



HAL
open science

Quelques méthodes de filtrage en Traitement d'Image

Maitine Bergounioux

► **To cite this version:**

| Maitine Bergounioux. Quelques méthodes de filtrage en Traitement d'Image. 2011. hal-00512280v2

HAL Id: hal-00512280

<https://hal.science/hal-00512280v2>

Preprint submitted on 24 Feb 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

QUELQUES MÉTHODES DE FILTRAGE EN TRAITEMENT D'IMAGE

par

Maitine Bergounioux

Résumé. — Nous présentons quelques méthodes « de base » en filtrage des images numériques. Un bref aperçu du filtrage unidimensionnel est donné puis les techniques linéaires et non linéaires sont abordées dans le cadre des images 2D. Nous terminons par une ouverture sur les méthodes variationnelles très utilisées actuellement pour la déconvolution et la restauration des images.

Abstract (Some filtering methods in image processing). — We present basic image processing denoising methods. We first recall briefly the main features of linear 1D filtering techniques. Then we present linear and non linear standard methods for (2D) images. We end with variational methods that are more and more used for deconvolution and image restoration.

Table des matières

1. Introduction.....	2
2. Les images numériques.....	3
3. Filtrage unidimensionnel.....	7
4. Débruitage par filtrage linéaire.....	10
5. Débruitage par filtrage non linéaire.....	33
6. Filtrage variationnel.....	40
7. Conclusion.....	52
Appendice A. Quelques outils mathématiques.....	52
Références.....	62

Classification mathématique par sujets (2000). — 68U10, 49J40.

Mots clefs. — Traitement d'image, filtrage, débruitage, segmentation.

1. Introduction

L'étude d'un signal nécessite de supprimer au maximum le bruit parasite dû aux conditions d'acquisition. L'un des buts du filtrage est de « nettoyer » le signal en éliminant le plus de bruit possible tout en préservant le maximum d'informations. En outre, l'information contenue dans un signal n'est pas forcément entièrement pertinente : il faut « sélectionner » l'information utile suivant l'usage que l'on veut en faire. Par exemple, à l'écoute d'un morceau de musique, on peut vouloir un renforcement des sons graves. Une autre finalité du filtrage est donc de sélectionner et renforcer certaines bandes de fréquences porteuses de l'information intéressante.

Le filtrage des images a la même finalité que celui des signaux 1D. Il s'agit essentiellement d'enlever le bruit (parasite) ou de sélectionner certaines fréquences. Si la notion de haute fréquence ou basse fréquence est naturelle en signal 1D (son aigu ou grave), la **fréquence spatiale** est un concept plus délicat qui découle du fait que les images appartiennent au domaine spatial. La fréquence temporelle est une grandeur qui caractérise le nombre de phénomènes qui se déroulent au cours d'un temps donné. Si en voiture, le long d'une route, on voit 2 bandes blanches PAR seconde : c'est une fréquence temporelle. Il est ensuite facile de comprendre que ce concept de fréquence « temporelle » peut aussi se traduire en fréquence spatiale en disant qu'il y a 200 bandes blanches PAR kilomètre.

Dans une image, les détails se répètent fréquemment sur un petit nombre de pixels, on dit qu'ils ont une fréquence élevée : c'est le cas pour les textures fines (comme les feuilles d'un arbre ou de l'herbe) et certains contours de l'image. Au contraire, les fréquences basses correspondent à de faibles variations diluées sur de grandes parties de l'image, par exemple des variations de fond de ciel.

Nous verrons dans la suite que la plupart des filtres agissent sélectivement sur ces fréquences pour les sélectionner, en vue de les amplifier ou de les réduire tout comme dans le cas 1D.

Les images peuvent être entâchées de dégradations de nature différente suivant les conditions d'acquisition. On s'intéressera ici plus particulièrement

- aux perturbations modélisés par un bruit additif, gaussien
- au flou (modélisé par un opérateur de convolution)
- au bruit impulsionnel dit « poivre et sel ».

Les bruits multiplicatifs et/ou poissoniens sont difficiles à appréhender : nous n'en parlerons pas ici. De la même façon, nous n'aborderons pas le filtrage par ondelettes qui nécessite des pré-requis importants.

2. Les images numériques

2.1. Principe de l'image numérique. — Une image est un signal bidimensionnel. Un image **analogique** est par exemple celle formée sur la rétine de l'œil ou l'images obtenue par la photographie argentique classique.

Une image **numérique** est un signal numérique composé d'unités élémentaires (appelées pixels) qui représentent chacun une portion de l'image. Contrairement au cas unidimensionnel, nous étudierons uniquement les images numériques (discrètes).

Une image numérique est définie par :

- le nombre de **pixels** qui la composent en largeur et en hauteur,
- l'étendue des teintes de gris ou des couleurs que peut prendre chaque pixel (on parle de **dynamique** de l'image).

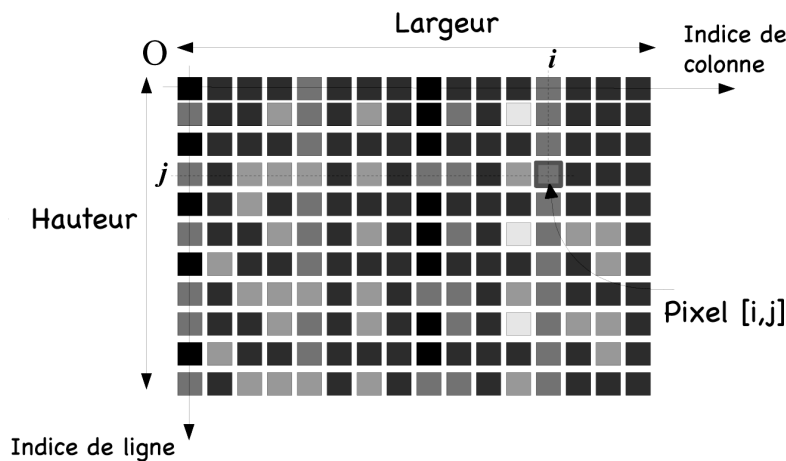


FIGURE 1. Pixels et niveaux de gris

I est l'image numérique : $I[i, j] = n$ est la valeur du niveau de gris. Lorsque $n \in [N_{min}, N_{max}]$, $N_{max} - N_{min}$ est le nombre de niveaux de gris . La **dynamique** de l'image est donnée par $\text{Log}_2(N_{max} - N_{min})$.

- **Les images binaires (noir ou blanc)** : ce sont les images les plus simples, un pixel peut prendre uniquement les valeurs noir ou blanc. C'est typiquement le type d'image que l'on utilise pour scanner du texte quand celui ci est composé d'une seule couleur.
- **Les images en niveaux de gris** : en général, les images en niveaux de gris renferment 256 teintes de gris. Par convention la valeur zéro représente le noir (intensité lumineuse nulle) et la valeur 255 le blanc (intensité lumineuse maximale). Le nombre 256 est lié à la **quantification** de l'image. En effet chaque entier représentant un niveau de gris est codé sur 8 bits. Il est donc compris entre

0 et $2^8 - 1 = 255$. C'est la quantification la plus courante. On peut coder une image en niveaux de gris sur 16 bits ($0 \leq n \leq 2^{16} - 1$) ou sur 2 bits : dans ce dernier cas le « niveau de gris » vaut 0 ou 1 : il s'agit alors d'une image binaire (Noir et Blanc).

– **Les images couleurs** : L'espace couleur est basé sur la synthèse additive des couleurs, c'est à dire que le mélange de trois composantes (par exemple (R, V, B)) donne une couleur. On garde l'information couleur, ou intensité lumineuse et chromaticité. Un pixel est codé par trois valeurs numériques. La signification de ces valeurs dépend du type de codage choisi. Le plus utilisé pour le maniement des images numériques est l'espace couleur Rouge, Vert, Bleu (R,V,B) (RGB en anglais). La restitution des couleurs sur écran utilise cette représentation (synthèse additive). Il en existe beaucoup d'autres : CMJ (ou CMY en anglais), TSL (ou HSL en anglais), YUV, YIQ, Lab, XYZ

Dans ce qui suit nous ne considérerons que des images en niveaux de gris. En effet, chaque calque couleur peut être considéré comme une image en niveaux de gris séparément et on peut lui appliquer les transformations et méthodes décrites dans ce livre. Toutefois, les techniques de recalage propres aux images couleur sont délicates et sortent largement du cadre que nous nous sommes fixés ici.

Terminons cette section par une notion statistique fort utile pour des images en niveaux de gris (y compris les images noir et blanc).

Définition 2.1 (Histogramme d'une image). — L'histogramme d'une image est l'histogramme de la série de données correspondant aux niveaux de gris des pixels. C'est une fonction définie par :

$$\forall p \in \{0, \dots, 255\} \quad h(p) = \text{Nombre des pixels ayant } p \text{ pour niveau de gris} .$$

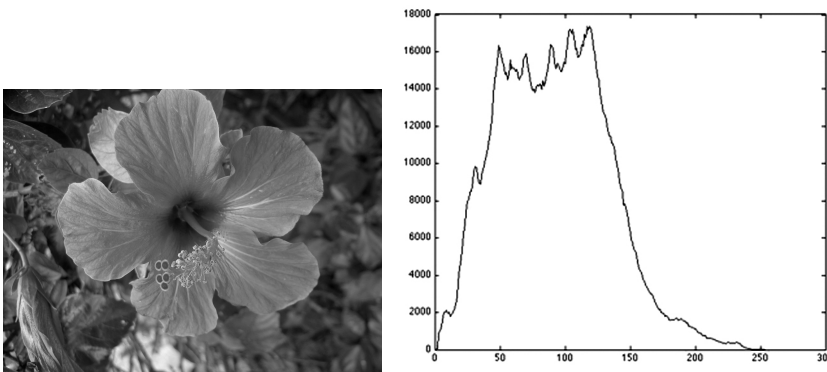


FIGURE 2. Histogramme

L'histogramme donne une indication de la dynamique de l'image (répartition des niveaux de gris) mais n'est, en aucun cas, une caractéristique de l'image.

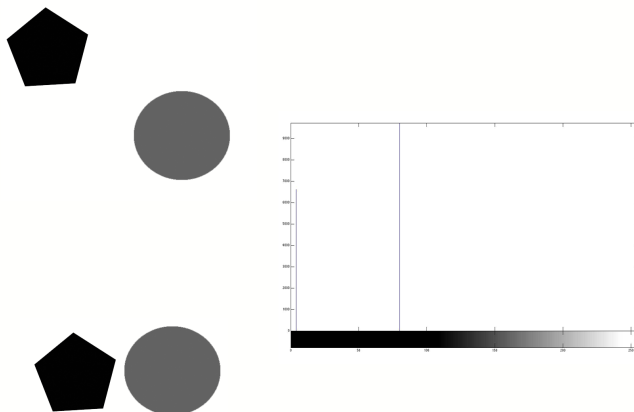


FIGURE 3. Deux images différentes peuvent avoir le même histogramme

2.2. Echantillonnage et quantification. — L'**échantillonnage** est le procédé de discrétisation spatiale d'une image consistant à associer à chaque zone rectangulaire $R(x, y)$ d'une image continue une unique valeur $I(x, y)$. On parle de **sous-échantillonnage** lorsque l'image est déjà discrétisée et qu'on diminue le nombre d'échantillons.

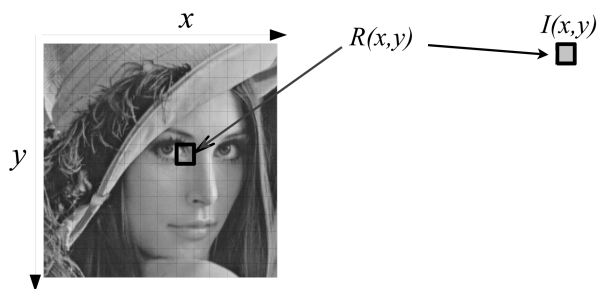


FIGURE 4. Echantillonnage

La **quantification** désigne la limitation du nombre de valeurs différentes que peut prendre $I(x, y)$, nombre déterminé en pratique par le nombre d'octets sur lequel on code la valeur numérique en question.



FIGURE 5. Quantification

Une image numérique est une image échantillonnée et quantifiée.

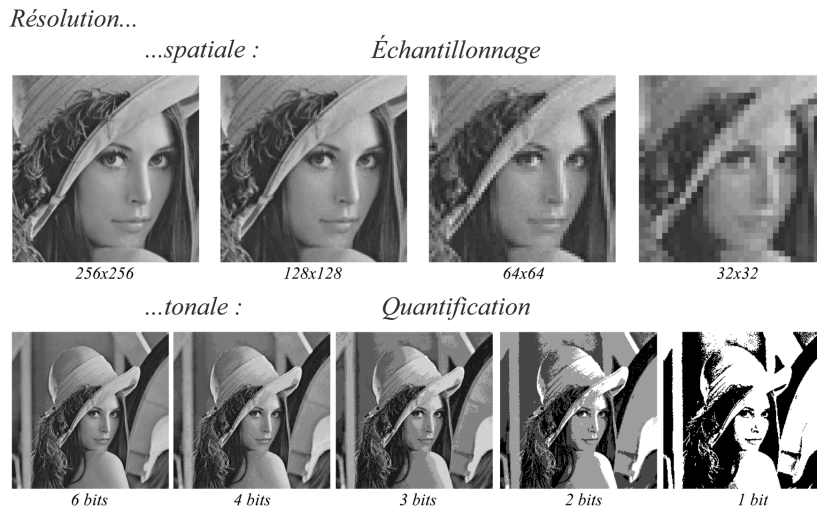


FIGURE 6. Échantillonnage et quantification

L'échantillonnage est une opération plus complexe qu'il n'y paraît. En effet, un sous-échantillonnage de l'image fait apparaître comme dans le cas unidimensionnel des fréquences parasites dues au repliement du spectre (ou **aliasing**). Le théorème d'échantillonnage de Shannon (voir [5] en 1D) s'applique aussi dans le cadre bidimensionnel.

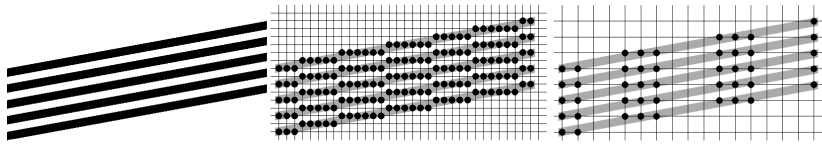


FIGURE 7. Apparition de fréquences parasites dans le cas du sous-échantillonnage

Le phénomène d'aliasing se traduit par un effet « moiré » sur les images. Nous renvoyons à [5] pour une étude détaillée de l'échantillonnage en 1D, qui s'étend sans difficulté au cas 2D.



FIGURE 8. Phénomène d'aliasing

3. Filtrage unidimensionnel

Avant de présenter les principales techniques de base pour le filtrage des images, nous rappelons brièvement le principe du filtrage unidimensionnel (pour plus de détails on peut se référer à [5]).

Pour définir un **filtre** linéaire mathématiquement, on se donne deux espaces vectoriels \mathcal{X} (entrée) et \mathcal{Y} (sortie) munis d'une topologie (par exemple séquentielle) et un opérateur \mathcal{A} linéaire qui, à un signal $e \in \mathcal{X}$ dit signal d'entrée, associe un signal $s \in \mathcal{Y}$ appelé signal de sortie :

$$\mathcal{A} : e \mapsto s := \mathcal{A}(e) .$$

Définition 3.1 (Filtre). — Un **filtre linéaire** est un **système** formé des espaces d'entrée et de sortie et d'un opérateur linéaire continu qui vérifie les deux propriétés suivantes :

1. Il est invariant dans le temps : si $\mathcal{T}_a : x \mapsto x(\cdot - a)$ est l'opérateur de translation alors

$$\mathcal{T}_a \mathcal{A} = \mathcal{A} \mathcal{T}_a .$$

2. Si $\lim_{k \rightarrow +\infty} e_k(t) = e(t)$ alors on a $\lim_{k \rightarrow +\infty} s_k(t) = s(t)$. Cette propriété est équivalente à la causalité.

Les espaces peuvent être de dimension infinie (signaux analogiques) ou finie (signaux discrets ou numériques).

Examinons l'effet d'un filtre linéaire \mathcal{A} sur un signal périodique de $\mathcal{X} = L_p^2(0, T)$ (où $T > 0$) de fréquence $\lambda = 1/T$. On sait que ce signal peut s'écrire sous la forme

$$(1) \quad x = \sum_{n \in \mathbb{Z}} c_n e_{n\lambda} ,$$

où on a posé

$$(2) \quad e_\lambda : t \mapsto \exp(2i\pi\lambda t),$$

de sorte que $e_{n\lambda} = (e_\lambda)^n$. On sait (grâce aux séries de Fourier) que la famille $(e_{n\lambda})_{n \in \mathbb{Z}}$ forme une base hilbertienne de \mathcal{X} .

Formellement ⁽¹⁾

$$\mathcal{A}x = \mathcal{A} \left(\sum_{n \in \mathbb{Z}} c_n e_{n\lambda} \right) = \sum_{n \in \mathbb{Z}} c_n (\mathcal{A}e_{n\lambda}).$$

On est donc ramené à examiner l'effet de \mathcal{A} sur $e_{n\lambda}$.

Proposition 3.2. — *Un filtre linéaire associé à tout signal exponentiel d'entrée le même signal multiplié par un facteur indépendant du temps, généralement complexe, appelé **fonction de transfert** ou **gain complexe** du filtre.*

Démonstration. — Cherchons l'image $f_\lambda = \mathcal{A}e_\lambda$ de e_λ par le filtre. On remarque que pour t fixé, on a $e_\lambda(t+u) = e_\lambda(t)e_\lambda(u)$. Donc

$$\forall u \in \mathbb{R} \quad f_\lambda(t+u) = \mathcal{A}(e_\lambda(t)e_\lambda(u)),$$

c'est-à-dire par linéarité de \mathcal{A}

$$f_\lambda(t+u) = e_\lambda(t) [\mathcal{A}e_\lambda(u)] = e_\lambda(t)f_\lambda(u).$$

Pour $u = 0$, on obtient $f_\lambda(t) = e_\lambda(t)f_\lambda(0)$; donc $\mathcal{A}e_\lambda = f_\lambda(0)e_\lambda$. Par conséquent, e_λ est une fonction propre de \mathcal{A} associée à la valeur propre $f_\lambda(0)$ qui ne dépend que de λ . La fonction $\lambda \mapsto H(\lambda) := f_\lambda(0)$ est appelée **fonction de transfert** du filtre. \square

Reprenons le signal périodique x défini par la relation (1); la sortie filtrée par \mathcal{A} est donc

$$y := \mathcal{A}x = \sum_{n \in \mathbb{Z}} c_n H(n\lambda) e_{n\lambda}.$$

En résumé, chaque coefficient de Fourier c_n de x est transformé en $\gamma_n := H(n\lambda)c_n$. Compte tenu des propriétés de la transformée de Fourier discrète \mathcal{F} , si x est un signal numérique échantillonné avec $2N$ échantillons, on peut traduire la relation précédente par

$$\forall n \in \{1, \dots, N\} \quad Y_n = H(n\lambda) X_n,$$

où $Y = \mathcal{F}(y)$ et $X = \mathcal{F}(x)$, c'est-à-dire $\mathcal{F}(y) = H \cdot \mathcal{F}(x)$, où \cdot désigne le produit composante par composante. Si on pose $h = \frac{1}{2N} \mathcal{F}^{-1}(H)$, cela donne $y = h \circledast x$ (où \circledast désigne ici la convolution circulaire discrète.) Nous obtenons donc le résultat suivant (que nous admettrons pour les signaux analogiques) :

1. On peut complètement justifier le calcul grâce à la linéarité et la continuité de l'opérateur \mathcal{A} de \mathcal{X} dans \mathcal{X}

Théorème 3.3. — *Un système linéaire continu est un filtre linéaire si et seulement si la relation entre l'entrée e et la sortie s est une **convolution** :*

$$s(t) = [h * e](t) = \int_{-\infty}^{+\infty} h(\theta) e(t - \theta) d\theta .$$

Pour les filtres à temps discret on a

$$s_n = \sum_{k \in \mathbb{Z}} h_k e_{n-k} .$$

En d'autres termes les filtres linéaires continus unidimensionnels sont et ne sont que des filtres de convolution (où la convolution est continue ou discrète).

Définition 3.4. — Le noyau de convolution h apparaissant dans le théorème 3.3 est la **réponse impulsionnelle** du filtre. La transformée de Fourier (continue ou discrète) $H := \hat{h}$ de h est la **fonction de transfert** du filtre.

Cette définition mérite d'être commentée :

- Dans le cas des signaux analogiques (d'énergie finie, i.e. dans $L^2(\mathbb{R})$ par exemple), le signal de sortie s est donné par : $s = h * e$, où e est le signal d'entrée. Si on applique la transformation de Fourier, on obtient $\hat{s} = \hat{h}\hat{e}$. En effet, si le signal d'entrée e est la mesure de Dirac δ on obtient $s = h * \delta = h$. C'est donc bien la sortie correspondant à une « impulsion ».
- Dans le cas des signaux discrets, on peut faire la même analyse. La transformation de Fourier est remplacée par la transformation de Fourier discrète et se calcule par FFT.

Généralement, on distingue les filtres suivant l'action qu'ils ont sur le spectre (c'est-à-dire par la forme de leur fonction de transfert) :

- un filtre **passé-bas** va éliminer ou atténuer fortement l'énergie des hautes fréquences d'un spectre en ne laissant « passer » que les basses fréquences ;
- un filtre **passé-haut** va éliminer ou atténuer fortement l'énergie des basses fréquences d'un spectre ;
- un filtre **passé-bande** ne conservera que l'énergie concentrée dans une bande de fréquences.
- un filtre **coupe-bande** ou filtre de **réjection** qui est le complémentaire du précédent.

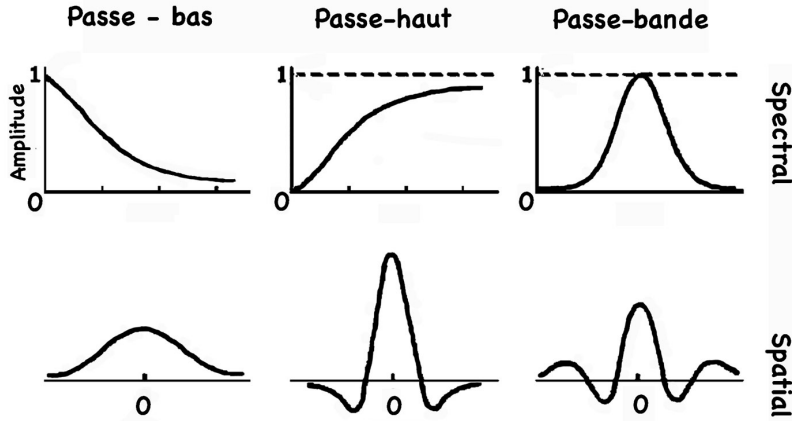


FIGURE 9. Différents types de filtrage (on n'a pas représenté le filtre coupe-bande)

4. Débruitage par filtrage linéaire

Comme nous l'avons déjà évoqué, les images peuvent être dégradées par différentes perturbations, dont nous donnons un aperçu avec la figure 10 p. 11. Le filtrage a pour but d'éliminer l'effet de ces perturbations en essayant de ne pas toucher aux informations essentielles de l'image (contours, dynamique, textures etc.)

Dans cette section, nous allons présenter les techniques **linéaires** de base pour effectuer un filtrage permettant de supprimer les effets d'un **bruit additif** que l'on supposera gaussien (centré).

Nous avons vu que, dans le cas d'un signal 1D, les filtres linéaires sont et ne sont que des filtres de convolution. Le filtrage linéaire **spatial** des images est aussi essentiellement une opération de convolution (2D).

Si f est l'image à filtrer (ou à **rehausser**) et κ la réponse impulsionnelle du filtre (spatial) on a :

$$f(x, y) * \kappa(x, y) = \mathcal{F}^{-1} \left\{ \mathcal{F}(f(x, y)) \cdot \underbrace{\mathcal{F}(g(x, y))}_{K(u, v)} \right\}.$$

où \mathcal{F} est la transformation de Fourier et K la fonction de transfert du filtre. Dans le contexte du traitement d'image le noyau κ est appelé **PSF** (*Point Spread Function*) ou **masque**)

Comme les images numériques sont des objets de dimension finie nous allons présenter les filtres dans le cas discret. Dans tout ce qui suit x et y sont des entiers (coordonnées des pixels) et f est à valeurs entières (dans $\{0, \dots, 255\}$). Comme dans le cas unidimensionnel, on peut distinguer trois types de filtrage :

- Le filtre **pas-se-bas** diminue le bruit mais atténue les détails de l'image (flou plus prononcé)
- Le filtre **pas-se-haut** accentue les contours et les détails de l'image mais amplifie le bruit
- Le filtre **pas-se-bande** élimine certaines fréquences indésirables présentes dans l'image

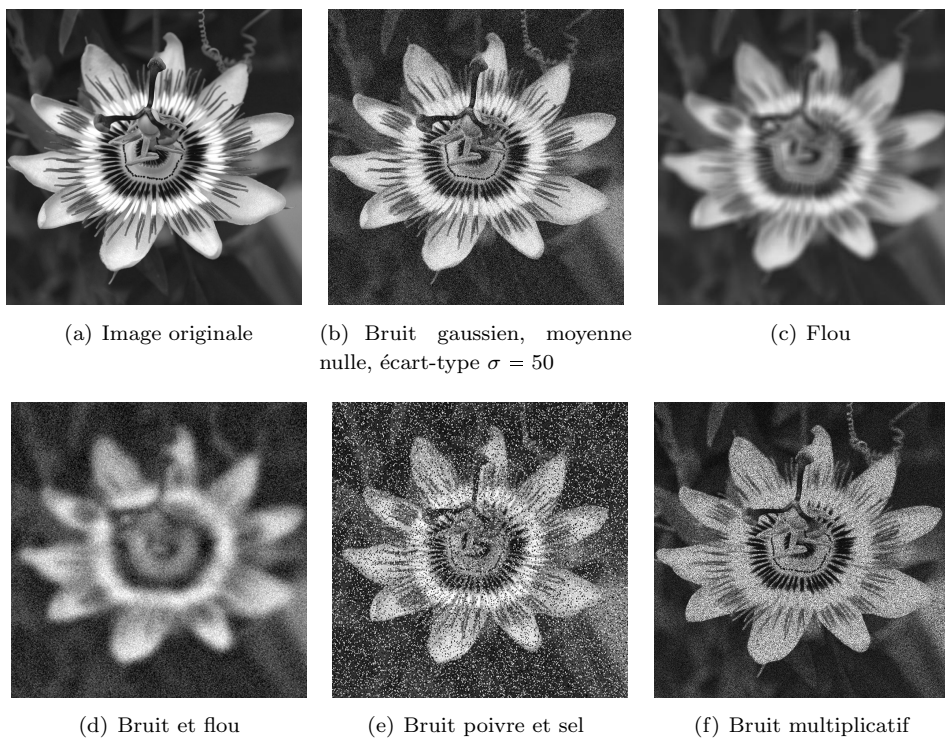


FIGURE 10. Exemples de perturbations d'une image (bruits et flou)

4.1. Filtrage spatial (bruit additif). — On ne fait pas en général une convolution globale qui serait d'une part coûteuse à implémenter, d'autre part régulariserait l'image dans sa globalité de manière trop prononcée. On choisit des noyaux dont le support est petit de façon à effectuer une convolution locale au voisinage de chaque pixel. Concrètement, cela revient à utiliser des masques (discrets) dont le support ne contient que quelques pixels au voisinage d'un pixel (x, y) :

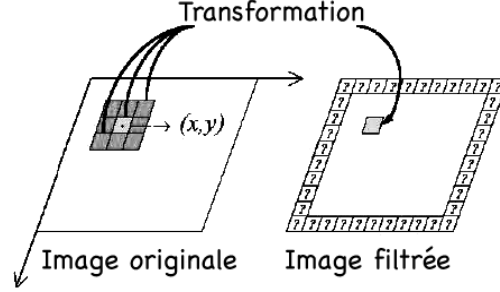


FIGURE 11. Convolution locale

Le masque du filtre κ est à support inclus dans $[x_1, x_2] \times [y_1, y_2]$:

$$g(x, y) = (f * \kappa)(x, y) = \sum_{i=x_1}^{x_2} \sum_{j=y_1}^{y_2} f(x-i, y-j)\kappa(i, j).$$

Généralement, le filtre est de dimensions d_i impaires et est symétrique. En effet, l'effet d'un filtre doit être (a priori) isotrope et il ne fait pas privilégier des directions par rapport à d'autres. Dans ce cas

$$[x_1, x_2] = \left[-\frac{d_1-1}{2}, \frac{d_1-1}{2}\right] \quad \text{et} \quad [y_1, y_2] = \left[-\frac{d_2-1}{2}, \frac{d_2-1}{2}\right],$$

$$(3) \quad (f * \kappa)(x, y) = \sum_{i=-(d_1-1)/2}^{(d_1-1)/2} \sum_{j=-(d_2-1)/2}^{(d_2-1)/2} f(x+i, y+j)\kappa(i, j).$$

Donnons un exemple avec $d_1 = d_2 = d = 3$:

w_1	w_2	w_3	$\leftarrow y-1$
w_4	w_5	w_6	$\leftarrow y$
w_7	w_8	w_9	$\leftarrow y+1$
\uparrow	\uparrow	\uparrow	
$x-1$	x	$x+1$	

TABLE 1. Représentation matricielle du filtre en (x, y) - $d_1 = d_2 = 3$

Ici $\kappa(0, 0) = w_5$. Sur cet exemple on a précisément

$$\begin{aligned} g(x, y) = & w_1 f(x-1, y-1) + w_2 f(x, y-1) + w_3 f(x+1, y-1) \\ & + w_4 f(x-1, y) + w_5 f(x, y) + w_6 f(x+1, y) \\ & + w_7 f(x-1, y+1) + w_8 f(x, y+1) + w_9 f(x+1, y+1). \end{aligned}$$

Afin de conserver la moyenne de l'image f (et donc sa luminosité), la somme des éléments du filtre est normalisée à 1 : $\sum_i w_i = 1$.

On constate tout de suite que la formule précédente ne permet pas de filtrer correctement le bord : en effet les bandes verticales de taille $(d_1 - 1)/2$ sur les bords gauche et droit de l'image ne sont pas « atteints ». De la même façon les bandes horizontales supérieures et inférieures de taille $(d_1 - 1)/2$ ne sont pas traitées non plus. Pour remédier à cet inconvénient, on effectue souvent une réflexion de l'image autour de ses bords ce qui permet de filtrer l'image réfléchie avant de la redimensionner.

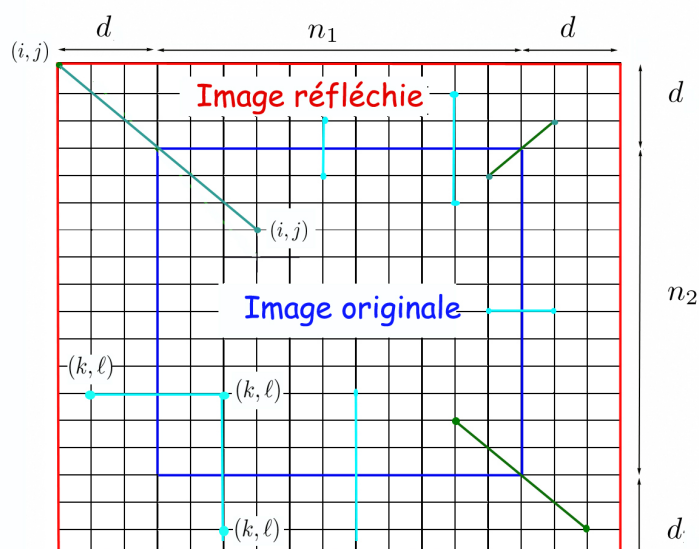


FIGURE 12. Réflexion d'une image par rapport à ses bords. Ici $(d_1 - 1)/2 = (d_2 - 1)/2 := d$ et la taille de l'image est $n_1 \times n_2$.

Un filtre 2D est dit **séparable** s'il est possible de décomposer le noyau de convolution h_{2D} en deux filtres 1D appliqués successivement en horizontal puis en vertical (ou inversement) :

$$h_{2D} = h_{1D}^V \otimes h_{1D}^H,$$

où le symbole \otimes désigne le produit tensoriel. On peut alors traiter séparément les lignes et les colonnes de l'image ce qui est un gros avantage pour l'implémentation. Pour qu'un filtre 2D soit séparable **il faut et il suffit que les coefficients de ses lignes et de ses colonnes soient proportionnels**.

Exemple 4.1 (Filtres séparables). — Ils sont obtenus comme suit, pour une dimension 3×3 par exemple. En pratique cela revient à faire un produit matriciel.

$$\begin{array}{|c|c|c|} \hline a & b & c \\ \hline \end{array} \otimes \begin{array}{|c|} \hline \alpha \\ \hline \beta \\ \hline \gamma \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline a\alpha & b\alpha & c\alpha \\ \hline a\beta & b\beta & c\beta \\ \hline a\gamma & b\gamma & c\gamma \\ \hline \end{array} = \begin{array}{|c|} \hline \alpha \\ \hline \beta \\ \hline \gamma \\ \hline \end{array} \times \begin{array}{|c|c|c|} \hline a & b & c \\ \hline \end{array}.$$

Exemple 4.2 (Filtre de moyenne passe-bas). —

$$\frac{1}{9} \cdot \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \quad \frac{1}{25} \cdot \begin{array}{|c|c|c|c|c|} \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline \end{array}$$

TABLE 2. Filtres de moyenne 3×3 et 5×5 respectivement

Nous présentons dans la figure 14. l'effet des filtres de moyenne sur une image bruitée par un bruit gaussien additif, de moyenne nulle et d'écart-type σ . Ce sont des filtres séparables.

On remarque aussi l'effet de lissage induit par l'opération de convolution qui régularise les contours en supprimant les sauts de la fonction. En effet, le contour d'une image est un endroit où le niveau de gris, c'est-à-dire la valeur de la fonction « image », varie brutalement par exemple de 0 à 255 pour un contour noir sur blanc. Cela correspond à une discontinuité de la fonction. Une opération de convolution par un noyau régulier (par exemple continu) donne comme résultat une fonction continue. Même, si la valeur de la fonction varie rapidement au voisinage d'un point, le contour qui était très marqué auparavant sera remplacé par un contour plus épais où le passage d'un niveau de gris à l'autre (par exemple du blanc au noir) se fera sur plusieurs pixels avec des niveaux de gris intermédiaires. Cela entraîne un floutage du contour. La figure 13 illustre le phénomène.

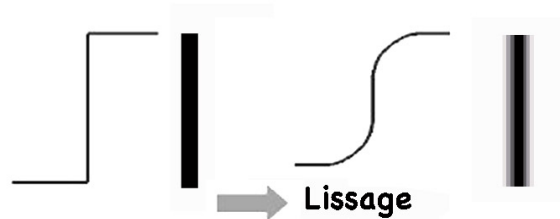


FIGURE 13. Effet de lissage



(a) Image originale

(b) Image bruitée (bruit gaussien $\sigma = 50$)



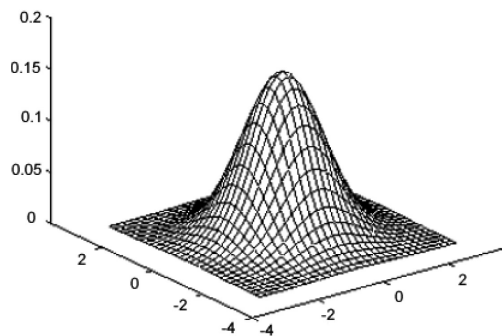
(c) Filtre de moyenne 3 x 3

(d) Filtre de moyenne 5 x 5

FIGURE 14. Effet d'un filtre de moyenne.

Exemple 4.3 (Filtre gaussien). — Un filtre gaussien est donné par discrétisation de la fonction gaussienne

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}},$$



sur un voisinage de $(0, 0)$. Ici σ est l'écart-type et la moyenne est nulle.

Si par exemple $\sigma = 0.8$ on a le filtre 3×3 suivant

$$\begin{array}{|c|c|c|} \hline G(-1, -1) & G(0, -1) & G(1, -1) \\ \hline G(-1, 0) & G(0, 0) & G(1, 0) \\ \hline G(-1, 1) & G(0, 1) & G(1, 1) \\ \hline \end{array} \simeq \frac{1}{16} \cdot \begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array}$$

et $\sigma = 1$ pour un filtre 5×5 donne environ

$$\frac{1}{300} \cdot \begin{array}{|c|c|c|c|c|} \hline 1 & 4 & 6 & 4 & 1 \\ \hline 4 & 18 & 30 & 18 & 4 \\ \hline 6 & 30 & 48 & 30 & 6 \\ \hline 4 & 18 & 30 & 18 & 4 \\ \hline 1 & 4 & 6 & 4 & 1 \\ \hline \end{array}.$$

La taille du filtre gaussien est gouvernée par σ qui doit être proportionnel à l'écart-type du bruit (s'il est connu!). En général un filtre gaussien avec $\sigma < 1$ est utilisé pour réduire le bruit, et si $\sigma > 1$ c'est dans le but de fabriquer une image qu'on va utiliser pour faire un « masque flou » personnalisé. Il faut noter que plus σ est grand, plus le flou appliqué à l'image sera important.

Les filtres présentés sont des filtres passe-bas : ils atténuent les détails de l'image (et donc le bruit additif) mais en érodant les contours ajoutent du flou à l'image. Nous verrons dans une section suivante comment atténuer le flou.

4.2. Filtrage fréquentiel (bruit additif). —

4.2.1. Filtre passe-bas. — On peut définir un filtre linéaire par sa formulation spatiale (comme convolution) mais aussi par sa formulation fréquentielle, c'est-à-dire par la façon dont il modifie les fréquences de l'image d'entrée, via sa fonction de transfert K . En effet les fréquences de l'entrée f et de la sortie g sont liées par

$$\mathcal{F}(g) = H \mathcal{F}(f).$$

Le passe-bas idéal est le filtre qui ne modifie pas les fréquences $\lambda = (\lambda_1, \lambda_2)$ telles que $\|\lambda\| \leq \delta_c$ (fréquence de coupure) et supprime les autres où $\|\cdot\|$ désigne (par exemple) la norme euclidienne de \mathbb{R}^2 .

En d'autres termes

$$H(\lambda) = \begin{cases} 1 & \text{si } \|\lambda\| \leq \delta_c \\ 0 & \text{sinon} \end{cases}$$

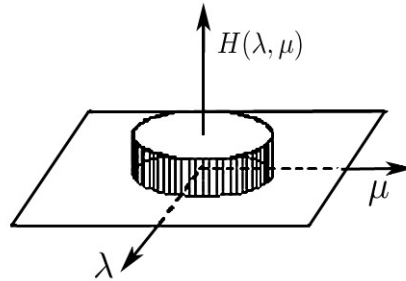
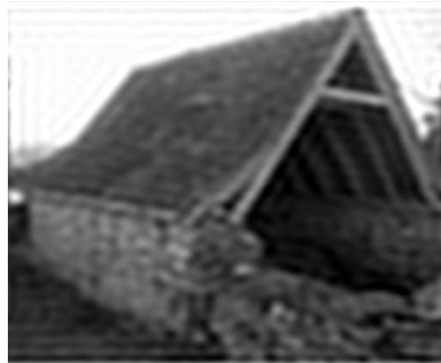


FIGURE 15. Fonction de transfert « idéale »

Pour des signaux unidimensionnels, la réponse impulsionnelle $h \in L^2(\mathbb{R})$ d'un tel filtre est un sinus cardinal $t \mapsto h(t) = \frac{\sin 2\pi\delta_c t}{\pi t}$. Un tel filtre est dit idéal car il n'est pas causal (ou réalisable [5]). En pratique, la convolution par le sinus cardinal provoque des ondulations (effet Gibbs) sur l'image filtrée.



(a) Image originale



(b) Image filtrée

FIGURE 16. Application d'un créneau « idéal » ($\delta_c \simeq 15\%$ de la taille de l'image) : on voit clairement les ondulations.



(a) Zoom (coin supérieur droit)- Original

(b) Zoom de l'image filtrée

FIGURE 17. Zoom sur le coin supérieur droit de la figure précédente : on voit les ondulations et un flou très important

Le filtre passe-bas idéal n'est pas réalisable du fait de la discontinuité de la fonction de transfert qui induit un effet Gibbs important. On fait donc une régularisation de la fonction H précédente qui aura pour effet de supprimer les discontinuités et, non pas de couper les hautes fréquences mais de les atténuer fortement. Le filtre suivant est le **filtre passe-bas de Butterworth**. La fonction de transfert est alors

$$H(\lambda) = \frac{1}{1 + \left(\frac{\|\lambda\|}{\delta_c}\right)^{2n}}$$

où δ_c est encore la fréquence de coupure et n un paramètre qui sert à régler l'approximation.

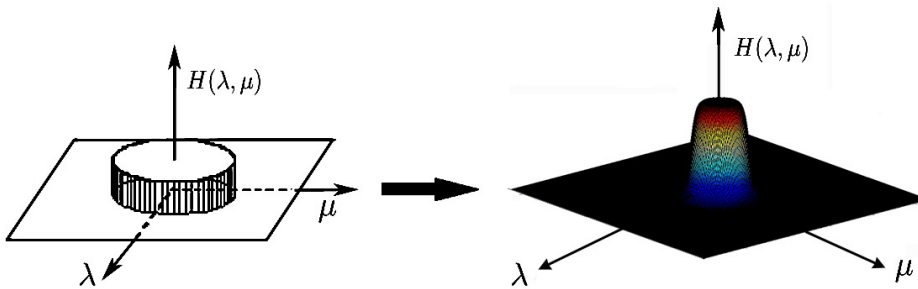
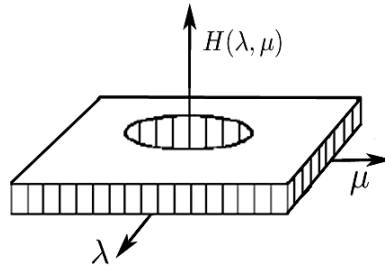


FIGURE 18. Fonction de transfert de Butterworth ($\delta_c = 0.2$, $n = 5$)

En traitement d'image, un filtre passe-bas atténue les hautes fréquences : le résultat obtenu après un tel filtrage est un adoucissement des détails et une réduction du bruit granuleux.

4.2.2. *Filtres passe-haut.* — Le filtre passe-haut idéal est obtenu de manière symétrique au passe-bas par

$$H(\lambda, \mu) = \begin{cases} 1 & \text{si } \sqrt{\lambda^2 + \mu^2} \geq \delta_c \\ 0 & \text{sinon} \end{cases}$$



Le filtre passe-haut de Butterworth est donné par

$$H(\lambda) = \frac{1}{1 + \left(\frac{\delta_c}{\|\lambda\|}\right)^{2n}}$$

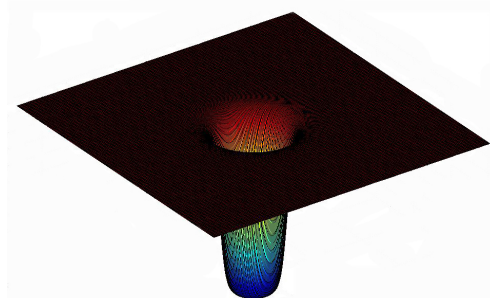


FIGURE 19. Fonction de transfert du filtre passe-haut de Butterworth ($\delta_c = 0.2, n = 5$)

Un filtre passe-haut favorise les hautes fréquences spatiales, comme les détails, et de ce fait, il améliore le contraste. Toutefois, il produit des effets secondaires :

- **augmentation du bruit** : dans les images avec un rapport Signal/ Bruit faible, le filtre augmente le bruit granuleux dans l'image.
- **effet de bord** : il est possible que sur les bords de l'image apparaisse un cadre qui correspond aux effets de bord observés lors de la convolution spatiale. Cet effet peut s'éliminer en faisant une réflexion de quelques pixels de l'image autour de son cadre.



(a) Original

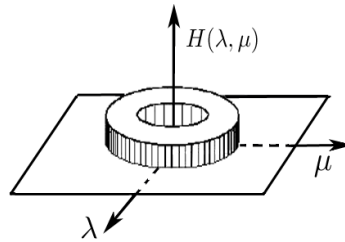
(b) Filtrage passe-haut

FIGURE 20. Filtrage passe-haut avec un filtre de Butterworth ($n = 2$ et $\delta_c \simeq 0.1 * \text{taille de l'image}$) sur une image non bruitée.

Si on filtre une image avec un filtre passe-bas, l'image obtenue par différence entre l'originale et l'image filtrée par le passe-bas correspond à un filtrage passe-haut.

4.2.3. *Filtres passe-bande.* — Mentionnons pour finir les filtres passe-bande (et coupe-bande) moins pertinents dans le cadre 2D que le cadre 1D. Ils permettent de ne garder que les fréquences comprises dans un certain intervalle :

$$H(\lambda) = \begin{cases} 1 & \text{si } \delta_c - \frac{\varepsilon}{2} \leq \|\lambda\| \leq \delta_c + \frac{\varepsilon}{2} \\ 0 & \text{sinon} \end{cases}$$



où ε est la largeur de bande et δ_c la fréquence de coupure.

4.3. Filtrage différentiel. — Dans les modèles différentiels, on considère l'image comme une fonction $f : \Omega \rightarrow [0, 255]$ où Ω est un ouvert de \mathbb{R}^2 , dont on étudie le comportement local à l'aide de ses dérivées. Une telle étude n'a de sens que si la fonction f est assez régulière. Ce n'est pas toujours le cas!! Par exemple, une image peut être continue par morceaux (comme un damier) et les points de discontinuités sont les points de contours.

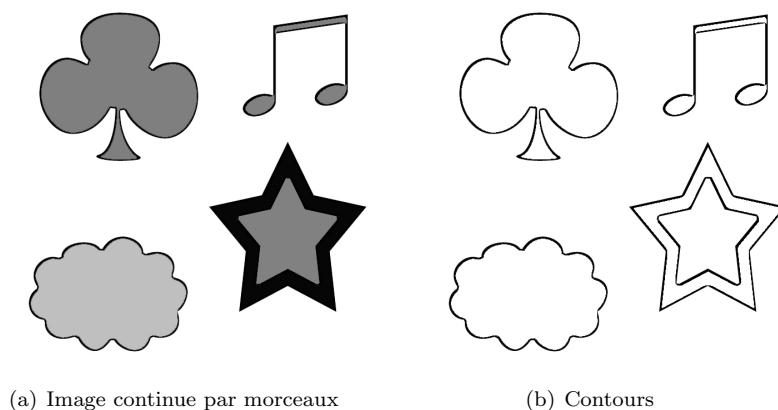


FIGURE 21. Exemple d'une image constante par morceaux

Il y a bien entendu beaucoup d'autres non-régularités dans une image, dues par exemple aux textures, au bruit etc.



FIGURE 22. Exemple d'une image texturée

Grâce au « plongement » dans « l'espace continu », un grand nombre d'opérations d'analyse peuvent s'exprimer en termes d'équations aux dérivées partielles. Ceci permet de donner un fondement mathématique satisfaisant aux traitements et aussi de fournir des méthodes pour les calculer, par des schémas numériques de résolution avant ou après discrétisation.

Les filtres différentiels permettent de mettre en évidence certaines variations spatiales de l'image. Ils sont utilisés comme traitements de base dans de nombreuses opérations comme le **rehaussement de contraste** ou la **détection de contours**.

Les filtres présentés dans cette section sont essentiellement des filtres passe-haut. Ils permettent d'isoler les détails d'une image (les contours et les textures sur une image non bruitée et le bruit en plus sur une image bruitée).

4.3.1. *Calcul des gradients discrets.* — Le gradient de l'image f en un point (pixel) $M(x, y)$, s'il existe est égal à :

$$\nabla f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right).$$

En pratique, il faut approcher ces gradients pour travailler avec des **gradients discrets** qui correspondent à des **taux de variation**, calculables même si l'image est discontinue (en particulier au voisinage d'un contour). Les approximations les plus simples des dérivées directionnelles se font par différences finies. On peut les calculer par exemple à l'aide de convolution avec des noyaux très simples : par exemple, l'approximation de $\frac{\partial f}{\partial x}$ se fait par convolution avec $[0 \ -1 \ 1]$. En effet, dans ce cas, la formule générale de convolution discrète (3) donne (avec $d_1 = 3$ et $d_2 = 0$) :

$$g(x, y) = \sum_{i=-1}^1 \sum_{j=0}^1 f(x+i, y+j) \kappa(i, j) = -f(x, y) + f(x+1, y) \simeq \frac{\partial f}{\partial x}(x, y).$$

De même l'approximation de $\frac{\partial f}{\partial y}$ se fait par convolution avec $\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$:

$$g(x, y) = -f(x, y) + f(x, y+1) \simeq \frac{\partial f}{\partial y}(x, y).$$

0	-1	1	← y
↑	↑	↑	
$x-1$	x	$x+1$	

0	← $y-1$
-1	← y
1	← $y+1$
↑	
x	

TABLE 3. Masques $(\kappa(i, j))$ des gradients par rapport à x (gauche) et y (droite)

On utilise plus souvent $[-1 \ 0 \ 1]$ et $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ qui produisent des frontières plus épaisses mais qui sont bien centrées. Ces opérations sont très sensibles au bruit et on les combine généralement avec un filtre lisseur dans la direction orthogonale à celle de dérivation, par exemple par le noyau suivant (ou sa transposée) : $[1 \ 2 \ 1]$. On obtient

alors des filtres séparables. Le calcul des dérivées directionnelles en x et y revient finalement à la convolution avec les noyaux suivants :

$$\frac{\partial f}{\partial x}(x, y) \simeq (f * h_x)(x, y) \text{ et } \frac{\partial f}{\partial y}(x, y) \simeq (f * h_y)(x, y)$$

avec

$$h_x = [-1 \ 0 \ 1] \otimes [1 \ 2 \ 1]^t = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}$$

et

$$h_y = [1 \ 2 \ 1] \otimes [-1 \ 0 \ 1]^t = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}.$$

Ce sont les **masques de Sobel**.



FIGURE 23. Image originale

(a) Noyau $[-1 \ 0 \ 1]$ (b) Noyau $[-1 \ 0 \ 1]^t$ 

(c) Gradient de Sobel horizontal



(d) Gradient de Sobel vertical



(e) Norme du gradient

FIGURE 24. Valeur absolue des gradients et des gradients de Sobel (contrastés)

TYPES DE MASQUE	NORME DU GRADIENT	DIRECTION																		
<p><i>Masques de Roberts</i></p> <table border="1" style="display: inline-table; margin-right: 20px;"> <tr><td>-1</td><td>0</td></tr> <tr><td>0</td><td>1</td></tr> </table> <table border="1" style="display: inline-table;"> <tr><td>0</td><td>-1</td></tr> <tr><td>1</td><td>0</td></tr> </table> <p style="text-align: center;">G_1, G_2</p>	-1	0	0	1	0	-1	1	0	$A = \sqrt{G_1^2 + G_2^2}$	$\theta = \frac{\pi}{4} + \arctan\left(\frac{G_2}{G_1}\right)$										
-1	0																			
0	1																			
0	-1																			
1	0																			
<p><i>Masques de Sobel</i></p> <table border="1" style="display: inline-table; margin-right: 20px;"> <tr><td>1</td><td>0</td><td>-1</td></tr> <tr><td>2</td><td>0</td><td>-2</td></tr> <tr><td>1</td><td>0</td><td>-1</td></tr> </table> <table border="1" style="display: inline-table;"> <tr><td>1</td><td>2</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>-2</td><td>-1</td></tr> </table> <p style="text-align: center;">G_x, G_y</p>	1	0	-1	2	0	-2	1	0	-1	1	2	1	0	0	0	-1	-2	-1	$A = \sqrt{G_x^2 + G_y^2}$	$\theta = \arctan\left(\frac{G_y}{G_x}\right)$
1	0	-1																		
2	0	-2																		
1	0	-1																		
1	2	1																		
0	0	0																		
-1	-2	-1																		
<p><i>Masques de Prewitt</i></p> <table border="1" style="display: inline-table; margin-right: 20px;"> <tr><td>1</td><td>0</td><td>-1</td></tr> <tr><td>1</td><td>0</td><td>-1</td></tr> <tr><td>1</td><td>0</td><td>-1</td></tr> </table> <table border="1" style="display: inline-table;"> <tr><td>1</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>-1</td><td>-1</td></tr> </table> <p style="text-align: center;">G_x, G_y</p>	1	0	-1	1	0	-1	1	0	-1	1	1	1	0	0	0	-1	-1	-1	$A = \sqrt{G_x^2 + G_y^2}$	$\theta = \arctan\left(\frac{G_y}{G_x}\right)$
1	0	-1																		
1	0	-1																		
1	0	-1																		
1	1	1																		
0	0	0																		
-1	-1	-1																		
<p><i>Masques de Kirsh</i></p> <table border="1" style="display: inline-table; margin-right: 20px;"> <tr><td>5</td><td>5</td><td>5</td></tr> <tr><td>-3</td><td>0</td><td>-3</td></tr> <tr><td>-3</td><td>-3</td><td>-3</td></tr> </table> <p>+ les 7 autres masques obtenus par permutation circulaire des coefficients G_i pour i de 1 à 8</p>	5	5	5	-3	0	-3	-3	-3	-3	<p>maximum des G_i</p>	<p>Direction correspondant au G_i sélectionné</p>									
5	5	5																		
-3	0	-3																		
-3	-3	-3																		
<p><i>Masques de Robison</i></p> <table border="1" style="display: inline-table; margin-right: 20px;"> <tr><td>1</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>-2</td><td>1</td></tr> <tr><td>-1</td><td>-1</td><td>-1</td></tr> </table> <p>+ les 7 autres masques obtenus par permutation circulaire des coefficients G_i pour i de 1 à 8</p>	1	1	1	1	-2	1	-1	-1	-1	<p>maximum des G_i</p>	<p>Idem</p>									
1	1	1																		
1	-2	1																		
-1	-1	-1																		

TABLE 4. Différents types de masques pour le gradient, sa norme et sa direction

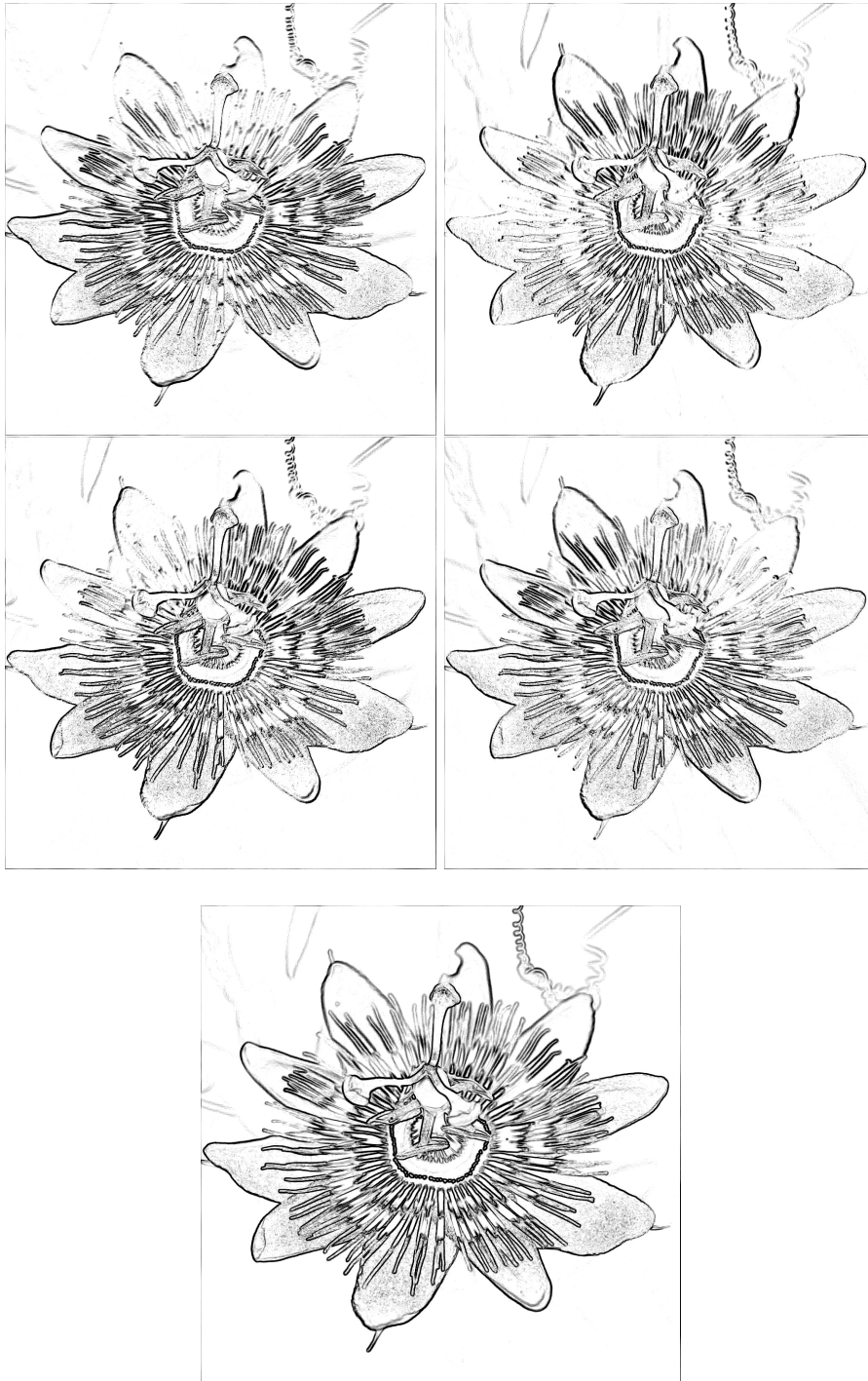


FIGURE 25. Gradients de Robinson (valeur absolue contrastée) dans 4 directions différentes (voir tableau 4) et norme ℓ^1 des gradients de Robinson (en bas)

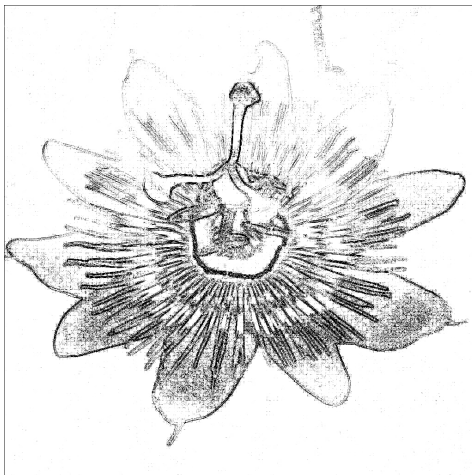
On remarque que la norme du gradient est un bon détecteur de contour : en effet un contour est un endroit où l'on observe de fortes variations de niveaux de gris, c'est-à-dire des gradients importants. Les points correspondant aux maxima de la norme du gradient sont donc des points appartenant à des contours. Toute la difficulté de la segmentation (c'est-à-dire de la recherche des contours dans une image) est alors de trouver des critères permettant de faire une sélection pertinente parmi ces points.

4.3.2. *Approximation de la dérivée seconde.* — De la même façon, l'approximation par différences finies la plus simple de la dérivée seconde est la convolution par le

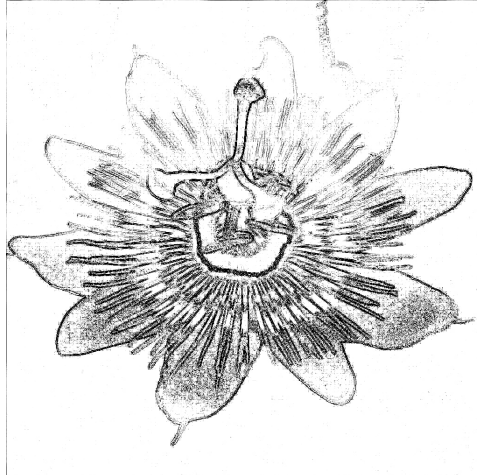
noyau $[1 \ -2 \ 1]$ pour l'approximation de $\frac{\partial^2 f}{\partial x^2}$ et $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ pour l'approximation de

$\frac{\partial^2 f}{\partial y^2}$. Le laplacien $\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$ peut donc être approché par l'un opérateurs linéaires suivants :

<i>Laplacien discret - 4</i>	<i>Laplacien discret - 8</i>	<i>Laplacien de Robinson</i>																											
<table border="1"> <tr><td>0</td><td>1</td><td>0</td></tr> <tr><td>1</td><td>-4</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>0</td></tr> </table>	0	1	0	1	-4	1	0	1	0	<table border="1"> <tr><td>1</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>-8</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>1</td></tr> </table>	1	1	1	1	-8	1	1	1	1	<table border="1"> <tr><td>1</td><td>-2</td><td>1</td></tr> <tr><td>-2</td><td>4</td><td>-2</td></tr> <tr><td>1</td><td>-2</td><td>1</td></tr> </table>	1	-2	1	-2	4	-2	1	-2	1
0	1	0																											
1	-4	1																											
0	1	0																											
1	1	1																											
1	-8	1																											
1	1	1																											
1	-2	1																											
-2	4	-2																											
1	-2	1																											



(a) Laplacien 4-connexe



(b) Laplacien 8-connexe

FIGURE 26. Calcul du laplacien (valeur absolue), avec rehaussement de contraste)

4.4. Filtrage par équations aux dérivées partielles. —

4.4.1. *Équation de la chaleur.* — Considérons un filtrage gaussien dans le cadre continu. On sait que si l'image de départ est une fonction $u_o \in L^\infty(\Omega)$ à support compact inclus dans $\Omega \subset \mathbb{R}^2$, l'image filtrée est la convolée de u_o avec un noyau gaussien

$$G_\sigma(x) = G_\sigma(x_1, x_2) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x_1^2 + x_2^2}{2\sigma^2}\right) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\|x\|^2}{2\sigma^2}\right).$$

On pose $u(t, x) = (h(t, \cdot) * u_o)(x)$ où $h(t, x) = G_{\sqrt{2t}}(x) = \frac{1}{4\pi t} \exp\left(-\frac{\|x\|^2}{4t}\right)$. Comme $h(t, \cdot) \in \mathcal{C}^\infty(\mathbb{R}^2)$ a ses dérivées bornées et $u_o \in L^1(\mathbb{R}^2)$, la convolée $u(t, \cdot)$ est aussi \mathcal{C}^∞ et on peut calculer Δu :

$$\forall t > 0, \forall x \in \mathbb{R}^2 \quad \Delta u(t, x) = \frac{\partial^2 u}{\partial x_1^2}(t, x) + \frac{\partial^2 u}{\partial x_2^2}(t, x) = (\Delta h(t, \cdot) * u_o)(x).$$

Un rapide calcul montre que

$$\Delta h(t, x) = \left(-\frac{1}{4\pi t^2} + \frac{\|x\|^2}{16\pi t^3}\right) \exp\left(-\frac{\|x\|^2}{4t}\right) = \left(-\frac{1}{t} + \frac{\|x\|^2}{4t^2}\right) h(t, x),$$

et on obtient

$$\Delta u(t, x) = \left(-\frac{1}{t} + \frac{\|x\|^2}{4t^2}\right) u(t, x).$$

D'autre part, pour $t > 0$ on peut dériver directement u par rapport à t :

$$\frac{\partial u}{\partial t}(t, x) = \iint_{\mathbb{R}^2} \frac{\partial h}{\partial t}(t, y) u_o(x - y) dy$$

ce qui donne :

$$\frac{\partial u}{\partial t}(t, x) - \Delta u(t, x) = 0 \text{ sur }]0, t[\times \mathbb{R}^2.$$

D'autre part, avec

$$u(t, x) = \iint_{\mathbb{R}^2} \frac{1}{4\pi t} \exp\left(-\frac{\|y\|^2}{4t}\right) u_o(x - y) dy$$

on obtient

$$\lim_{t \rightarrow 0} u(t, x) = \langle \delta_x, u_o \rangle = u_o(x),$$

car la famille de gaussiennes converge au sens des distributions vers la mesure de Dirac.

La fonction « filtrée » u vérifie l'équation aux dérivées partielles suivante (équation de la chaleur) :

$$(4) \quad \begin{cases} \frac{\partial u}{\partial t}(t, x) - \Delta u(t, x) = 0 \text{ dans }]0, T[\times \Omega \\ u(0, x) = u_o(x) \quad \forall x \in \Omega \end{cases}$$

où Ω est le « cadre » de l'image, i.e. l'ouvert de \mathbb{R}^2 où la fonction u est définie. On peut alors imposer soit des conditions aux limites au bord de Ω (niveau de gris fixé) soit des conditions aux limites périodiques en périodisant la fonction u (si le cadre est rectangle par exemple).

On peut alors utiliser un schéma aux différences finies pour calculer la solution de l'EDP ou une convolution par le noyau $G_{\sqrt{2t}}$ en utilisant une FFT . Suivant le temps d'évolution, on obtient une version plus ou moins lissée de l'image de départ.

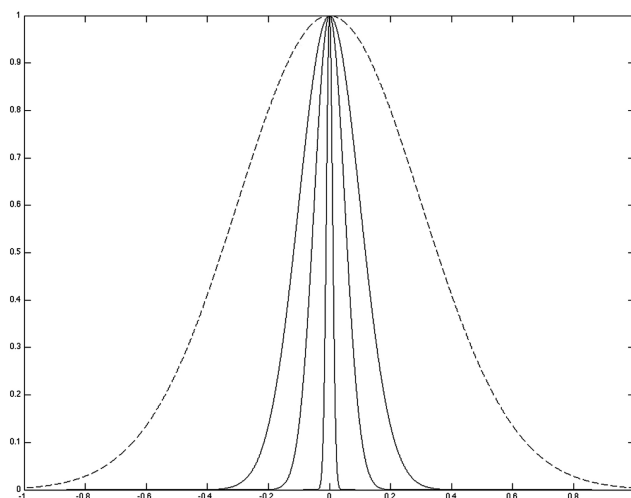


FIGURE 27. Gaussiennes



(a) Original

(b) Image bruitée $\sigma \simeq 0.1$

(c) Image filtrée

FIGURE 28. Filtrage par EDP de la chaleur avec pas de temps $dt = 0.2$ et 15 itérations

4.4.2. *Mise en œuvre numérique.* — La mise en œuvre numérique se fait avec une discrétisation en différences finies, la plupart du temps explicite en raison de la très grande taille des images (et donc des matrices associées). La condition de Neumann est assurée grâce à une réflexion de l'image par rapport à ses bords. En traitement d'image on considère souvent que la taille de l'image est donnée par le nombre de pixels de sorte que le pas de discrétisation est $h = 1$. On peut discrétiser le gradient de différentes manières (centrée, à droite, à gauche)

$$(5) \quad \begin{aligned} \delta_x u_{i,j} &= \frac{u_{i+1,j} - u_{i-1,j}}{2}, \quad \delta_y u_{i,j} = \frac{u_{i,j+1} - u_{i,j-1}}{2}, \\ \delta_x^+ u_{i,j} &= u_{i+1,j} - u_{i,j}, \quad \delta_y^+ u_{i,j} = u_{i,j+1} - u_{i,j}, \\ \delta_x^- u_{i,j} &= u_{i,j} - u_{i-1,j}, \quad \delta_y^- u_{i,j} = u_{i,j} - u_{i,j-1}. \end{aligned}$$

La norme du gradient peut se calculer par un schéma ENO (Essentially Non Oscillatory). Deux approximations possibles de $|\nabla u|$ sont

$$\nabla^+ u_{i,j} = \sqrt{\max(\delta_x^- u_{i,j}, 0)^2 + \min(\delta_x^+ u_{i,j}, 0)^2 + \max(\delta_y^- u_{i,j}, 0)^2 + \min(\delta_y^+ u_{i,j}, 0)^2},$$

ou

$$\nabla^- u_{i,j} = \sqrt{\max(\delta_x^+ u_{i,j}, 0)^2 + \min(\delta_x^- u_{i,j}, 0)^2 + \max(\delta_y^+ u_{i,j}, 0)^2 + \min(\delta_y^- u_{i,j}, 0)^2}.$$

Si l'opérateur gradient est discrétisé par différences finies à droite, alors une discrétisation possible de la divergence d'un couple $p = (p^1, p^2)$ est donnée par

$$(6) \quad (\operatorname{div} p)_{i,j} = \begin{cases} p_{i,j}^1 - p_{i-1,j}^1 & \text{si } 1 < i < N \\ p_{i,j}^1 & \text{si } i = 1 \\ -p_{i-1,j}^1 & \text{si } i = N \end{cases} + \begin{cases} p_{i,j}^2 - p_{i,j-1}^2 & \text{si } 1 < j < M \\ p_{i,j}^2 & \text{si } j = 1 \\ -p_{i,j-1}^2 & \text{si } j = M \end{cases}$$

avec (N, M) est la taille des l'images p^1 et p^2 .

4.5. **Déconvolution (cas d'un flou).** — Une autre source de perturbation d'une image est le « flou ». Nous avons vu qu'un filtre de convolution passe-bas permettait d'enlever le bruit additif mais que l'image filtrée était floutée. Cela s'explique par le fait que l'opérateur de convolution est régularisant. Un tel opérateur est souvent modélisé par un produit de convolution, de noyau positif symétrique h (qui est la plupart du temps gaussien). Il n'est pas nécessairement inversible (et même lorsqu'il est inversible, son inverse est souvent numériquement difficile à calculer).

4.5.1. *Approche « spatiale » : équation de la chaleur rétrograde (inverse).* — Nous avons constaté précédemment que faire une convolution par un noyau gaussien revient à résoudre une équation de la chaleur. Le temps final joue le rôle de l'écart type de la gaussienne. Pour faire l'opération inverse, la déconvolution on peut donc imaginer de résoudre une équation de la chaleur « rétrograde » en partant de l'état « final » qui est l'image floutée et en ajustant le temps final au rapport signal sur bruit.

$$(7) \quad \begin{cases} \frac{\partial u}{\partial t}(t, x) + \Delta u(t, x) = 0 \text{ dans }]0, T[\times \Omega \\ u(T, x) = u_o(x) \quad \forall x \in \Omega \end{cases}$$

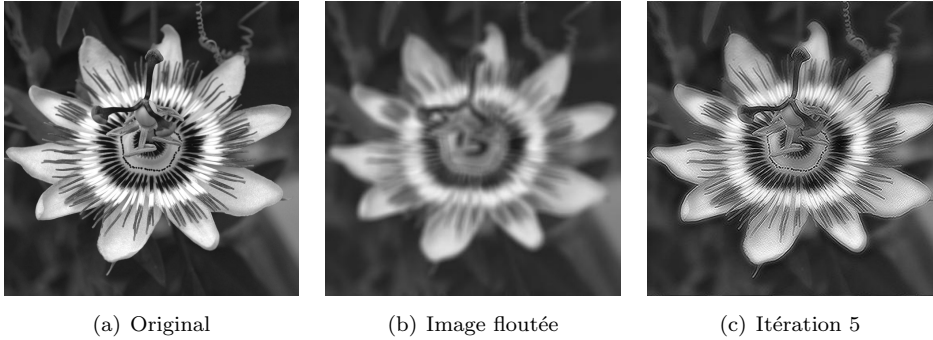


FIGURE 29. Déconvolution par équation de la chaleur inverse : $dt = 1.5$

Cette équation est notoirement mal posée (on ne peut assurer ni l'existence d'une solution, ni la stabilité d'un schéma numérique) et il convient de ne faire qu'un petit nombre d'itérations.

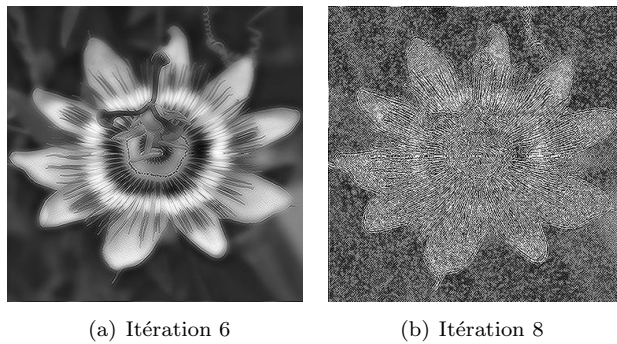


FIGURE 30. Déconvolution par équation de la chaleur inverse : $dt = 1.5$

4.5.2. *Filtre inverse et algorithme de Van Cittert.* — L'algorithme de Van Cittert repose sur la formulation fréquentielle d'une convolution. Supposons que h soit l'opérateur de flou (inconnu ou donné par étalonnage des appareils de mesure). On ne prend pas en compte le bruit. L'image floutée f vérifie $f = h * u$, où u est l'image originale (qu'on veut retrouver). Une formulation équivalente est $\hat{f} = \hat{h}\hat{u}$ c'est-à-dire

$\hat{u} = \frac{\hat{f}}{\hat{h}}$. Le filtre inverse est le plus simple des filtres. Dans certaines conditions, il peut donner de très bons résultats. Il consiste donc à calculer $1/\hat{h}$ et à l'appliquer à l'image floutée. C'est le meilleur filtre pour déconvoluer une image non bruitée. Toutefois il n'est pas toujours possible de calculer $1/\hat{h}$ car \hat{h} peut s'annuler. Une alternative est l'algorithme de Van Cittert.

Posons $\hat{g} = 1 - \hat{h}$ de sorte que **formellement** on obtient

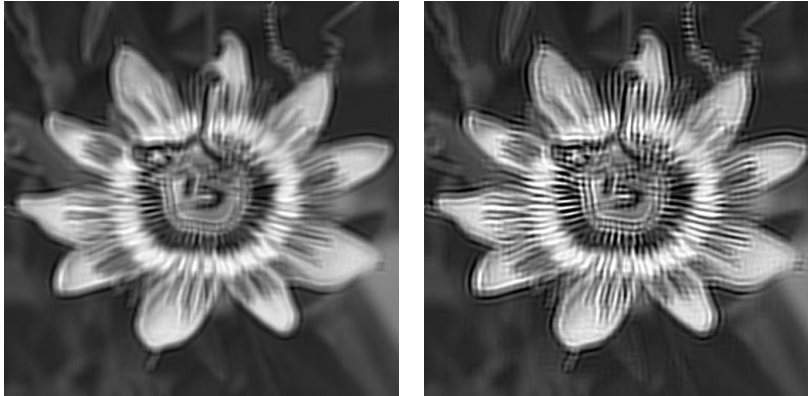
$$\hat{u} = \frac{\hat{f}}{1 - \hat{g}} = \left(\sum_{k=0}^{+\infty} \hat{g}^k \right) \hat{f}.$$

Si on pose $u_0 = f$ et $\hat{u}_n = \left(\sum_{k=0}^n \hat{g}^k \right) \hat{f}$ pour tout $n \geq 1$ on obtient

$$\hat{u}_{n+1} = \hat{f} + \hat{g}\hat{u}_n = \hat{f} + (1 - \hat{h})\hat{u}_n,$$

ou de manière équivalente

$$u_{n+1} = f + u_n - h * u_n.$$



(a) Itération 4

(b) Itération 8

FIGURE 31. Déconvolution par algorithme de Van Cittert : h est un masque gaussien de taille 9 et d'écart-type $\sigma = 1$ (connu)

La principale difficulté dans l'utilisation de cet algorithme est le choix a priori du filtre h . Le filtre inverse est une technique très utile pour la déconvolution mais il ne prend pas en compte le bruit. Si f est une image floutée et bruitée : $f = h * u + b$, où u est l'image originale à restaurer et b un bruit blanc gaussien, le passage à un filtre inverse donne $\hat{u} = \frac{\hat{f}}{\hat{h}} - \frac{\hat{b}}{\hat{h}}$. Le bruit blanc chargeant uniformément les fréquences on

a $\hat{b} \simeq 1$ et si h est un filtre passe-bas $\hat{h} \simeq 0$ au voisinage de l'infini (c'est-à-dire au dessus d'une fréquence de coupure λ_c). Il s'ensuit que le filtrage est efficace dans une bande de fréquences inférieures à λ_c mais que le bruit est amplifié au delà. Pour traiter des images à la fois floutées et bruitées on préfère utiliser le filtre de Wiener (voir section 5.3).

5. Débruitage par filtrage non linéaire

5.1. Filtres médians. — Ce sont des filtres non linéaires et donc ce ne sont pas des filtres de convolution. Ces filtres sont utiles pour contrer l'effet « Poivre et Sel » (P&S) c'est-à-dire des faux « 0 » et « 255 » dans l'image.

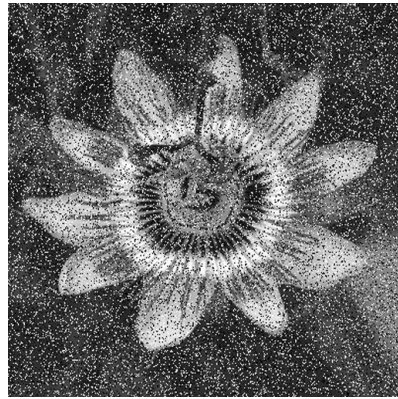
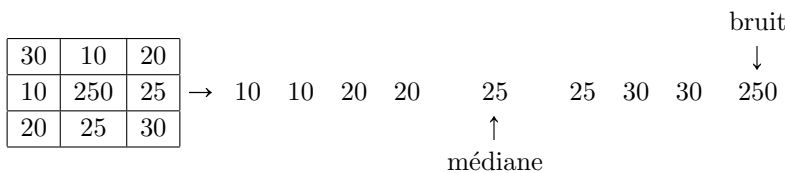


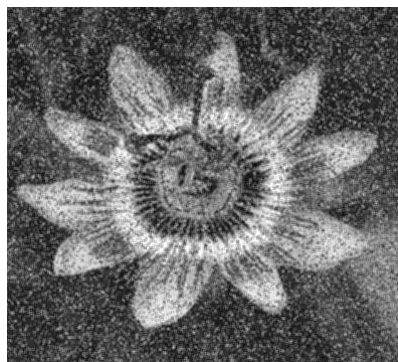
FIGURE 32. Image bruitée « Poivre et Sel »

En effet un filtre de convolution affecte au pixel traité un barycentre des valeurs des niveaux de gris des pixels situés dans un voisinage. Un pixel dont le niveau de gris est très différent des autres va donc affecter le résultat de la convolution. On préfère, dans ce cas remplacer la valeur du pixel par la valeur médiane et non la valeur moyenne (par exemple).

$$g(x, y) = \text{médiane}\{f(n, m) \mid (n, m) \in S(x, y)\},$$

où $S(x, y)$ est un voisinage de (x, y) .

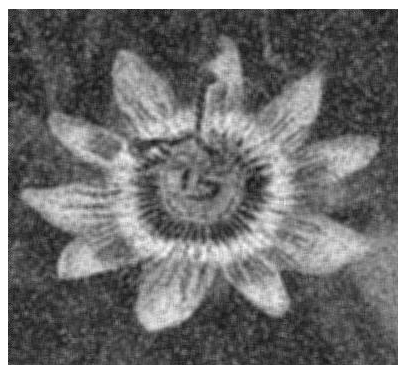




(a) Filtre de moyenne - taille 3



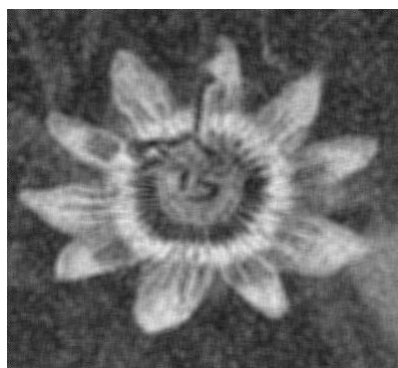
(b) Filtre médian - taille 3



(c) Filtre de moyenne - taille 5



(d) Filtre médian - taille 5



(e) Filtre de moyenne - taille 7



(f) Filtre médian - taille 7

FIGURE 33. Comparaison filtres de moyenne et filtres médians

Si le bruit P&S recouvre plus de la moitié de la dimension du filtre, le filtrage est inefficace. La localisation du bruit étant aléatoire il convient de prendre un voisinage suffisamment grand.

5.2. Modèle de Perona-Malik. — Pour améliorer les résultats obtenus par l'EDP de la chaleur, Perona et Malik ont proposé de modifier l'équation en y intégrant un processus de détection des bords :

$$(8) \quad \begin{cases} \frac{\partial u}{\partial t}(t, x) = \operatorname{div} (c(|\nabla u|)\nabla u)(t, x) & \text{dans }]0, T[\times \Omega \\ \frac{\partial u}{\partial n} = 0 & \text{sur }]0, T[\times \partial\Omega \\ u(0, x) = u_o(x) & \forall x \in \Omega \end{cases}$$

où c est une fonction décroissante de \mathbb{R}^+ dans \mathbb{R}^+ .

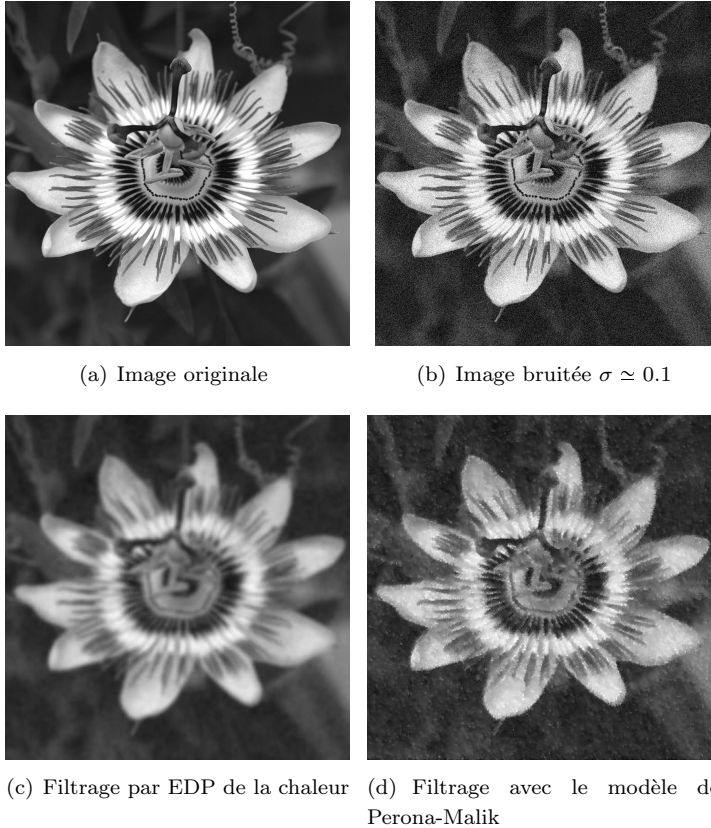


FIGURE 34. Filtrage par EDP de la chaleur avec pas de temps $dt = 0.2$ et 15 itérations et modèle de Peronna-Malik avec $dt = 0.4$, $\alpha = 1$ et 10 itérations

Si $c = 1$, on retrouve l'équation de la chaleur. On impose souvent $\lim_{t \rightarrow +\infty} c(t) = 0$ et $c(0) = 1$. Ainsi, dans les régions de faible gradient, l'équation agit essentiellement comme l'EDP de la chaleur, et dans les régions de fort gradient, la régularisation est stoppée ce qui permet de préserver les bords. Un exemple d'une telle fonction c est :

$$c(t) = \frac{1}{\sqrt{1 + (t/\alpha)^2}} \quad \text{ou} \quad c(t) = \frac{1}{1 + (t/\alpha)^2} \quad \text{où } \alpha \text{ est un paramètre positif.}$$

5.3. Approche fréquentielle : Filtre de Wiener. — Le filtre de Wiener lui, ne caractérise pas le signal et le bruit par leur forme analytique mais par leurs propriétés statistiques. On considère que les images sont des réalisations d'un processus aléatoire stationnaire : on cherche alors à minimiser la moyenne du carré de la différence entre l'image initiale et l'image restaurée. Ce filtre est très efficace pour traiter des images dégradées **à la fois par du flou et du bruit**. On considère donc une image dégradée $f = h * u + b$, où u est l'image originale à restaurer, h un noyau de convolution symétrique positif (réponse impulsionnelle du filtre « flou ») et b est une bruit de loi de probabilité (identique pour chaque pixel) μ . L'hypothèse généralement faite est celle d'un bruit plan gaussien d'écart-type σ .

On modélise donc le problème par une formulation au sens des moindres carrés en considérant l'erreur quadratique $\|f - h * u\|_{L^2}^2$ que l'on va minimiser. Le problème de minimisation n'admettant pas nécessairement de solution convenable (si \hat{h} s'annule par exemple, voir p. 32), on ajoute un terme de régularisation quadratique de la forme $\|q * u\|_{L^2}^2$. Le noyau q sera fixé ultérieurement en fonction du rapport signal sur bruit.

En appliquant la transformation de Fourier le problème de minimisation s'écrit alors

$$\min_{U \in L^2(\mathbb{R})} \|F - HU\|_2^2 + \|UQ\|_2^2,$$

où $U = \hat{u}$, $F = \hat{f}$, $Q = \hat{q}$ et $H = \hat{h}$. La solution U de ce problème est obtenue par dérivation :

$$\forall V \in L^2(\mathbb{R}) \quad (HU - F, HV)_{L^2} + (QU, QV)_{L^2} = 0.$$

On obtient

$$\bar{H}(HU - F) + |Q|^2 U = 0,$$

c'est-à-dire

$$(9) \quad U = \frac{\bar{H}}{|H|^2 + |Q|^2} F.$$

Le principe du filtrage de Wiener est de fixer $|Q|^2$ en fonction d'une estimation du rapport signal sur bruit. Lorsque $Q = 0$ on retrouve le filtrage inverse (pas de bruit). Idéalement, il faudrait choisir

$$Q(\omega) = \frac{|\hat{b}(\omega)|}{|\hat{u}(\omega)|}.$$

En effet dans ce cas le terme de régularisation $\|q * u\|_{L^2}^2$ pénalise exactement le bruit :

$$\|q * u\|_{L^2} = \|Q\hat{u}\|_{L^2} = \|\hat{b}\|_{L^2} = \|b\|_{L^2} .$$

Le choix le plus courant est de prendre l'inverse du rapport signal sur bruit :

$$Q^2 = \frac{\langle |\hat{b}|^2 \rangle}{\langle |\hat{u}|^2 \rangle},$$

où $\langle |\hat{b}|^2 \rangle$ (resp. $\langle |\hat{u}|^2 \rangle$) est la puissance spectrale moyenne de b (resp. u) c'est-à-dire

$$\langle |\hat{b}|^2 \rangle = \frac{1}{|D|} \iint_D |\hat{b}|^2(\omega) d\omega ,$$

où D est un domaine de \mathbb{R}^2 suffisamment grand et $|D|$ son aire.

Pour implémenter le filtre de Wiener, nous devons être en mesure d'estimer correctement la puissance spectrale de l'image d'origine et du bruit. Pour un bruit blanc gaussien, la puissance spectrale moyenne est égale à la variance σ^2 du bruit. La puissance spectrale de u est difficile à obtenir puisqu'on ne connaît pas u !! Toutefois

$$f = h * u + b \implies \hat{f} = \hat{h}\hat{u} + \hat{b} \implies |\hat{f}|^2 \simeq |\hat{h}|^2|\hat{u}|^2 + |\hat{b}|^2,$$

puisque le bruit b et l'image sont indépendants. Donc, pour un bruit gaussien

$$|\hat{u}|^2 \simeq \frac{|\hat{f}|^2 - |\hat{b}|^2}{|\hat{h}|^2} \simeq \frac{|\hat{f}|^2 - \sigma^2}{|\hat{h}|^2}.$$

En résumé, la fonction de transfert du filtre de Wiener est donnée par

$$(10) \quad W = \frac{\overline{\hat{h}}}{|\hat{h}|^2 + Q^2},$$

avec dans le cas où b est un bruit blanc gaussien d'écart-type σ

$$(11) \quad Q^2 = \frac{\sigma^2 \langle |\hat{h}| \rangle^2}{\langle |\hat{f}| \rangle^2 - \sigma^2}.$$

Remarque 5.1. — Rappelons que le spectre de puissance d'un signal est obtenu par la transformée de Fourier de sa fonction d'autocorrélation (Théorème de Wiener-Khinchine).

On peut bien sûr tester « à l'aveugle » différentes valeurs de Q que l'on peut supposer constant.

Dans l'exemple suivant, l'image a été floutée par un filtre gaussien de taille 15 et d'écart-type 45. Elle a aussi été réfléchiée pour minimiser les effets de bord.



FIGURE 35. Déconvolution par filtre de Wiener d'une image floutée et un masque gaussien de taille 11 et d'écart-type 30 ($Q = 0.15$)

Dans l'exemple suivant, l'image a été floutée par un filtre gaussien de taille 15 et d'écart-type 45 et bruitée par un bruit additif gaussien d'écart type 15.

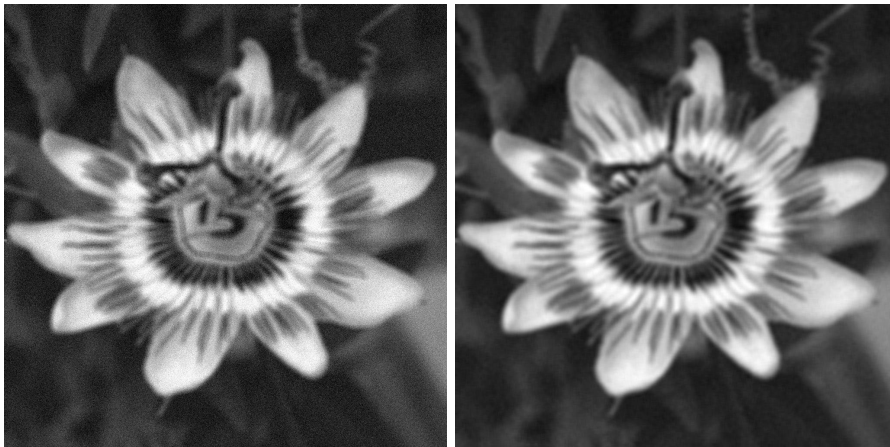


FIGURE 36. Déconvolution par filtre de Wiener d'une image floutée et bruitée, même masque et $Q = 1$

On peut dériver d'autres filtres sur le modèle ci-dessus. Citons

- le **filtre de Wiener paramétrique** $Q = \gamma Q_o$ où Q_o est donné par la relation (11) et $0 < \gamma < 1$.

Si $\gamma = 0$ on retrouve le filtre inverse et $\gamma = 1$ le filtre de Wiener.

– le filtre de Cannon :

$$W = \left[\frac{1}{|\hat{h}|^2 + Q^2} \right]^{\frac{1}{2}},$$

où Q est donné par (11).

– le filtre de moyenne géométrique :

$$W = \frac{1}{|\hat{h}|^s} \left[\frac{\bar{h}}{|\hat{h}|^2 + Q^2} \right]^{1-s},$$

où Q est donné par (11). Si $s = 1$ on retrouve le filtre inverse, $s = 0$ le filtre de Wiener et $s = 0.5$ le filtre de Cannon.

5.4. Déconvolution de Richardson-Lucy. — L'algorithme de Richardson-Lucy est un algorithme itératif spatial basé sur une approche bayésienne. Comme précédemment on considère une image dégradée par du flou (noyau h connu ou estimé) et du bruit b : $f = h * u + b$. On veut identifier u . On suppose que le masque h est connu et le bruit est inconnu. L'itération courante est

$$u_{k+1} = u_k \left[\frac{f}{(u_k * h)} * \check{h} \right],$$

avec (par exemple) $u_o = f$ et $\check{h}(x, y) = h(-x, -y)$.

Lorsque la réponse impulsionnelle du flou h n'est pas connue on fait une déconvolution **aveugle** (« blind deconvolution ») avec l'algorithme suivant :

Algorithme de Richardson-Lucy aveugle

(1) Initialisation : choix de u_o et h_o

(2) (a) Estimation de h

$$h_{k+1} = \frac{h_k}{\sum_{i,j} u_k(i, j)} \left[\frac{f}{u_k * h_k} * \check{u}_k \right]$$

(b) Estimation de u

$$u_{k+1} = u_k \left[\frac{f}{(u_k * h_{k+1})} * \check{h}_{k+1} \right],$$

(3) $u_{k+1} = \max(u_k, 0)$.



(a) Image dégradée comme dans la figure 36

(b) Itération 3



(c) Itération 5

(d) Itération 10

FIGURE 37. Déconvolution par l'algorithme de Richardson-Lucy aveugle - Initialisation avec un masque gaussien de taille 5 et d'écart-type $\sigma = 10$

6. Filtrage variationnel

La définition du filtre de Wiener suggère fortement l'utilisation de méthodes variationnelles puisqu'il s'agit de minimiser des erreurs tout en se donnant un a priori sur l'image. Nous allons préciser cette idée. Nous nous plaçons maintenant dans un cadre « dimension infinie », et nous effectuerons une discrétisation ensuite.

Etant donnée une image originale, on suppose qu'elle a été dégradée par un bruit additif v , et éventuellement par un opérateur R de flou. Un tel opérateur est souvent

modélisé par un produit de convolution. A partir de l'image observée $u_d = Ru + v$ (qui est donc une version dégradée de l'image originale u), on cherche à reconstruire u . Si on suppose que le bruit additif v est gaussien, la méthode du Maximum de Vraisemblance nous conduit à chercher u comme solution du problème de minimisation

$$\inf_u \|u_d - Ru\|_2^2,$$

où $\|\cdot\|_2$ désigne la norme dans L^2 . Il s'agit d'un problème inverse **mal posé** : l'opérateur n'est pas nécessairement inversible (et même lorsqu'il est inversible, son inverse est souvent numériquement difficile à calculer). En d'autres termes, l'existence et/ou l'unicité de solutions n'est pas assurée et si c'est le cas, la solution n'est pas stable (continue par rapport aux données). Pour le résoudre numériquement, on est amené à introduire un terme de régularisation, et à considérer le problème

$$\inf_u \underbrace{\|u_d - Ru\|_2^2}_{\text{ajustement aux données}} + \underbrace{L(u)}_{\text{Régularisation}}.$$

Dans toute la suite, nous ne considérerons que le cas où est R est l'opérateur identité ($Ru = u$). Commençons par un procédé de régularisation classique : celui de Tychonov.

6.1. Régularisation de Tychonov. — C'est un procédé de régularisation très classique mais trop sommaire dans le cadre du traitement d'image. Nous le présentons sur un exemple.

Soit $V = H^1(\Omega)$ et $H = L^2(\Omega)$: on considère le problème de minimisation originel (ajustement aux données) :

$$(\mathcal{P}) \quad \min_{u \in V} \|u - u_d\|_H^2,$$

où u_d est l'image observée (données) et le problème régularisé suivant : pour tout $\alpha > 0$

$$(\mathcal{P}_\alpha) \quad \min_{u \in V} \|u - u_d\|_H^2 + \alpha \|\nabla u\|_H^2.$$

Non seulement on veut ajuster u à la donnée u_d , mais on impose également que le gradient soit « assez petit » (cela dépend du paramètre α).

Il est facile de voir sur l'exemple suivant que la fonctionnelle $J(u) = \|u - u_d\|_H^2$ n'est pas coercive sur V :

$$\Omega =]0, 1[, \quad u_n(x) = x^n, \quad u_d = 0.$$

On voit que $\|u_n\|_2 = \frac{1}{\sqrt{2n}}$, $\|u'_n\|_2 = \frac{n}{\sqrt{2n-1}}$. On a donc

$$\lim_{n \rightarrow +\infty} \|u_n\|_V = +\infty \quad \text{et} \quad \lim_{n \rightarrow +\infty} J(u_n) = 0.$$

Il n'est donc même pas clair (a priori) que le problème (\mathcal{P}) ait une solution.

Proposition 6.1. — Supposons que (\mathcal{P}) admet au moins une solution \tilde{u} . Le problème (\mathcal{P}_α) admet une solution unique u_α . De la famille (u_α) on peut extraire une sous-suite qui converge (faiblement) dans V vers une solution u^* de (\mathcal{P}) lorsque $\alpha \rightarrow 0$.

Démonstration. — Le problème (\mathcal{P}_α) admet une solution unique u_α car la fonctionnelle

$$J_\alpha = \|u - u_d\|_H^2 + \alpha \|\nabla u\|_H^2,$$

est coercive et strictement convexe (c'est en gros la norme de V à une partie affine près). Montrons maintenant que la famille (u_α) est bornée dans V uniformément par rapport à α .

$$\forall u \in V \quad J_\alpha(u_\alpha) \leq J_\alpha(u).$$

En particulier pour $u = \tilde{u}$ solution de (\mathcal{P}) :

$$(12) \quad \underbrace{J(\tilde{u}) \leq J(u_\alpha)}_{\tilde{u} \text{ solution de } (\mathcal{P})} \leq \underbrace{J_\alpha(u_\alpha) = J(u_\alpha) + \alpha \|\nabla u_\alpha\|_H^2}_{u_\alpha \text{ solution de } (\mathcal{P}_\alpha)} \leq J_\alpha(\tilde{u}) = J(\tilde{u}) + \alpha \|\nabla \tilde{u}\|_H^2,$$

où on a posé $J(u) := \|u - u_d\|_H^2$.

Par conséquent, pour $\alpha \leq \alpha_o$, $J_\alpha(u_\alpha)$ est borné indépendamment de α . Ceci entraîne que la famille $(u_\alpha)_{\alpha \leq \alpha_o}$ est bornée dans H . D'autre part, avec (12) on a aussi

$$\alpha \|\nabla u_\alpha\|_H^2 \leq J(\tilde{u}) + \alpha \|\nabla \tilde{u}\|_H^2 - J(u_\alpha) \leq J(\tilde{u}) + \alpha \|\nabla \tilde{u}\|_H^2 - J(\tilde{u}) = \alpha \|\nabla \tilde{u}\|_H^2 ;$$

par conséquent la famille $(u_\alpha)_{\alpha \leq \alpha_o}$ est bornée dans V . On peut donc en extraire une sous-suite qui converge (faiblement) dans V vers u^* . D'autre part l'équation (12) montre que

$$\lim_{\alpha \rightarrow 0} J_\alpha(u_\alpha) = J(\tilde{u}) = \inf(\mathcal{P}).$$

Par semi-continuité inférieure de J il vient

$$J(u^*) \leq \liminf_{\alpha \rightarrow 0} J(u_\alpha) = \liminf_{\alpha \rightarrow 0} J_\alpha(u_\alpha) \leq \inf(\mathcal{P}).$$

Par conséquent u^* est une solution de (\mathcal{P}) . □

Cherchons maintenant le moyen de calculer u_α . Comme la fonctionnelle est strictement convexe, une condition nécessaire et suffisante d'optimalité est

$$J'_\alpha(u_\alpha) = 0.$$

Un calcul assez standard montre que

$$\begin{aligned} \forall u \in V \quad J'_\alpha(u_\alpha) \cdot u &= \int_{\Omega} (u_\alpha - u_d)(x)u(x)dx + \int_{\Omega} \nabla u_\alpha(x) \nabla u(x)dx \\ &= \int_{\Omega} (u_\alpha - u_d - \Delta u_\alpha)(x)u(x)dx. \end{aligned}$$

Par conséquent l'équation d'Euler qui fournit la solution u_α est la suivante :

$$u_\alpha - u_d - \Delta u_\alpha = 0, \quad u_\alpha \in H_o^1(\Omega).$$

On peut se contenter d'approcher la solution u_α en écrivant la formulation dynamique :

$$\frac{\partial u}{\partial t} - \Delta u + u = u_d.$$

Remarque 6.2. — L'approche dynamique revient ici à calculer une suite minimisante par une méthode de descente de gradient : en effet, l'algorithme du gradient à pas constant donne

$$\frac{u_{t+\delta t} - u_t}{\delta t} = -J'_\alpha(u_t);$$

Par passage à la limite quand $\delta t \rightarrow 0$, on obtient

$$\frac{\partial u}{\partial t} = -J'_\alpha(u) = \Delta u - u + u_d.$$

Le terme de régularisation classique $L(u) := \|\nabla u\|_2^2$ (régularisation de Tychonov) n'est pas adapté au problème de restauration d'images : l'image restaurée u est alors beaucoup trop lissée car le Laplacien est un opérateur de diffusion isotrope. En particulier, les bords sont érodés. Une approche beaucoup plus efficace consiste à considérer la variation totale, c'est à dire à prendre $L(u) = \int |Du|$. Nous allons préciser la définition dans la section suivante.

Cela conduit à une minimisation de fonctionnelle dans un espace de Banach particulier, mais bien adapté au problème : l'espace des fonctions à variation bornée.

6.2. Le modèle continu de Rudin-Osher-Fatemi. —

6.2.1. L'espace des fonctions à variation bornée. — Dans tout ce qui suit Ω un ouvert borné de \mathbb{R}^2 de frontière Lipschitz et $\mathcal{C}_c^1(\Omega, \mathbb{R}^2)$ est l'espace des fonctions \mathcal{C}^1 à support compact dans Ω à valeurs dans \mathbb{R}^2 .

Définition 6.3. — Une fonction f de $L^1(\Omega)$ (à valeurs dans \mathbb{R}) est à **variation bornée** dans Ω si $J(f) < +\infty$ où

$$(13) \quad J(f) = \sup \left\{ \int_{\Omega} f(x) \operatorname{div} \varphi(x) \, dx \mid \varphi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2), \|\varphi\|_{\infty} \leq 1 \right\}.$$

On note

$$BV(\Omega) = \{f \in L^1(\Omega) \mid J(f) < +\infty\}$$

l'espace de telles fonctions.

On rappelle que si $\varphi = (\varphi_1, \varphi_2) \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2)$ alors

$$\operatorname{div} \varphi(x) = \frac{\partial \varphi_1}{\partial x_1}(x) + \frac{\partial \varphi_2}{\partial x_2}(x).$$

Définition 6.4 (Périmètre). — Un ensemble E mesurable (pour la mesure de Lebesgue) dans \mathbb{R}^2 est de périmètre (ou de **longueur**) fini si sa fonction indicatrice χ_E est dans $BV(\Omega)$.

On rappelle qu'une mesure de Radon est une mesure finie sur tout compact et que grâce au théorème de Riesz, toute forme \mathcal{L} linéaire continue sur $\mathcal{C}_c^0(\Omega)$ (fonctions continues à support compact) est de la forme

$$\mathcal{L}(f) = \int_{\Omega} f(x) d\mu ,$$

où μ est une (unique) mesure de Radon associée à \mathcal{L} . Nous pouvons donc donner une propriété structurelle des fonctions BV.

Théorème 6.5. — Soit $f \in BV(\Omega)$. Alors il existe une mesure de Radon positive μ sur Ω et une fonction μ -mesurable $\sigma : \Omega \rightarrow \mathbb{R}$ telle que

(i) $|\sigma(x)| = 1$, μ p.p. , et

(ii) $\int_{\Omega} f(x) \operatorname{div} \varphi(x) dx = - \int_{\Omega} \varphi(x) \sigma(x) d\mu$ pour toute fonction $\varphi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2)$

La relation (ii) est une formule d'intégration par parties « faible ». Ce théorème indique que la dérivée faible (au sens des distributions) d'une fonction BV est une mesure de Radon. $J(u) := \mu(\Omega) \geq 0$ est la **variation totale** de f . L'application

$$\begin{aligned} BV(\Omega) &\rightarrow \mathbb{R}^+ \\ u &\mapsto \|u\|_{BV(\Omega)} = \|u\|_{L^1} + J(u) . \end{aligned}$$

est une norme. L'espace $BV(\Omega)$ muni de cette norme est un espace de Banach. Donnons sans démonstration quelques propriétés importantes de $BV(\Omega)$. Pour plus de détail, on pourra se reporter à [10].

Théorème 6.6 (Semi-continuité inférieure de la variation totale)

L'application $u \mapsto J(u)$ de $BV(\Omega)$ dans \mathbb{R}^+ est semi-continue inférieurement (sci) pour la topologie séquentielle de $L^1(\Omega)$.

Plus précisément, si (u_k) est une suite de fonctions de $BV(\Omega)$ qui converge vers u fortement dans $L^1(\Omega)$ alors

$$J(u) \leq \liminf_{k \rightarrow +\infty} J(u_k) .$$

Théorème 6.7 (Approximation régulière). — Pour toute fonction $u \in BV(\Omega)$, il existe une suite de fonctions $(u_k)_{k \in \mathbb{N}}$ de $BV(\Omega) \cap \mathcal{C}^\infty(\Omega)$ telle que

(i) $u_k \rightarrow u$ dans $L^1(\Omega)$ et

(ii) $J(u_k) \rightarrow J(u)$ (dans \mathbb{R}).

Terminons par un résultat important de compacité :

Théorème 6.8 (Compacité, [10] p.176.). — L'espace $BV(\Omega)$ s'injecte dans $L^1(\Omega)$ de manière compacte. Plus précisément : si $(u_n)_{n \in \mathbb{N}}$ une suite bornée de $BV(\Omega)$

$$\sup_{n \in \mathbb{N}} \|u_n\|_{BV(\Omega)} < +\infty ,$$

alors, il existe une sous-suite $(u_{n_k})_{k \in \mathbb{N}}$ et une fonction $u \in BV(\Omega)$ telle que u_{n_k} converge fortement vers u dans $L^1(\Omega)$.

Plus généralement on a des injections de type Sobolev :

Théorème 6.9 (Injection dans les espaces L^p). — On suppose que Ω est régulier de \mathbb{R}^N . Alors l'espace $BV(\Omega)$ s'injecte dans $L^p(\Omega)$ de manière

- continue pour $1 \leq p \leq \frac{N}{N-1}$
- compacte pour $1 \leq p < \frac{N}{N-1}$

En particulier pour $N = 2$ l'espace $BV(\Omega)$ est contenu dans $L^2(\Omega)$ avec injection continue.

6.2.2. Présentation du modèle de Rudin-Osher et Fatemi (ROF). — Il faut trouver une alternative à la régularisation de Tychonov qui est trop violente. La première idée consiste à remplacer le terme de régularisation $\|\nabla u\|_H^2$ (qui est en réalité une pénalisation par une norme V), par une norme moins contraignante et régularisante. Rudin-Osher et Fatemi [15] ont proposé un modèle où l'image est décomposée en deux parties : $u_d = u + v$ où v est le bruit et u la partie « régulière ». On va donc a priori chercher la solution du problème sous la forme $u + v$ avec $u \in BV(\Omega)$ et $v \in L^2(\Omega)$. Cela conduit à :

$$(\mathcal{P}_{ROF}) \quad \min \left\{ J(u) + \frac{1}{2\varepsilon} \|v\|_2^2 \mid u \in BV(\Omega), v \in L^2(\Omega), u + v = u_d \right\}.$$

Ici le terme de régularisation $L(u)$ est $J(u)$ la variation totale de u et $\varepsilon > 0$ (ici $\alpha = 2\varepsilon$, si on se réfère aux notations du début de la section).

Théorème 6.10. — Le problème (\mathcal{P}_{ROF}) admet une solution unique.

Démonstration. — Soit $(u_n, v_n) \in BV(\Omega) \times L^2(\Omega)$ une suite minimisante. Comme v_n est bornée dans $L^2(\mathbb{R})$ on peut en extraire une sous-suite (notée de la même façon) faiblement convergente vers v^* dans $L^2(\mathbb{R})$. Comme la norme de $L^2(\mathbb{R})$ est convexe, sci, il vient

$$\|v^*\|_2^2 \leq \liminf_{n \rightarrow +\infty} \|v_n\|_2^2.$$

De la même façon $u_n = u_d - v_n$ est bornée dans $L^2(\mathbb{R})$ et donc dans $L^1(\Omega)$ puisque Ω est borné. Comme $J(u_n)$ est borné, il s'ensuit que u_n est bornée dans $BV(\Omega)$. Grâce à la compacité de l'injection de $BV(\Omega)$ dans $L^1(\Omega)$ (voir [2, 13]), cela entraîne que u_n converge (à une sous-suite près) **fortement** dans $L^1(\Omega)$ vers $u^* \in BV(\Omega)$.

D'autre part J est semi-continue inférieurement (sci), donc

$$J(u^*) \leq \liminf_{n \rightarrow +\infty} J(u_n),$$

et finalement

$$J(u^*) + \frac{1}{2\varepsilon} \|v^*\|_2^2 \leq \liminf_{n \rightarrow +\infty} J(u_n) + \frac{1}{2\varepsilon} \|v_n\|_2^2 = \inf(\mathcal{P}_{ROF}).$$

Comme $u_n + v_n = u_d$ pour tout n , on a $u^* + v^* = u_d$. Par conséquent u^* est une solution du problème (\mathcal{P}_{ROF}) .

La fonctionnelle est strictement convexe par rapport au couple (u, v) et la contrainte est affine. On a donc unicité. \square

Nous aurons besoin d'établir des conditions d'optimalité pour la ou les solutions optimales des modèles proposés. Toutefois les fonctionnelles considérées (en particulier J) ne sont en général pas Gâteaux-différentiables et nous devons utiliser des notions d'analyse non lisse (voir annexe).

6.2.3. Condition d'optimalité du premier ordre. — Le problème (\mathcal{P}_{ROF}) peut s'écrire de la manière (équivalente) suivante

$$(14) \quad \min_{u \in BV(\Omega)} \mathcal{F}(u) := J(u) + \frac{1}{2\varepsilon} \|u - u_d\|_2^2.$$

La fonctionnelle \mathcal{F} est convexe et \bar{u} est solution de (\mathcal{P}_{ROF}) si et seulement si $0 \in \partial\mathcal{F}(\bar{u})$. La relation : $0 \in \partial\mathcal{F}(\bar{u}) \iff \bar{u} \in \partial\mathcal{F}^*(0)$, est vraie même si l'espace n'est pas réflexif (voir annexe et [2] Théorème 9.5.1 p 333). Ici \mathcal{F}^* désigne la conjuguée de Fenchel de \mathcal{F} (la définition et les propriétés de cette conjuguée sont données en annexe, section A.2.6).

On peut utiliser le théorème A.25 pour calculer $\partial\mathcal{F}(u)$. L'application $u \mapsto \|u - u_d\|_2^2$ est continue sur $L^2(\Omega)$ et J est finie sur $BV(\Omega)$ à valeurs dans $\mathbb{R} \cup \{+\infty\}$. D'autre part $u \mapsto \|u - u_d\|_2^2$ est Gâteaux-différentiable sur $L^2(\Omega)$. Comme nous allons utiliser la dualité (convexe) nous devons dans un premier temps calculer la conjuguée de Fenchel de J .

Théorème 6.11. — *La transformée de Fenchel J^* de la fonctionnelle J « variation totale » est l'indicatrice de l'adhérence \bar{K} de l'ensemble K , où*

$$K := \{ \xi = \operatorname{div} \varphi \mid \varphi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2), \|\varphi\|_\infty \leq 1 \}.$$

Démonstration. — Comme J est positivement homogène la conjuguée J^* de J est l'indicatrice d'un ensemble convexe fermé \tilde{K} (proposition A.34).

Montrons d'abord que $K \subset \tilde{K}$: soit $u \in K$. Par définition de J

$$J(u) = \sup_{\xi \in K} \langle \xi, u \rangle,$$

où $\langle \cdot, \cdot \rangle$ désigne le produit de dualité entre $BV(\Omega)$ et son dual. Par conséquent $\langle \xi, u \rangle - J(u) \leq 0$ pour tous $\xi \in K$ et $u \in BV(\Omega)$. On déduit donc que pour tout $u^* \in K$

$$J^*(u^*) = \sup_{u \in K} \langle u^*, u \rangle - J(u) \leq 0.$$

Comme J^* ne prend qu'une seule valeur finie on a $J^*(u^*) = 0$, et donc $u^* \in \tilde{K}$. Par conséquent $K \subset \tilde{K}$ et comme \tilde{K} est fermé :

$$\bar{K} \subset \tilde{K}.$$

En particulier

$$J(u) = \sup_{\xi \in \tilde{K}} \langle u, \xi \rangle \leq \sup_{\xi \in \bar{K}} \langle u, \xi \rangle \leq \sup_{\xi \in \tilde{K}} \langle u, \xi \rangle = \sup_{\xi \in \bar{K}} \langle u, \xi \rangle - J^*(\xi) = J^{**}(u).$$

Comme $J^{**} = J$, il vient

$$\sup_{\xi \in K} \langle u, \xi \rangle \leq \sup_{\xi \in \bar{K}} \langle u, \xi \rangle \leq \sup_{\xi \in K} \langle u, \xi \rangle,$$

et donc

$$(15) \quad \sup_{\xi \in K} \langle u, \xi \rangle = \sup_{\xi \in \bar{K}} \langle u, \xi \rangle = \sup_{\xi \in \tilde{K}} \langle u, \xi \rangle.$$

Supposons maintenant qu'il existe $u^* \in \tilde{K}$ tel que $u^* \notin \bar{K}$. On peut alors séparer strictement u^* et le convexe fermé \bar{K} . Il existe $\alpha \in \mathbb{R}$ et u_o tels que

$$\langle u_o, u^* \rangle > \alpha \geq \sup_{v \in \bar{K}} \langle u_o, v \rangle.$$

D'après (15) il vient

$$\sup_{\xi \in \tilde{K}} \langle u_o, \xi \rangle \geq \langle u_o, u^* \rangle > \alpha \geq \sup_{v \in \bar{K}} \langle u_o, v \rangle = \sup_{v \in \tilde{K}} \langle u_o, v \rangle.$$

On a donc une contradiction : $\tilde{K} = \bar{K}$. □

D'après la définition du sous-différentiel (voir Annexe, Définition A.22, p.57), u est une solution de (\mathcal{P}_{ROF}) si et seulement si

$$0 \in \partial \left(J(u) + \frac{1}{2\varepsilon} \|u - u_d\|_2^2 \right) = \frac{u - u_d}{\varepsilon} + \partial J(u).$$

Comme J est convexe, sci, propre on peut appliquer le corollaire A.41. Donc

$$\frac{u_d - u}{\varepsilon} \in \partial J(u) \iff u \in \partial J^* \left(\frac{u_d - u}{\varepsilon} \right) \iff 0 \in -u + \partial J^* \left(\frac{u_d - u}{\varepsilon} \right).$$

Nous aimerions alors appliquer (par exemple) la proposition A.28 mais elle fait intervenir la notion de projection qui n'est définie que dans le cadre hilbertien. Or J^* est la conjuguée de J pour la dualité $\langle BV, BV' \rangle$. Nous allons donc conclure et calculer $\partial J^*(u)$ une fois le problème discrétisé : le cadre fonctionnel est alors hilbertien, puisque que l'espace de travail est de dimension finie.

6.3. Modèle discret. — On va maintenant considérer des images discrètes (ce qui est le cas en pratique). Une image discrète est une matrice $N \times M$ que nous identifierons à un vecteur de taille NM (par exemple en la rangeant ligne par ligne). On note X l'espace euclidien $\mathbb{R}^{N \times M}$ et $Y = X \times X$. On munit X du produit scalaire usuel

$$(u, v)_X = \sum_{1 \leq i \leq N} \sum_{1 \leq j \leq M} u_{ij} v_{ij},$$

et de la norme associée : $\|\cdot\|_X$.

Nous allons donner une formulation discrète de ce qui a été fait auparavant et en particulier définir une variation totale discrète que nous noterons de la même façon. Pour cela nous introduisons une version discrète de l'opérateur gradient. Si $u \in X$, le gradient ∇u est un vecteur de Y donné par

$$(\nabla u)_{i,j} = ((\nabla u)_{i,j}^1, (\nabla u)_{i,j}^2),$$

avec

$$(16) \quad (\nabla u)_{i,j}^1 = \begin{cases} u_{i+1,j} - u_{i,j} & \text{si } i < N \\ 0 & \text{si } i = N \end{cases}$$

$$(17) \quad (\nabla u)_{i,j}^2 = \begin{cases} u_{i,j+1} - u_{i,j} & \text{si } j < M \\ 0 & \text{si } j = M \end{cases}$$

La variation totale discrète est alors donnée par la norme ℓ^1 du gradient discret :

$$(18) \quad J(u) = \sum_{1 \leq i \leq N} \sum_{1 \leq j \leq M} |(\nabla u)_{i,j}|,$$

où $|(\nabla u)_{i,j}| = \sqrt{|(\nabla u)_{i,j}^1|^2 + |(\nabla u)_{i,j}^2|^2}$. On introduit également une version discrète de l'opérateur de divergence défini par analogie avec le cadre continu en posant

$$\operatorname{div} = -\nabla^*,$$

où ∇^* est l'opérateur adjoint de ∇ , c'est-à-dire

$$\forall p \in Y, \forall u \in X \quad (-\operatorname{div} p, u)_X = (p, \nabla u)_Y = (p^1, \nabla^1 u)_X + (p^2, \nabla^2 u)_X.$$

On peut alors vérifier que la divergence discrète est donnée par la relation (6) p. 30.

On utilisera aussi une version discrète du laplacien définie par

$$\Delta u = \operatorname{div} \nabla u.$$

On va remplacer le problème (\mathcal{P}_{ROF}) (dans la version (14)) par le problème obtenu après discrétisation suivant

$$(19) \quad \min_{u \in X} J(u) + \frac{1}{2\varepsilon} \|u - u_d\|_X^2.$$

Il est facile de voir que ce problème a une solution unique que nous allons caractériser. On rappelle que $|g_{i,j}| = \sqrt{(g_{i,j}^1)^2 + (g_{i,j}^2)^2}$ et que la version discrète de la variation totale est donnée (de manière analogue au cas continu) par

$$J(u) = \sup_{\xi \in K} \langle u, \xi \rangle.$$

où

$$(20) \quad K := \{ \xi = \operatorname{div} (g) \mid g \in Y, |g_{i,j}| \leq 1, \forall (i,j) \in \{1, \dots, N\} \times \{1, \dots, M\} \},$$

est la version « discrète » de $K := \{ \xi = \operatorname{div} \varphi \mid \varphi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2), \|\varphi\|_\infty \leq 1 \}$. Nous pouvons, comme dans le cas de la dimension infinie donner la caractérisation de la conjuguée de Fenchel de J (la démonstration est analogue).

Théorème 6.12. — La transformée de Fenchel J^* de la fonctionnelle J « variation totale » approchée définie sur X , est l'indicatrice de l'ensemble \bar{K} , où K est donné par (20).

Le résultat suivant donne la caractérisation attendue de la solution [7] :

Théorème 6.13. — La solution de (19) est donnée par

$$(21) \quad u = u_d - P_{\varepsilon K}(u_d),$$

où P_K est le projecteur orthogonal sur K .

Démonstration. — Nous avons vu que u est solution de (19) si et seulement si

$$u \in J^* \left(\frac{u_d - u}{\varepsilon} \right).$$

Comme J^* est l'indicatrice de K , en utilisant la proposition A.28 on en déduit que pour tout $c > 0$

$$u = c \left[\frac{u_d - u}{\varepsilon} + \frac{u}{c} - P_K \left(\frac{u_d - u}{\varepsilon} + \frac{u}{c} \right) \right].$$

Le choix de $c = \varepsilon$ conduit au résultat car $\varepsilon P_K \left(\frac{u_d}{\varepsilon} \right) = P_{\varepsilon K}(u_d)$. □

Tout revient maintenant à calculer

$$P_{\varepsilon K}(u_d) = \operatorname{argmin} \{ \|\varepsilon \operatorname{div}(p) - u_d\|_X^2 \mid |p_{i,j}| \leq 1, i, j = 1, \dots, N \}.$$

On peut le résoudre par une méthode de point fixe : $p^o = 0$.

$$(22) \quad p_{i,j}^{n+1} = \frac{p_{i,j}^n + \rho (\nabla[\operatorname{div} p^n - u_d/\varepsilon])_{i,j}}{1 + \rho |(\nabla[\operatorname{div} p^n - u_d/\varepsilon])_{i,j}|}.$$

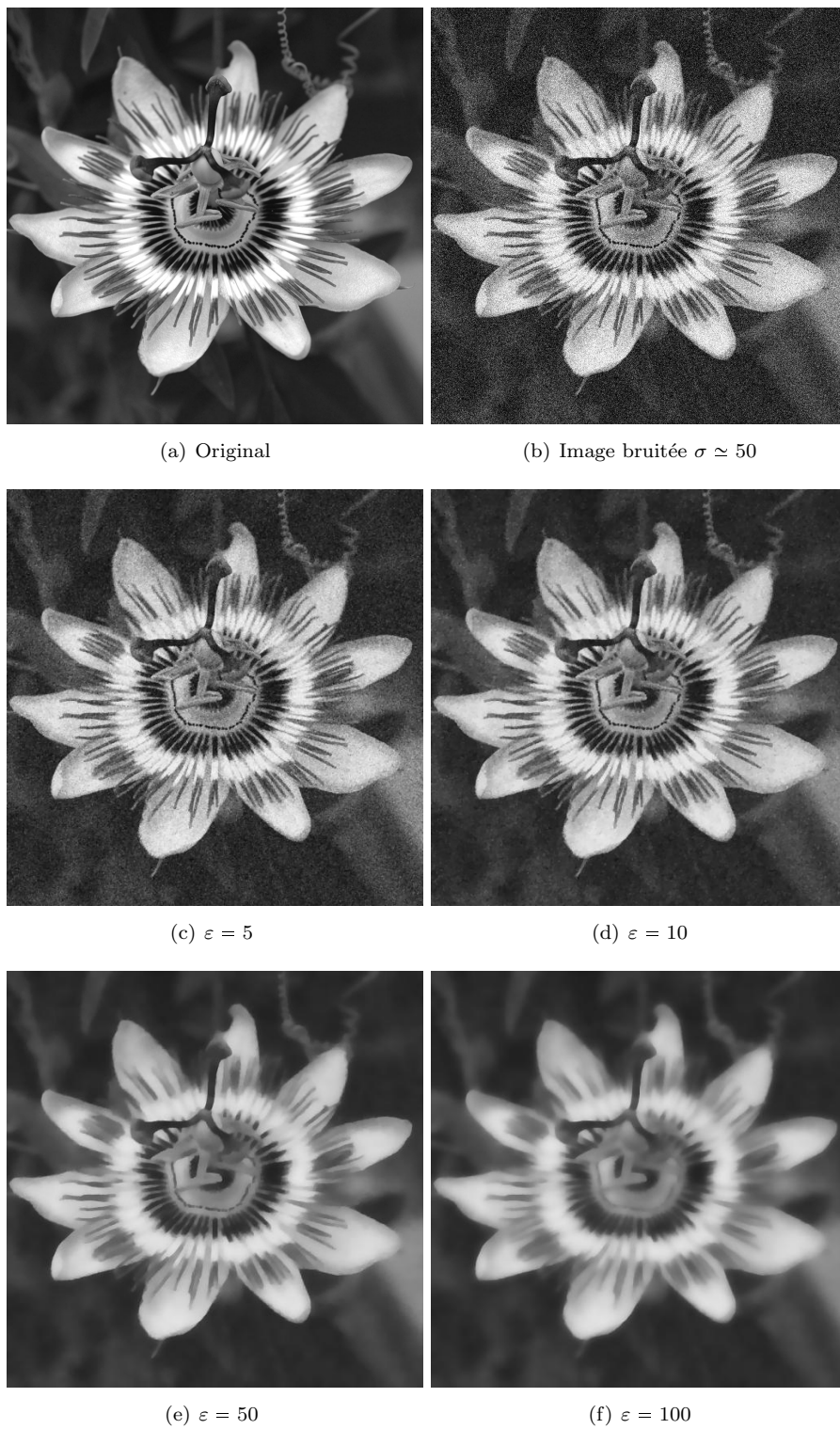
Théorème 6.14 (Algorithme de projection de Chambolle [7])

Si le paramètre ρ dans (22) vérifie $\rho \leq 1/8$, alors $\varepsilon \operatorname{div} p^n \rightarrow P_{\varepsilon K}(u_d)$.

La solution du problème est donc donnée par

$$u = u_d - \varepsilon \operatorname{div} p^\infty \quad \text{où} \quad p^\infty = \lim_{n \rightarrow +\infty} p^n.$$

En pratique le paramètre ε permet de « régler » le niveau de régularisation. Plus ε est grand, plus la variation totale de la solution, c'est-à-dire la longueur des contours de l'image est petite. Le bruit est éliminé au prix d'un lissage de l'image.

FIGURE 38. Filtrage ROF : sensibilité au paramètre ε (50 itérations)

6.4. Déconvolution. — Nous évoquons très succinctement dans cette section une généralisation possible de la méthode précédente dans la cas où l'image est à la fois floutée et bruitée. L'image observée est donc de la forme : $R(f) + b$ avec $R(f) = h * f$. Nous avons déjà traité dans la section 4 la question de la déconvolution. Nous généralisons donc ici le principe du filtre de Wiener.

Le problème (\mathcal{P}_{ROF}) se généralise donc de la manière suivante

$$(23) \quad \min_{u \in BV(\Omega)} \mathcal{F}(u) := J(u) + \frac{1}{2\varepsilon} \|Ru - u_d\|_2^2.$$

où R est un opérateur de convolution à noyau symétrique et positif. La fonctionnelle \mathcal{F} est convexe et \bar{u} est solution de (\mathcal{P}_{ROF}) si et seulement si $0 \in \partial\mathcal{F}(\bar{u})$. Nous avons un résultat analogue à celui de la section précédente. Nous nous plaçons dans le cas **discret** :

Théorème 6.15. — La solution u_ε de (23) est caractérisée par l'existence de μ_ε vérifiant

$$(24) \quad \mu_\varepsilon = \frac{R^*(u_d - Ru)}{\varepsilon},$$

$$(25) \quad u_\varepsilon = u_\varepsilon + \mu_\varepsilon - P_K(u_\varepsilon + \mu_\varepsilon),$$

où P_K est le projecteur orthogonal sur K donné par (20) et R^* est l'opérateur adjoint de R .

Démonstration. — Une condition nécessaire et suffisante pour que u_ε soit une solution de (23) est

$$0 \in \partial \left(J(u_\varepsilon) + \frac{1}{2\varepsilon} \|Ru_\varepsilon - u_d\|_X^2 \right) = \frac{R^*(Ru_\varepsilon - u_d)}{\varepsilon} + \partial J(u_\varepsilon).$$

Ceci est équivalent à

$$u_\varepsilon \in \partial J^*(\mu_\varepsilon),$$

où on posé

$$\mu_\varepsilon = \frac{R^*(u_d - Ru_\varepsilon)}{\varepsilon}.$$

Comme $J^* = 1_K$, la proposition A.28 indique que

$$u_\varepsilon \in \partial 1_K(\mu_\varepsilon) \iff \mu_\varepsilon = P_K(u_\varepsilon + \mu_\varepsilon),$$

d'où le résultat. □

On peut maintenant résoudre le système (24-25) par exemple par une méthode de point fixe, qui conduit à l'algorithme :

Algorithme

- (1) Initialisation : $u_o = u_d, n = 0$.
- (2) Etape n : u_n est connu

- Calcul de $\mu_n = \frac{R^*(u_d - Ru_n)}{\varepsilon}$
- Calcul de $u_{n+1} = u_n + \mu_n + P_K(u_n + \mu_n)$

Il est clair que si la suite u_n converge vers u_ε alors μ_n converge vers μ_ε et $(u_\varepsilon, \mu_\varepsilon)$ vérifie le système (24-25). C'est donc une solution du problème (discret).

7. Conclusion

Nous avons abordé très rapidement les méthodes génériques de base pour effectuer le filtrage des images. Nous nous sommes volontairement limités au cas d'un bruit additif. Il existe des méthodes de plus en plus sophistiquées pour traiter des bruits de nature très différente. Enfin, nous n'avons pas parlé des méthodes par ondelettes faute d'espace et de compétence qui constituent des outils puissants de débruitage. Le lecteur intéressé pourra se référer à [12]

Appendice A

Quelques outils mathématiques

A.1. Optimisation dans les espaces de Banach. — Sauf mention du contraire, on considère dans toute la section un espace de Banach réflexif V de dual (topologique) V' . On note $\|\cdot\|_V$ la norme de V et $\langle \cdot, \cdot \rangle$ le crochet de dualité entre V et V' .

A.1.1. Semi-continuité et convexité de fonctionnelles sur V . —

Définition A.1. — Une fonction J de V dans $\mathbb{R} \cup \{+\infty\}$ est semi-continue inférieurement (sci) sur V si elle satisfait aux conditions équivalentes :

- $\forall a \in \mathbb{R}, \quad \{ u \in V \mid J(u) \leq a \}$ est fermé
- $\forall \bar{u} \in V, \quad \liminf_{u \rightarrow \bar{u}} J(u) \geq J(\bar{u}).$

Théorème A.2. — Toute fonction **convexe sci** pour la topologie forte (celle de la norme) de V est encore sci pour la topologie faible de V .

En pratique ce résultat s'utilise sous la forme du corollaire suivant :

Corollaire A.3. — Soit J une fonctionnelle convexe de V dans $\mathbb{R} \cup \{+\infty\}$ sci (par exemple continue) pour la topologie forte. Si v_n est une suite de V **faiblement convergente** vers v alors

$$J(v) \leq \liminf_{n \rightarrow +\infty} J(v_n).$$

A.1.2. *Gâteaux-différentiabilité des fonctionnelles convexes.* —

Définition A.4. — Soit J une fonctionnelle de V dans $\mathbb{R} \cup \{+\infty\}$. On dit que J est **Gâteaux-différentiable** en $u \in \text{dom}(J)$ si la dérivée directionnelle

$$J'(u; v) = \lim_{t \rightarrow 0^+} \frac{J(u + tv) - J(u)}{t},$$

existe dans toute direction v de V et si l'application

$$v \mapsto J'(u; v)$$

est linéaire continue.

De manière générale on notera $\nabla J(u)$ la Gâteaux-différentielle de J en u . C'est un élément du dual V' .

Si V est un espace de Hilbert, avec le théorème de représentation de Riesz (voir [6] par exemple) on identifie V et son dual; on note alors

$$J'(u; v) = \langle \nabla J(u), v \rangle,$$

où $(,)$ désigne le produit scalaire de V . L'élément $\nabla J(u)$ de V est le **gradient** de J en u .

Il est clair que si J est différentiable au sens classique en u (on dit alors **Fréchet** - différentiable), alors J est Gâteaux-différentiable en u , et la dérivée classique et la dérivée au sens de Gâteaux coïncident.

La réciproque est fautive comme le montre le contre-exemple suivant : soit f de \mathbb{R}^2 dans \mathbb{R} définie par :

$$f(x, y) = \begin{cases} y & \text{si } x = y^2, \\ 0 & \text{sinon} \end{cases}$$

La fonction f est continue en $(0,0)$ et Gâteaux-différentiable en $(0,0)$ mais pas Fréchet - différentiable en $(0,0)$.

Théorème A.5. — Soit $J : \mathcal{C} \subset V \rightarrow \mathbb{R}$, Gâteaux différentiable sur \mathcal{C} , avec \mathcal{C} convexe. J est convexe si et seulement si

$$(26) \quad \forall (u, v) \in \mathcal{C} \times \mathcal{C} \quad J(v) \geq J(u) + \langle \nabla J(u), v - u \rangle$$

Théorème A.6. — Soit $J : \mathcal{C} \subset V \rightarrow \mathbb{R}$, Gâteaux différentiable sur \mathcal{C} , avec \mathcal{C} convexe. J est convexe si et seulement si ∇J est un opérateur **monotone**, c'est-à-dire

$$(27) \quad \forall (u, v) \in \mathcal{C} \times \mathcal{C} \quad \langle \nabla J(u) - \nabla J(v), u - v \rangle \geq 0.$$

Remarque A.7. — Supposons que ∇ soit un opérateur **strictement** monotone :

$$(28) \quad \forall (u, v) \in \mathcal{C} \times \mathcal{C}, \quad u \neq v, \quad \langle \nabla J(u) - \nabla J(v), u - v \rangle > 0.$$

alors J est **strictement** convexe.

De manière analogue, on définit la (Gâteaux) dérivée seconde de J en u , comme étant la dérivée de la fonction (vectorielle) $u \rightarrow \nabla J(u)$. On la note $D^2J(u)$ et on l'appellera Hessien par analogie avec le Hessien au sens de Fréchet ; ce Hessien est identifiable à une matrice carrée $n \times n$ lorsque $V = \mathbb{R}^n$.

A.1.3. Minimisation dans un Banach réflexif. — Commençons par un résultat général de minimisation d'une fonctionnelle semi-continue sur un ensemble fermé de V .

Définition A.8. — On dit que $J : V \rightarrow \mathbb{R}$ est **coercive** si

$$\lim_{\|x\|_V \rightarrow +\infty} J(x) = +\infty.$$

Théorème A.9. — *On suppose que V est un Banach réflexif. Soit J une fonctionnelle de V dans $\mathbb{R} \cup \{+\infty\}$, semi-continue inférieurement (sci) pour la topologie faible de V . Soit K un sous-ensemble non vide et faiblement fermé de V . On suppose que J est **propre** (c'est-à-dire qu'il existe un élément v_o de K tel que $J(v_o) < +\infty$). Alors le problème de minimisation suivant :*

$$(29) \quad (\mathcal{P}) \quad \left\{ \begin{array}{l} \text{Trouver } u \text{ tel que} \\ J(u) = \inf \{ J(v) \mid v \in K \}, \end{array} \right.$$

admet au moins une solution dans l'un des cas suivants :

- soit J est coercive i.e. $\lim_{\|v\|_V \rightarrow +\infty} J(v) = +\infty$,
- soit K est borné.

Démonstration. — On pose $d = \inf \{ J(v) \mid v \in K \}$; $d < +\infty$, sinon J serait identiquement égale à $+\infty$ sur K .

Soit une suite minimisante, c'est-à-dire une suite u_n de K telle que $J(u_n) \rightarrow d$. Montrons que cette suite est bornée dans V . Si K est borné, c'est clair. Sinon, on suppose que J est coercive. Si u_n n'est pas bornée, on peut en extraire une sous-suite encore notée u_n telle que $\lim_{n \rightarrow +\infty} \|u_n\|_V = +\infty$. La coercivité de J implique alors que $J(u_n)$ converge vers $+\infty$ ce qui est en contradiction avec le fait que $d < +\infty$.

V est un Banach réflexif, donc sa boule unité est faiblement compacte ; on peut donc extraire de la suite u_n une sous-suite encore notée u_n qui converge faiblement vers u dans V . On va montrer que u est solution de (29).

Comme K est un ensemble faiblement fermé, la limite faible u de la suite u_n de K est bien un élément de K .

D'autre part, on sait que J est sci pour la topologie faible ; on a donc

$$u_n \rightharpoonup u \text{ dans } V \Rightarrow J(u) \leq \liminf_{n \rightarrow +\infty} J(u_n),$$

ce qui donne

$$J(u) \leq \liminf_{n \rightarrow +\infty} J(u_n) = \lim_{n \rightarrow +\infty} J(u_n) = d \leq J(u).$$

On a donc $J(u) = d$ et $u \in K$: u est bien solution. □

Un corollaire important concerne le cas convexe.

Corollaire A.10. — *On suppose que V est un Banach réflexif. Soit J une fonctionnelle de V dans $\mathbb{R} \cup \{+\infty\}$, convexe sci, propre et K un sous-ensemble convexe non vide et fermé de V . Si J est coercive ou si K est borné, le problème de minimisation admet une solution. Si, de plus, J est strictement convexe la solution est unique.*

Rappelons enfin le théorème donnant une condition nécessaire d'optimalité du premier ordre.

Théorème A.11. — *Soient K un sous-ensemble convexe, non vide de V et J une fonctionnelle de K vers \mathbb{R} Gâteaux-différentiable sur K . Soit u dans V une solution du problème (\mathcal{P}) . Alors*

$$(30) \quad \forall v \in K, \quad \langle \nabla J(u), v - u \rangle \geq 0.$$

A.1.4. *Exemple : Projection sur un convexe fermé.* — Dans ce qui suit V est un espace de Hilbert muni d'un produit scalaire (\cdot, \cdot) et de la norme associée $\|\cdot\|$ et C est un sous-ensemble **convexe** et **fermé** de V .

Théorème A.12. — *Etant donné C un sous-ensemble convexe, fermé et non vide de V et x un élément quelconque de V . Alors le problème*

$$\min \{ \|x - y\|^2, y \in C \}$$

a une solution unique $x^ \in C$. De plus $x^* \in C$ est caractérisé par :*

$$(31) \quad \forall y \in C \quad (x - x^*, y - x^*) \leq 0.$$

Nous avons une seconde caractérisation de x^* de manière immédiate :

Corollaire A.13. — *Sous les hypothèses du théorème (A.12) on peut caractériser le projeté $x^* \in C$ de x par :*

$$(32) \quad \forall y \in C \quad (x^* - y, y - x) \leq 0.$$

Le point x^* est le **projeté** de x sur C . L'application $P_C : V \rightarrow C$ qui à x associe son projeté x^* est la **projection** sur C . Le projeté $P_C(x)$ est donc le point de C qui est le « plus près » de x . On définit de manière standard la fonction **distance** d'un point x à l'ensemble C par

$$(33) \quad d(x, C) = \inf_{y \in C} \|x - y\|.$$

Dans le cas où C est un convexe fermé, on vient donc de démontrer que

$$d(x, C) = \|x - P_C(x)\|.$$

Proposition A.14. — La projection P_C est continue de V dans C . Plus précisément on a

$$\forall (x, y) \in V \times V \quad \|P_C(x) - P_C(y)\| \leq \|x - y\|,$$

c'est-à-dire P_C est une contraction.

Remarque A.15. — 1. Si $x \in C$ alors $P_C(x) = x$. Plus généralement si $C = V$ alors $P_C = Id_V$.

2. Le théorème A.12 est faux si C n'est pas convexe ou si C n'est pas fermé.

3. La projection P_C n'est pas différentiable en général, mais l'application $x \mapsto \|x - P_C(x)\|^2$ l'est.

Exemple A.16 (Projection sur un sous-espace vectoriel)

Dans le cas où C est un sous-espace vectoriel fermé de V , c'est bien sûr un convexe fermé non vide. L'opérateur de projection est dans ce cas **linéaire** (c'est **faux** dans le cas général). Le projeté x^* d'un élément x , sur C , est caractérisé par

$$\forall y \in C \quad (x - x^*, y) = 0.$$

Cela signifie que $x - x^* \in C^\perp$ (l'orthogonal de C). On retrouve ainsi la classique projection orthogonale sur un sous-espace vectoriel fermé.

A.2. Analyse convexe et analyse non lisse. —

A.2.1. Théorème de Hahn-Banach. — Dans ce qui suit \mathcal{X} est un espace de Banach réel de dual \mathcal{X}' (pas nécessairement réflexif).

Le théorème de Hahn-Banach, sous sa forme géométrique, permet de séparer des ensembles convexes. Il est très important en analyse convexe et sert en particulier à exhiber des multiplicateurs en optimisation. Nous rappelons ici la forme géométrique de ce théorème qui est la seule utile dans notre cas, ainsi que des corollaires importants. Pour les démonstrations et plus de précisions nous renvoyons à [6].

Définition A.17 (Hyperplan affine). — Un **hyperplan affine fermé** est un ensemble de la forme

$$H = \{ x \in \mathcal{X} \mid \alpha(x) + \beta = 0 \},$$

où $\alpha \in \mathcal{X}'$ est une forme linéaire continue non nulle sur \mathcal{X} et $\beta \in \mathbb{R}$.

Dans le cas où \mathcal{X} est un espace de Hilbert V (en particulier si $V = \mathbb{R}^n$), on peut identifier V à son dual et tout hyperplan affine fermé est de la forme

$$H = \{ x \in V \mid (\alpha, x) + \beta = 0 \},$$

où on rappelle que (\cdot, \cdot) désigne le produit scalaire de V , $\alpha \in V$, $\alpha \neq 0$ et $\beta \in \mathbb{R}$.

Définition A.18 (Séparation). — Soient A et B deux sous-ensembles non vides de \mathcal{X} . On dit que l'hyperplan affine H d'équation : $\alpha(x) + \beta = 0$, sépare A et B **au sens large** si

$$\forall x \in A \quad \alpha(x) + \beta \leq 0 \quad \text{et} \quad \forall y \in B \quad \alpha(y) + \beta \geq 0.$$

On dit que H sépare A et B **strictement** s'il existe $\varepsilon > 0$ tel que

$$\forall x \in A \quad \alpha(x) + \beta \leq -\varepsilon \quad \text{et} \quad \forall y \in B \quad \alpha(y) + \beta \geq \varepsilon.$$

Donnons à présent la première forme géométrique du théorème de Hahn-Banach :

Théorème A.19. — Soient A et B deux sous-ensembles de \mathcal{X} convexes, non vides et disjoints. On suppose que A est ouvert. Alors, il existe un hyperplan affine fermé qui sépare A et B au sens large.

Corollaire A.20. — Soit C un convexe non vide de \mathbb{R}^n et fermé et $x^* \in C$. Alors : $x^* \in \text{Int}(C)$ si et seulement si il n'existe aucune forme linéaire séparant x^* et C .

Citons aussi la deuxième forme géométrique du théorème de Hahn-Banach :

Théorème A.21. — Soient A et B deux sous-ensembles de \mathcal{X} convexes, non vides et disjoints. On suppose que A est fermé et que B est compact. Alors, il existe un hyperplan affine fermé qui sépare A et B strictement.

A.2.2. Sous-différentiel. —

Définition A.22. — Soit $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ et $u \in \text{dom } f$ (i.e. $f(u) < +\infty$). Le sous-différentiel de f en u est l'ensemble $\partial f(u)$ (éventuellement vide) des $u^* \in \mathcal{X}'$ tels que

$$\forall v \in \mathcal{X} \quad f(v) \geq f(u) + \langle u^*, v - u \rangle.$$

Les éléments u^* sont appelés **sous-gradients**.

Remarque A.23. — 1. $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ atteint son minimum en $u \in \text{dom } f$ si et seulement si

$$0 \in \partial f(u).$$

2. Si $f, g : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ et $u \in \text{dom } f \cap \text{dom } g$, on a

$$\partial f(u) + \partial g(u) \subset \partial(f + g)(u).$$

3. Comme

$$\partial f(u) = \bigcap_{v \in \mathcal{X}} \{u^* \in \mathcal{X}' \mid \langle u^*, v - u \rangle \leq f(v) - f(u)\},$$

$\partial f(u)$ est un sous-ensemble convexe, fermé pour la topologie faible*, comme intersection de convexes fermés.

4. Pour tout $\lambda > 0$ on a $\partial(\lambda f)(u) = \lambda \partial f(u)$.

A.2.3. *Lien avec la Gâteaux-différentiabilité.* —

Théorème A.24. — Soit $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ **convexe**. Si f est Gâteaux-différentiable en $u \in \text{dom } f$, elle est sous-différentiable et $\partial f(u) = \{f'(u)\}$.

Réciproquement, si f est finie, continue en u et ne possède qu'un seul sous-gradient, alors f est Gâteaux-différentiable en u et $\partial f(u) = \{f'(u)\}$.

A.2.4. *Sous-différentiel d'une somme de fonctions.* —

Théorème A.25. — Soient f et g convexes, sci à valeurs dans $\mathbb{R} \cup \{+\infty\}$. On suppose qu'il existe $u_o \in \text{dom } f \cap \text{dom } g$ tel que f soit continue en u_o . Alors

$$\forall u \in \mathcal{X} \quad \partial(f + g)(u) = \partial f(u) + \partial g(u).$$

Démonstration. — La démonstration est laissée en exercice. Elle se fait suivant le schéma suivant : si $u^* \in \partial(f + g)(u)$, on définit

$$C_1 = \{(v, a) \mid f(v) - < v - u, u^* > - f(u) \leq a\},$$

et

$$C_2 = \{(v, a) \mid a \leq g(u) - g(v)\}.$$

Ce sont deux convexes. On sépare alors int C_1 et C_2 grâce au théorème de Hahn-Banach. \square

Le résultat suivant s'en déduit mais on peut le montrer directement sans hypothèse de convexité sur les deux fonctions :

Théorème A.26. — Supposons que f est Gâteaux-différentiable, φ est convexe et que le problème suivant

$$\min_{u \in \mathcal{X}} f(u) + \varphi(u),$$

a (au moins) une solution \bar{u} . Alors

$$-f'(u) \in \partial\varphi(\bar{u}).$$

Nous admettrons enfin le résultat suivant :

Théorème A.27. — Soit Λ linéaire continue de V dans Y (espaces de Banach). Soit f convexe, sci de V dans $\mathbb{R} \cup \{+\infty\}$ continue en au moins un point de son domaine (supposé non vide). Alors

$$\forall u \in V \quad \partial(f \circ \Lambda)(u) = \Lambda^* \partial f(\Lambda u),$$

où Λ^* est l'opérateur adjoint de Λ .

Pour plus de détails sur ces notions on peut consulter [3, 9]. Nous terminons par un exemple important.

A.2.5. *Application à l'indicatrice d'un ensemble.* — Dans le cas où f est la fonction indicatrice d'un sous-ensemble non vide K de \mathcal{X} :

$$f(u) \stackrel{\text{def}}{=} 1_K(u) = \begin{cases} 0 & \text{si } u \in K, \\ +\infty & \text{sinon} \end{cases}$$

le sous-différentiel de f en u est le **cône normal** de K en u :

$$\partial 1_K(u) = N_K(u) = \{ u^* \in \mathcal{X}' \mid \langle u^*, v - u \rangle \leq 0 \text{ pour tout } v \in K \}.$$

Dans le cas où \mathcal{X} est un espace de Hilbert identifié à son dual, et K un sous-ensemble fermé, convexe non vide de \mathcal{X} , nous allons préciser le sous-différentiel de 1_K en u (c'est-à-dire le cône normal à K en u) :

Proposition A.28. — *Soit $u \in K$, où K est un sous-ensemble fermé, convexe non vide de \mathcal{X} espace de Hilbert. Alors*

$$\lambda \in \partial 1_K(u) \iff \lambda = c \left[u + \frac{\lambda}{c} - P_K \left(u + \frac{\lambda}{c} \right) \right]$$

pour tout c réel strictement positif, où P_K est la projection de \mathcal{X} sur le convexe K .

Démonstration. — Remarquons tout d'abord que $\partial 1_K(u)$ est un sous-ensemble de \mathcal{X} . On rappelle également que si P_K est la projection de \mathcal{X} sur le convexe fermé K , l'image $P_K(w)$ d'un élément quelconque w de \mathcal{X} est caractérisée par

$$\forall v \in K \quad (w - P_K(w), v - P_K(w))_{\mathcal{X}} \leq 0,$$

où $(\cdot, \cdot)_{\mathcal{X}}$ désigne le produit scalaire de \mathcal{X} . Soit $\lambda \in \partial 1_K(u)$: λ est caractérisé par

$$\forall v \in K \quad (\lambda, v - u)_{\mathcal{X}} \leq 0$$

c'est-à-dire, pour tout $c > 0$

$$\forall v \in K \quad \left(u + \frac{\lambda}{c} - u, v - u \right)_{\mathcal{X}} \leq 0.$$

D'après ce qui précède (en posant $w = u + \frac{\lambda}{c}$)

$$\lambda \in \partial 1_K(u) \iff u = P_K \left(u + \frac{\lambda}{c} \right) \iff \lambda = c \left[u + \frac{\lambda}{c} - P_K \left(u + \frac{\lambda}{c} \right) \right].$$

□

A.2.6. *Calcul « pratique » : transformation de Legendre-Fenchel.* —

Définition A.29. — Soit $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction. La transformée de **Legendre-Fenchel** ou **conjuguée** de f est la fonction $f^* : \mathcal{X}' \rightarrow \bar{\mathbb{R}}$ définie par

$$(34) \quad \forall \ell \in \mathcal{X}' \quad f^*(\ell) = \sup_{u \in \mathcal{X}} \{ \ell(u) - f(u) \}.$$

Remarque A.30. — (a) Si f « prend » la valeur $-\infty$, alors $f^* \equiv +\infty$. Si f est propre (c'est-à-dire non indumentiquement égale à $+\infty$) alors f^* est à valeurs dans $\mathbb{R} \cup \{+\infty\}$.
 (a) On notera désormais $\ell(u) = \langle \ell, u \rangle$, où $\langle \cdot, \cdot \rangle$ désigne le produit de dualité entre \mathcal{X} et \mathcal{X}' . Ce produit de dualité coïncide avec le produit scalaire dans le cas où \mathcal{X} est un espace de Hilbert, après identification de \mathcal{X} et de son dual.

L'équation (34) s'écrit alors

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

Définition A.31. — Soit $A \subset \mathcal{X}$ un ensemble (non vide). La **fonction d'appui** de l'ensemble A est la fonction $\sigma_A : \mathcal{X}' \rightarrow \mathbb{R} \cup \{+\infty\}$ définie par $\sigma_A = (1_A)^*$ où 1_A désigne l'**indicatrice** de A

$$1_A(x) = \begin{cases} 0 & \text{si } x \in A, \\ +\infty & \text{sinon.} \end{cases}$$

Exemple A.32. — Soit A un ensemble et $f(x) = d(x, A)$. Alors $f^* = \sigma_A + 1_{B^*}$ où B est la boule unité de \mathcal{X}' .

Si $f : u \mapsto \|u\|_{\mathcal{X}}$ (où $\|\cdot\|_{\mathcal{X}}$ désigne la norme de \mathcal{X} , alors $f^* = 1_{B^*}$

Proposition A.33. — Pour toute fonction $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$, la fonction f^* est convexe et sci pour la topologie faible*.

Démonstration. — Par définition

$$f^* = \sup_{u \in \text{dom} f} \varphi_u,$$

où $\text{dom} f$ est le domaine de f (i.e. l'ensemble des éléments $u \in \mathcal{X}$ tels que $f(u)$ est finie) et $\varphi_u : \mathcal{X}' \rightarrow \mathbb{R}$ est définie par

$$\varphi_u(u^*) = \langle u^*, u \rangle - f(u).$$

Chacune des fonctions φ_u est affine et continue, donc convexe et sci pour la topologie faible de \mathcal{X}' . Il en est de même pour le sup. \square

Plus généralement

Proposition A.34. — Soit f une fonction positivement homogène (propre) de \mathcal{X} dans $\mathbb{R} \cup \{+\infty\}$, c'est-à-dire vérifiant

$$\forall \lambda \in \mathbb{R}, \forall x \in \mathcal{X} \quad f(\lambda x) = |\lambda| f(x).$$

Alors sa conjuguée f^* est l'indicatrice d'un sous-ensemble K convexe et fermé de \mathcal{X}' .

Démonstration. — Soit f une fonction positivement homogène (propre) de \mathcal{X} dans $\mathbb{R} \cup \{+\infty\}$. Soit $u^* \in \mathcal{X}'$. Deux cas se présentent :

- $\exists u_o \in \mathcal{X}$ tel que $\langle u^*, u_o \rangle - f(u_o) > 0$. Alors par homogénéité, pour tout $\lambda > 0$

$$\langle u^*, \lambda u_o \rangle - f(\lambda u_o) = \lambda [\langle u^*, u_o \rangle - f(u_o)] \leq f^*(u^*).$$

Donc, en passant à la limite pour $\lambda \rightarrow +\infty$ on obtient $f^*(u^*) = +\infty$.

• Dans le cas contraire

$$\forall u \in \mathcal{X} \quad \langle u^*, u \rangle - f(u) \leq 0,$$

et donc $f^*(u^*) \leq 0$. Or par définition de f^* ,

$$\langle u^*, 0 \rangle - f(0) \leq f^*(u^*) ;$$

comme f est positivement homogène $f(0) = f(n \cdot 0) = nf(0)$ pour tout $n \in \mathbb{N}$ et donc $f(0) = 0$. On a donc finalement : $f^*(u^*) = 0$.

Posons $K = \{u^* \in \mathcal{X}^* \mid f^*(u^*) = 0\}$. On vient de montrer que $f^* = 1_K$. Comme f^* est convexe et sci, K est fermé, convexe. \square

Nous allons maintenant donner un résultat reliant $f + g$ et $f^* + g^*$ qui est le fondement de la théorie de la dualité en analyse convexe :

Théorème A.35. — Soient $f, g : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ des fonctions convexes telles qu'il existe $u_o \in \text{dom } g$ avec f continue en u_o . Alors

$$\inf_{u \in \mathcal{X}} (f(u) + g(u)) = \max_{u^* \in \mathcal{X}'} (-f^*(u^*) - g^*(-u^*)).$$

Remarque A.36. — Notons que dans le théorème le « sup » dans le terme de droite est toujours atteint (c'est un « max ») ce qui n'est pas le cas dans le terme de gauche. où l'inf n'est pas nécessairement atteint.

Corollaire A.37. — Soit $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe continue en $u \in \mathcal{X}$. Alors

$$f(u) = \max_{u^* \in \mathcal{X}'} (\langle u^*, u \rangle - f^*(u^*)) .$$

Démonstration. — Posons $g = 1_{\{u\}}$. On a $g^*(u^*) = \langle u^*, u \rangle$ pour tout $u^* \in \mathcal{X}'$. Les fonctions f et g sont convexes et f est continue en $u \in \text{dom } g$. On a donc d'après le théorème précédent :

$$f(u) = \inf_{u \in \mathcal{X}} (f + g)(u) = \max_{u^* \in \mathcal{X}'} (-f^*(u^*) - g^*(-u^*)) = \max_{u^* \in \mathcal{X}'} (\langle u^*, u \rangle - f^*(u^*)).$$

\square

On peut généraliser ce résultat à des fonctions convexes sci.

Théorème A.38. — Soit $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe semi-continue inférieurement. Alors, pour tout $u \in \mathcal{X}$

$$f(u) = \max_{u^* \in \mathcal{X}'} (\langle u^*, u \rangle - f^*(u^*)).$$

Démonstration. — Voir [3] p. 89. C'est une démonstration similaire à celle du théorème (A.35). \square

Terminons par un résultat de bi-dualité qui est un corollaire du théorème précédent dans le cas où \mathcal{X} est réflexif. Ce résultat est encore vrai si même si \mathcal{X} n'est pas réflexif.

Théorème A.39. — Soit f une fonction propre, convexe et sci de \mathcal{X} dans $\mathbb{R} \cup \{+\infty\}$. Alors $f^{**} = f$.

A.2.7. Lien avec le sous-différentiel. —

Théorème A.40. — Soit $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ et f^* sa conjuguée. Alors

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

Démonstration. — Soit $u^* \in \partial f(u)$:

$$\forall v \in \mathcal{X} \quad f(v) \geq f(u) + \langle u^*, v - u \rangle.$$

Donc

$$f^*(u^*) \geq \langle u^*, u \rangle - f(u) \geq \sup\{\langle u^*, v \rangle - f(v) \mid v \in \mathcal{X}\} = f^*(u^*).$$

On obtient : $f(u) + f^*(u^*) = \langle u^*, u \rangle$.

Réciproquement, si $f(u) + f^*(u^*) = \langle u^*, u \rangle$ on a pour tout $v \in \mathcal{X}$

$$\langle u^*, u \rangle - f(u) = f^*(u^*) \geq \langle u^*, v \rangle - f(v),$$

$$\langle u^*, v - u \rangle \leq f(v) - f(u),$$

c'est-à-dire $u^* \in \partial f(u)$. □

Corollaire A.41. — Si $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ est convexe, propre et sci, alors

$$u^* \in \partial f(u) \iff u \in \partial f^*(u^*).$$

Démonstration. — Il suffit d'appliquer le théorème précédent à f^* et d'utiliser le fait que lorsque f est convexe, propre et sci alors $f = f^{**}$. □

Références

- [1] Aubert G. et Kornprobst P., *Mathematical Problems in Image Processing*, Applied Mathematical Science 147, Springer 2006
- [2] Attouch H., Buttazzo G. et Michaille G., *Variational analysis in Sobolev and BV spaces*, MPS-SIAM Series on Optimization, 2006
- [3] Azé D., *Éléments d'analyse convexe et variationnelle*, Ellipses, Paris, 1997.
- [4] Barbu V. et Precupanu Th., *Convexity and Optimization in Banach Spaces*, Sijthoff & Noordhoff, Bucarest 1978.
- [5] Bergounioux M. *Méthodes mathématiques pour le traitement du signal*, Collection « Mathématiques Appliquées pour le Master/SMAI », Dunod, Paris, 2010
- [6] Brézis H., *Analyse Fonctionnelle*, Masson, Paris, 1987.
- [7] Chambolle A., *An algorithm for total variation minimization and applications*, JMIV, 20, 89-97, 2004.
- [8] Dautray R. et Lions J.L., *Analyse mathématique et calcul numérique*, 9 volumes, Masson, Paris, 1987.
- [9] Ekeland I. et Temam R., *Analyse convexe et problèmes variationnels*, Dunod, Paris, 1973.

- [10] Evans L. et Gariepy R., *Measure theory and fine properties of functions*, CRC Press, Boca Raton, Floride, 1992.
- [11] Gasquet C. et Witomski P., *Analyse de Fourier et applications*, Masson, 1995.
- [12] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1998.
- [13] Pallara D., Ambrosio L., Fusco N., *Functions of bounded variations and free discontinuity problems*, Oxford Mathematical Monographs, 2000.
- [14] Rudin W., *Analyse réelle et complexe*, Masson, 1978.
- [15] Rudin L., Osher S. et Fatemi E., *Non linear total variation based noise removal algorithms*, Physica D, 20, pp. 259-268, 1992

MAÏTINE BERGOUNIOUX, Université d'Orléans, Département de mathématique, BP 6759, F-45067 Orléans Cedex 2 • *E-mail* : maitine.bergounioux@univ-orleans.fr
Url : <http://web.mac.com/maitine.bergounioux/iWeb/PagePro/Accueil.html>