W^{1,p}(\Omega)$ una successió de Cauchy. Aleshores per definició de l'espai de Sobolev $\{u_n\} \subset L^p(\Omega)$ i $\{\nabla u_n\} \subset (L^p(\Omega))^N$ són successions de Cauchy i, en concret, $\{\frac{\partial u_n}{\partial x_i}\} \subset L^p(\Omega)$ també és de Cauchy.

Pel teorema de Riesz-Fischer (3.2.4), $(L^p(\Omega), \|\cdot\|_p)$ és un espai de Banach, i per tant tota successió de Cauchy hi convergeix. Així, tenim que $\exists u \in L^p(\Omega)$ tal que $\|u_n - u\|_p \rightarrow 0$, i que $\exists g_i \in L^p(\Omega)$ tal que $\|\frac{\partial u_n}{\partial x_i} - g_i\|_p \rightarrow 0$ per tot $i = 1, \dots, N$. Ara $\forall \varphi \in C_c^\infty(\Omega)$, $\forall i = 1, 2, \dots, N$, usant la desigualtat de Hölder, per una banda tenim que:

$$\left| \int_{\Omega} u_n \frac{\partial \varphi}{\partial x_i} - \int_{\Omega} u \frac{\partial \varphi}{\partial x_i} \right| \leq \int_{\Omega} |u_n - u| \cdot \left| \frac{\partial \varphi}{\partial x_i} \right| \leq \|u_n - u\|_p \left\| \frac{\partial \varphi}{\partial x_i} \right\|_q \xrightarrow{n \rightarrow \infty} 0,$$

i per l'altra que:

$$\left| \int_{\Omega} \frac{\partial u_n}{\partial x_i} \varphi - \int_{\Omega} g_i \varphi \right| \leq \int_{\Omega} \left| \frac{\partial u_n}{\partial x_i} - g_i \right| \cdot |\varphi| \leq \left\| \frac{\partial u_n}{\partial x_i} - g_i \right\|_p \|\varphi\|_q \xrightarrow{n \rightarrow \infty} 0.$$

Per tant, usant que $\{u_n\} \subset W^{1,p}(\Omega)$, per definició:

$$\int_{\Omega} u_n \frac{\partial \varphi}{\partial x_i} = - \int_{\Omega} \frac{\partial u_n}{\partial x_i} \varphi \quad \forall \varphi \in C_c^\infty(\Omega), \forall i = 1, 2, \dots, N;$$

i si prenem límits a ambdós costats quan $n \rightarrow \infty$, arribam a:

$$\int_{\Omega} u \frac{\partial \varphi}{\partial x_i} = - \int_{\Omega} g_i \varphi \quad \forall \varphi \in C_c^\infty(\Omega), \forall i = 1, 2, \dots, N;$$

Per tant $u \in W^{1,p}(\Omega)$ i, a més, $\frac{\partial u}{\partial x_i} = g_i$ per tot $i = 1, \dots, N$. Així concloem que $W^{1,p}(\Omega)$ és un espai de Banach. \square

Quan estudiàvem els espais L^p , vam veure que en L^2 la norma provenia d'un producte escalar. En el cas dels espais de Sobolev, per $p = 2$ també podem definir un producte escalar i veure que és un espai de Hilbert.

Exemple 3.3.1. *L'espai de Sobolev $W^{1,2}$ és un espai de Hilbert dotat del producte escalar*

$$\langle u, v \rangle = \int_{\Omega} u(x) v(x) dx + \sum_{i=1}^N \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i},$$

que induceix la norma

$$\|u\|_{W^{1,2}} = \left(\|u\|_2^2 + \sum_{i=1}^N \left\| \frac{\partial u}{\partial x_i} \right\|_2^2 \right)^{1/2} = \sqrt{\langle u, u \rangle}.$$

Normalment l'espai de Hilbert $W^{1,2}(\Omega)$ es sol denotar per $H^1(\Omega)$.

El nostre objectiu és desenvolupar la teoria necessària dels espais de Sobolev per a poder provar l'existència de solució de problemes del Càcul de Variacions. Per tant, ens serà especialment útil provar que aquests espais són reflexius, perquè així pel teorema 2.1.4 tendrem que tota successió fitada tendrà una subsuccessió feblement convergent.

Proposició 3.3.2. *L'espai de Sobolev $W^{1,p}$ és un espai reflexiu per $1 < p < \infty$.*

Demostració. Considerem l'operador lineal i fitat

$$\begin{aligned} T : W^{1,p} &\longrightarrow L^p \times L_N^p \\ u &\longmapsto (u, \nabla u) \end{aligned}$$

Aquest operador és una isometria, ja que

$$\|(u, \nabla u)\|_{L_{N+1}^p} = \left(\|u\|_p^p + \sum_{j=1}^N \left\| \frac{\partial u}{\partial x_j} \right\|_p^p \right)^{1/p} = \|u\|_{W^{1,p}}.$$

Així, com $W^{1,p}$ és un espai de Banach, la seva imatge $T(W^{1,p}) \equiv W^{1,p}$, també ho serà. Llavors, tenim que és tancat per ser un subconjunt complet d'un espai complet. Per tant, com $L^p \times L_N^p = L_{N+1}^p$ és un espai normat reflexiu i $W^{1,p}$ és un subconjunt tancat, per la proposició 2.1.7 tenim que aquest és reflexiu. \square

3.3.2 Desigualtats de Sobolev. L'espai $W_0^{1,p}(\Omega)$

Anem a introduir una nova definició d'inclusió de conjunts.

Definició 3.3.2. *Siguin X, Y espais normats amb $X \subset Y$. Direm que la injecció de X en Y és compacta si la injecció canònica és contínua, és a dir, si*

$$\|u\|_Y \leq C\|u\|_X \quad \forall u \in X,$$

i, a més, tota successió fitada en X té una subsuccessió de Cauchy en Y .

Observació 9. *En el cas on Y és un espai complet, l'injecció compacta serà un operador compacte, és a dir, que tota successió fitada en X té una subsuccessió convergent en Y*

En general, la injecció $W^{1,p} \subset L^p$ és certa. El següent teorema ens dirà sota quines condicions serà compacta.

Teorema 3.3.1 (Rellich-Kondrachov). *Siguui $\Omega \subset \mathbb{R}^N$ un obert fitat amb frontera Lipschitz. Llavors les següents injeccions són compactes:*

$$\begin{aligned} W^{1,p}(\Omega) &\subset L^r(\Omega) & \forall r \in [1, \frac{pN}{N-p}), & \text{si } p < N \\ W^{1,p}(\Omega) &\subset L^r(\Omega) & \forall r \in [p, \infty), & \text{si } p = N \\ W^{1,p}(\Omega) &\subset C(\overline{\Omega}) & & \text{si } p > N \end{aligned}$$

A més, en general l'injecció de $W^{1,p}(\Omega)$ en $L^p(\Omega)$ és compacta per a tota $1 \leq p \leq \infty$ i per a tota N .

Demostració. Consultar Theorem 9.16 del capítol 9 de [7]. \square

A continuació definirem un nou espai que ens serà útil en tractar amb problemes variacionals.

Definició 3.3.3. *Siguui $1 \leq p < \infty$. Denotam per $W_0^{1,p}(\Omega)$ la clausura de $C_c^\infty(\Omega)$ en $W^{1,p}(\Omega)$. Denotarem $H_0^1(\Omega) = W_0^{1,2}(\Omega)$.*

3. ELS ESPAIS DE SOBOLEV, $W^{1,p}$

Fent un abús de notació, podríem dir que si Ω és compacte, les funcions $u \in W_0^{1,p}(\Omega)$ són “aproximadament” les funcions $u \in W^{1,p}(\Omega)$ que s’anulen a la frontera, $\partial\Omega$. En general, quan escrivim $u \in u_0 + W_0^{1,p}(\Omega)$ significa que $u, u_0 \in W^{1,p}(\Omega)$ i que $u - u_0 \in W_0^{1,p}$.

Proposició 3.3.3. *L’espai $W_0^{1,p}$, dotat de la norma $\|\cdot\|_{W^{1,p}}$, és un espai de Banach per $1 \leq p \leq \infty$, i és reflexiu per $1 < p < \infty$. A més, $W_0^{1,2} = H_0^1$ dotat del producte escalar de H^1 també és un espai de Hilbert.*

Una propietat especialment interessant en els espais $W_0^{1,p}$ és la coneguda com la desigualtat de Poincaré, que estableix relacions entre les normes de u i ∇u .

Teorema 3.3.2 (Desigualtat de Poincaré). *Sigui $\Omega \subset \mathbb{R}^N$ un obert fitat, i $1 \leq p < \infty$. Llavors existeix una constant $C > 0$ que depèn de Ω , N i p tal que es verifica*

$$\|u\|_p \leq C\|\nabla u\|_p, \quad \forall u \in W_0^{1,p}(\Omega)$$

o equivalentment

$$\|u\|_{W^{1,p}} \leq C\|\nabla u\|_p, \quad \forall u \in W_0^{1,p}(\Omega)$$

Demostració. Per ser $W_0^{1,p}(\Omega)$ la clausura de $C_c^\infty(\Omega)$ en $W^{1,p}(\Omega)$, considerem $u \in C_c^\infty(\Omega)$. Per ser Ω fitat podem suposar que $\Omega \subset Q = \{x \in \mathbb{R}^N : a \leq x_i \leq b, i = 1, 2, \dots, N\}$, és a dir, Q és l’hipercub $[a, b]^N$. Considerem

$$v(x) = \begin{cases} u(x) & x \in \Omega \\ 0 & x \in Q \setminus \Omega \end{cases}$$

Per a cada $x \in Q$ tenim que $x = (x_1, y)$ amb $y = (x_2, x_3, \dots, x_N)$. Així podem escriure:

$$v(x) = v(x_1, y) = v(a, y) + \int_a^{x_1} v_{x_1}(t, y) dt.$$

Notem que $v(a, y) = 0$ perquè v té suport compacte. Llavors aplicant la desigualtat de Hölder (3.2.2):

$$|v(x_1, y)| \leq \int_a^{x_1} |\nu_{x_1}(t, y)| dt \leq \left(\int_a^{x_1} 1^q dt \right)^{1/q} \left(\int_a^{x_1} |\nu_{x_1}(t, y)|^p dt \right)^{1/p}.$$

on q és l’exponent conjugat de p . Llavors

$$|\nu(x_1, y)|^p \leq (x_1 - a)^{p/q} \int_a^{x_1} |\nu_{x_1}(t, y)|^p dt \leq (x_1 - a)^{p/q} \int_a^b |\nu_{x_1}(t, y)|^p dt.$$

Sigui $Q' = \{(x_2, \dots, x_N) : a \leq x_i \leq b, i = 2, \dots, N\} = [a, b]^{N-1}$. Si ara integrarem els dos costats respecte de y sobre Q' , tenim

$$\int_{Q'} |\nu(x_1, y)|^p dy \leq (x_1 - a)^{p/q} \int_{Q'} \int_a^b |\nu_{x_1}(t, y)|^p dt dy = (x_1 - a)^{p/q} \int_Q |\nu_{x_1}(x)|^p dx.$$

Ara integrant els dos costats respecte de x_1 sobre $[a, b]$, ens queda

$$\int_Q |\nu(x)|^p dx \leq \frac{(b-a)^{p/q+1}}{p/q+1} \int_Q |\nu_{x_1}(x)|^p dx = \frac{(b-a)^p}{p} \int_Q |\nu_{x_1}(x)|^p dx$$

3.4. Introducció a l'espai de funcions de Variació Fitada

Repetint el mateix raonament per a cada x_i amb $i = 2, 3, \dots, N$, arribam

$$\int_Q |v(x)|^p dx \leq \frac{(b-a)^p}{p} \int_Q |v_{x_i}(x)|^p dx, \quad i = 1, 2, \dots, N.$$

Sumant ambdós costats respecte de i tenim

$$\int_Q |v(x)|^p dx \leq \frac{(b-a)^p}{Np} \int_Q \sum_{i=1}^N |v_{x_i}(x)|^p dx,$$

que ho podem reescriure com

$$\|v\|_p^p \leq \frac{(b-a)^p}{Np} \sum_{i=1}^N \|v_{x_i}\|_p^p = \frac{(b-a)^p}{Np} \|\nabla v\|_p^p,$$

Com que v i les seves derivades parcials valen 0 fora de Ω , si denotam $C \equiv C(\Omega, p, N) = (\frac{(b-a)^p}{Np})^{1/p}$, tenim el resultat desitjat per a tota $u \in C_c^\infty(\Omega)$:

$$\|u\|_p \leq C \|\nabla u\|_p$$

Ara bé, com $\overline{C_c^\infty(\Omega)} = W_0^{1,p}(\Omega)$, tenim que el resultat també és cert per a tota $u \in W_0^{1,p}(\Omega)$, ja que existirà una successió $\{u_n\} \subset C_c^\infty(\Omega)$ tal que $u_n \rightarrow u$ en $W_0^{1,p}(\Omega)$, i per la continuïtat de la norma tenim

$$\|u\|_p = \|\lim_{n \rightarrow \infty} u_n\|_p = \lim_{n \rightarrow \infty} \|u_n\|_p \leq C \lim_{n \rightarrow \infty} \|\nabla u_n\|_p = C \|\lim_{n \rightarrow \infty} \nabla u_n\|_p = C \|\nabla u\|_p$$

A més, podem enunciar la desigualtat de forma equivalent com:

$$\begin{aligned} \|u\|_p \leq C \|\nabla u\|_p &\iff \|u\|_p^p \leq C^p \|\nabla u\|_p^p + \|\nabla u\|_p^p - \|\nabla u\|_p^p \\ &\iff \|u\|_p^p + \|\nabla u\|_p^p \leq (C^p + 1) \|\nabla u\|_p^p \\ &\iff \|u\|_{W^{1,p}}^p \leq K \|\nabla u\|_p^p, \end{aligned}$$

on $K = (C^p + 1)^{1/p}$. □

La desigualtat de Poincaré es pot presentar de diverses formes diferents. A part de la que hem demostrat, també pot ser enunciada de la següent forma:

Corol·lari 3.3.2.1 (Desigualtat de Poincaré, v2). *Siguin $u_0 \in W^{1,p}(\Omega)$ i $u \in u_0 + W_0^{1,p}(\Omega)$. Llavors existeixen $C_1, C_2 > 0$ tals que*

$$\|\nabla u\|_p \geq C_1 \|u\|_{W^{1,p}} - C_2 \|u_0\|_{W^{1,p}}$$

3.4 Introducció a l'espai de funcions de Variació Fitada

Abans d'estudiar l'espai de les funcions de variació fitada, necessitam saber què és la variació total d'una funció.

Definició 3.4.1. *Donada una funció $u \in L^1(\Omega)$, la seva variació total és*

$$\int_\Omega |Du| = \sup_\varphi \left\{ \int_\Omega u \cdot \operatorname{div} \varphi dx \mid \varphi = (\varphi_1, \dots, \varphi_N) \in C_0^\infty(\Omega; \mathbb{R}^N), |\varphi(x)| \leq 1 \forall x \in \Omega \right\}.$$

Darem que una funció $u \in L^1(\Omega)$ és de variació fitada si la seva variació total és finita

3. ELS ESPAIS DE SOBOLEV, $W^{1,p}$

Les funcions de variació fitada també es poden definir com les funcions de L^1 les derivades parcials de les quals, en sentit distribucional, són mesures. Molts mètodes de processament d'imatges es basen en la minimització de funcionals. L'ús de la variació total com a terme de regularització es va emprar originalment per eliminar el renou de les imatges [8] i s'ha convertit en una de les regularitzacions més populars per a molts problemes en aquest camp gràcies a la seva capacitat per recuperar les discontinuïtats de la imatge, és a dir, els contorns. Això és la base de les seves nombroses aplicacions com l'eliminació de renou, segmentació, o interpolació, entre d'altres. L'espai de les funcions de variació fitada juga un paper clau quan es requereix una estimació de les discontinuïtats de la solució, i el podem definir de la següent forma:

Definició 3.4.2. *Definim l'espai de les funcions de variació fitada com*

$$BV(\Omega) = \left\{ u \in L^1(\Omega) \mid \int_{\Omega} |Du| < +\infty \right\}$$

Denotarem l'espai de funcions de Variació Fitada per $BV(\Omega)$ pel seu nom en anglès (*Bounded Variation*). Notem que aquest espai està construït sobre l'espai normat $L^1(\Omega)$, el qual hem vist que és un espai de Banach. Es pot comprovar, que dotat d'una norma adequada, l'espai $BV(\Omega)$ també és complet.

Proposició 3.4.1. *L'espai $BV(\Omega)$ és un espai de Banach respecte de la norma*

$$\|u\|_{BV} = \|u\|_1 + \int_{\Omega} |Du|.$$

Una propietat bàsica dels espais de Sobolev ens permet establir una relació entre ells i els espais de Variació Fitada. Si tenim una funció $u \in W^{1,p}$, aleshores:

$$\int_{\Omega} u \cdot \operatorname{div} \varphi = \int_{\Omega} u \sum_{i=1}^N \frac{\partial \varphi_i}{\partial x_i} = - \int_{\Omega} \sum_{i=1}^N \frac{\partial u}{\partial x_i} \varphi_i = - \int_{\Omega} \langle \nabla u, \varphi \rangle,$$

i fins i tot, es pot demostrar que

$$\int_{\Omega} |Du| = \int_{\Omega} |\nabla u| dx, \quad (3.1)$$

però això no és una tasca senzilla, perquè requereix entrar en l'estudi de la teoria de distribucions, que es troba fora de l'abast d'aquest treball. A més, si $u \in W^{1,p}(\Omega)$, on $\Omega \subset \mathbb{R}^N$ és un domini fitat, u serà una funció de variació fitada. Per definició, tendrem que $u, \frac{\partial u}{\partial x_i} \in L^p(\Omega)$ per a tot $i = 1, \dots, N$. I així, en general

$$\begin{aligned} \int_{\Omega} u \cdot \operatorname{div} \varphi &\leq \left| \int_{\Omega} u \sum_{i=1}^N \frac{\partial \varphi_i}{\partial x_i} \right| = \left| - \int_{\Omega} \sum_{i=1}^N \frac{\partial u}{\partial x_i} \varphi_i \right| \leq \sum_{i=1}^N \int_{\Omega} \left| \frac{\partial u}{\partial x_i} \right| |\varphi_i| \leq \sum_{i=1}^N \int_{\Omega} \left| \frac{\partial u}{\partial x_i} \right| \\ &\leq \sum_{i=1}^N \left(\int_{\Omega} \left| \frac{\partial u}{\partial x_i} \right|^p \right)^{1/p} \left(\int_{\Omega} |1|^q \right)^{1/q} = \mu(\Omega)^{1/q} \sum_{i=1}^N \left\| \frac{\partial u}{\partial x_i} \right\|_p < +\infty, \end{aligned}$$

i prenent supремs a ambdós costats tenim que

$$\int_{\Omega} |Du| = \sup_{\varphi} \left\{ \int_{\Omega} u \cdot \operatorname{div} \varphi dx \right\} \leq \mu(\Omega)^{1/q} \sum_{i=1}^N \left\| \frac{\partial u}{\partial x_i} \right\|_p < +\infty.$$

És a dir, que funcions de $W^{1,p}(\Omega)$ són funcions de variació fitada.

EL MÈTODE DIRECTE DEL CÀLCUL DE VARIACIONS

4.1 El problema de minimització

La branca del Càcul de Variacions, estudia problemes del següent tipus:

$$\inf \left\{ J(u) = \int_{\Omega} f(x, u(x), \nabla u(x)) dx \mid u \in X \right\} = m \quad (P)$$

on $\Omega \subset \mathbb{R}^N$ amb $N \geq 1$ és un subconjunt obert fitat; $u : \Omega \rightarrow \mathbb{R}^n$ amb $n \geq 1$ és una funció tal que $\nabla u \in \mathbb{R}^{N \times n}$; $f : \bar{\Omega} \times \mathbb{R}^n \times \mathbb{R}^{N \times n} \rightarrow \bar{\mathbb{R}}$ és una funció contínua; i X és un espai de funcions. El nostre objectiu serà trobar una funció $\hat{u} \in X$ que minimitzi la integral del problema (P) . Tendrem dues formes de resoldre el problema: els mètodes clàssics o els mètodes directes.

Els mètodes clàssics són l'equivalent al procés de derivar i igualar a 0 que es fa en el càlcul diferencial en dimensió finita. En dimensió infinita, calculam i resolem el que es coneix com a l'*equació d'Euler-Lagrange*:

$$\sum_{i=1}^n \frac{\partial}{\partial x_i} \left[f_{u_{x_i}}(x, u, \nabla u) \right] = f_u(x, u, \nabla u).$$

Resoldre l'equació no sol ser una tasca senzilla, i en cas de resoldre-la, per poder determinar si les solucions \hat{u} de l'equació són mínims o no, haurem de recórrer a estudiar la convexitat de f .

Per evitar les complicacions que poden sorgir durant els mètodes clàssics, introduïm els mètodes directes. En aquests mètodes, els espais de Sobolev $W^{1,p}$ introduïts al capítol 3 juguen un paper clau. En aquests mètodes, consideram una successió minimitzant $\{u_n\} \subset W^{1,p}$, és a dir, una successió tal que

$$J(u_n) \rightarrow \inf \left\{ J(u) \mid u|_{\partial\Omega} = u_0 \text{ i } u \in W^{1,p}(\Omega) \right\} = m, \quad \text{quan } n \rightarrow \infty.$$

Llavors hem de demostrar que aquesta successió té una subsuccessió convergent i gràcies a un cert tipus de continuïtat de J , podrem veure que el límit és el mínim. Això però, no és una tasca senzilla i haurem de recórrer a les nocions de convergència feble introduïdes en el capítol (2)

4.2 El mètode directe

4.2.1 Convexitat, coercivitat i semi-continuïtat

Per a poder demostrar la convergència de la successió minimitzant utilitzada en el mètode directe, s'han de complir certes condicions. Primer introduirem unes nocions bàsiques de convexitat.

Definició 4.2.1. *Un subconjunt U d'un espai vectorial X és convex si $\forall x, y \in U$ es satisfà que $\{tx + (1 - t)y \mid t \in [0, 1]\} \subset U$.*

Definició 4.2.2. *Un funcional $J : U \rightarrow \bar{\mathbb{R}}$, on U és un subconjunt convex d'un espai vectorial X , es diu convex si*

$$J(tx + (1 - t)y) \leq tJ(x) + (1 - t)J(y)$$

per a tot $x, y \in U$ i per $t \in [0, 1]$. Si la desigualtat és estricta, direm que el funcional és estrictament convex.

D'aquí en endavant, quan ens referirem a un funcional convex, suposarem que el seu domini de definició és tot l'espai vectorial X o bé un subconjunt convex de X .

Definició 4.2.3. *El domini efectiu d'un funcional $J : X \rightarrow \bar{\mathbb{R}}$ es defineix com*

$$\text{dom } J = \{u \in X : J(u) < \infty\}.$$

Si $\text{dom } J \neq \emptyset$ direm que J és un funcional propi.

A continuació, donarem la definició de coercivitat, una propietat que ens ajudarà a veure que una successió està fitada.

Definició 4.2.4. *Un funcional $J : X \rightarrow \bar{\mathbb{R}}$ és coerciu sobre X si*

$$\lim_{\|u\| \rightarrow \infty} J(u) = \infty$$

Finalment, per veure que un funcional té mínim, necessitam alguna condició de continuïtat. Ara bé, com estam cercant un mínim, realment és suficient amb veure que un funcional és semi-continu inferiorment.

Definició 4.2.5. *Sigui X un espai de Banach, $u_0 \in X$ i un funcional $J : X \rightarrow \bar{\mathbb{R}}$.*

(i) *Direm que J és (seqüencialment) semi-continu inferiorment en u_0 si per a tota successió $\{u_n\} \subset X$ convergent a u_0 (és a dir, $u_n \rightarrow u_0$), es té que*

$$J(u_0) \leq \liminf_{n \rightarrow \infty} J(u_n)$$

(ii) Direm que J és (seqüencialment) feblement semi-continu inferiorment en u_0 si per a tota successió $\{u_n\} \subset X$ convergent feblement a u_0 (és a dir, $u_n \rightharpoonup u_0$), es té que

$$J(u_0) \leq \liminf_{n \rightarrow \infty} J(u_n)$$

Clarament, si J és continu, llavors serà semi-continu, encara que el recíproc no sempre és cert. A més, si J és un funcional feblement semi-continu inferiorment, també serà semi-continu inferiorment, però el recíproc no és cert en general.

4.2.2 El principi variacional. Teorema general d'existència

Una vegada sabem què és un funcional convex, coerciu i semi-continu, podem enunciar el resultat conegut com el *Principi Variacional*.

Teorema 4.2.1 (Principi Variacional). Sigui X un espai de Banach reflexiu. Sigui $J : X \rightarrow \overline{\mathbb{R}}$ un funcional feblement semi-continu inferiorment. Suposem què:

- J és propi,
- J està fitat inferiorment,
- i J és coerciu.

Aleshores existeix un $\hat{u} \in X$ tal que $J(\hat{u}) = \inf_{u \in X} J(u)$. A més, si J és estrictament convex, aquest \hat{u} és únic.

Demostració. Per ser J un funcional fitat inferiorment, existeix una constant $m > -\infty$ tal que $m = \inf_{u \in X} J(u)$. Prenem $\{u_n\}$ una successió minimitzant de J , és a dir,

$$J(u_n) \xrightarrow{n \rightarrow \infty} m.$$

Com, a més, J és coerciu, necessàriament $\{u_n\}$ ha de ser una successió fitada, ja que si no ho fos,

$$m = \lim_{n \rightarrow \infty} J(u_n) = \infty,$$

i això contradiria el fet que J sigui propi. Llavors al ser $\{u_n\}$ una successió fitada en un espai reflexiu, té una subsuccessió feblement convergent a un $\hat{u} \in X$. Per tant, per la semicontinuitat inferior de J , tendrem que

$$J(\hat{u}) \leq \liminf_{k \rightarrow \infty} J(u_{n_k}) = m,$$

i això només pot passar si

$$J(\hat{u}) = m = \inf_{u \in X} J(u).$$

A més, si el funcional és estrictament convex, aquesta solució \hat{u} és única. Si existís \hat{v} una altra solució, és a dir, que satisfés $J(\hat{v}) = m$, podríem construir $\hat{w} = \frac{1}{2}\hat{u} + \frac{1}{2}\hat{v}$. Aquest \hat{w} satisfaria

$$m \leq J(\hat{w}) = J\left(\frac{1}{2}\hat{u} + \frac{1}{2}\hat{v}\right) < \frac{1}{2}J(\hat{u}) + \frac{1}{2}J(\hat{v}) = m,$$

que és una contradicció. Per tant la solució és única. \square

Ara bé, no sempre és una tasca senzilla comprovar la semi-continuitat inferior feble i la coercivitat d'un funcional del tipus (P) . Per això, enunciam el següent teorema, que ens dona una condició suficient (però no necessària) per a la semi-continuitat inferior d'un funcional.

Teorema 4.2.2. *Siguin $\Omega \subset \mathbb{R}^N$ un obert i $r, p \geq 1$. Sigui q l'exponent conjugat de p . Sigui $f : \Omega \times \mathbb{R}^n \times \mathbb{R}^M$ una funció de Carathéodory que satisfà*

$$f(x, u, \xi) \geq \langle a(x), \xi \rangle + b(x) + c|u|^r$$

qpt $x \in \Omega$ i $\forall (u, \xi) \in \mathbb{R}^n \times \mathbb{R}^M$; per $a \in L^q(\Omega; \mathbb{R}^M)$, $b \in L^1(\Omega)$ i $c \in \mathbb{R}$. Sigui

$$J(u, \xi) := \int_{\Omega} f(x, u(x), \xi(x)) dx.$$

Suposem, a més, que $f(x, u, \cdot)$ és convexa, que $u_n \rightarrow \hat{u}$ en $L^r(\Omega; \mathbb{R}^n)$ i que $\xi_n \rightarrow \hat{\xi}$ en $L^p(\Omega; \mathbb{R}^M)$. Aleshores

$$\liminf_{n \rightarrow \infty} J(u_n, \xi_n) \geq J(\hat{u}, \hat{\xi}).$$

Demostració. Es pot consultar al *Theorem 3.23* del capítol 3 de [9]. □

Aquest teorema és un resultat molt general, per funcionals on la tercera variable és una funció arbitrària. Ara bé, nosaltres treballam amb problemes on la tercera variable del funcional és el gradient de la funció que es troba a la segona variable. Per tant, el que farem a continuació serà afinar una mica aquest teorema per adaptar-lo al nostre problema en concret.

Corol·lari 4.2.2.1. *Siguin $p \geq 1$, $\Omega \subset \mathbb{R}^N$ un obert fitat amb frontera Lipschitz. Sigui q l'exponent conjugat de p . Sigui $f : \Omega \times \mathbb{R}^n \times \mathbb{R}^{N \times n} \rightarrow \bar{\mathbb{R}}$ una funció de Carathéodory que satisfà*

$$f(x, u, \xi) \geq \langle a(x), \xi \rangle + b(x) + c|u|^r$$

qpt $x \in \Omega$, $\forall (u, \xi) \in \mathbb{R}^n \times \mathbb{R}^{N \times n}$; per $a \in L^q(\Omega; \mathbb{R}^{N \times n})$, $b \in L^1(\Omega)$, $c \in \mathbb{R}$, $1 \leq r < \frac{Np}{N-p}$ si $p < N$ i $1 \leq r < \infty$ si $p \geq N$. Sigui

$$J(u) = \int_{\Omega} f(x, u(x), \nabla u(x)) dx.$$

Suposem, a més, que $f(x, u, \cdot)$ és convexa i que $u_n \rightarrow \hat{u}$ en $W^{1,p}(\Omega; \mathbb{R}^n)$. Aleshores

$$\liminf_{n \rightarrow \infty} J(u_n) \geq J(\hat{u})$$

Demostració. Pel teorema de Rellich-Kondrachov (3.3.1) tenim que la injecció $W^{1,p}(\Omega) \subset L^r(\Omega)$ és compacta, és a dir, que donada $\{u_n\}$ una successió fitada en $W^{1,p}(\Omega; \mathbb{R}^n)$, aquesta té una subsuccessió convergent. A més, com la injecció és contínua, tenim la següent desigualtat:

$$\|u_n\|_r \leq C \|u_n\|_{W^{1,p}} \quad \forall n \in \mathbb{N}.$$

Així, la successió també estarà fitada en $L^r(\Omega)$ i la subsuccessió $\{u_{n_k}\}$ serà convergent a un $\hat{u} \in L^r(\Omega; \mathbb{R}^n)$. Així, podem usar el teorema anterior (4.2.2) per arribar a

$$\liminf_{n \rightarrow \infty} J(u_n) \geq J(\hat{u}).$$

□

Ara que tenim les eines necessàries, podem demostrar el següent teorema. Aquest serà molt important, ja que ens garanteix l'existència i (amb hipòtesis de convexitat estricta) unicitat de solució dels problemes del tipus (P).

Teorema 4.2.3. *Sigui $\Omega \subset \mathbb{R}^N$ un obert fitat amb frontera Lipschitz. Considerem ara $f : \Omega \times \mathbb{R}^n \times \mathbb{R}^{N \times n} \rightarrow \bar{\mathbb{R}}$ una funció de Carathéodory que satisfà la condició de coercivitat*

$$f(x, u, \xi) \geq \alpha_1 |\xi|^p + \alpha_2 |u|^q + \alpha_3(x) \quad (4.1)$$

qpt $x \in \Omega$ i $\forall (u, \xi) \in \mathbb{R}^n \times \mathbb{R}^{N \times n}$; on $\alpha_3 \in L^1(\Omega)$, $\alpha_2 \in \mathbb{R}$, $\alpha_1 > 0$ i $p > q \geq 1$. Suposem, a més, que $f(x, u, \cdot)$ és convexa. Sigui el problema

$$\inf \left\{ J(u) = \int_{\Omega} f(x, u(x), \nabla u(x)) dx \mid u \in u_0 + W_0^{1,p}(\Omega; \mathbb{R}^n) \right\}, \quad (4.2)$$

on suposam $J(u_0) < \infty$. Aleshores el problema (4.2) assoleix el seu mínim. A més, si $f(x, \cdot, \cdot)$ és estrictament convexa qpt $x \in \Omega$, aleshores el mínim és únic.

Demostració. Dividirem la demostració del teorema en dues parts: l'existència de mínim, i la seva unicitat.

1. Existència. Denotarem

$$\inf \left\{ J(u) = \int_{\Omega} f(x, u(x), \nabla u(x)) dx \mid u \in u_0 + W_0^{1,p}(\Omega; \mathbb{R}^n) \right\} = m.$$

Com $J(u_0) < +\infty$, podem assegurar que $m < +\infty$. A més, per la condició de coercivitat (4.1) podem veure que f està fitada inferiorment. Primerament:

$$\begin{aligned} J(u) &= \int_{\Omega} f(x, u, \nabla u) dx \geq \int_{\Omega} \alpha_1 |\nabla u(x)|^p dx + \int_{\Omega} \alpha_2 |u(x)|^q dx + \int_{\Omega} \alpha_3(x) dx \\ &\geq \alpha_1 \|\nabla u\|_p^p - |\alpha_2| \|u\|_q^q - \|\alpha_3\|_1 \end{aligned}$$

Per una banda, notem que $|u|^q = (|u|^p)^{q/p} \in L^{p/q}$ ja que

$$\int_{\Omega} |(|u|^p)^{q/p}|^{p/q} = \int_{\Omega} |u|^p < \infty,$$

al ser $u \in L^p$. Ara com p/q i $p/(p-q)$ són exponents conjugats i podem considerar $1 \in L^{p/(p-q)}$, emprant la desigualtat de Hölder tenim que:

$$\|u\|_q^q = \int_{\Omega} |u|^q \leq \left(\int_{\Omega} (|u|^q)^{p/q} \right)^{q/p} \left(\int_{\Omega} 1^{p/(p-q)} \right)^{(p-q)/p} = \mu(\Omega)^{(p-q)/p} \|u\|_p^q$$

D'ara en endavant denotarem $\gamma_k > 0$ constants. Existeixen γ_1, γ_2 tals que

$$\begin{aligned} J(u) &\geq \alpha_1 \|\nabla u\|_p^p - \gamma_1 \|u\|_p^q - \gamma_2 \\ &\geq \alpha_1 \|\nabla u\|_p^p - \gamma_1 \|u\|_{W^{1,p}}^q - \gamma_2 \end{aligned}$$

Usant la segona versió de la desigualtat de Poincaré (3.3.2.1) sobre el terme amb $\|\nabla u\|_p^p$, tenim que existeixen $\gamma_3, \gamma_4, \gamma_5$ tals que:

$$J(u) \geq \gamma_3 \|u\|_{W^{1,p}}^p - \gamma_4 \|u_0\|_{W^{1,p}}^p - \gamma_1 \|u\|_{W^{1,p}}^q - \gamma_5.$$

4. EL MÈTODE DIRECTE DEL CÀLCUL DE VARIACIONS

Al ser $\|u_0\|_{W^{1,p}}^p$ constant, podem trobar una constant γ_6 tal que

$$J(u) \geq \gamma_3 \|u\|_{W^{1,p}}^p - \gamma_1 \|u\|_{W^{1,p}}^q - \gamma_6.$$

Com $1 \leq q < p$, podem trobar constants γ_7 i γ_8 suficientment gran tal que

$$J(u) \geq \gamma_7 \|u\|_{W^{1,p}}^p - \gamma_8, \quad (4.3)$$

fet que prova que J està fitat inferiorment, ja que tendrem que $J(u) \geq -\gamma_8$ i, per tant, $m > -\infty$.

Prenem ara $\{u_n\}$ una successió minimitzant de J , és a dir, $J(u_n) \rightarrow m$. Per a n prou gran, tenim

$$m + 1 \geq J(u_n).$$

Emprant la fita obtinguda en (4.3), tenim

$$m + 1 \geq \gamma_7 \|u_n\|_{W^{1,p}}^p - \gamma_8,$$

d'on es dedueix que $\{u_n\}_n$ està fitada

$$\|u_n\|_{W^{1,p}} \leq \left(\frac{m + 1 + \gamma_8}{\gamma_7} \right)^{1/p} = \gamma_9$$

Aleshores, com $1 < p < \infty$ i $W^{1,p}$ és reflexiu, podem extreure una subsuccessió feblement convergent, diem-li $\{u_{n_k}\}$. És a dir, existirà un $\hat{u} \in u_0 + W_0^{1,p}(\Omega; \mathbb{R}^N)$ tal que $u_{n_k} \rightharpoonup \hat{u}$.

Emprant el resultat anterior (corol·lari 4.2.2.1), tenim que

$$m = \liminf_{n \rightarrow \infty} J(u_n) \geq J(\hat{u}).$$

Per tant, \hat{u} minimitza (4.2).

2. Unicitat. Suposem que la solució no és única, és a dir, que existeixen $\hat{u}, \hat{v} \in u_0 + W_0^{1,p}(\Omega)$ tals que $J(\hat{u}) = m = J(\hat{v})$. Si definim $\hat{w} = \frac{1}{2}\hat{u} + \frac{1}{2}\hat{v}$, tenim que $\hat{w} \in u_0 + W_0^{1,p}(\Omega)$. Usant que $f(x, \cdot, \cdot)$ és estrictament convexa, tenim que

$$m \leq J(\hat{w}) \leq \frac{1}{2}J(\hat{u}) + \frac{1}{2}J(\hat{v}) = m, \quad (4.4)$$

és a dir, \hat{w} minimitza J . A més, (4.4) implica que

$$\frac{1}{2}J(\hat{u}) + \frac{1}{2}J(\hat{v}) - J(\hat{w}) = 0.$$

És a dir,

$$\int_{\Omega} \frac{1}{2}f(x, \hat{u}, \nabla \hat{u}) + \frac{1}{2}f(x, \hat{v}, \nabla \hat{v}) - f(x, \frac{1}{2}\hat{u} + \frac{1}{2}\hat{v}, \frac{1}{2}\nabla \hat{u} + \frac{1}{2}\nabla \hat{v}) dx = 0.$$

Ara bé, per la convexitat estricta de $f(x, \cdot, \cdot)$ tenim que l'integrand és no-negatiu, i l'únic cas on la integral podrà ser 0 serà si

$$\frac{1}{2}f(x, \hat{u}, \nabla \hat{u}) + \frac{1}{2}f(x, \hat{v}, \nabla \hat{v}) - f(x, \frac{1}{2}\hat{u} + \frac{1}{2}\hat{v}, \frac{1}{2}\nabla \hat{u} + \frac{1}{2}\nabla \hat{v}) = 0 \quad \text{qpt } x \in \Omega.$$

Ara bé, $f(x, \cdot, \cdot)$ és estrictament convexa i per tant

$$\frac{1}{2}f(x, \hat{u}, \nabla \hat{u}) + \frac{1}{2}f(x, \hat{v}, \nabla \hat{v}) - f(x, \frac{1}{2}\hat{u} + \frac{1}{2}\hat{v}, \frac{1}{2}\nabla \hat{u} + \frac{1}{2}\nabla \hat{v}) > 0 \quad \text{qpt } x \in \Omega,$$

que és una contradicció amb que l'integrand sigui 0 qpt $x \in \Omega$. Per tant la solució és única. \square

OPTIMITZACIÓ CONVEXA I ALGORISMES PRIMAL-DUAL

Al càlcul diferencial, quan treballam en \mathbb{R}^n amb funcions convexes i pròpies, sabem que si f pren un mínim en \hat{x} i, a més, f és diferenciable, es satisfà que $\nabla f(\hat{x}) = 0$. Sota aquestes hipòtesis, si volem trobar el mínim de la funció, un dels algorismes més emprats és el de descens del gradient. Sabem que el gradient d'una funció en un punt ens diu la direcció de màxim creixement. El mètode del descens del gradient es basa en aquesta idea i a cada passa es va movent en la direcció oposada fins a arribar al mínim. És a dir,

$$x^{k+1} = x^k - \tau \nabla f(x^k).$$

Però en els problemes del dia a dia, les funcions que descriuen com es mou el món no solen ser diferenciables, i necessitam nous mètodes per poder treballar amb elles. En aquesta secció considerarem X un espai de Hilbert, normalment (però no necessàriament) de dimensió finita.

5.1 Nocions bàsiques i propietats

5.1.1 Subdiferenciabilitat

Si tenim una funció f sobre \mathbb{R}^n que és convexa, pròpia i diferenciable en $x_0 \in \mathbb{R}^n$, l'únic hiperplà que conté $(x_0, f(x_0))$ i està sota f , és l'hiperplà tangent a f en $(x_0, f(x_0))$, que té per pendent $\nabla f(x_0)$. És a dir,

$$f(x) \geq f(x_0) + \langle p, x - x_0 \rangle \quad \forall x \in \mathbb{R}^n \iff p = \nabla f(x_0).$$

Si f no és diferenciable, això no és necessàriament cert. Ara bé, podem trobar punts p que sí que satisfacin la desigualtat desitjada, i per tant ens poden ser útils. Arrel d'aquesta idea neixen les nocions de *subgradient* i *subdiferencial*.

Definició 5.1.1. Sigui f una funció convexa.

5. OPTIMITZACIÓ CONVEXA I ALGORISMES PRIMAL-DUAL

(i) Un $p \in X$ és el subgradient de f en el punt $x \in X$ si

$$f(y) \geq f(x) + \langle p, y - x \rangle \quad \forall y \in X.$$

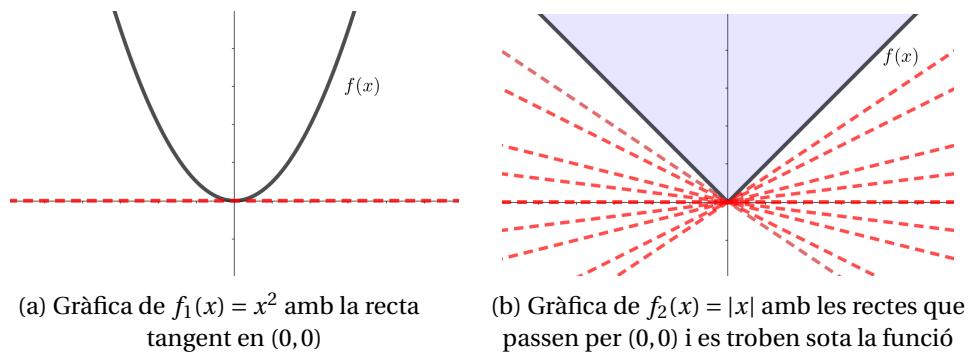
(ii) La subdiferencial de f en el punt x és el conjunt

$$\partial f(x) = \{p \in X \mid f(y) \geq f(x) + \langle p, y - x \rangle, \forall y \in X\}.$$

Observació 10. Si f és una funció convexa i diferenciable en X (o un subconjunt convex de X), llavors $\nabla f(x) \in \partial f(x)$.

Abans de continuar, vegem un exemple intuïtiu de què és la subdiferencial en el cas d'una funció real de variable real.

Exemple 5.1.1. Considerem les funcions $f_1(x) = x^2$ i $f_2(x) = |x|$. Són exemples clàssics de funcions diferenciable i no diferenciable, respectivament. De la primera podem calcular la diferencial en el 0, i la respectiva recta tangent, però de la segona no, així que vegem quina és la subdiferencial:



La subdiferencial del valor absolut en $x = 0$ és el conjunt format per totes les pendents de les rectes que passen per el $(0,0)$ que estan sota la funció. Per tant, $\partial f_2(0) = [-1, 1]$, que són totes les pendents d'aquestes rectes.

Una vegada sabem què és la subdiferencial, podem enunciar un dels teoremes més importants d'aquesta secció, que ens permetrà caracteritzar quan tendrà mínim o no una funció convexa.

Teorema 5.1.1 (Principi de Fermat). Un punt $\hat{x} \in X$ és un mínim global d'una funció f convexa si, i només si $0 \in \partial f(\hat{x})$.

Demostració. Usant la definició de subdiferencial, $0 \in \partial f(\hat{x})$ si i només si

$$f(y) \geq f(\hat{x}) + \langle 0, y - \hat{x} \rangle \quad \forall y \in X \iff f(y) \geq f(\hat{x}) \quad \forall y \in X.$$

És a dir, \hat{x} és un mínim global de f . □

Quan treballam amb la subdiferencial és molt important saber que aquesta no té les mateixes propietats de linealitat que la diferencial, com per exemple que la diferencial d'una suma és suma de diferencials. En els següent teorema, recopilam algunes de les propietats de la subdiferencial que ens seran més útils.

Teorema 5.1.2. *Siguin $f, g : X \rightarrow (-\infty, +\infty]$ funcions pròpies i convexes. Aleshores:*

(i) *Si existeixen un $x \in X$ tal que $\partial f(x) \neq \emptyset$ i $\alpha \geq 0$, aleshores $\partial(\alpha f)(x) = \alpha \partial f(x)$.*

(ii) *En general, per a tot $x \in X$*

$$\partial f(x) + \partial g(x) \subset \partial(f + g)(x)$$

(iii) *Si existeix $x \in \text{dom}(f)$ on g és contínua, aleshores*

$$\partial f(x) + \partial g(x) = \partial(f + g)(x)$$

(iv) *Si f és diferenciable en $x \in X$, aleshores $\partial f(x) = \{\nabla f(x)\}$*

(v) *Si f és diferenciable en $x \in X$ i $x \in \text{dom}(g)$, llavors*

$$\partial(f + g)(x) = \nabla f(x) + \partial g(x)$$

Demostració. Per (i), si $\alpha = 0$, $\partial(\alpha f)(x) = \{0\} = \alpha \partial f(x)$. Altrament, vegem les dues inclusions. Sigui $q \in \alpha \partial f(x)$. Llavors existeix un $p \in \partial f(x)$ tal que $\alpha p = q$. Llavors per a tot $y \in X$:

$$f(y) \geq f(x) + \langle p, y - x \rangle \implies \alpha f(y) \geq \alpha f(x) + \langle q, y - x \rangle \implies q \in \partial(\alpha f)(x).$$

En canvi, si prenem $q \in \partial(\alpha f)(x)$, per a tot $y \in X$ tenim

$$\alpha f(y) \geq \alpha f(x) + \langle q, y - x \rangle \implies f(y) \geq f(x) + \langle \frac{q}{\alpha}, y - x \rangle \implies \frac{q}{\alpha} \in \partial f(x) \implies q \in \alpha \partial f(x).$$

Anem a veure (ii). Sigui $p \in \partial f(x)$ i $q \in \partial g(x)$. Llavors per a tot $y \in X$

$$f(y) + g(y) \geq f(x) + g(x) + \langle p + q, y - x \rangle \implies p + q \in \partial(f + g)(x).$$

La demostració de (iii) es troba fora de l'abast d'aquest treball. Es pot consultar al punt 4.1.3 *Subgradient* de [10]. Per veure (iv), sabem que $\nabla f(x) \in \partial f(x)$. Falta veure l'altra inclusió. Si prenem $p \in \partial f(x)$, $t > 0$ i $y = x + th$, llavors per a tot $h \in X$ i per a tot $t > 0$:

$$f(x + th) \geq f(x) + \langle p, th \rangle \implies \frac{f(x + th) - f(x)}{t} \geq \langle p, h \rangle.$$

Si prenem límits quan $t \rightarrow 0$: $\langle \nabla f(x), h \rangle \geq \langle p, h \rangle \implies \langle \nabla f(x) - p, h \rangle \geq 0$. Això és cert per a tot $h \in X$, i en particular, per $-(\nabla f(x) - p)$. Per tant, tendrem que $\|\nabla f(x) - p\|^2 \leq 0$ i per tant $\nabla f(x) = p$.

Per a (v) basta aplicar (iii), ja que $x \in \text{dom}(g)$ i f és diferenciable en x , per tant és contínua en x , i (iv), ja que $\partial f(x) = \{\nabla f(x)\}$.

□

5.1.2 La conjugada de Legendre-Fenchel

A continuació introduïm la conjugada de Legendre-Fenchel. Les demostracions de les seves propietats es poden trobar al capítol 5 de [11].

Definició 5.1.2. *Donat X un espai de Banach qualsevol i $f : X \rightarrow [-\infty, +\infty]$, la conjugada convexa (o de Legendre-Fenchel) de f és una aplicació $f^* : X' \rightarrow [-\infty, +\infty]$ tal que*

$$f^*(\varphi) = \sup_{x \in X} \{\varphi(x) - f(x)\}, \quad \forall \varphi \in X'.$$

En concret, si X és un espai de Hilbert, fent ús del Teorema de Representació de Riesz-Fréchet (2.2.2), es té que

$$f^*(y) = \sup_{x \in X} \{\langle y, x \rangle - f(x)\}, \quad \forall y \in X.$$

Aquesta conjugada, satisfà les següents propietats:

Proposició 5.1.1. *Sigui X un espai de Hilbert, $f : X \rightarrow [-\infty, +\infty]$. Aleshores:*

- (i) f^* és convexa i semi-continua inferiorment.
- (ii) La biconjugada és $f^{**} = (f^*)^*$, i per tant $f^{**}(x) \leq f(x)$.

Però a l'hora de treballar problemes primal-duals, ens seran realment útils els següents teoremes:

Teorema 5.1.3 (Fenchel-Moreau). *Si f és convexa, pròpia i semi-contínua inferiorment, aleshores $f^{**}(x) = f(x)$.*

Teorema 5.1.4 (Desigualtat de Fenchel-Young). *Per a tot $x, y \in X$, $f(x) + f^*(y) \geq \langle y, x \rangle$.*

Teorema 5.1.5. *Sigui $f : X \rightarrow [-\infty, +\infty]$ convexa, pròpia i semi-continua inferiorment.*

- (i) *Aleshores existeix un $x \in X$ tal que*

$$f^*(y) = \langle y, x \rangle - f(x) \iff y \in \partial f(x) \text{ i } f(x) + f^*(y) = \langle x, y \rangle.$$

- (ii) *Es satisfà la identitat de Legendre-Fenchel:*

$$y \in \partial f(x) \iff x \in \partial f^*(y) \iff f(x) + f^*(y) = \langle x, y \rangle.$$

En últim lloc, veiem quines són les conjugades convexes d'una norma i del quadrat d'una norma arbitrària.

Proposició 5.1.2. *Sigui $f(x) = \frac{1}{2}\|x\|^2$. Llavors $f^*(y) = \frac{1}{2}\|y\|^2$*

Demostració. Sigui $f^*(y) = \sup_{x \in X} \{\langle y, x \rangle - \frac{1}{2}\|x\|^2\}$. Per la desigualtat de Cauchy-Schwarz:

$$\langle y, x \rangle - \frac{1}{2}\|x\|^2 \leq \|y\|\|x\| - \frac{1}{2}\|x\|^2.$$

El terme de la dreta és un polinomi de grau 2 en $\|x\|$, que pren el seu màxim en $\|x\| = \|y\|$. Així per a tot $y \in X$ es satisfà

$$\langle y, x \rangle - \frac{1}{2} \|x\|^2 \leq \frac{1}{2} \|y\|^2,$$

i prenent suprems a ambdós costats respecte x , tenim que $f^*(y) \leq \frac{1}{2} \|y\|^2$. Ara bé, per definició

$$f^*(y) \geq \langle y, y \rangle - \frac{1}{2} \|y\|^2 = \frac{1}{2} \|y\|^2.$$

Per tant $f^*(y) = \frac{1}{2} \|y\|^2$. □

Proposició 5.1.3. *Sigui $f(x) = \|x\|$. Llavors la seva conjugada és la funció indicatriu sobre la bolla unitària dual, és a dir:*

$$f^*(y) = \delta_{\{\|\cdot\|_* \leq 1\}} = \begin{cases} 0 & \text{si } \|y\|_* \leq 1 \\ \infty & \text{si } \|y\|_* > 1 \end{cases}$$

Demostració. Distingirem els dos casos. Primer, suposem que $\|y\|_* \leq 1$. Per la pròpia definició de norma dual:

$$\sup_{\|x\|_X \neq 0} \frac{\langle x, y \rangle}{\|x\|_X} = \|y\|_* \implies \langle x, y \rangle \leq \|x\|_X \|y\|_* \implies \langle x, y \rangle \leq \|x\|_X.$$

La darrera desigualtat és certa per a tot $x \in X$, en particular per $x = 0$, pel qual es satisfà la igualtat. Per tant,

$$f^*(y) = \sup_{x \in X} \{\langle x, y \rangle - \|x\|_X\} = 0.$$

Ara, suposant que $\|y\|_* > 1$, si usam una de les definicions equivalents de la norma dual:

$$\|y\|_* = \sup_{\|x\| \leq 1} \langle x, y \rangle > 1.$$

Llavors existeix un $x_0 \in X$ amb $\|x_0\|_X \leq 1$ tal que $\langle x_0, y \rangle > \|x_0\|_X$. Si definim $z = |\lambda|x_0$ per $\lambda \in \mathbb{R}$, es té que

$$\langle z, y \rangle - \|z\|_X = |\lambda| (\langle x_0, y \rangle - \|x_0\|_X) > 0,$$

que no està fitat per ser λ una constant arbitrària. Aleshores, com

$$f^*(y) = \sup_{x \in X} \{\langle x, y \rangle - \|x\|_X\} \geq |\lambda| (\langle x_0, y \rangle - \|x_0\|_X),$$

si $|\lambda| \rightarrow \infty$, llavors $f^*(y) \rightarrow \infty$. Per tant:

$$f^*(y) = \delta_{\{\|\cdot\|_* \leq 1\}} = \begin{cases} 0 & \text{si } \|y\|_* \leq 1 \\ \infty & \text{si } \|y\|_* > 1 \end{cases}$$

□

5.1.3 L'operador proximal

Quan tractam problemes de minimització, si la funció és derivable, el més usual és derivar la funció i igualar a zero per trobar punts crítics, i després estudiar el comportament de les derivades d'ordre superior. Però aquest mètode no sempre es pot emprar, i de vegades no és el més bo. En tals casos, es pot fer servir l'operador proximal.

Definició 5.1.3. Sigui $f : \mathbb{R}^n \rightarrow (-\infty, +\infty]$ pròpia, convexa i semi-contínua inferiorment. L'operador proximal de f amb paràmetre $\tau > 0$ és una aplicació $\text{prox}_{\tau f} : \mathbb{R}^n \rightarrow \mathbb{R}$ definida com

$$\text{prox}_{\tau f}(x) = \arg \min_y \left\{ f(y) + \frac{1}{2\tau} \|y - x\|_2^2 \right\}$$

L'operador proximal no cerca el mínim com a tal de la funció f , sinó que minimitzant la funció amb un cert terme de regularitat, es troba a mig camí entre minimitzar la funció i no allunyar-se massa del punt x . Usant paraules planeres, es podria dir que l'operador proximal cerca el mínim de f , però “fora fer llarg”. Ara bé, com podem saber quan trobarem un mínim emprant l'operador proximal?

Teorema 5.1.6. Direm que $\hat{x} \in X$ minimitza f si, i només si \hat{x} és un punt fix de $\text{prox}_{\tau f}$, és a dir,

$$\hat{x} = \text{prox}_{\tau f}(\hat{x})$$

Demostració. Sigui $h(x) = f(x) + \frac{1}{2\tau} \|x - \hat{x}\|_2^2$. Aleshores \hat{x} és un mínim de f si i només si

$$0 \in \partial f(\hat{x}) = \partial f(\hat{x}) + \frac{1}{2\tau}(\hat{x} - \hat{x}) = \partial h(\hat{x}).$$

És a dir, si i només si \hat{x} minimitza h , o equivalentment, si $\hat{x} = \text{prox}_{\tau f}(\hat{x})$. □

Per acabar, anem a calcular l'operador proximal de la funció indicatriu, que ens serà de gran utilitat a l'hora de trobar algorismes de minimització.

Proposició 5.1.4. Sigui $M \subset X$ un subespai tancat, convex i no buit. La funció indicatriu ve donada per

$$\delta_M(x) = \begin{cases} 0 & \text{si } x \in M \\ \infty & \text{si } x \notin M \end{cases}$$

Aleshores $\text{prox}_{\tau \delta_M}(x) = \Pi_M(x)$, on Π_M denota la projecció ortogonal sobre M , definida a (2.2.4).

Demostració. L'operador proximal ve donat per

$$\text{prox}_{\tau \delta_M}(x) = \arg \min_{y \in X} \left\{ \delta_M(y) + \frac{1}{2\tau} \|y - x\|_2^2 \right\} = \arg \min_{y \in M} \|x - y\|_2^2,$$

que és la definició de la projecció ortogonal sobre M en un espai de Hilbert. □

5.2 Algorismes Primal-Dual

5.2.1 Dualitat convexa

Siguin $f : \mathbb{R}^m \rightarrow \mathbb{R}$ i $g : \mathbb{R}^n \rightarrow \mathbb{R}$ funcions convexes i semi-contínues inferiorment. Sigui $K : \mathbb{R}^n \rightarrow \mathbb{R}^m$ un operador lineal. Definim el problema primal com

$$\min_{x \in \mathbb{R}^n} f(Kx) + g(x). \quad (5.1)$$

Per (5.1.3) tenim que $f = f^{**}$, i usant la definició de la conjugada convexa tendrem que

$$f(Kx) = f^{**}(Kx) = \sup_{y \in \mathbb{R}^m} \langle y, Kx \rangle - f^*(y).$$

Substituïnt-ho a l'equació anterior arribam al problema primal-dual, donat per

$$\min_{x \in \mathbb{R}^n} \sup_{y \in \mathbb{R}^m} \langle y, Kx \rangle - f^*(y) + g(x). \quad (5.2)$$

Volem estudiar la relació entre el problema primal i el problema primal-dual. Per això, primer definim el lagrangià associat al problema primal-dual.

Definició 5.2.1. *El lagrangià associat al problema primal-dual, denotat per \mathcal{L} , es defineix com*

$$\mathcal{L}(x, y) = \langle y, Kx \rangle - f^*(y) + g(x).$$

La definició bàsica de punt de sella vista a les assignatures de càlcul es pot extender als espais de dimensió infinita amb el lagrangià.

Definició 5.2.2. *Direm que (x^*, y^*) és un punt de sella de \mathcal{L} si $\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^m$*

$$\mathcal{L}(x^*, y) \leq \mathcal{L}(x^*, y^*) \leq \mathcal{L}(x, y^*).$$

Quan passam del problema primal al primal-dual, normalment es fa per tenir una major facilitat per resoldre el problema. Per tant, ens interessa saber quan les solucions de cada un dels problemes estan relacionades. Per això introduim el següent resultat.

Teorema 5.2.1. *Sota les hipòtesis establertes sobre f i g , si (x^*, y^*) és un punt de sella de \mathcal{L} , llavors x^* és una solució del problema primal.*

Demostració. Sigui (x^*, y^*) un punt de sella. Per una part, tenim que

$$\mathcal{L}(x^*, y^*) = \sup_{y \in \mathbb{R}^m} \mathcal{L}(x^*, y) = \sup_{y \in \mathbb{R}^m} \langle y, Kx^* \rangle - f^*(y) + g(x^*) = f(Kx^*) + g(x^*),$$

i per definició, $f(Kx^*) + g(x^*) \geq \inf_{x \in \mathbb{R}^n} f(Kx) + g(x)$. Ara bé, també tenim que

$$\mathcal{L}(x^*, y^*) = \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, y^*) = \inf_{x \in \mathbb{R}^n} \langle y^*, Kx \rangle - f^*(y^*) + g(x) \leq \inf_{x \in \mathbb{R}^n} f(Kx) + g(x).$$

Per tant, necessàriament $f(Kx^*) + g(x^*) = \inf_{x \in \mathbb{R}^n} f(Kx) + g(x)$, és a dir, x^* és un mínim del problema primal. \square

5.2.2 L'algorisme Primal-Dual de Chambolle-Pock

Quan tenim (x^*, y^*) un punt de sella de \mathcal{L} , aleshores $\mathcal{L}(x^*, y^*) = \min_{x \in \mathbb{R}^n} \mathcal{L}(x, y^*)$. Usant el Principi de Fermat (5.1.1), tenim que

$$0 \in \partial(\mathcal{L}(\cdot, y^*))(x^*) = \partial g(x^*) + K^T y^* \implies -K^T y^* \in \partial g(x^*). \quad (5.3)$$

Però per una altra banda, també tenim que $\mathcal{L}(x^*, y^*) = \max_{y \in \mathbb{R}^m} \mathcal{L}(x^*, y)$, i el Principi de Fermat ens diu que

$$0 \in \partial(\mathcal{L}(x^*, \cdot))(y^*) = Kx^* - \partial f^*(y^*) \implies Kx^* \in \partial f^*(y^*). \quad (5.4)$$

Ara considerem el següent lemma:

Lema 5.2.2. Sigui $f : X \rightarrow [-\infty, +\infty]$ propi, convex i semi-continu inferiorment. Siguin $x, x^* \in X$. Aleshores per qualsevol $\tau > 0$ es té que

$$x^* \in \partial f(x) \iff x = \text{prox}_{\tau f}(x + \tau x^*)$$

Demostració.

$$\begin{aligned} x &= \text{prox}_{\tau f}(x + \tau x^*) = \arg \min_y \left\{ f(y) + \frac{1}{2\tau} \|y - x - \tau x^*\|_2^2 \right\} \\ &\iff 0 \in \partial f \left(f(\cdot) + \frac{1}{2\tau} \|\cdot - x - \tau x^*\|_2^2 \right)(x) \iff 0 \in \partial f(x) + \frac{1}{\tau} (x - x - \tau x^*) \iff x^* \in \partial f(x) \end{aligned}$$

□

Si utilitzam el lema per $\sigma > 0$ i $\tau > 0$ qualssevol, obtenim

$$\begin{cases} x^* = \text{prox}_{\tau g}(x^* - \tau K^T y^*) \\ y^* = \text{prox}_{\sigma f^*}(y^* + \sigma Kx^*) \end{cases}$$

La qual cosa ens diu que (x^*, y^*) és un punt fix de l'operador T definit com

$$T(x, y) = (\text{prox}_{\tau g}(x - \tau K^T y), \text{prox}_{\sigma f^*}(y + \sigma Kx))$$

L'algorisme primal-dual de Chambolle-Pock es basa en aquesta expressió de punt fix. L'algorisme és el següent:

1. Escollim uns $\tau, \sigma > 0$, $\theta \in [0, 1]$, $x^0 \in \mathbb{R}^n$, $y^0 \in \mathbb{R}^m$. Fixam $\hat{x}^0 = x^0$.
2. Per a $k = 0, 1, \dots$ calculam

$$\begin{cases} y^{k+1} = \text{prox}_{\sigma f^*}(y^k + \sigma K\hat{x}^k) \\ x^{k+1} = \text{prox}_{\tau g}(x^k - \tau K^T y^{k+1}) \\ \hat{x}^{k+1} = x^{k+1} + \theta(x^{k+1} - x^k) \end{cases}$$

Com veurem en el següent teorema, l'algorisme de Chambolle-Pock és convergent. La demostració es pot trobar a [3].

Teorema 5.2.3. Suposem que el problema primal-dual

$$\min_{x \in \mathbb{R}^n} \max_{y \in \mathbb{R}^m} \langle y, Kx \rangle - f^*(x) + g(x)$$

té un punt de sella. Sigui (x^k, \hat{x}^k, y^k) la seqüència definida a l'algorisme de Chambolle-Pock. Prenem $\theta = 1$ i $\tau \sigma \|K\| < 1$. Aleshores existeix un punt de sella (x^*, y^*) tal que

$$x^k \rightarrow x^* \quad y^k \rightarrow y^*$$

TRACTAMENT D'IMATGES

Les imatges són part del nostre dia a dia, estam en contacte amb elles constantment: quan encenem el televisor, quan miram pel·lícules i sèries a través de plataformes de *streaming*, quan miram les xarxes socials... Les imatges són tot allò que veiem amb els ulls. Un podria arribar a dir que la vida no és més que una successió contínua d'imatges en constant canvi i moviment. Per tant, és lògic demanar-se com podem interactuar amb elles matemàticament per tal de poder-les entendre i millorar-les.

6.1 Imatges com a objectes matemàtics. Problemes mal posats

En general, podem pensar una imatge com una graella de píxels, i aquesta graella de píxels forma un domini rectangular discret. A cada píxel li associam un valor que depèn del tipus d'imatge: si és en escala de grisos o si és a color. En una imatge en escala de grisos, cada un dels píxels pren un valor entre 0 i 255, on 0 representa el valor més fosc i 255 el més clar. En canvi, en el cas d'imatges a color, cada píxel és un vector RGB (*Red, Green and Blue*), on cada component és un valor entre 0 i 255 que representa la quantitat de cada un dels tres colors.

Ara bé, les imatges no sempre són perfectes. Depenen de les condicions sota les quals s'ha pres, pot presentar renou, boira, borrositat... Així, podem modelar una imatge matemàticament com

$$f = Au + \eta, \quad (6.1)$$

on f és la imatge observada, A és un operador lineal entre espais de Hilbert (per exemple un nucli de convolució que causi borrositat), η és un renou aleatori que normalment es suposa gaussià, i u és la imatge real que es vol estimar a partir de la imatge observada f .

Quan un fa una fotografia aquesta pot ser borrosa, tenir renou... Si es vol recuperar la imatge original, veient el model (6.1) un podria pensar que el millor pla d'atac és aïllar

6. TRACTAMENT D'IMATGES

la u , però això no sempre és possible. Aquest procés de recuperació i restauració d'una imatge és el que ens duu a xerrar de *problemes inversos*. Aquests tipus de problemes consisteixen, en el context del tractament d'imatges, en recuperar una imatge a partir d'una observació degradada fent suposicions (*o priors*) sobre els factors que poden haver influït en la seva degradació.

Abans d'enfrontar-nos a aquests problemes, cal xerrar de problemes ben posats i mal posats. Jacques Hadamard, matemàtic del segle XX, va introduir el concepte de problema ben posat [12]. Perquè un problema estigui *ben posat* en el sentit de Hadamard s'ha de satisfet:

- (i) existeix una solució del problema,
- (ii) la solució és única,
- (iii) la solució depèn contínuament de les condicions inicials, és a dir, petites pertorbacions a les dades inicials no poden causar grans pertorbacions en la solució.

Si un problema no satisfà alguna de les tres condicions, direm que està *mal posat*. Així, podem dir que, en general, el problema (6.1) està mal posat, perquè pot ser que A no sigui invertible, i mai podem saber amb certesa la distribució de η . Per poder resoldre aquest problema introduïm mètodes del Càlcul de Variacions, com per exemple, minimitzar l'error quadràtic del renou:

$$\min_{u \in X} \|f - Au\|_2^2. \quad (6.2)$$

Si volem resoldre aquest problema amb el mètode clàssic, tornam a trobar problemes: pot ser que l'equació d'Euler-Lagrange no tengui solució. Per poder fer que el nostre problema estigui ben posat, necessitam imposar condicions fetes a partir de suposicions de les causes de degradació de la imatge i a partir dels nostres objectius. En afegir noves condicions, necessitam cercar un equilibri entre aconseguir una solució prou regular i la fidelitat de la solució.

Per exemple, Tikhonov [13] va proposar incorporar directament informació prèvia sobre la imatge u mitjançant la inclusió d'un terme addicional a la funció de cost de mínims quadrats (6.2). En particular, l'estimació regularitzada de Tikhonov es defineix com la solució del següent problema de minimització:

$$\min_{u \in X} \|\nabla u\|_2^2 + \frac{\lambda}{2} \|f - Au\|_2^2, \quad (6.3)$$

on el primer terme rep el nom de terme de regularització, i el segon, de terme de fidelitat. La versió integral de (6.3) seria

$$\min_{u \in X} \frac{1}{2} \int_{\Omega} |\nabla u|^2 + \frac{\lambda}{2} \int_{\Omega} |f - Au|^2, \quad (6.4)$$

Aquest funcional és diferenciable, i per tant podríem resoldre el problema de minimització resolent les equacions d'Euler-Lagrange i implementant-ho directament, o bé aplicant l'algorisme Primal-Dual de Chambolle-Pock. Però cal dir que el model de

Tikhonov té alguns problemes: el fet que la norma del gradient sigui un terme quadràtic no permet recuperar correctament discontinuïtats molt pronunciades, com les dels contorns. Per això, Rudin, Osher i Fatemi [8], van introduir el model de la TV: la Variació Total o *Total Variation* en anglès. Ara el problema esdevé:

$$\min_{u \in X} \|\nabla u\|_1 + \frac{\lambda}{2} \|f - Au\|_2^2, \quad (6.5)$$

la versió integral del qual seria

$$\min_{u \in X} \frac{1}{2} \int_{\Omega} |\nabla u| + \frac{\lambda}{2} \int_{\Omega} |f - Au|^2. \quad (6.6)$$

Aquí ja no podem emprar els mètodes clàssics perquè el funcional no és diferenciable. Per tant, haurem d'usar el mètode directe i implementar-ho mitjançant tècniques d'optimització convexa. A la figura 6.1d es pot veure com el model de Tikhonov crea una imatge molt regular on es perd la informació dels contorns. En canvi, emprant la TV (figura 6.1f) com a terme de regularització, es preserven millor els contorns dels objectes de l'escena.

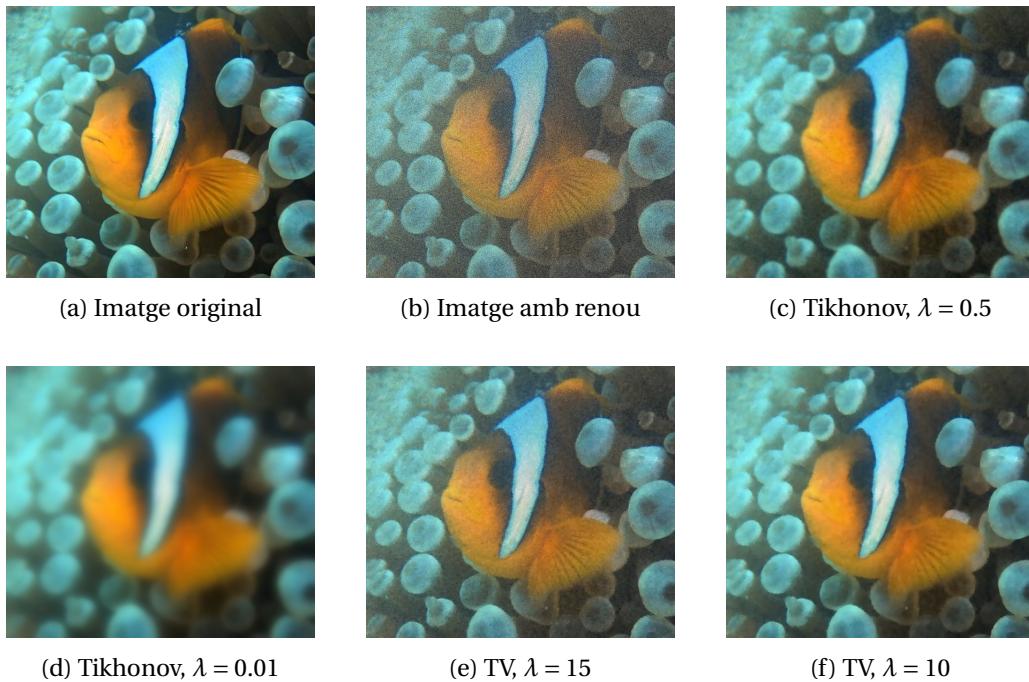


Figura 6.1: Comparació dels mètodes de Tikhonov i la Variació Total

6.2 Dehazing o reducció de boira

6.2.1 Causes de la boira i el *dark channel prior*

Segons el Diccionari de l'Institut d'Estudis Catalans, la boira es defineix com una *extensió més o menys gran d'aire que porta en suspensió partícules petitíssimes d'aigua provinents de la condensació del vapor d'aigua de l'atmosfera, especialment quan forma una capa extensa tocant a terra que redueix la visibilitat horitzontal*. En imatges

6. TRACTAMENT D'IMATGES

submarines, l'aigua participa físicament a la imatge, fent que la llum perdi intensitat i es dispersi, fent que aparegui aquesta boira, que a més es veu accentuada a causa de la presència de grans d'arena o altres partícules en suspensió. El principal problema és que no es disposa d'una imatge vertadera o *ground truth*, perquè no és possible saber com és una imatge submarina si no es troba sota l'aigua. Per això, la restauració d'imatges submarines és un procés difícil i bastant limitat, però a la vegada interessant per tasques de biologia marina, arqueologia o inspecció submarina.

A l'hora d'enfrontar-se al problema de *dehazing* o reducció de boira, necessitam considerar unes hipòtesis fortes, com per exemple suposar que les imatges sense boira han de tenir un major contrast. Ara bé, aquestes suposicions en alguns casos donen bons resultats, però en altres no, ja que no tenen una explicació física consistent vàlida per tots els casos. Per evitar aquests problemes, He, Sun i Tang [1] van introduir el que es coneix com a *dark channel prior*, amb el qual es pot estimar directament l'espessor de la boira i recuperar la imatge original.

En imatges amb boira i imatges submarines, la llum que arriba a la càmera és generada principalment per tres components [14]: una component directa, una de dispersió frontal o *forward-scattering*, i una de retrodispersió o *backscattering*. La component directa és la llum que arriba a la càmera, que es pot definir com $\mathbf{J}t$, on \mathbf{J} és la radiància de les superfícies i t és la transmissió del medi, és a dir, el percentatge de la radiància que arriba a la càmera. La component de retrodispersió és el resultat de la interacció entre les fonts de llum i les partícules del medi, i pot ser modelat de forma simplificada com $\mathbf{A}(1-t)$, on \mathbf{A} és la llum atmosfèrica. La component de dispersió frontal es pot obviar, ja que se suposa que la degradació de la imatge és causada principalment la retrodispersió. Així, una imatge amb boira es pot modelar de la següent manera:

$$\mathbf{I}(x) = t(x) \cdot \mathbf{J}(x) + (1 - t(x)) \cdot \mathbf{A}, \quad (6.7)$$

on $\mathbf{I}: \Omega \rightarrow [0, 255]^3$ és la imatge amb boira; Ω és un obert de \mathbb{R}^2 amb frontera compacta Lipschitz; $x \in \Omega$ és un píxel; i $\mathbf{A} \in [0, 255]^3$, $\mathbf{J}: \Omega \rightarrow [0, 255]^3$ i $t: \Omega \rightarrow [0, 1]$ les hem definides anteriorment. L'objectiu del *dark channel prior* és recuperar \mathbf{J} , t i \mathbf{A} a partir de \mathbf{I} , que és l'única imatge donada. El mètode parteix del següent principi: en una imatge exterior sense boira, a la majoria d'entorns on no es veu el cel, almanco un canal de color té molt poca intensitat. És a dir,

$$\mathbf{J}_{dark}(x) = \min_{c \in \{R, G, B\}} \left(\min_{y \in w(x)} \mathbf{J}_c(y) \right) \approx 0, \quad (6.8)$$

on $w(x)$ és un entorn de píxels centrat en x . Aquesta poca intensitat en algun dels canals pot ser causada per tres factors[15]: la presència d'ombres a la imatge, objectes colorits o objectes foscos. Amb aquestes suposicions, si usam un mètode de *dehazing* per a la restauració d'imatges submarines, té sentit basar-nos en el *dark channel prior* per estimar t i \mathbf{A} , ja que sota l'aigua hi continuarà havent ombres, els peixos i els coralls són objectes colorits, i les roques i les restes de posidònia són objectes foscos. A més, si en aquest esquema se suposa t constant a trossos a cada $w(x)$, la podem estimar a partir de (6.7) i (6.8) com

$$t_0(x) = 1 - \nu \min_{c \in \{R, G, B\}} \left(\min_{y \in w(x)} \left(\frac{\mathbf{I}_c(y)}{\mathbf{A}_c} \right) \right), \quad (6.9)$$

on $\nu \in (0, 1]$ és una constant, normalment fixada a $\nu = 0.95$. Per calcular la \mathbf{A} , el que fem és seleccionar el 0.1% de píxels més brillants de \mathbf{J}_{dark} . Així, He, Sun i Tang [1] recuperen la imatge sense boira com

$$\mathbf{J}_0(x) = \frac{\mathbf{I}(x) - \mathbf{A}}{\max\{t_0(x), \epsilon_0\}} + \mathbf{A}, \quad (6.10)$$

on $\epsilon_0 > 0$ és una constant positiva per evitar dividir per 0. Ara bé, aquest mètode no és molt precís a l'hora de reduir la boira sobre els contorns, no evita l'aparició de renou, i a l'hora d'implementar-lo pot arribar a ser computacionalment molt costós. Podria existir un model que tengués en compte l'aparició de renou i el problema dels contorns?

6.2.2 Model proposat

El problema de l'estimació trobada de t_0 és que el resultat és una funció constant a trossos, i en alguns casos s'obté un resultat massa irregular. Per això, He, Sun i Tang [1] proposen aplicar un filtre per regularitzar l'estimació obtinguda de t_0 . Però encara així, amb aquesta millora, s'aconsegueixen imatges fosques, amb poc contrast i amb renou. L'esquema proposat per Fang, Li, i Zeng [2], parteix de reescriure (6.7) canal a canal com

$$\mathbf{A}_c - \mathbf{I}_c = t(\mathbf{A}_c - \mathbf{J}_c). \quad (6.11)$$

Aplicant logaritmes a ambdós costats per linealitzar el model, obtenim

$$\log(\mathbf{A}_c - \mathbf{I}_c) = \log t + \log(\mathbf{A}_c - \mathbf{J}_c). \quad (6.12)$$

Quan l'atmosfera és homogènia, la transmissió t ve donada per

$$t(x) = e^{-\eta d(x)},$$

on η és una constant positiva que expressa característiques del medi¹ i $d(x)$ és la profunditat, que aïllant-la a l'expressió, la podem expressar com

$$d(x) = -\frac{1}{\eta} \log t(x).$$

Així, definint $f_c = \frac{1}{\eta} \log(\mathbf{A}_c - \mathbf{I}_c)$ i $g_c = \frac{1}{\eta} \log(\mathbf{A}_c - \mathbf{J}_c)$, podem reescriure (6.12) com

$$g_c = f_c + d. \quad (6.13)$$

Si vectoritzam i definim $\mathbf{g} = (g_R, g_G, g_B)$, $\mathbf{f} = (f_R, f_G, f_B)$ i $\mathbf{d} = (d, d, d)$, arribam a

$$\mathbf{g} = \mathbf{f} + \mathbf{d}. \quad (6.14)$$

El model variacional que estudiarem es basa en les següents suposicions:

- Primer es suposa que \mathbf{g} és constant a trossos. Per això usarem una WVT (Weighted Vectorial Total Variation) com a terme de regularització:

$$\int_{\Omega} h(x) |D\mathbf{g}| = \sup \left\{ \int_{\Omega} \sum_{c \in \{R, G, B\}} g_c \operatorname{div} \varphi_c : \boldsymbol{\varphi} \in \mathcal{C}_0^1(\Omega; \mathbb{R}^{3 \times 2}), |\boldsymbol{\varphi}(x)| \leq h(x), x \in \Omega \right\}$$

on $h(x)$ és una funció positiva i contínua que descriu la gruixa de la boira.

¹La constant η també rep el nom de coeficient d'atenuació.

6. TRACTAMENT D'IMATGES

- També es suposa que d és regular a trossos, i per tant el seu terme de regularització vendrà donat per la seva variació total definida a (3.4.2):

$$\int_{\Omega} |Dd|.$$

Una vegada reescrit (6.7) com a (6.14) i haver fet les suposicions necessàries per poder definir els termes de regularització, Fang, Li i Zeng [2] proposen el següent funcional per recuperar una imatge sense boira i sense renou a l'hora:

$$E(\mathbf{g}, d) := \int_{\Omega} h|D\mathbf{g}| + \lambda \int_{\Omega} |Dd| + \frac{1}{2} \int_{\Omega} |\mathbf{g} - \mathbf{f} - \mathbf{d}|^2 + \frac{\gamma}{2} \int_{\Omega} |d - d_0|^2, \quad (6.15)$$

on $d_0 = -\frac{1}{\eta} \log t_0$ i λ, γ són constants positives. La h és una funció relacionada amb la boira. A fotografies on la boira és més densa, necessitarem majors valors de h ; i si la boira és més clara, podrà tenir valors més petits. Com d_0 és una estimació de la densitat de la boira, és lògic pensar que h sigui una funció de d_0 , és a dir:

$$h(x) := \frac{1}{1 + ke^{-5d_0(x)}} \quad (6.16)$$

Notem que té sentit que la funció tengui aquesta forma, ja que satisfà el que cercam: a mesura que augmenta d_0 , augmenta el valor de la funció (veure fig. 6.2), és a dir, augmenta el pes que li donam al terme de regularització per a la \mathbf{g} .



Figura 6.2: Funció $h(x)$ per $k = 1$ i $d_0(x) = x$

El problema de treballar amb (6.15) és que amb el que hem vist al llarg d'aquest treball, no tendríem eines suficients per a poder estudiar-ne l'existència de solució. Per tant, per poder estudiar l'existència i unicitat de solució en espais de Sobolev, proposam considerar el següent funcional:

$$E_p(\mathbf{g}, d) := \int_{\Omega} h|\nabla \mathbf{g}|^p + \lambda \int_{\Omega} |\nabla d|^p + \frac{1}{2} \int_{\Omega} |\mathbf{g} - \mathbf{f} - \mathbf{d}|^2 + \frac{\gamma}{2} \int_{\Omega} |d - d_0|^2, \quad (6.17)$$

per $p > 1$ com a funcional a minimitzar per provar l'existència i unicitat de solució del problema de *dehazing*, centrant-nos en el cas $p = 2$ per poder treballar a l'espai de Hilbert H^1 . Notem que hem usat la relació (3.1) per reescriure els termes de regularització. El nostre objectiu és, coneixent el problema (6.15), que seria el nostre equivalent a la TV per *denoising* (6.6), plantejar en l'àmbit de *dehazing* un model equivalent al model de Tikhonov (6.4), provar-ne l'existència i unicitat de solució, i implementar-lo.

6.2.3 Model proposat, cas $p = 2$

Existència i unicitat de solució

En aquest apartat ens proposam demostrar l'existència i unicitat de solució del problema de minimitzar el funcional anterior pel cas $p = 2$. Així, proposam el següent resultat:

Teorema 6.2.1. Sigui $\mathbf{f} \in L^2(\Omega; \mathbb{R}^3)$ i $d_0 \in L^2(\Omega; \mathbb{R})$. Definim $\mathcal{W} = W^{1,2}(\Omega; \mathbb{R}^3) \times W^{1,2}(\Omega; \mathbb{R})$. I denotem $u = (\mathbf{g}, d)$. Aleshores el problema

$$\min_{u \in \mathcal{W}} E_2(\mathbf{g}, d) := \int_{\Omega} h|\nabla \mathbf{g}|^2 + \lambda \int_{\Omega} |\nabla d|^2 + \frac{1}{2} \int_{\Omega} |\mathbf{g} - \mathbf{f} - \mathbf{d}|^2 + \frac{\gamma}{2} \int_{\Omega} |d - d_0|^2, \quad (6.18)$$

té una única solució en \mathcal{W} .

Demostració. Ho demostrarem usant el teorema 4.2.3. Vegem que satisfà les seves hipòtesis. Considerem la funció

$$f(x, u, \nabla u) = h|\nabla \mathbf{g}|^2 + \lambda|\nabla d|^2 + \frac{1}{2}|\mathbf{g} - \mathbf{f} - \mathbf{d}|^2 + \frac{\gamma}{2}|d - d_0|^2 \quad (6.19)$$

La funció f és contínua per ser suma de termes quadràtics, i a més, usant la σ -àlgebra de Borel, tota funció contínua és mesurable. Com f és contínua i mesurable respecte totes les variables, en concret serà una funció de Carathéodory.

També volem veure que la funció satisfà la condició de coercivitat (4.1), és a dir, que $\exists x \in \Omega$ i $\forall (u, \nabla u) \in \mathbb{R}^4 \times \mathbb{R}^8$ existeixen una $\alpha_3 \in L^1(\Omega)$, $\alpha_2 \in \mathbb{R}$, $\alpha_1 > 0$ i $p > q \geq 1$ tals que:

$$f(x, u, \nabla u) \geq \alpha_1|\nabla u|^p + \alpha_2|u|^q + \alpha_3(x).$$

Vegem primer que la funció h està fitada inferiorment per una constant positiva.

$$0 \leq d_0 \implies 0 < e^{-5d_0} \leq 1 \implies 1 < 1 + ke^{-5d_0} \leq 1 + k \implies 1 > \frac{1}{1 + ke^{-5d_0}} \geq \frac{1}{1 + k} > 0.$$

Vegem ara que f satisfà la condició de coercivitat:

$$\begin{aligned} f(x, u, \nabla u) &= h|\nabla \mathbf{g}|^2 + \lambda|\nabla d|^2 + \frac{1}{2}|\mathbf{g} - \mathbf{f} - \mathbf{d}|^2 + \frac{\gamma}{2}|d - d_0|^2 \\ &= h \sum_{c \in \{R, G, B\}} |\nabla g_c|^2 + \lambda \sum_{c \in \{R, G, B\}} |\nabla d|^2 + \frac{1}{2} \sum_{c \in \{R, G, B\}} |g_c - f_c - d|^2 + \frac{\gamma}{2} |d - d_0|^2 \\ &\geq h \sum_{c \in \{R, G, B\}} |\nabla g_c|^2 + \lambda \sum_{c \in \{R, G, B\}} |\nabla d|^2 \geq \frac{1}{1+k} \sum_{c \in \{R, G, B\}} |\nabla g_c|^2 + \lambda |\nabla d|^2 \\ &\geq \alpha_1 \left(\sum_{c \in \{R, G, B\}} |\nabla g_c|^2 + |\nabla d|^2 \right) = \alpha_1 |\nabla u|^2, \end{aligned}$$

per $\alpha_1 = \min\{\lambda, \frac{1}{1+k}\}$. Llavors satisfà la condició de coercivitat per l' α_1 definida, $\alpha_2 = 0$, $\alpha_3 \equiv 0$, $p = 2$ i q arbitrària tal que $1 \leq q < 2$.

A continuació comprovarrem que f és estrictament convexa respecte la tercera variable usant que és suma de termes quadràtics, que són estrictament convexos. Volem veure que

$$f(x, u, \alpha \nabla u_a + (1-\alpha) \nabla u_b) < \alpha f(x, u, \nabla u_a) + (1-\alpha) f(x, u, \nabla u_b)$$

Vegem-ho. Sigui $F(u) = \frac{1}{2}|\mathbf{g} - \mathbf{f} - \mathbf{d}|^2 + \frac{\gamma}{2}|d - d_0|^2$. Llavors

$$\begin{aligned} &f(x, u, \alpha \nabla u_a + (1-\alpha) \nabla u_b) \\ &= F(u) + h \sum_{c \in \{R, G, B\}} \left| \alpha \nabla g_c^a + (1-\alpha) \nabla g_c^b \right|^2 + \lambda \left| \alpha \nabla d^a + (1-\alpha) \nabla d^b \right|^2 \\ &< \alpha F(u) + (1-\alpha) F(u) + h \sum_{c \in \{R, G, B\}} \left(\alpha |\nabla g_c^a|^2 + (1-\alpha) |\nabla g_c^b|^2 \right) + \lambda \left(\alpha |\nabla d^a|^2 + (1-\alpha) |\nabla d^b|^2 \right) \\ &= \alpha f(x, u, \nabla u_a) + (1-\alpha) f(x, u, \nabla u_b). \end{aligned}$$

6. TRACTAMENT D'IMATGES

Per tant el funcional és estrictament convex.

Per últim, falta veure que el funcional és propi, és a dir, que existeix un $u_0 \in \mathcal{W}$ tal que $E(u_0) < \infty$. Prenent $u_0 = (\mathbf{0}, 0)$:

$$E(\mathbf{0}, 0) = \frac{1}{2} \int_{\Omega} |\mathbf{f}|^2 + \frac{\gamma}{2} \int_{\Omega} |d_0|^2 = \frac{1}{2} \|\mathbf{f}\|_2^2 + \frac{\gamma}{2} \|d_0\|_2^2 < +\infty,$$

que és finit per ser $\mathbf{f} \in L^2(\Omega; \mathbb{R}^3)$ i $d_0 \in L^2(\Omega; \mathbb{R})$.

Per tant, com el problema (6.18) satisfà les hipòtesis del teorema (4.2.3) amb convexitat estricta, podem garantir que té una única solució.

□

Discretització i algorisme Primal-Dual

En aquesta secció considerem el funcional (6.18) afegint unes constants tècniques per poder desenvolupar l'algorisme còmodament. Notem que podem reescriure el funcional en continu de la següent forma:

$$\begin{aligned} E_2(\mathbf{g}, d) &= \frac{1}{2} \int_{\Omega} h |\nabla \mathbf{g}|^2 + \frac{\lambda}{2} \int_{\Omega} |\nabla d|^2 + \frac{1}{2} \int_{\Omega} |\mathbf{g} - \mathbf{f} - \mathbf{d}|^2 + \frac{\gamma}{2} \int_{\Omega} |d - d_0|^2 \\ &= \frac{1}{2} \int_{\Omega} |\sqrt{h} \nabla \mathbf{g}|^2 + \frac{\lambda}{2} \int_{\Omega} |\nabla d|^2 + \frac{1}{2} \int_{\Omega} |\mathbf{g} - \mathbf{f} - \mathbf{d}|^2 + \frac{\gamma}{2} \int_{\Omega} |d - d_0|^2. \end{aligned} \quad (6.20)$$

Llavors, siguin

$$\nabla \mathbf{g} = (\nabla g_R, \nabla g_G, \nabla g_B) = \begin{pmatrix} g_{Rx} & g_{Gx} & g_{Bx} \\ g_{Ry} & g_{Gy} & g_{By} \end{pmatrix}, \quad \nabla d = \begin{pmatrix} d_x \\ d_y \end{pmatrix}. \quad (6.21)$$

En el cas discret, treballarem amb $X = \mathbb{R}^{MN}$ i $Y = X \times X$, i considerarem la imatge definida sobre el domini discret $\Omega_d = \{(i, j) \mid 1 \leq i \leq M, 1 \leq j \leq N\}$, on (i, j) és un píxel. Si usam les següents normes:

$$\|\nabla \mathbf{g}\|_2 = \left(\sum_i^M \sum_j^N \sqrt{\sum_{c \in \{R, G, B\}} (g_{cx})_{i,j}^2 + (g_{cy})_{i,j}^2} \right)^{1/2} \quad (6.22)$$

$$\|\nabla d\|_2 = \left(\sum_i^M \sum_j^N \sqrt{(d_x)_{i,j}^2 + (d_y)_{i,j}^2} \right)^{1/2}, \quad (6.23)$$

la versió discreta del nostre problema serà:

$$E_2^d(\mathbf{g}, d) = \frac{1}{2} \|\sqrt{h} \nabla \mathbf{g}\|_2^2 + \frac{\lambda}{2} \|\nabla d\|_2^2 + \frac{1}{2} \|\mathbf{g} - \mathbf{f} - \mathbf{d}\|_2^2 + \frac{1}{2} \|d - d_0\|_2^2. \quad (6.24)$$

Si, a més, definim els operadors:

$$\begin{array}{ll} K_1 : X^3 \longrightarrow Y^3 & K_2 : X \longrightarrow Y \\ \mathbf{g} \longmapsto \sqrt{h} \nabla \mathbf{g} & d \longmapsto \nabla d \end{array}$$

el problema de minimització s'escriu com

$$\min_{\mathbf{g}, d} F(K_1 \mathbf{g}) + G(K_2 d) + \frac{1}{2} \|\mathbf{g} - \mathbf{f} - \mathbf{d}\|_2^2 + \frac{\gamma}{2} \|d - d_0\|_2^2,$$

on $F(\mathbf{y}) = \frac{1}{2} \|\mathbf{y}\|_2^2$ per a $\mathbf{y} \in Y^3$; i $G(z) = \frac{\lambda}{2} \|z\|_2^2$ per a $z \in Y$. Sabem que F, G són convexes, pròpies i semi-contínues inferiorment, llavors $F^{**}(\mathbf{y}) = F(\mathbf{y}) = \sup_{\mathbf{q} \in Y^3} \langle \mathbf{y}, \mathbf{q} \rangle - F^*(\mathbf{q})$ i $G^{**}(z) = G(z) = \sup_{p \in Y} \langle z, p \rangle - G^*(p)$. A més tal com varem veure a (5.1.2), tenim que $F^*(\mathbf{q}) = \frac{1}{2} \|\mathbf{q}\|_2^2$ i $G^*(p) = \frac{1}{2\lambda} \|p\|_2^2$. Així, el nostre problema de minimització esdevé:

$$\begin{aligned} & \min_{\mathbf{g}, d} F(K_1 \mathbf{g}) + G(K_2 d) + \frac{1}{2} \|\mathbf{g} - \mathbf{f} - \mathbf{d}\|_2^2 + \frac{\gamma}{2} \|d - d_0\|_2^2 \\ &= \min_{\mathbf{g}, d} \max_{\mathbf{q}, p} \langle \mathbf{q}, K_1 \mathbf{g} \rangle - \frac{1}{2} \|\mathbf{q}\|_2^2 + \langle p, K_2 d \rangle - \frac{1}{2\lambda} \|p\|_2^2 + \frac{1}{2} \|\mathbf{g} - \mathbf{f} - \mathbf{d}\|_2^2 + \frac{\gamma}{2} \|d - d_0\|_2^2 \\ &= \min_{\mathbf{g}, d} \max_{\mathbf{q}, p} \left\langle \begin{pmatrix} \mathbf{q} \\ p \end{pmatrix}, \begin{pmatrix} K_1 & 0 \\ 0 & K_2 \end{pmatrix} \begin{pmatrix} \mathbf{g} \\ d \end{pmatrix} \right\rangle - \frac{1}{2} \|\mathbf{q}\|_2^2 - \frac{1}{2\lambda} \|p\|_2^2 + \frac{1}{2} \|\mathbf{g} - \mathbf{f} - \mathbf{d}\|_2^2 + \frac{\gamma}{2} \|d - d_0\|_2^2. \end{aligned}$$

Ara si definim l'operador

$$K = \begin{pmatrix} K_1 & 0 \\ 0 & K_2 \end{pmatrix},$$

i les funcions

$$\Phi^*(y) = \Phi^*(\mathbf{q}, p) = \frac{1}{2} \|\mathbf{q}\|_2^2 + \frac{1}{2\lambda} \|p\|_2^2$$

$$\Psi(x) = \Psi(\mathbf{g}, d) = \frac{1}{2} \|\mathbf{g} - \mathbf{f} - \mathbf{d}\|_2^2 + \frac{\gamma}{2} \|d - d_0\|_2^2,$$

el problema primal de minimització que havíem de resoldre, es converteix en el problema Primal-Dual

$$\boxed{\min_x \max_y \langle y, Kx \rangle - \Phi^*(y) + \Psi(x)} \quad (6.25)$$

Usant el lema (5.2.2), sabem que la solució (x^*, y^*) vendrà donada, per unes constants $\tau, \sigma > 0$, per l'expressió

$$\begin{cases} x^* = \text{prox}_{\tau\Psi}(x^* - \tau K^T y^*) \\ y^* = \text{prox}_{\sigma\Phi}(y^* + \sigma K x^*) \end{cases}.$$

Usant que $x = (\mathbf{g}, d)$ i $y = (\mathbf{q}, p)$, i que els operadors adjunts de K_1 i K_2 són, respectivament, $K_1^T = -\nabla^T \sqrt{h}$ i $K_2^T = -\nabla^T$, on ∇^T és la divergència, podrem calcular la solució del problema usant l'algorisme Primal-Dual de Chambolle-Pock, que ens dona els següents iterats:

$$\begin{cases} \mathbf{q}^{n+1} = \text{prox}_{\sigma\Phi^*(\cdot, p)}(\mathbf{q}^n + \sigma \sqrt{h} \nabla \hat{\mathbf{g}}^n) \\ p^{n+1} = \text{prox}_{\sigma\Phi^*(\mathbf{q}, \cdot)}(p^n + \sigma \nabla \hat{d}^n) \\ \mathbf{g}^{n+1} = \text{prox}_{\tau\Psi(\cdot, d)}(\mathbf{g}^n + \tau \nabla^T (\sqrt{h} \mathbf{q}^n + 1)) \\ d^{n+1} = \text{prox}_{\tau\Psi(\mathbf{g}, \cdot)}(d^n + \tau \nabla^T p^{n+1}) \\ \hat{\mathbf{g}}^{n+1} = 2\mathbf{g}^{n+1} - \mathbf{g}^n \\ \hat{d}^{n+1} = 2d^{n+1} - d^n \end{cases}. \quad (6.26)$$

En aquest cas, com la norma $\|\cdot\|_2$ és diferenciable, per calcular els operadors proximals on es veu involucrada basta derivar l'expressió i igualar a 0. Si calculam els operadors

6. TRACTAMENT D'IMATGES

proximals, arribam a:

$$\begin{cases} \mathbf{q}^{n+1} &= \frac{\mathbf{q}^n + \sigma \sqrt{\mathbf{h}} \nabla \hat{\mathbf{g}}^n}{\sigma + 1} \\ p^{n+1} &= \frac{p^n + \sigma \nabla \hat{d}^n}{1 + \sigma / \lambda} \\ \mathbf{g}^{n+1} &= \frac{\mathbf{g}^n + \tau \nabla^T (\sqrt{h} \mathbf{q}^{n+1}) + \tau (\mathbf{f} + \mathbf{d}^n)}{\tau + 1} \\ d^{n+1} &= \frac{d^n + \tau \nabla^T p^{n+1} + \tau \sum_{c \in \{R, G, B\}} (g_c - f_c) + \tau \gamma d_0}{1 + \gamma \tau + 3\tau} \\ \hat{\mathbf{g}}^{n+1} &= 2\mathbf{g}^{n+1} - \mathbf{g}^n \\ \hat{d}^{n+1} &= 2d^{n+1} - d^n \end{cases}. \quad (6.27)$$

Experimentació

Una vegada implementat l'algorisme (6.27) amb codi Python (disponible a (A.1.6)), deixam alguns dels resultats obtinguts més significatius.

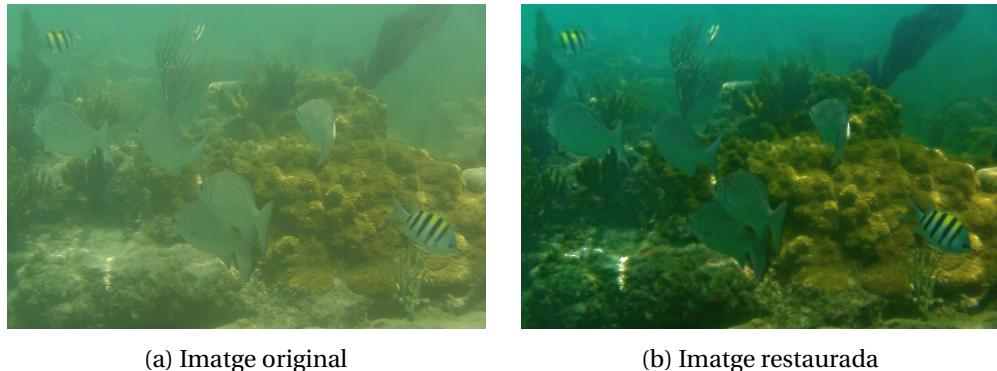


Figura 6.3: 100 iteracions amb $\eta = 1$, $k = 50$, $\lambda = 2$, $\gamma = 0.5$, $\sigma = 1$ i $\tau = 0.11$

A la figura (6.3) podem veure com tot el tel de boira desapareix i es poden veure amb claredat els reflexos del sol a terra.



Figura 6.4: 100 iteracions amb $\eta = 1$, $k = 500$, $\lambda = 3$, $\gamma = 0.1$, $\sigma = 1$ i $\tau = 0.11$

A la figura (6.4) podem veure com es recupera la imatge sense boira, però el model penalitza molt els contorns de les figures del fons.

6.2.4 Model proposat, cas $p = 1$

Existència i unicitat de solució. Discretització i algorisme Primal-Dual

Els resultats obtinguts amb el cas $p = 2$ són bons, però són millorables, com podem veure sobretot en la figura (6.4). Hem adaptat el funcional perquè satisfés el teorema 4.2.3, però el cas interessant és el cas amb $p = 1$, on el funcional no és diferenciable. Per desgràcia, no podem estudiar aquest cas sobre els espais de Sobolev i requerim l'ús de l'espai de funcions de Variació Fitada.

Malgrat no puguem emprar el teorema 4.2.3, es pot comprovar que el problema té una única solució en $\Lambda = BV(\Omega, \mathbb{R}^3) \times BV(\Omega; \mathbb{R})$, però la prova, la qual es pot consultar a l'article de Fang, Li i Zeng [2], es troba fora de l'àbast d'aquest treball. En aquest cas, la versió contínua del nostre funcional és la següent:

$$E_1(\mathbf{g}, d) := \int_{\Omega} h|D\mathbf{g}| + \lambda \int_{\Omega} |Dd| + \frac{1}{2} \int_{\Omega} |\mathbf{g} - \mathbf{f} - \mathbf{d}|^2 + \frac{\gamma}{2} \int_{\Omega} |d - d_0|^2, \quad (6.28)$$

Una altra vegada, per la versió discreta treballarem amb $X = \mathbb{R}^{MN}$ i $Y = X \times X$, i considerarem la imatge definida sobre el domini discret $\Omega_d = \{(i, j) \mid 1 \leq i \leq M, 1 \leq j \leq N\}$. Si consideram $\nabla \mathbf{g}$ i ∇d com a 6.21, i empram ara les normes

$$\|\nabla \mathbf{g}\|_1 = \sum_i^M \sum_j^N \sqrt{\sum_{c \in \{R, G, B\}} (g_{cx})_{i,j}^2 + (g_{cy})_{i,j}^2} \quad (6.29)$$

$$\|\nabla d\|_1 = \sum_i^M \sum_j^N \sqrt{(d_x)_{i,j}^2 + (d_y)_{i,j}^2}, \quad (6.30)$$

la corresponent versió discreta serà

$$E_1^d(\mathbf{g}, d) = h\|\nabla \mathbf{g}\|_1 + \lambda\|\nabla d\|_1 + \frac{1}{2}\|\mathbf{g} - \mathbf{f} - \mathbf{d}\|_2^2 + \frac{1}{2}\|d - d_0\|_2^2. \quad (6.31)$$

Ara, definim els operadors:

$$\begin{array}{ll} K_1 : X^3 \longrightarrow Y^3 & K_2 : X \longrightarrow Y \\ \mathbf{g} \longmapsto \nabla \mathbf{g} & d \longmapsto \nabla d \end{array}$$

Si en aquest cas denotam $F(\mathbf{y}) = h\|\mathbf{y}\|_1$ per a $\mathbf{y} \in Y^3$, sabem que per a cada $x \in \Omega$ la seva conjugada convexa és $F^*(\mathbf{q}) = \delta_{h(x)}(\mathbf{q})$ per un $\mathbf{q} \in Y^3$, on $\delta_{h(x)}$ denota la funció indicatriu sobre la bolla dual de radi $h(x)$. Per l'altra banda, denotant $G(z) = \lambda\|z\|_1$, tenim que la seva conjugada convexa és $G^*(p) = \delta_\lambda(p)$ per un $p \in Y$, és a dir, la funció indicatriu sobre la bolla dual de radi λ . Així, usant un raonament anàleg que en el cas $p = 2$, el nostre problema de minimització esdevé:

$$\begin{aligned} & \min_{\mathbf{g}, d} F(K_1 \mathbf{g}) + G(K_2 d) + \frac{1}{2}\|\mathbf{g} - \mathbf{f} - \mathbf{d}\|_2^2 + \frac{\gamma}{2}\|d - d_0\|_2^2 \\ &= \min_{\mathbf{g}, d} \max_{\mathbf{q}, p} \left\langle \begin{pmatrix} \mathbf{q} \\ p \end{pmatrix}, \begin{pmatrix} K_1 & 0 \\ 0 & K_2 \end{pmatrix} \begin{pmatrix} \mathbf{g} \\ d \end{pmatrix} \right\rangle - \delta_{h(x)}(\mathbf{q}) - \delta_\lambda(p) + \frac{1}{2}\|\mathbf{g} - \mathbf{f} - \mathbf{d}\|_2^2 + \frac{\gamma}{2}\|d - d_0\|_2^2. \end{aligned}$$

6. TRACTAMENT D'IMATGES

Ara si definim l'operador

$$K = \begin{pmatrix} K_1 & 0 \\ 0 & K_2 \end{pmatrix},$$

i les funcions

$$\begin{aligned}\Phi^*(y) &= \Phi^*(\mathbf{q}, p) = \delta_{h(x)}(\mathbf{q}) + \delta_\lambda(p) \\ \Psi(x) &= \Psi(\mathbf{g}, d) = \frac{1}{2} \|\mathbf{g} - \mathbf{f} - \mathbf{d}\|_2^2 + \frac{\gamma}{2} \|d - d_0\|_2^2,\end{aligned}$$

el problema primal de minimització que havíem de resoldre, es converteix en el problema Primal-Dual

$$\boxed{\min_x \max_y \langle y, Kx \rangle - \Phi^*(y) + \Psi(x)} \quad (6.32)$$

Tornant a usar el lema (5.2.2), sabem que la solució (x^*, y^*) vendrà donada, per unes constants $\tau, \sigma > 0$, per l'expressió

$$\begin{cases} x^* = \text{prox}_{\tau\Psi}(x^* - \tau K^T y^*) \\ y^* = \text{prox}_{\sigma\Phi^*}(y^* + \sigma K x^*) \end{cases}.$$

Usant que $x = (\mathbf{g}, d)$ i $y = (\mathbf{q}, p)$, i que els operadors adjunts de K_1 i K_2 són, respectivament, $K_1^T = -\nabla^T$ i $K_2^T = -\nabla^T$, podrem calcular la solució del problema usant l'algorisme Primal-Dual de Chambolle-Pock, que ens dona els següents iterats:

$$\begin{cases} \mathbf{q}^{n+1} = \text{prox}_{\sigma\Phi^*(\cdot, p)}(\mathbf{q}^n + \sigma \nabla \hat{\mathbf{g}}^n) \\ p^{n+1} = \text{prox}_{\sigma\Phi^*(\mathbf{q}, \cdot)}(p^n + \sigma \nabla \hat{d}^n) \\ \mathbf{g}^{n+1} = \text{prox}_{\tau\Psi(\cdot, d)}(\mathbf{g}^n + \tau \nabla^T \mathbf{q}^n + 1) \\ d^{n+1} = \text{prox}_{\tau\Psi(\mathbf{g}, \cdot)}(d^n + \tau \nabla^T p^{n+1}) \\ \hat{\mathbf{g}}^{n+1} = 2\mathbf{g}^{n+1} - \mathbf{g}^n \\ \hat{d}^{n+1} = 2d^{n+1} - d^n \end{cases}. \quad (6.33)$$

Si calculam els operadors proximals, arribam a:

$$\begin{cases} \mathbf{q}^{n+1} = \Pi_{h(x)}(\mathbf{q}^n + \sigma \nabla \hat{\mathbf{g}}^n) \\ p^{n+1} = \Pi_\lambda(p^n + \sigma \nabla \hat{d}^n) \\ \mathbf{g}^{n+1} = \frac{\tau(\mathbf{f} + \mathbf{d}^{n+1}) + \mathbf{g}^n - \tau \nabla^T \mathbf{q}^{n+1}}{1 + \tau} \\ d^{n+1} = \frac{d^n + \tau \nabla^T p^{n+1} + \tau \sum_{c \in \{R, G, B\}} (g_c - f_c) + \tau \gamma d_0}{1 + \gamma \tau + 3\tau} \\ \hat{\mathbf{g}}^{n+1} = 2\mathbf{g}^{n+1} - \mathbf{g}^n \\ \hat{d}^{n+1} = 2d^{n+1} - d^n \end{cases}. \quad (6.34)$$

Experimentació

Una vegada implementat l'algorisme (6.34) amb codi Python (disponible a (A.1.6)), deixam alguns dels resultats obtinguts més significatius.



Figura 6.5: 100 iteracions amb $\eta = 1$, $k = 50$, $\lambda = 0.01$, $\gamma = 0.1$, $\sigma = 1$ i $\tau = 0.11$

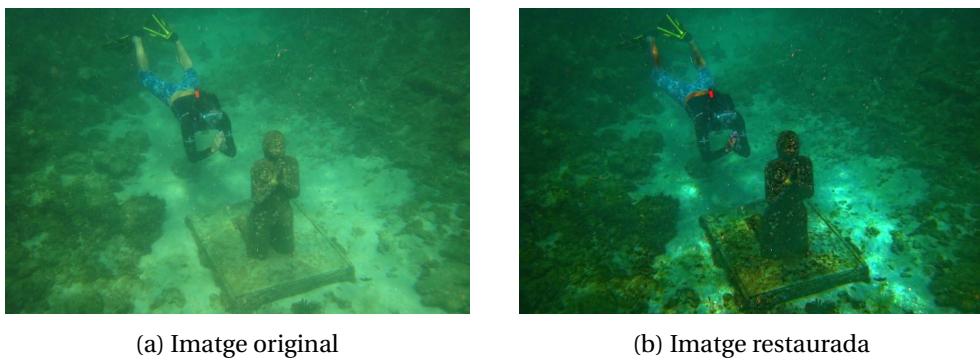


Figura 6.6: 100 iteracions amb $\eta = 1$, $k = 50$, $\lambda = 0.01$, $\gamma = 0.1$, $\sigma = 1$ i $\tau = 0.11$

Ara en ambdós casos els contorns queden ben definits, es recupera el color, i tornen a sortir els reflexos del sol en el fons marí.

6.2.5 Comparativa: cas $p = 1$ vs. cas $p = 2$

Una vegada hem proposat, estudiat i implementat els dos models de *dehazing*, anem a realitzar una comparació. Matemàticament les diferències són òbves: a l' hora de discretitzar-los, les normes que usam no són les mateixes i això ja fa que els algorismes Primal-Dual obtinguts en cada un dels casos siguin diferents. Sobre el paper tot això està clar, però ens interessa veure les diferències visuals sobre imatges.

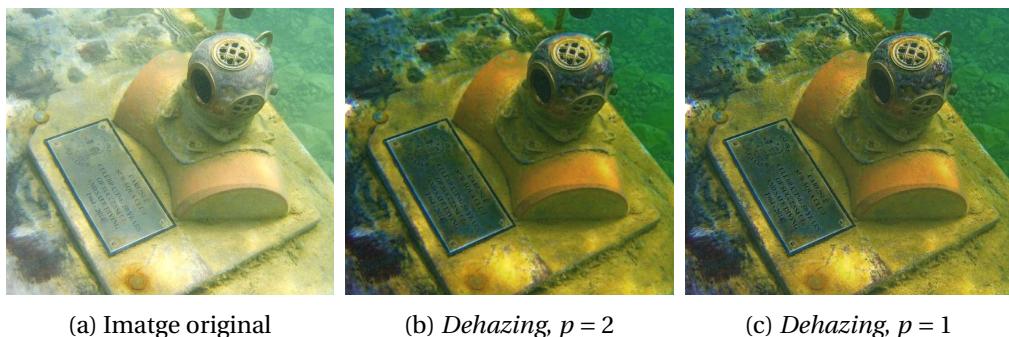


Figura 6.7: Estàtua commemorativa d'un club de buceig

6. TRACTAMENT D'IMATGES

En aquesta primera imatge, a primera vista sembla que els resultats són gairebé iguals. Però si ens fixam, en el cas $p = 2$ totes les superfícies són més suaus. Les porositats de la pedra sobre la qual es troba l'escultura han desaparegut quasi per complet i els caires queden més difuminats. En canvi, en el cas $p = 1$ totes les imperfeccions es mantenen. El que més destaca és la llegibilitat de la placa. En el cas $p = 2$ podem llegir bona part de les paraules de forma clara, però els traçats d'algunes de les lletres no són prou clars i poden dur a confusió. En canvi, en el cas $p = 1$ podem llegir: "Carlile Sub Aqua Club. Celebrating 50 years of successful and safe diving. 1964-2014." Així i tot, *a grosso modo*, les imatges són semblants. Anem a veure el següent cas.



(a) Imatge original

(b) Dehazing, $p = 2$

(c) Dehazing, $p = 1$

Figura 6.8: Bitllet d'un dòlar estatunidenc

En aquest cas les diferències són més pronunciades. En ambdós casos desapareix la capa de boira causada per les partícules en suspensió. Ara bé, en el cas $p = 2$ els detalls del bitllet són pràcticament il·legibles, mentre que en el cas $p = 1$ es poden llegir fins i tot millor que a la imatge original.

CONCLUSIONS

Prop d'un 80% d'aquest treball ha consistit en la introducció i desenvolupament de les eines necessàries per poder demostrar l'existència de solució del problema de *dehazing* (6.18) a l'espai de Sobolev $H^1(\Omega)$ mitjançant el mètode directe del Càlcul de Variacions, així com de les eines necessàries d'Optimització Convexa per a poder implementar l'algorisme Primal-Dual de Chambolle-Pock (6.27) i arribar a la solució explícita del problema. Durant la resta del treball, ens hem pogut endinsar dins el tractament d'imatges des d'un punt de vista matemàtic, i entendre perquè el procés de restauració d'una imatge és un problema mal posat que requereix de l'ús de tècniques variacionals basades en suposicions per a poder-lo resoldre. Una vegada hem implementat els nostres algorismes amb codi Python, hem vist com considerant termes de regularització quadràtics, hem obtingut un problema resoluble a l'espai de Sobolev, però que en casos com el de la figura (6.4) no és capaç de control·lar els contorns dels objecetes que no es troben en primer plà. Per resoldre aquest problema, hem vist que considerant el model (6.15), la solució del qual es troba en l'espai de funcions de Variació Fitada, $BV(\Omega)$, obtenim imatges més clares amb un contorns més marcats.

El principal problema, però, d'aquest treball, ha estat la limitació d'espai. La feina començada en aquest projecte, podria ser continuada en moltes direccions diferents. Per exemple, podríem decidir estudiar una versió no local del problema, similar a l'estudiada a [16], que parteix de la suposició de que pots agrupar els principals colors d'una imatge amb boira mitjançant *clusters*; o fins i tot seguir usant el *dark channel prior* però incloent l'ús de xarxes neuronals com a [17] per a minimitzar un funcional d'energia. Com podem veure, la restauració d'imatges submarines és un tema de gran importància i interès tractat de moltes formes diferents. Hi ha molts de models diferents que ens ajuden a saber com és realment el fons marí, i com va dir Hilbert,

Wir müssen wissen — wir werden wissen! (Hem de saber-ho. Ho sabrem!)



APÈNDIX

A.1 Codi Python implementat

A.1.1 Paquets emprats

Per el correcte funcionament del codi Python implementat, serà necessari importar el següent:

```
1 import imageio.v2 as imageio
2 import time
3 import numpy as np
4 from skimage.morphology import erosion
5 import os
```

A.1.2 Funcions auxiliars

En aquest apartat, adjuntarem les funcions necessàries per al correcte funcionament dels algorismes implementats de *denoising* a (6.1) i (6.2.2).

Primer, definim el que seran el nostre operador, el gradient, i el seu adjunt, la divergència. I a continuació, el programa principal:

```
1 def grad(img):
2     """
3     Parameters
4     - img: 2d image
5     Computes the discrete gradient of img
6     """
7     grad_1 = np.zeros(np.shape(img)).astype(np.float32)
8     grad_2 = np.zeros(np.shape(img)).astype(np.float32)
9
10    img = img.astype(np.float32)
11    grad_1[:, 0:-1] = img[:, 1:] - img[:, :-1]
12    grad_2[0:-1, :] = img[1:, :] - img[:-1, :]
13
14    return np.stack([grad_1, grad_2], axis=2).astype(np.float32)
```

A. APÈNDIX

```
15
16
17 def channel_grad(img):
18     """
19     Parameters
20     - img: 3d image
21     Computes the discrete gradient of img
22     """
23     grad_1 = np.zeros(np.shape(img)).astype(np.float32)
24     grad_2 = np.zeros(np.shape(img)).astype(np.float32)
25
26     img = img.astype(np.float32)
27     grad_1[:, 0:-1, :] = img[:, 1:, :] - img[:, :-1, :]
28     grad_2[0:-1, :, :] = img[1:, :, :] - img[:-1, :, :]
29
30     return np.stack([grad_1, grad_2], axis=3).astype(np.float32)
31
32
33 def back_grad(img):
34     """
35     Parameters:
36     - img: 2d image
37     Computes the backward discrete gradient of img
38     """
39     grad_1 = np.zeros(np.shape(img)).astype(np.float32)
40     grad_2 = np.zeros(np.shape(img)).astype(np.float32)
41     img = img.astype(np.float32)
42     # interior
43     grad_1[:, 1:] = img[:, 1:] - img[:, :-1]
44     grad_2[1:, :] = img[1:, :] - img[:-1, :]
45     # boundary
46     grad_1[:, 0] = img[:, 0]
47     grad_2[0, :] = img[0, :]
48     grad_1[:, -1] = -img[:, -2]
49     grad_2[-1, :] = -img[-2, :]
50     return np.stack([grad_1, grad_2], axis=2).astype(np.float32)
51
52
53 def channel_back_grad(img):
54     """
55     Parameters:
56     - img: 3d image
57     Computes the backward discrete gradient of img
58     """
59     grad_1 = np.zeros(np.shape(img)).astype(np.float32)
60     grad_2 = np.zeros(np.shape(img)).astype(np.float32)
61     img = img.astype(np.float32)
62     # interior
63     grad_1[:, 1:, :] = img[:, 1:, :] - img[:, :-1, :]
64     grad_2[1:, :, :] = img[1:, :, :] - img[:-1, :, :]
65     # boundary
66     grad_1[:, 0, :] = img[:, 0, :]
67     grad_2[0, :, :] = img[0, :, :]
68     grad_1[:, -1, :] = -img[:, -2, :]
69     grad_2[-1, :, :] = -img[-2, :, :]
70     return np.stack([grad_1, grad_2], axis=2).astype(np.float32)
71
```

```

72
73 def div(img):
74     """
75     Parameters:
76     - img: 2d image
77     Computes the divergence of img
78     """
79     g1 = back_grad(img[..., 0])[..., 0]
80     g2 = back_grad(img[..., 1])[..., 1]
81     div = g1 + g2
82     return div
83
84
85 def channel_div(img):
86     """
87     Parameters:
88     - img: 3d image
89     Computes the divergence of img
90     """
91     g1 = channel_back_grad(img[..., 0])[..., 0, :]
92     g2 = channel_back_grad(img[..., 1])[..., 1, :]
93     div = g1 + g2
94     return div

```

A continuació, adjuntam el codi emprat per calcular la norma $\|\cdot\|_2$, per vectoritzar arrays, i per calcular projeccions sobre la bola de radi λ i de radi $h(x)$, on h és la funció definida a (6.16).

```

1 def norma_dos_dos(x):
2     """
3     Parameters:
4     - x: 4d array
5     Computes the 2-2 norm of x
6     """
7     return np.sum((x[:, :, :, 0] ** 2 + x[:, :, :, 1] ** 2) ** 0.5)
8
9
10 def vectorize(x):
11     """
12     Parameters:
13     - x: array
14     Transforms x into an array with 3 channels
15     """
16     return np.stack([x, x, x], axis=-1)
17
18 def lambda_projection(x, l):
19     """
20     Parameters:
21     - x: 3d array
22     - l: lambda
23     Computes the projection of x on the l-radius ball
24     """
25     x = x.astype(np.float32)
26     x_norm = (x[:, :, 0] ** 2 + x[:, :, 1] ** 2) ** 0.5
27     maximum = np.maximum(x_norm, l * np.ones(x.shape[0:2])).astype(
28         np.float32)
29     x[:, :, 0] = np.divide(l * x[:, :, 0], maximum).astype(np.

```

A. APÈNDIX

```
29     x[:, :, 1] = np.divide(l * x[:, :, 1], maximum).astype(np.
30                             float32)
31
32
33
34 def h_projection(x, d0, k):
35
36     d0 = d0.astype(np.float32)
37     h = np.divide(1.0, 1.0 + float(k) * np.exp(-0.1 * d0)).astype(np.
38                             float32)
39
40     x = x.astype(np.float32)
41     x_norm = (np.sum(x[:, :, :, 0] ** 2 + x[:, :, :, 1] ** 2, axis
42                      =2)) ** 0.5
43     maximum = np.maximum(x_norm, h).astype(np.float32)
44     x[:, :, :, 0] = np.divide(np.multiply(vectorize(h), x[:, :, :, 0]),
45                               vectorize(maximum)).astype(np.float32)
46     x[:, :, :, 1] = np.divide(np.multiply(vectorize(h), x[:, :, :, 1]),
47                               vectorize(maximum)).astype(np.float32)
48
49
50     return x
```

Per acabar, dues funcions especialment útils són una per normalitzar imatges i una per separar els 3 canals de color d'una imatge.

```
1 def normalize(img):
2     img = np.array(img).astype(np.float32)
3     min = np.min(img)
4     max = np.max(img)
5
6     return (img - min) * (1.0 / (max - min))
7
8
9 def RGB(img):
10    """
11        Parameters:
12        - img: RGB image
13        Separates the image channels
14    """
15    R = img[:, :, 0]
16    G = img[:, :, 1]
17    B = img[:, :, 2]
18    return R, G, B
```

A.1.3 Funcions auxiliars pròpies del *dehazing*

Ara adjuntarem les funcions auxiliars pròpies de l'algorisme de *dehazing* proposat, és a dir, les funcions per calcular la transmissió, $t(x)$; la llum atmosfèrica, \mathbf{A} , ... Algunes de les següents funcions estan basades en el codi de [18].

```
1 def DarkChannel(img, patch_size):
2    """
3        Parameters:
4        - img: 3d image
5        - patch_size: size of the patches for the dark channel prior
6    """
```

```

7     # Separate Channels
8     r, g, b = RGB(img)
9     # Minimum from each channel
10    dc = np.minimum(np.minimum(r, g), b)
11    # Rectangular patch
12    patch = np.ones([patch_size, patch_size])
13    # Get dark channel
14    dark = erosion(dc, patch)
15
16    return dark
17
18
19 def AtmosphericLight(img, dc_img):
20     """
21     Parameters:
22     - img: 3d image
23     - dc_img: img's dark channel
24     """
25
26     # Total of pixels
27     size = img.shape[0]*img.shape[1]
28     # Pixels that won't be used
29     pixels = int(max(math.floor(size/1000), 1))
30     # Vectorize the image and its dark channel
31     dc_vec = dc_img.reshape(size)
32     img_vec = img.reshape(size, 3)
33
34     # Order indices from dark (low) to bright (high)
35     dc_order = dc_vec.argsort()
36     # Select brighter indices
37     dc_brighter = dc_order[size-pixels::]
38
39     A = np.zeros([1, 3])
40     for i in range(pixels):
41         A = A + img_vec[dc_brighter[i]]
42
43     A = A[0]/pixels
44
45     return A
46
47 def Transmission(img, A, patch_size):
48     """
49     Parameters:
50     - img: 3d image
51     - A: img's atmospheric light
52     - patch_size: dark channel's patch size
53     """
54     # Divide each channel by its atmospheric light
55     atm_mean_img = np.zeros(img.shape)
56     for i in range(3):
57         atm_mean_img[:, :, i] = img[:, :, i]/A[i]
58     # Compute the rough transmission map
59     t0 = 1 - 0.95*DarkChannel(atm_mean_img, patch_size)
60
61     return t0
62
63

```

A. APÈNDIX

```
64 def BoxFilter(img, r):
65     """
66     Parameters:
67     - img: 3d image
68     - r: height/width of the filtering box
69     """
70     area = (r*2+1)**2.0
71     ny, nx = img.shape
72     filtered_img = np.zeros((ny-2*r, nx-2*r), dtype=float)
73     for k in range(-r, r+1):
74         for l in range(-r, r+1):
75             filtered_img += img[r+k:ny-r+k, r+l:nx-r+l] / area
76
77     return filtered_img
78
79
80 def GuidedFilter(guide, img, r):
81     """
82     Parameters:
83     - guide: filtering image guide
84     - img: image to be filtered
85     - r: height/width of the filtering box
86     """
87     eps = 0.0000001
88     guide = guide/255
89
90     IG = np.multiply(guide, img)
91     G2 = np.multiply(guide, guide)
92
93     FI = BoxFilter(img, r)
94     FG = BoxFilter(guide, r)
95     FIG = BoxFilter(IG, r)
96     FG2 = BoxFilter(G2, r)
97
98     A = (FIG - np.multiply(FI, FG)) / (FG2 - np.multiply(FG, FG) + eps)
99     B = FI - np.multiply(A, FG)
100
101    FI = BoxFilter(A, r)
102    FG = BoxFilter(B, r)
103
104    output = np.multiply(FI, guide[2*r:-2*r, 2*r:-2*r]) + FG
105
106    return output
107
108
109 def GFTransmission(img, A, guide, dc_patch, r):
110     """
111     Parameters:
112     - img: 3d image
113     - A: 3d array, atmospheric light
114     - guide: filtering image guide
115     - dc_patch: dark channel patch size
116     - r: height/width of the filtering box
117     """
118     ny, nx, c = img.shape
119
```

```

120     TF_brut = Transmission(img, A, dc_patch)
121     TF = np.zeros([ny, nx])
122     TF[2 * r:-2 * r, 2 * r:-2 * r] = GuidedFilter(guide, TF_brut, r)
123
124     # Edge corrections
125     TF[0:2 * r + 2, :] = TF_brut[0:2 * r + 2, :]
126     TF[:, 0:2 * r + 2] = TF_brut[:, 0:2 * r + 2]
127     TF[(ny - 1 - 2 * r - 2):, :] = TF_brut[(ny - 1 - 2 * r - 2):, :]
128     TF[:, (nx - 1 - 2 * r - 2):] = TF_brut[:, (nx - 1 - 2 * r - 2):]
129
130     return TF

```

A.1.4 Funcions de correcció de color

Adjuntam les funcions de correcció de color emprades per l'output de l'algorisme de *dehazing*.

```

1 def MinMaxSat(img, s1, s2):
2     """
3     Parameters:
4     - img: 3d image
5     - s1: inferior boundary for saturation
6     - s2: upper boundary for saturation
7     Selects the range of saturation of our image
8     """
9
10    shape = img.shape
11    flat_img = img.flatten()
12    dim = 1
13    for i in shape:
14        dim = dim*i
15    dim = np.intc(dim)
16
17    pct1 = np.intc(s1 * dim / 100)
18    pct2 = np.intc(s2 * dim / 100)
19
20    ordered = np.sort(flat_img)
21
22    minsat = ordered[pct1]
23    maxsat = ordered[dim - 1 - pct2]
24
25    return minsat, maxsat, flat_img, dim
26
27 def simplest_color_balance(img, s1, s2):
28     """
29     Parameters:
30     - img: 3d image
31     - s1: inferior boundary for saturation
32     - s2: upper boundary for saturation
33     Balances the image colour
34     """
35    minim, maxim, data, dim = MinMaxSat(img, s1, s2)
36
37    if maxim <= minim:
38        for i in range(dim):
39            data[i] = maxim

```

A. APÈNDIX

```
41     else:
42         scale = 255. / (maxim - minim)
43         for i in range(dim):
44             if data[i] < minim:
45                 data[i] = 0
46             elif data[i] > maxim:
47                 data[i] = 255
48             else:
49                 data[i] = scale * (data[i] - minim)
50
51 colored = data.reshape(img.shape)
52
53 return colored
```

A.1.5 Algorismes de *denoising* o reducció de renou

Model de Tikhonov

El model de *denoising* de Tikhonov obtingut a partir de l'algorisme Primal-Dual de Chambolle-Pock implementat a Python és el següent:

```
1 def tikhonov_denoise(img, iterations, sigma, tau, lmb):
2
3     p = np.zeros_like(channel_grad(img))
4     u1 = np.zeros_like(img)
5     s = sigma
6     t = tau
7     su = np.zeros_like(img)
8     for i in range(iterations):
9         print("Tikhonov. Iteracio", i)
10        u0 = np.copy(u1)
11        p = 2 * (p + s * channel_grad(su)) / (2 + s)
12        u1 = (u0 + t * channel_div(p) + lmb * t * img)
13            / (1 + lmb * t)
14        su = 2 * u1 - u0
15
16    return u1
```

Model de la *Total Variation*

El model de *denoising* de la TV obtingut a partir de l'algorisme Primal-Dual de Chambolle-Pock implementat a Python és el següent:

```
1 def total_variation_denoise(img, iterations, sigma, tau, lmb):
2     p = np.zeros_like(channel_grad(img))
3     u1 = np.zeros_like(img)
4     s = sigma
5     t = tau
6     su = np.zeros_like(img)
7     for i in range(iterations):
8         print("TV. Iteracio", i)
9         u0 = np.copy(u1)
10        p = projection(p + s * channel_grad(su))
11        u1 = (t * lmb * img + u0 + t * channel_div(p))
12            / (1 + t * lmb)
13        su = 2 * u1 - u0
```

```

14
15     return u1

```

Programa principal de *denoising*

El programa principal per llegir la imatge, afegir-li renou gaussià, i aplicar l'algorisme de denoising és el següent:

```

1 # Read image
2 filename = r'IMAGE DIRECTORY'
3 img = imageio.imread(filename)
4
5 # Normalize [0,1]
6 img = normalize(img)
7
8 # Add noise
9 mean = 0
10 sigma = (3/255) ** 0.5
11 gaussian = np.random.normal(mean, sigma,
12                               (img.shape[0], img.shape[1]))
13 img_noisy = img + vectorize(gaussian)
14
15 # normalize [0,1]
16 img_noisy = normalize(img_noisy)
17 imageio.imsave(r'IMAGE SAVE DIRECTORY', img_noisy)
18 img_noisy2 = np.copy(img_noisy)
19
20 # Apply denoising algorithms
21 img_denoised_tik = tikhonov_denoise(img_noisy, iter,
22                                       sigma, tau, lambda)
23 img_denoised_tik = normalize(img_denoised_tik)
24 imageio.imsave(r'IMAGE SAVE DIRECTORY', img_denoised_tik)
25
26 img_denoised_tv = total_variation_denoise(img_noisy, iter,
27                                             sigma, tau, lambda)
28 img_denoised_tv = normalize(img_denoised_tv)
29 imageio.imsave(r'IMAGE SAVE DIRECTORY', img_denoised_tv)

```

A.1.6 Algorismes de *dehazing* o reducció de boira

Model proposat, $p = 2$

El model proposat de *dehazing* per el cas $p = 2$ a partir de l'algorisme Primal-Dual de Chambolle-Pock és el següent:

```

1 def dehazing2(f, t0, eta, sigma, l, k, gamma, tau, imgA, img,
2               iterations):
3
4     tt = time.time_ns()
5
6     d0 = (-(1.0 / eta) * np.log(t0))
7
8     d = d0
9     d_line = d
10
11    p = np.zeros_like(grad(d0))

```

A. APÈNDIX

```

11
12     g = f + vectorize(d0)
13     g_line = f + vectorize(d0)
14
15     q = np.zeros_like(channel_grad(g))
16
17     h = np.divide(1.0, 1.0 + k * np.exp(-5 * d0)).astype(np.float32)
18     h = vectorize(h)
19     h = np.stack([h, h], axis=-1)
20
21     for i in range(iterations):
22         print('iter: %s' % (str(i).zfill(4)))
23         t1 = time.time_ns()
24
25         pn = (p + sigma * grad(d_line)) / (sigma / l + 1)
26         qn = (q + sigma * channel_grad(g_line)) / (sigma + 1)
27         dn = (d + tau * div(p) + gamma * tau * d0
28                + tau * np.sum(g - f, axis=2))
29                / (3 * tau + gamma * tau + 1)
30         vdn = vectorize(dn)
31         gn = (tau * (f + vectorize(d)) + g
32                + tau * channel_div(np.multiply(np.sqrt(h), q)))
33                / (1 + tau)
34
35         d_line = 2 * dn - d
36         g_line = 2 * gn - g
37
38         p, q, d, vd, g = pn, qn, dn, vdn, gn
39
40         t2 = time.time_ns()
41         print('iter time: ', (t2 - t1) / (10 ** 9), 'seconds')
42
43         image = imgA - np.sign(imgA - img) * np.exp(g)
44
45         tf = time.time_ns()
46         print('process finished! :)')
47         print('total time: ', (tf - tt) / (10 ** 9), 'seconds')
48
49     return image

```

Model proposat, $p = 1$

El model proposat de *dehazing* per el cas $p = 1$ a partir de l'algorisme Primal-Dual de Chambolle-Pock és el següent:

```

1 def dehazing1(f, t0, eta, sigma, l, depth, gamma, tau, imgA, img,
2               iterations):
3
4     tt = time.time_ns()
5
6     d0 = (-(1.0 / eta) * np.log(t0))
7
8     d = d0
9     d_line = d
10
11    p = np.zeros_like(grad(d0))

```

```

12     g = f + vectorize(d0)
13     g_line = f + vectorize(d0)
14
15     q = np.zeros_like(channel_grad(g))
16
17     for i in range(iterations):
18         print('iter: %s' % (str(i).zfill(4)))
19         t1 = time.time_ns()
20
21         pn = lambda_projection(p + sigma * grad(d_line), 1)
22         qn = h_projection(q + sigma * channel_grad(g_line),
23                             d0, depth)
24         dn = (d + tau * div(p) + gamma * tau * d0
25               + tau * np.sum(g - f, axis=2))
26               / (3 * tau + gamma * tau + 1)
27         vdn = vectorize(dn)
28         gn = (tau * (f + vectorize(d)) + g + tau * channel_div(q))
29               / (1 + tau)
30
31         d_line = 2 * dn - d
32         g_line = 2 * gn - g
33
34         p, q, d, vd, g = pn, qn, dn, vdn, gn
35
36         t2 = time.time_ns()
37         print('iter time: ', (t2 - t1) / (10 ** 9), 'seconds')
38
39         image = imgA - np.sign(imgA - img) * np.exp(g)
40
41         tf = time.time_ns()
42         print('process finished! :)')
43         print('total time: ', (tf - tt) / (10 ** 9), 'seconds')
44
45     return image

```

Programa principal de *dehazing*

El programa principal de *dehazing*, per estimar t_0 i \mathbf{A} , i posteriorment restaurar l’imatge amb la corresponent correcció de color, és el següent:

```

1 # Set constants
2 eta = 1.0
3 k = 50
4 l = 0.01
5 gamma = 0.1
6 sigma = 1
7 tau = 0.11
8
9 # Read image
10 nom_img = 'IMAGE NAME'
11 dir_img = os.path.join('IMAGE FOLDER', nom_img)
12 img = imageio.imread(dir_img)
13 img = img
14 print(img.shape)
15
16 # Compute dark channel, A, guide and t0
17 dc_img = DarkChannel(img, 3)

```

A. APÈNDIX

```
18
19 A = AtmosphericLight(img, dc_img)
20 print(A)
21
22 guide = np.zeros(img.shape)
23 for i in range(3):
24     guide[:, :, i] = img[:, :, i]/A[i]
25 guide = np.mean(guide, -1)
26
27 t0 = GFTransmission(img, A, guide, 3, 5)
28
29 imageio.imwrite(os.path.join('SAVE FOLDER', nom_img),
30                  (normalize(dc_img)*255).astype(np.uint8))
31 imageio.imwrite(os.path.join('SAVE FOLDER', nom_img),
32                  (normalize(t0)*255).astype(np.uint8))
33
34 imgA = np.ones(np.shape(img))
35 imgA[:, 0] *= A[0]
36 imgA[:, 1] *= A[1]
37 imgA[:, 2] *= A[2]
38
39 # Compute f
40 f = (1.0 / eta) * np.log(np.abs(imgA - img) + 0.000001)
41
42 # Dehaze image
43 dehazed = dehazing2(f, t0, eta, sigma, l, k, gamma, tau,
44                      imgA, img, 100)
45 imageio.imwrite(os.path.join('SAVE FOLDER', nom_img),
46                  (normalize(dehazed)*255).astype(np.uint8))
47
48 dehazed = dehazing1(f, t0, eta, sigma, l, k, gamma, tau,
49                      imgA, img, 100)
50 imageio.imwrite(os.path.join('SAVE FOLDER', nom_img),
51                  (normalize(dehazed)*255).astype(np.uint8))
52
53 # Color balance
54 colored = simplest_color_balance(dehazed,
55                                   lower_percent, upper_percent)
56 imageio.imwrite(os.path.join('SAVE FOLDER', nom_img),
57                  (normalize(colored)*255).astype(np.uint8))
```

Cal dir que aquest programa principal està fet per guardar cada resultat a carpetes diferents amb el mateix nom de l'arxiu

BIBLIOGRAFIA

- [1] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior." *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, pp. 2341–2353, 01 2011. 1, 6.2.1, 6.2.1, 6.2.2
- [2] F. Fang, F. Li, and T. Zeng, "Single image dehazing and denoising: A fast variational approach," *SIAM Journal on Imaging Sciences [electronic only]*, vol. 7, 04 2014. 1, 6.2.2, 6.2.2, 6.2.4
- [3] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, pp. 120–145, 2011. 1, 5.2.2
- [4] A. Vera and P. Alegría, *Un Curso de Análisis Funcional: Teoría y Problemas*. AVL, 1997. 2, 6
- [5] J. D. Martín-Gómez, "Apuntes de clase de análisis funcional y optimización," 2010. [Online]. Available: <http://departamento.us.es/edan/php/asig/LICMAT/LMAFO/ApuntesAFO1011.pdf> 2, 2.1.2
- [6] R. C. James, "A non-reflexive banach space isometric with its second conjugate space," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 37, no. 3, pp. 174–177, 1951. [Online]. Available: <http://www.jstor.org/stable/88062> 4
- [7] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer New York, NY, 2010. 3.3.2
- [8] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, pp. 259–268, 11 1992. 3.4, 6.1
- [9] B. Dacorogna, *Direct Methods in the Calculus of Variations*. Springer New York, NY, 2010. 4.2.2
- [10] A. Chambolle and et. al, "Continuous optimization, an introduction," 2022. [Online]. Available: <http://www.cmap.polytechnique.fr/~antonin/Opti/CoursOpti2020.pdf> 5.1.1
- [11] C. Clason, "Lecture notes on nonsmooth analysis and optimization," 2022. [Online]. Available: <https://imsc.uni-graz.at/clason/skripte/NonsmoothNotes.pdf> 5.1.2

BIBLIOGRAFIA

- [12] J. Hadamard, “Sur les problèmes aux dérivés partielles et leur signification physique,” *Princeton University Bulletin*, vol. 13, pp. 49–52, 1902. 6.1
- [13] A. N. Tikhonov, “Solution of incorrectly formulated problems and the regularization method,” *Soviet Math. Dokl.*, vol. 4, pp. 1035–1038, 1963. 6.1
- [14] C. Li, J. Guo, R. Cong, Y. Pang, and B. Wang, “Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior,” *IEEE Transactions on Image Processing*, vol. PP, pp. 1–1, 09 2016. 6.2.1
- [15] P. Drews-Jr, E. Nascimento, F. Codevilla, S. Botelho, and M. Campos, “Transmission estimation in underwater single images,” *Proceedings of the IEEE International Conference on Computer Vision*, 12 2013. 6.2.1
- [16] D. Berman, T. Treibitz, and S. Avidan, “Non-local image dehazing,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 1674–1682. 7
- [17] A. Golts, D. Freedman, and M. Elad, “Unsupervised single image dehazing using dark channel prior loss,” *IEEE Transactions on Image Processing*, vol. 29, pp. 2692–2701, 2020. 7
- [18] Z. He, “Single image dehaze,” https://github.com/He-Zhang/image_dehaze, 2021. A.1.3