

# Mathématiques pour l'Ingénieur

Thomas Cluzeau

École Nationale Supérieure d'Ingénieurs de Limoges

16 rue d'atlantis, Parc ester technopole

87068 Limoges CEDEX

cluzeau@ensil.unilim.fr

<http://www.ensil.unilim.fr/~cluzeau>





# Table des matières

<b>1</b>	<b>Introduction aux distributions</b>	<b>7</b>
1.1	Fonctionnelle . . . . .	10
1.2	L'espace de fonctions tests $\mathcal{D}$ . . . . .	10
1.2.1	Définition . . . . .	10
1.2.2	Exemples . . . . .	11
1.2.3	Topologie de $\mathcal{D}$ . . . . .	13
1.3	L'espace $\mathcal{D}'$ des distributions . . . . .	13
1.3.1	Définition . . . . .	13
1.3.2	Exemples, distributions régulières et singulières . . . . .	14
1.3.3	Support d'une distribution . . . . .	15
1.4	Opérations sur les distributions . . . . .	15
1.4.1	Translation . . . . .	16
1.4.2	Transposition . . . . .	16
1.4.3	Dilatation (homothétie ou changement d'unité) . . . . .	16
1.4.4	Multiplication des distributions . . . . .	17
1.4.5	Dérivation des distributions . . . . .	17
1.4.6	Dérivation d'une fonction discontinue . . . . .	18
1.4.7	Convergence (faible) dans l'espace $\mathcal{D}'$ des distributions . . . . .	19
1.4.8	Sous-espaces de $\mathcal{D}'$ . . . . .	19
1.5	Distributions à plusieurs dimensions . . . . .	20
<b>2</b>	<b>La Convolution</b>	<b>21</b>
2.1	Produit tensoriel . . . . .	23
2.1.1	De deux fonctions . . . . .	23
2.1.2	De deux distributions . . . . .	24
2.2	Produit de convolution . . . . .	24
2.2.1	Convolution de deux fonctions . . . . .	24
2.2.2	Notion de mesure floue . . . . .	25
2.2.3	Convolution de deux distributions . . . . .	25
2.2.4	Propriétés du produit de convolution de deux distributions . . . . .	27
2.3	Algèbre de convolution et résolution d'équations différentielles . . . . .	28
2.3.1	Définition . . . . .	28
2.3.2	Calcul algébrique . . . . .	28

2.3.3	Résolution d'une équation différentielle avec conditions initiales . . . . .	29
2.4	Interprétation physique de la convolution . . . . .	30
2.4.1	Systèmes décrits par un opérateur de convolution . . . . .	30
2.4.2	Système causal . . . . .	31
2.4.3	Réponse à une excitation exponentielle . . . . .	31
<b>3</b>	<b>La Transformation de Fourier</b>	<b>33</b>
3.1	Transformée de Fourier des fonctions . . . . .	33
3.1.1	Définition et existence . . . . .	33
3.1.2	Inversion . . . . .	34
3.1.3	Transformée de Fourier en sinus et cosinus . . . . .	34
3.1.4	Propriétés . . . . .	35
3.1.5	Dérivation . . . . .	36
3.1.6	Transformée de Fourier et convolution . . . . .	37
3.1.7	Formule de Parseval-Plancherel . . . . .	38
3.2	Transformée de Fourier des distributions . . . . .	39
3.2.1	Définition . . . . .	39
3.2.2	Espace $\mathcal{S}$ et transformée de Fourier . . . . .	40
3.2.3	Transformée de Fourier des distributions tempérées . . . . .	40
3.2.4	Propriétés . . . . .	41
3.2.5	Transformée de Fourier de la distribution peigne de Dirac . . . . .	42
3.3	Séries de Fourier et Échantillonnage . . . . .	42
3.3.1	Transformée de Fourier des fonctions périodiques . . . . .	42
3.3.2	Échantillonnage . . . . .	44
<b>4</b>	<b>La Transformation de Laplace</b>	<b>47</b>
4.1	Transformée de Laplace des fonctions . . . . .	47
4.1.1	Définition . . . . .	47
4.1.2	Lien entre transformées de Laplace et de Fourier . . . . .	47
4.1.3	Domaine de définition, abscisse de sommabilité . . . . .	48
4.1.4	Formule d'inversion . . . . .	48
4.1.5	Propriétés et Exemples . . . . .	49
4.2	Transformée de Laplace des distributions . . . . .	49
4.2.1	Lien entre transformée de Laplace et de Fourier . . . . .	50
4.2.2	Exemples . . . . .	50
4.3	Application à la résolution d'équations de convolution . . . . .	50
4.4	Utilisation de la transformée de Laplace en physique . . . . .	51
4.4.1	Calcul des fonctions de transfert en électronique . . . . .	51
4.4.2	En mécanique . . . . .	52

<b>5</b>	<b>Introduction à l'Optimisation</b>	<b>55</b>
5.1	Rappels et compléments de calcul différentiel . . . . .	55
5.1.1	Dérivées premières . . . . .	55
5.1.2	Dérivation de fonctions composées . . . . .	57
5.1.3	Dérivées d'ordre supérieurs . . . . .	58
5.2	Extrema d'une fonction de plusieurs variables réelles . . . . .	58
5.2.1	Définitions et premiers résultats . . . . .	58
5.2.2	Extrema relatifs . . . . .	60
5.2.3	Conditions nécessaires et suffisantes en utilisant la dérivée seconde . . . . .	62
5.3	Optimisation sous contrainte dans une partie de $\mathbb{R}^n$ . . . . .	63
5.3.1	Optimisation avec contrainte d'égalité . . . . .	63
5.3.2	Optimisation sous d'autres contraintes . . . . .	65



# Chapitre 1

## Introduction aux distributions

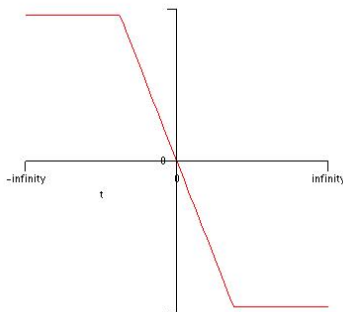
### Introduction

Les distributions sont utilisées depuis longtemps par les physiciens (distributions de Dirac, ...) mais une théorie mathématique rigoureuse n'est apparue que récemment dans les travaux de Sobolev (1936) et surtout L. Schwartz (1950) (en parallèle : Gelfand (1964)). C'est la théorie la plus adaptée à l'étude de nombreux systèmes physiques et notamment à celle des systèmes linéaires continus. Avec la notion de distribution, la convolution (voir Chapitre 2) et la transformée de Fourier (voir Chapitre 3) deviennent des outils mathématiques d'une grande puissance.

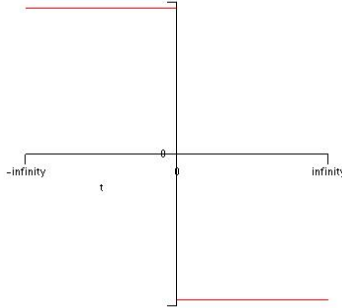
Intuitivement, les distributions sont des outils mathématiques utilisés pour représenter des phénomènes physiques que les fonctions classiques s'avèrent incapables de transcrire.

**Premier exemple introductif** : Choc élastique entre deux objets

Considérons une partie de squash. On suppose que la balle arrive sur un mur (perpendiculairement à la surface pour simplifier) à la vitesse  $v_0$  et rebondit. La balle s'écrase quelque peu ce qui fait que le choc dure un temps  $\Delta t$  non nul, puis elle repart avec une vitesse  $-v_0$ . Le graphe de la vitesse en fonction du temps est donc le suivant :



La loi de la mécanique Newtonienne stipule que, tout au long du mouvement, la force  $F$  exercée sur la balle est telle que  $F = m \dot{v}$ ; elle est donc proportionnelle à la dérivée de la fonction représentée ci-dessus. Maintenant, si l'on veut modéliser un choc dur (partie de pétanque), le graphe de la vitesse devient alors :



La force exercée devrait toujours être proportionnelle à la dérivée de cette fonction donc  $F$  devrait être nulle pour tout  $t \neq 0$  et vérifier

$$\frac{1}{m} \int_{-\infty}^{+\infty} F(t) dt = v(+\infty) - v(-\infty) = -2v_0,$$

ce qui est absurde car l'intégrale d'une fonction presque partout nulle est nulle. Par conséquent, ni cette intégrale, ni la dérivée précédente ne peuvent être traitées au sens des fonctions; on a besoin d'objets plus généraux, *i.e.*, les distributions.

### Deuxième exemple introductif : Distributions de charges en électrostatique

Trois notions de charges apparaissent en électrostatique :

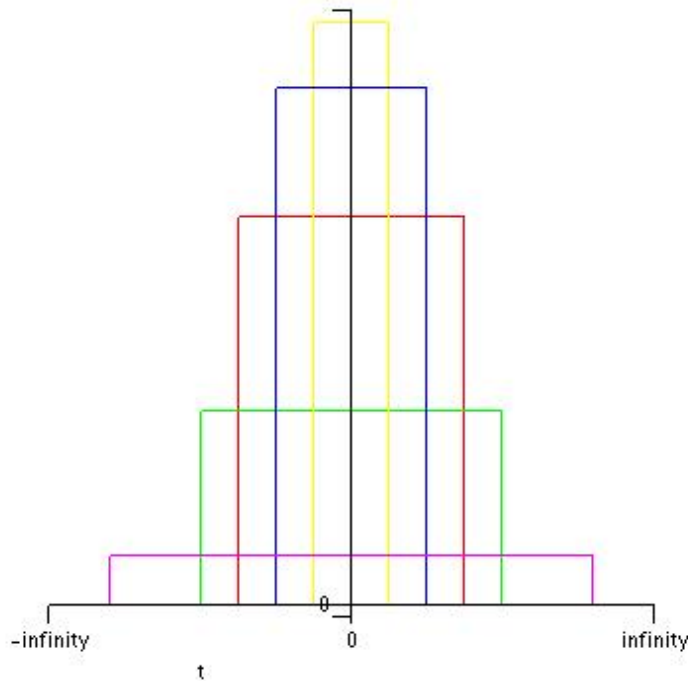
- les charges ponctuelles  $q_i$ ,
- les densités de charges superficielles  $\sigma$  (ou charge par unité de surface) sur les conducteurs,
- les densités de charges volumiques  $\rho$  (par unité de volume).

On peut remarquer qu'une distribution ponctuelle de charge peut s'obtenir comme limite d'une distribution volumique en faisant tendre, à charge constante, le volume vers zéro. Pour simplifier, prenons un exemple à une dimension. Soit  $\Pi(x)$  la fonction "porte" définie par :

$$\Pi(x) = \begin{cases} 1 & \text{pour } |x| \leq 1/2, \\ 0 & \text{pour } |x| > 1/2. \end{cases}$$

Considérons, sur une droite, une suite de densités de charges  $\rho_k(x) = k \Pi(kx)$  donnée par la figure suivante :





L'intégrale, qui représente la charge totale en électrostatique, est indépendante de  $k$  :

$$\int \rho_k(x) dx = 1.$$

Lorsque  $k$  tend vers l'infini, la charge totale, qui reste égale à 1, est entièrement concentrée à l'origine. On a obtenu une charge unité ponctuelle à l'origine. On a donc envie de représenter cette charge par une fonction  $\delta(x)$  qui vaudrait :

$$\delta(x) = \begin{cases} 0 & \text{pour } x \neq 0, \\ +\infty & \text{pour } x = 0, \end{cases}$$

et telle que  $\int \delta(x) dx = 1$  ce qui est absurde car l'intégrale d'une fonction presque partout nulle est nulle. Les fonctions ne permettent pas de représenter ce phénomène : on a besoin d'objets plus généraux, *i.e.*, les distributions.

### Autres exemples :

- En mécanique, dans le cadre de l'application du Principe Fondamental de la Dynamique, comment écrire l'équation du mouvement d'un solide lorsque le système est soumis à une force intense appliquée pendant un intervalle de temps très court à partir de l'instant  $t = t_0$  ?
- En électricité, comment va se comporter un circuit dont l'entrée varie brusquement ; par exemple par fermeture d'un interrupteur sur une source de tension continue ?

- En hydraulique, comment va se comporter un système dont on ouvre brusquement une vanne à l'instant  $t = t_0$  ?

## 1.1 Fonctionnelle

**Définition 1.1.1.** *On dit que l'on a une fonctionnelle sur un ensemble de fonctions appelées fonctions tests, si à chacune de ces fonctions on peut associer un nombre complexe. Autrement dit une fonctionnelle  $T$  sur un espace de fonctions  $\mathcal{F}$  est une application de  $\mathcal{F}$  dans  $\mathbb{C}$ . Le nombre associé par  $T$  à  $\varphi \in \mathcal{F}$  est noté  $\langle T, \varphi \rangle$ .*

Une grande variété de fonctions tests peuvent être utilisées et plus les conditions de régularité imposées aux fonctions tests sont sévères, plus les fonctionnelles définies sont générales. Les distributions seront définies comme fonctionnelles sur un certain espace, noté  $\mathcal{D}$ , que nous allons présenter maintenant.

## 1.2 L'espace de fonctions tests $\mathcal{D}$

Dans ce chapitre, nous nous restreindrons au cas à une dimension, c'est-à-dire que les fonctions considérées seront des fonctions à une seule variable réelle.

### 1.2.1 Définition

**Définition 1.2.1.** *Soit  $f$  une fonction à valeurs complexes définie sur  $\mathbb{R}$ . Le support de  $f$ , noté  $\text{Supp}(f)$ , est l'adhérence des  $x \in \mathbb{R}$  tels que  $f(x) \neq 0$ .*

$$\text{Supp}(f) = \overline{\{x \in \mathbb{R}; f(x) \neq 0\}}.$$

**Rappel :** l'adhérence d'un ensemble est le plus petit fermé contenant cet ensemble. Dans le cas des fonctions d'une seule variable, l'adhérence est un intervalle compact du type  $[a, b]$ .

Le support de  $f$  est donc un ensemble fermé en dehors duquel  $f$  est nulle et en outre c'est le plus petit ensemble possédant cette propriété.

**Définition 1.2.2.** *On définit l'ensemble  $\mathcal{D}$  comme l'espace des fonctions à valeurs complexes définies sur  $\mathbb{R}$ , indéfiniment dérivables et à support borné.*

**Remarque :** C'est un espace vectoriel de dimension infinie.

Le support étant fermé par définition, on peut remplacer dans la définition précédente support borné par support compact. En effet, pour qu'un sous-ensemble de  $\mathbb{R}$  soit compact, il faut et il suffit qu'il soit fermé et borné.

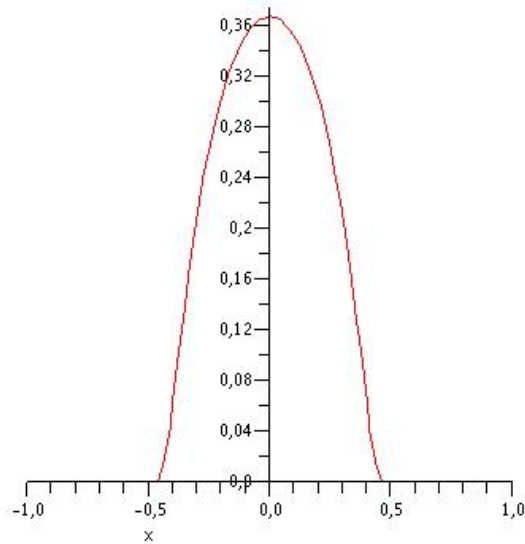
## 1.2.2 Exemples

Des exemples de fonctions appartenant à  $\mathcal{D}$  ne viennent pas immédiatement à l'esprit ; les fonctions analytiques ne peuvent pas convenir.

**Exemple fondamental** : Soit  $\xi_a$  la fonction définie par :

$$\xi_a(x) = \begin{cases} 0 & \text{pour } |x| \geq 1/a, \\ \exp\left(\frac{-1}{1-a^2x^2}\right) & \text{pour } |x| < 1/a, \end{cases}$$

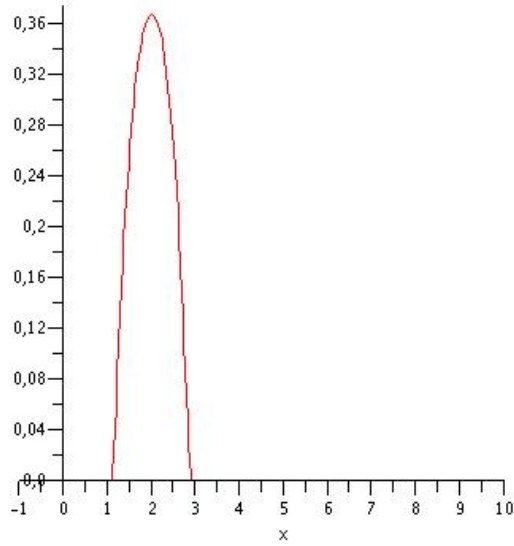
avec  $a > 0$ . Elle est indéfiniment dérivable, son support est  $[-1/a, 1/a]$  et il est facile de vérifier que toutes ses dérivées sont nulles en  $x = 1/a$  et  $x = -1/a$ . Voici le graphe de  $\xi_2$  :



Plus généralement, toute fonction  $\xi_{ab}$  définie par

$$\xi_{ab}(x) = \begin{cases} 0 & \text{pour } x \notin ]a, b[, \\ \exp\left(\frac{1}{2} \left[ \frac{1}{x-b} - \frac{1}{x-a} \right]\right) & \text{pour } x \in ]a, b[, \end{cases}$$

est une fonction de  $\mathcal{D}$  de support  $[a, b]$ . Voici le graphe de  $\xi_{13}$  :



Une autre famille de fonctions de  $\mathcal{D}$  est définie par

$$\gamma_k(x) = \frac{\xi_1(kx)}{\int \xi_1(kx) dx}.$$

Ces fonctions permettent d'en construire beaucoup d'autres grâce au théorème suivant :

**Théorème 1.2.1.** *Si  $\varphi \in \mathcal{D}$  et si  $f$  est une fonction sommable à support borné, alors*

$$\psi(x) = \int f(t) \varphi(x - t) dt$$

*est une fonction de  $\mathcal{D}$ .*

Considérons maintenant la suite de fonctions

$$\psi_k(x) = \int f(t) \gamma_k(x - t) dt.$$

On démontre que si  $f$  est continue, alors cette suite converge uniformément vers  $f$ . D'où le théorème admis :

**Théorème 1.2.2** (Théorème d'approximation). *Toute fonction continue à support borné peut être approchée uniformément par une suite  $(\varphi_n)_{n>0}$  de fonctions de  $\mathcal{D}$ .*

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, \text{ tel que, } \forall n \geq N, \forall x, |f(x) - \varphi_n(x)| \leq \epsilon.$$

### 1.2.3 Topologie de $\mathcal{D}$

Elle sera définie par un critère de convergence pour les suites.

**Définition 1.2.3.** Une suite  $(\varphi_n)_{n>0}$  de fonctions de  $\mathcal{D}$  converge vers une fonction  $\varphi$  lorsque  $n$  tend vers l'infini si :

1. Il existe un ensemble borné  $B$  (indépendant de  $n$ ) de  $\mathbb{R}$  tel que pour tout  $n > 0$ ,  $\text{Supp}(\varphi_n) \subset B$  ;
2. Pour tout entier  $k \geq 0$ , la suite des dérivées  $(\varphi_n^{(k)})_n$  converge uniformément sur  $\mathbb{R}$  vers  $\varphi^{(k)}$ .

On peut montrer que la limite  $\varphi$  appartient alors à  $\mathcal{D}$ .

## 1.3 L'espace $\mathcal{D}'$ des distributions

### 1.3.1 Définition

**Définition 1.3.1.** On appelle distribution toute fonctionnelle linéaire continue sur l'espace vectoriel  $\mathcal{D}$ .

Soit  $T$  une distribution. Par définition,  $T$  est une fonctionnelle sur  $\mathcal{D}$  donc  $T$  associe à toute fonction  $\varphi \in \mathcal{D}$  un complexe noté  $\langle T, \varphi \rangle$  (ou parfois  $T(\varphi)$ ).

La définition d'une distribution implique les deux points suivants :

#### 1. Linéarité

- $\langle T, \varphi_1 + \varphi_2 \rangle = \langle T, \varphi_1 \rangle + \langle T, \varphi_2 \rangle$ ,
- $\langle T, \lambda \varphi_1 \rangle = \lambda \langle T, \varphi_1 \rangle$ .

2. Si  $(\varphi_k)_{k>0}$  converge dans  $\mathcal{D}$  vers  $\varphi$ , alors la suite  $(\langle T, \varphi_k \rangle)_{k>0}$  converge au sens usuel vers  $\langle T, \varphi \rangle$ , i. e.,

$$\forall \epsilon > 0, \exists N \in \mathbb{N} \text{ tel que, } \forall k > N, |\langle T, \varphi \rangle - \langle T, \varphi_k \rangle| \leq \epsilon.$$

L'ensemble des distributions est un espace vectoriel noté  $\mathcal{D}'$ . La somme de deux distributions et le produit d'une distribution par un scalaire sont définis comme suit :

- $\langle S + T, \varphi \rangle = \langle S, \varphi \rangle + \langle T, \varphi \rangle$ ,
- $\langle \lambda T, \varphi \rangle = \lambda \langle T, \varphi \rangle$ .

### 1.3.2 Exemples, distributions régulières et singulières

**Définition 1.3.2.** Une fonction  $f : \mathbb{R} \rightarrow \mathbb{C}$  est dite localement sommable si elle est intégrable sur tout intervalle borné. À toute fonction  $f$  localement sommable, on associe la distribution  $T_f$  définie par

$$\forall \varphi \in \mathcal{D}, \langle T_f, \varphi \rangle = \int f(x) \varphi(x) dx.$$

Une telle distribution est dite *régulière*. Les autres (celles qui ne s'écrivent pas  $T_f$  pour  $f$  localement sommable) sont dites *singulières*.

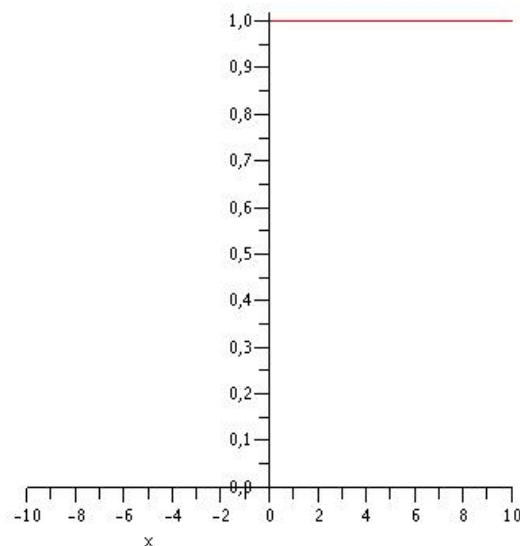
**Remarque :** Deux fonctions localement sommables définissent la même distribution si et seulement si elle sont égales presque partout.

Un premier exemple de distribution régulière est la distribution *valeur principale de Cauchy* de  $1/x$  notée  $\text{vp} \frac{1}{x}$  et définie par :

$$\forall \varphi \in \mathcal{D}, \langle \text{vp} \frac{1}{x}, \varphi \rangle = \lim_{\epsilon \rightarrow 0^+} \int_{|x| > \epsilon} \frac{\varphi(x)}{x} dx.$$

Un second exemple est la distribution de Heaviside. La fonction  $H$  de Heaviside est définie par

$$H(x) = \begin{cases} 1 & \text{pour } x \geq 0, \\ 0 & \text{pour } x < 0. \end{cases}$$



La distribution de Heaviside, notée  $W = T_H$ , est définie par

$$\forall \varphi \in \mathcal{D}, \langle W, \varphi \rangle = \int_0^{+\infty} \varphi(x) dx.$$

L'exemple le plus usuel de distribution singulière est la *distribution de Dirac* notée  $\delta$  et définie par :

$$\forall \varphi \in \mathcal{D}, \langle \delta, \varphi \rangle = \varphi(0).$$

Plus généralement, on définit la *distribution de Dirac au point  $a$*  et on note  $\delta_a$  la distribution définie par

$$\forall \varphi \in \mathcal{D}, \langle \delta_a, \varphi \rangle = \varphi(a).$$

**Attention** : En physique, on écrit souvent  $\delta(x)$  ou  $\delta(x - a)$  au lieu de  $\delta$  et  $\delta_a$ . Cette écriture laisse croire que  $\delta$  est une fonction, ce qui est faux !

La distribution de Dirac  $\delta_a$  est souvent interprétée comme représentant la masse (ou la charge)  $+1$  au point  $a$ .

Toute combinaison linéaire de distributions de Dirac est une distribution singulière. En particulier la distribution  $\sum_{n=-\infty}^{+\infty} \delta_n$  ( $n$  entier) a des propriétés intéressantes et joue un rôle important en physique. On l'appelle *distribution peigne de Dirac* et on la note  $\text{III}$ .

**Remarque** : Ceux sont les généralisations à trois dimensions des distributions de Dirac qui donnent une représentation mathématique correcte des charges ponctuelles et superficielles en électrostatique.

### 1.3.3 Support d'une distribution

**Définition 1.3.3.** On dit que deux distributions  $S$  et  $T$  sont égales si  $\langle S, \varphi \rangle = \langle T, \varphi \rangle$  quel que soit  $\varphi \in \mathcal{D}$ . On dit qu'elles sont égales sur un ouvert  $\Omega \subset \mathbb{R}$  si  $\langle S, \varphi \rangle = \langle T, \varphi \rangle$  quel que soit  $\varphi \in \mathcal{D}$  ayant son support dans  $\Omega$ .

**Exemples** : Les distributions régulières  $T_1$  et  $W$  sont égales sur  $]0, +\infty[$ . Les distributions  $\delta$  et  $\text{III}$  sont égales sur  $] -\frac{1}{2}, \frac{1}{2}[$ .

**Définition 1.3.4.** Considérons la réunion de tous les ouverts sur lesquels une distribution  $T$  est nulle. Cet ensemble est alors le plus grand ouvert sur lequel  $T$  est nulle (admis). Son complémentaire (qui est un fermé) est appelé support de la distribution  $T$  ; on le note  $\text{Supp}(T)$ .

**Exemples** :  $\text{Supp}(\delta_a) = \{a\}$  et  $\text{Supp}(\text{III}) = \mathbb{Z}$ .

## 1.4 Opérations sur les distributions

**Méthodologie** : on souhaite définir un certain nombre d'opérations sur les distributions. Pour ceci, on va étudier comment ces opérations sont définies pour une fonction localement sommable, traduire ceci avec le langage des distributions sur la distribution régulière associée et généraliser.

### 1.4.1 Translation

Si  $f$  est localement sommable et si  $a \in \mathbb{R}$ , alors la translatée  $f_a$  de  $f$  est la fonction donnée par  $f_a(x) = f(x - a)$ . La distribution régulière associée à  $f_a$  vérifie donc

$$\forall \varphi \in \mathcal{D}, \langle T_{f_a}, \varphi \rangle = \int f(x - a)\varphi(x)dx = \int f(y)\varphi(y + a)dy = \langle T_f, \varphi_{-a} \rangle .$$

**Définition 1.4.1.** La translatée d'une distribution  $T$ , notée  $T_a$  est la distribution définie par :

$$\forall \varphi \in \mathcal{D}, \langle T_a, \varphi \rangle = \langle T, \varphi_{-a} \rangle .$$

**Exemple :** La translatée de la valeur principale de Cauchy de  $1/x$  est la valeur principale de  $1/(x - a)$ .

### 1.4.2 Transposition

Soit  $f$  une fonction localement sommable et cherchons la distribution associée à la fonction  $\check{f}$  qui à  $x$  associe  $f(-x)$ . On a

$$\forall \varphi \in \mathcal{D}, \langle T_{\check{f}}, \varphi \rangle = \int f(-x)\varphi(x)dx = \int f(x)\varphi(-x)dx = \langle T_f, \check{\varphi} \rangle .$$

**Définition 1.4.2.** La transposée d'une distribution  $T$ , notée  $\check{T}$  est la distribution définie par :

$$\forall \varphi \in \mathcal{D}, \langle \check{T}, \varphi \rangle = \langle T, \check{\varphi} \rangle .$$

**Remarque :** Ceci permet de définir des distributions paires et impaires comme pour les fonctions.

### 1.4.3 Dilatation (homothétie ou changement d'unité)

Si  $f$  est localement sommable et si  $a \in \mathbb{R}$ , alors la dilatée de la fonction  $f$  est définie par  $x \mapsto f(ax)$ . Sa distribution régulière associée vérifie

$$\forall \varphi \in \mathcal{D}, \langle "T_{f(ax)}", \varphi \rangle = \int f(ax)\varphi(x)dx = \int f(y)\varphi\left(\frac{y}{a}\right)\frac{dy}{|a|} = \frac{1}{|a|} \langle T_f, " \varphi\left(\frac{x}{a}\right) " \rangle .$$

**Définition 1.4.3.** La dilatée d'une distribution  $T$  est la distribution définie par :

$$\forall \varphi \in \mathcal{D}, \langle "T(ax)", \varphi \rangle = \frac{1}{|a|} \langle T, " \varphi\left(\frac{x}{a}\right) " \rangle .$$

**Exemple :** " $\delta(ax)$ " =  $\frac{1}{|a|} \delta$ .



### 1.4.4 Multiplication des distributions

Il n'existe pas de moyen de multiplier entre elles deux distributions. En outre, si  $f$  et  $g$  sont deux fonctions localement sommables, alors leur produit ne l'est pas nécessairement ( $f(x) = g(x) = 1/\sqrt{x}$ ).

Cependant si  $\psi$  est une fonction indéfiniment dérivable, alors le produit par  $\psi$  d'une fonction test de  $\mathcal{D}$  est encore dans  $\mathcal{D}$ . Soit  $f$  une fonction localement sommable et  $\psi$  une fonction indéfiniment dérivable. On a alors

$$\forall \varphi \in \mathcal{D}, \langle T_{\psi f}, \varphi \rangle = \int (\psi(x) f(x)) \varphi(x) dx = \int f(x) (\psi(x) \varphi(x)) dx = \langle T_f, \psi \varphi \rangle .$$

**Définition 1.4.4.** Soit  $\psi$  une fonction indéfiniment dérivable. Le produit  $\psi T$  d'une distribution  $T$  par  $\psi$  est la distribution définie par :

$$\forall \varphi \in \mathcal{D}, \langle \psi T, \varphi \rangle = \langle T, \psi \varphi \rangle .$$

À partir de cette définition, on peut définir le produit d'une distribution quelconque  $T$  par une distribution régulière  $T_\psi$  associée à une fonction indéfiniment dérivable  $\psi$  de la manière suivante :

$$\forall \varphi \in \mathcal{D}, \langle T_\psi T, \varphi \rangle = \langle T, \psi \varphi \rangle .$$

**Lemme 1.4.1.** Soit  $\psi$  une fonction indéfiniment dérivable. On a :

$$\psi \delta = \psi(0) \delta .$$

*Démonstration.*

$$\forall \varphi \in \mathcal{D}, \langle \psi \delta, \varphi \rangle = \langle \delta, \psi \varphi \rangle = \psi(0) \varphi(0) = \psi(0) \langle \delta, \varphi \rangle = \langle \psi(0) \delta, \varphi \rangle ,$$

d'où le résultat. □

En particulier,  $x \delta = 0$ . L'équation  $x T = 0$ , de distribution inconnue  $T$ , a pour solutions les multiples de la distribution de Dirac.

### 1.4.5 Dérivation des distributions

Soit  $f$  une fonction localement sommable que nous supposons de plus dérivable. Dans ce cas,  $f'$  est localement sommable et sa distribution régulière associée vérifie :

$$\forall \varphi \in \mathcal{D}, \langle T_{f'}, \varphi \rangle = \int f'(x) \varphi(x) dx = - \int f(x) \varphi'(x) dx = - \langle T_f, \varphi' \rangle .$$

Notons que ceci s'obtient par intégration par partie en utilisant le fait que  $\varphi$  est à support borné. C'est la raison principale du choix restrictif des fonctions tests, *i.e.*, de l'espace  $\mathcal{D}$ .

**Définition 1.4.5.** La dérivée  $T'$  d'une distribution  $T$  de  $\mathcal{D}'$  est la distribution définie par :

$$\forall \varphi \in \mathcal{D}, \langle T', \varphi \rangle = - \langle T, \varphi' \rangle .$$

De même on pourra définir les dérivées successives  $T^{(m)}$  par :

$$\forall \varphi \in \mathcal{D}, \langle T^{(m)}, \varphi \rangle = (-1)^m \langle T, \varphi^{(m)} \rangle .$$

**Exemple :**  $W' = \delta$ . En effet

$$\forall \varphi \in \mathcal{D}, \langle W', \varphi \rangle = - \langle W, \varphi' \rangle = - \int_0^{+\infty} \varphi'(x) dx = -[\varphi(x)]_0^{+\infty} = \varphi(0) = \langle \delta, \varphi \rangle .$$

**Lemme 1.4.2.** Soit  $T$  une distribution quelconque et  $\psi$  une fonction indéfiniment dérivable. On a alors la règle de Leibniz suivante :

$$(T\psi)' = T'\psi + T\psi' .$$

*Démonstration.* Voir TD 1. □

## 1.4.6 Dérivation d'une fonction discontinue

On a vu que la dérivée au sens des distributions de la distribution de Heaviside était égale à la distribution de Dirac. Maintenant si on considère la fonction de Heaviside, sa dérivée est nulle partout sauf en 0 où elle n'est pas définie et la distribution associée n'est pas  $\delta$ . Par conséquent, les opérations "prendre la distribution associée" et "dérivation" ne commutent pas, ou, autrement dit,  $(T_f)' \neq T_{f'}$ . Cela sera ainsi pour toute fonction présentant une discontinuité en un point.

Soit  $f$  une fonction  $\mathcal{C}^1$  par morceaux. Soient  $a_1, \dots, a_n$  les points de discontinuité de  $f$  (que nous supposons en nombre fini) et  $\sigma_i^{(0)} = f(a_i^+) - f(a_i^-)$  le saut de discontinuité de  $f$  en  $a_i$ . La fonction  $f$  peut alors s'écrire comme la somme d'une fonction continue  $g$  et de fonctions de Heaviside.

$$f(x) = g(x) + \sum_{i=0}^n \sigma_i^{(0)} H(x - a_i) .$$

De plus, on a  $T_{f'} = T_{g'}$  où  $f'$  désigne la dérivée de  $f$  là où elle est bien définie c'est-à-dire en dehors des points de discontinuité.

**Théorème 1.4.1.** Soit  $f$  une fonction de classe  $\mathcal{C}^1$  par morceaux. Avec les notations précédentes, on a alors

$$(T_f)' = T_{f'} + \sum_i \sigma_i^{(0)} \delta_{a_i} .$$

On notera plus simplement

$$(T_f)' = T_{f'} + \sigma^{(0)} \delta .$$

De même, soit  $f$  une fonction  $\mathcal{C}^\infty$  par morceaux. Si l'on note  $\sigma^{(j)}$  les sauts de discontinuité de  $f^{(j)}$ , on a

$$(T_f)^{(m)} = T_{f^{(m)}} + \sigma^{(m-1)} \delta + \sigma^{(m-2)} \delta' + \dots + \sigma^{(0)} \delta^{(m-1)} .$$

### 1.4.7 Convergence (faible) dans l'espace $\mathcal{D}'$ des distributions

**Théorème et Définition 1.4.1.** Soit  $(T_n)_{n \in \mathbb{N}}$ <sup>1</sup> une suite de distributions. On dit que  $(T_n)_{n \in \mathbb{N}}$  converge dans  $\mathcal{D}'$  si, pour tout  $\varphi \in \mathcal{D}$ , la suite  $\langle T_n, \varphi \rangle$  converge au sens ordinaire. Si on appelle  $\langle T, \varphi \rangle = \lim_n \langle T_n, \varphi \rangle$  cette limite, alors l'application  $\varphi \mapsto \langle T, \varphi \rangle$  est une distribution.

**Théorème 1.4.2.** Soit  $(T_n)_{n \in \mathbb{N}}$  une suite de distributions. Si  $(T_n)_{n \in \mathbb{N}}$  converge dans  $\mathcal{D}'$  vers une distribution  $T$ , alors, pour tout  $m \in \mathbb{N}$ , la suite de distributions  $(T_n^{(m)})_{n \in \mathbb{N}}$  converge dans  $\mathcal{D}'$  vers  $T^{(m)}$ .

**Convergence vers  $\delta$ .** Si la suite de fonctions localement sommables  $(f_k)_k$  vérifie :

1.  $\exists A > 0$  tel que pour tout  $|x| \leq A$ ,  $f_k(x) \geq 0$ ,
2.  $\forall a > 0$ ,  $\int_{|x| \leq a} f_k(x) dx \rightarrow 1$  lorsque  $k \rightarrow +\infty$ ,
3.  $f_k(x) \rightarrow 0$  uniformément dans tout ensemble  $0 < a < |x| < \frac{1}{a} < \infty$  ( $a \in ]0, 1[$ ),

alors la suite des distributions régulières  $(T_{f_k})_k$  converge dans  $\mathcal{D}'$  vers  $\delta$ .

Une suite de fonctions localement sommables satisfaisant les conditions précédentes est souvent appelée *suite de fonctions de Dirac*. Un exemple de suite de fonctions de Dirac est la suite  $(g_n)_n$  des gaussiennes définies par  $g_n(x) = \frac{n}{\sqrt{\pi}} \exp(-n^2 x^2)$ .

### 1.4.8 Sous-espaces de $\mathcal{D}'$

Si on considère un espace de fonctions tests plus grand que  $\mathcal{D}$ , alors les fonctionnelles linéaires continues sur ce nouvel espace forment un sous-espace vectoriel de  $\mathcal{D}'$ . Deux espaces de fonctions tests couramment utilisés sont :

1. L'espace  $\mathcal{E}$  des fonctions indéfiniment dérivables quelconques,
2. L'espace  $\mathcal{S}$  des fonctions indéfiniment dérivables et qui décroissent, ainsi que leurs dérivées, plus vite que toute puissance de  $1/x$  à l'infini c'est-à-dire que pour tout  $k > 0$  et pour tout  $h > 0$ ,  $x^k f^{(h)}(x)$  est bornée.

On obtient ainsi l'espace  $\mathcal{E}'$  des distributions à support compact et l'espace  $\mathcal{S}'$  des distributions dites *tempérées* (ou à croissance lente).

$$\begin{array}{l} \text{Espaces de fonctions tests : } \mathcal{D} \subset \mathcal{S} \subset \mathcal{E} \\ \text{Espaces de distributions : } \mathcal{D}' \supset \mathcal{S}' \supset \mathcal{E}' \end{array}$$

---

<sup>1</sup>Attention : ici  $T_n$  désigne une distribution quelconque et non pas la distribution régulière associée à une certaine fonction  $n$ .

Dans la pratique, la plupart des distributions sont tempérées ( $\delta_a, \text{vp} \frac{1}{x}$ ). Cependant, en général, si  $f$  est une fonction localement sommable, alors  $Tf$  n'est pas tempérée.

**Théorème 1.4.3** (Caractérisation des distributions tempérées). *Pour qu'une forme linéaire continue  $T$  sur  $\mathcal{S}$  soit tempérée, il faut et il suffit qu'il existe  $A > 0$  et  $p \in \mathbb{N}^+$  tels que, pour tout  $\varphi \in \mathcal{S}$ , on ait :*

$$|\langle T, \varphi \rangle| \leq A \|\varphi\|_p,$$

où

$$\|\varphi\|_p = \left( \int |\varphi(t)|^p dt \right)^{\frac{1}{p}}.$$

## 1.5 Distributions à plusieurs dimensions

D'une façon analogue, on peut définir des *distributions à  $n$  dimensions* comme fonctionnelles sur l'espace  $\mathcal{D}(\mathbb{R}^n)$  des fonctions de  $\mathbb{R}^n$  dans  $\mathbb{C}$  indéfiniment dérivables sur  $\mathbb{R}^n$  et à support borné. Par exemple, la distribution régulière associée à une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{C}$  localement sommable est définie par :

$$\forall \varphi \in \mathcal{D}(\mathbb{R}^n), \quad \langle T_f, \varphi \rangle = \int \cdots \int f(x_1, \dots, x_n) \varphi(x_1, \dots, x_n) dx_1 \dots dx_n.$$

# Chapitre 2

## La Convolution

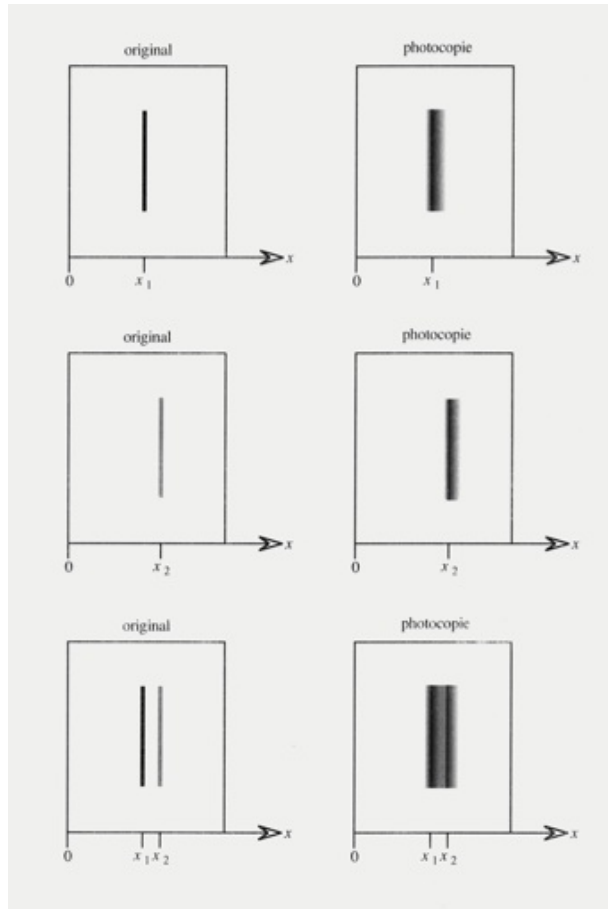
### Introduction

Le produit de convolution est un outil d'une grande importance en physique. On le rencontre, par exemple, à chaque fois que l'on étudie :

- La transmission d'un signal par un appareil,
- Une impulsion électrique fonction du temps,
- Une image représentée par une fonction d'une ou deux variables,
- En diffraction.

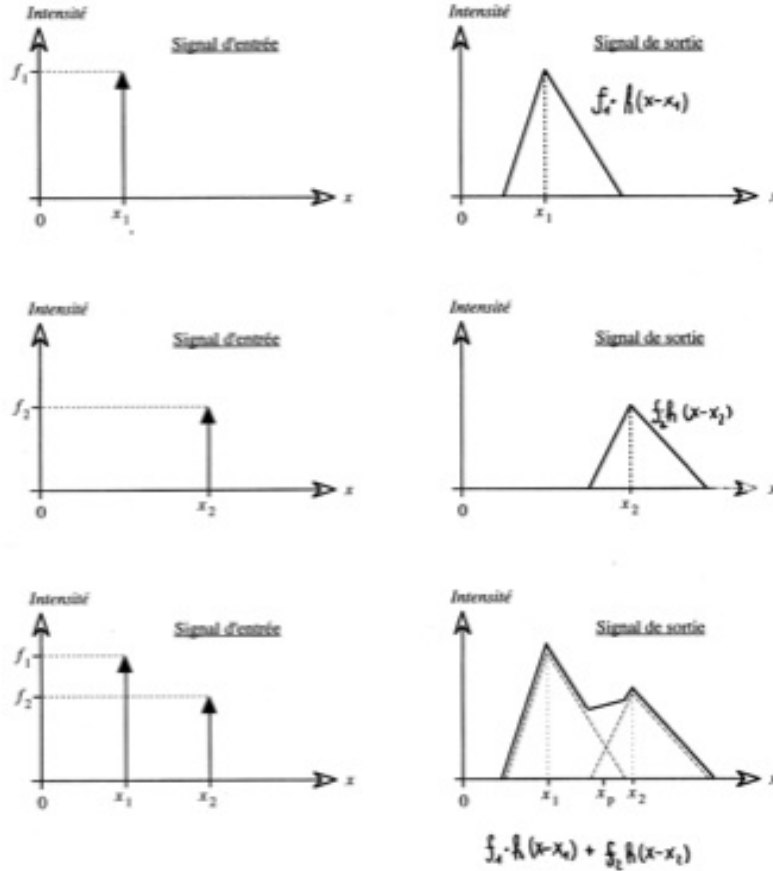
#### **Exemple introductif** : Le photocopieur

Considérons une machine photocopieuse imparfaitement réglée : un trait fin en position  $x_1$  donne sur la photocopie un trait étalé, centré en  $x_1$ . De même un trait d'une intensité moindre situé en  $x_2$  donnera sur la photocopie un trait étalé, centré en  $x_2$ . On admet que l'étalement de l'encre sur la photocopie est identique pour les deux traits. Cet étalement est donc une fonction caractéristique de l'appareil que nous appelons fonction d'étalement (ou fonction de réponse) et que nous noterons  $h(x)$ .



Pour un trait placé en  $x_1$  et d'intensité  $f_1$ , la photocopie donne donc  $f_1 h(x - x_1)$ . Si  $n$  traits placés en  $x_1, \dots, x_n$  et d'intensité respectives  $f_1, \dots, f_n$  sont présents sur l'original, la photocopie sera constituée de la superposition des  $n$  traits étalés. Le signal de sortie  $S(x)$  sera donc

$$S(x) = \sum_{i=1}^n f_i h(x - x_i).$$



Maintenant si le signal d'entrée est une fonction continue de  $x$  (à la place d'une fonction discrète), on remplace la somme par une intégrale et on obtient :

$$S(x) = \int f(u) h(x - u) du,$$

ce qui sera défini comme le produit de convolution des fonctions  $f$  et  $h$ .

## 2.1 Produit tensoriel

### 2.1.1 De deux fonctions

**Définition 2.1.1.** Soient  $f$  et  $g$  deux fonctions. On appelle produit tensoriel (ou produit direct) de  $f$  par  $g$  la fonction  $h : \mathbb{R}^2 \rightarrow \mathbb{R}$  définie par  $h(x, y) = f(x)g(y)$  pour tout  $(x, y)$  appartenant à  $\mathbb{R}^2$ . On note alors  $h = f \otimes g$ .

**Exemple :** Produit tensoriel de la fonction  $H$  de Heaviside par la fonction « porte »  $\Pi$  donnée par

$$\Pi(x) = \begin{cases} 1 & \text{si } x \in [-\frac{1}{2}, \frac{1}{2}], \\ 0 & \text{sinon.} \end{cases}$$

On a

$$(H \otimes \Pi)(x, y) = H(x) \Pi(y) = \begin{cases} 1 & \text{si } x \in [0, +\infty[ \text{ et } y \in [-\frac{1}{2}, \frac{1}{2}], \\ 0 & \text{sinon.} \end{cases}$$

## 2.1.2 De deux distributions

Cherchons maintenant à définir un produit tensoriel de deux distributions. Soient  $f$  et  $g$  deux fonctions localement sommables et soit  $h$  leur produit tensoriel. On a alors pour tout  $\varphi \in \mathcal{D}(\mathbb{R}^2)$  (ensemble des fonctions indéfiniment dérivables et à support borné sur  $\mathbb{R}^2$ ) :

$$\begin{aligned} \langle T_{f \otimes g}, \varphi \rangle &= \int \int f(x) g(y) \varphi(x, y) dx dy \\ &= \int f(x) \left( \int g(y) \varphi(x, y) dy \right) dx \\ &= \langle T_f(x), \langle T_g(y), \varphi(x, y) \rangle \rangle . \end{aligned}$$

**Définition 2.1.2.** Soient  $S$  et  $T$  deux distributions sur  $\mathbb{R}$ . On appelle produit tensoriel (ou produit direct) de  $S$  par  $T$  la distribution notée  $S \otimes T$  définie sur  $\mathcal{D}(\mathbb{R}^2)$  par

$$\langle S(x) \otimes T(y), \varphi(x, y) \rangle = \langle S(x), \langle T(y), \varphi(x, y) \rangle \rangle .$$

Il n'est pas du tout évident de montrer que ce produit est bien défini. En particulier, on doit vérifier que la fonction  $x \mapsto \langle T(y), \varphi(x, y) \rangle$  appartient bien à l'espace  $\mathcal{D}$  et que la fonctionnelle ainsi défini est linéaire et continue.

**Exemple :** Produit tensoriel des distributions  $\delta$  et  $W$  :  $\forall \varphi \in \mathcal{D}(\mathbb{R}^2)$ ,

$$\langle \delta(x) \otimes W(y), \varphi(x, y) \rangle = \langle \delta(x), \langle W(y), \varphi(x, y) \rangle \rangle = \langle \delta(x), \int_0^{+\infty} \varphi(x, y) dy \rangle = \int_0^{+\infty} \varphi(0, y) dy .$$

**Lemme 2.1.1.** Le produit tensoriel de deux distributions est commutatif, continu, associatif. De plus si  $S$  et  $T$  sont deux distributions sur  $\mathbb{R}$ , alors

$$\partial_x^i (S(x) \otimes T(y)) = (\partial_x^i S(x)) \otimes T(y), \quad \partial_y^i (S(x) \otimes T(y)) = S(x) \otimes (\partial_y^i T(y)),$$

où  $\partial_x$  (resp.  $\partial_y$ ) désigne la dérivée partielle par rapport à  $x$  (resp.  $y$ ) et  $i \in \mathbb{N}$ .

## 2.2 Produit de convolution

### 2.2.1 Convolution de deux fonctions

**Définition 2.2.1.** Soient  $f$  et  $g$  deux fonctions localement sommables. On définit, s'il existe, le produit de convolution  $h$  de  $f$  et  $g$  par

$$h(x) = \int f(t) g(x - t) dt,$$

pour tout  $x$  appartenant à  $\mathbb{R}$ . On note alors  $h = f * g$ .



**Remarque** : Ce produit de convolution n'existe pas toujours. Ce produit est commutatif dès lors qu'il est défini.

Donnons une interprétation graphique de ce produit de convolution. Soit  $k$  le produit tensoriel de  $f$  par  $g$ ;  $k = f \otimes g$ . On a alors

$$(f * g)(x) = \int f(t) g(x - t) dt = \int k(t, x - t) dt,$$

c'est-à-dire qu'on intègre la fonction  $k$  sur le chemin formé par la droite  $D_x$  de pente  $-1$  et passant par  $(0, x)$ .

**Exemple** : Posons  $f(x) = \Pi(x)$  et  $g(x) = \Pi(x - \frac{1}{2})$ . Notons  $k = f \otimes g$  et  $h = f * g$ . La valeur  $h(x)$  est donnée par l'intégrale de  $k$  sur la droite  $D_x$ .

Si la droite  $D_x$  rencontre le support de  $f \otimes g$  de manière finie, le produit de convolution est alors bien défini.

**Théorème 2.2.1.** *Le produit de convolution de deux fonctions localement sommables  $f$  et  $g$  existe dès lors que l'une des conditions suivantes est satisfaite :*

1. *Les fonctions  $f$  et  $g$  sont toutes les deux à support borné,*
2. *Les fonctions  $f$  et  $g$  sont toutes les deux à support borné à gauche,*
3. *Les fonctions  $f$  et  $g$  sont toutes les deux à support borné à droite.*

## 2.2.2 Notion de mesure floue

Soit  $\Pi$  la fonction porte. Soit  $a > 0$  et  $f$  une fonction. Notons  $f_{(a)}$  le produit de convolution de  $f$  et de la fonction porte de largeur  $a$  et de hauteur  $1/a$ . On a

$$f_{(a)}(x) = f(x) * \frac{1}{a} \Pi\left(\frac{x}{a}\right) = \frac{1}{a} \int_{x-a/2}^{x+a/2} f(t) dt.$$

Ceci représente la moyenne de  $f$  autour de  $x$  sur une largeur  $a$ . On peut ainsi modéliser une mesure imparfaite de la fonction  $f$ , pour laquelle on a un "flou" de largeur  $a$ . Les détails de largeur  $l \ll a$  disparaissent, les autres restent visibles.

## 2.2.3 Convolution de deux distributions

On cherche à étendre le produit de convolution de deux fonctions aux distributions. Soient  $f$  et  $g$  deux fonctions localement sommables et soit  $\varphi \in \mathcal{D}$ . On a

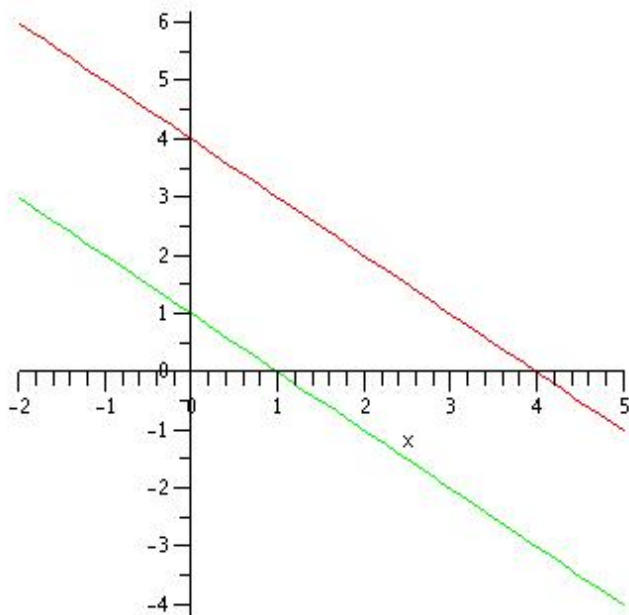
$$\begin{aligned} \langle T_{f * g}, \varphi \rangle &= \int (f * g)(t) \varphi(t) dt \\ &= \int \varphi(t) \left( \int f(s) g(t - s) ds \right) dt \\ &= \int \int f(x) g(y) \varphi(x + y) dx dy \\ &= \langle T_{f \otimes g}, \varphi(x + y) \rangle \end{aligned}$$

**Définition 2.2.2.** Soient  $S$  et  $T$  deux distributions. On définit le produit de convolution de  $S$  et  $T$  comme la distribution notée  $S * T$  définie par

$$\forall \varphi \in \mathcal{D}, \quad \langle S * T, \varphi \rangle = \langle S \otimes T, \varphi(x + y) \rangle .$$

Le produit de convolution de deux distributions appliqué à une fonction test  $\varphi$  de  $\mathcal{D}$  est donc égal au produit tensoriel des deux distributions appliqué à  $\varphi(x + y)$ .

Le produit de convolution de deux distributions n'existe pas toujours. En effet, si la fonction  $\varphi$  a pour support le segment  $[a, b]$ , alors le support de  $\varphi(x + y)$  est la bande comprise entre les deux droites  $x + y = a$  et  $x + y = b$ .



Puisque  $\varphi(x + y)$  n'est pas à support compact, il suffit que l'une des deux distributions soit à support compact.

D'une manière générale, on a le résultat suivant :

**Théorème 2.2.2.** 1. Les distributions à support borné à gauche (resp. à droite) peuvent toujours être convoluées entre elles.

2. Une distribution à support borné peut être convoluée avec n'importe quelle autre distribution.

## 2.2.4 Propriétés du produit de convolution de deux distributions

### Commutativité et Associativité

1. Soient  $S$  et  $T$  deux distributions. On suppose que leur produit de convolution  $S * T$  existe. Alors  $T * S$  existe et  $S * T = T * S$ .
2. Soient  $S, T$  et  $U$  trois distributions. Si  $S * T, T * U$  et  $S * U$  existent, alors

$$S * (T * U) = (S * T) * U = S * T * U.$$

**Exemple :**  $(T_{\mathbb{1}} * \delta') * W \neq T_{\mathbb{1}} * (\delta' * W)$ . On peut vérifier que  $T_{\mathbb{1}} * \delta' = 0$ . En effet,  $T_{\mathbb{1}} * \delta' = T'_{\mathbb{1}}$  et

$$\forall \varphi \in \mathcal{D}, \langle T'_{\mathbb{1}}, \varphi \rangle = - \langle T_{\mathbb{1}}, \varphi' \rangle = - \int_{-\infty}^{+\infty} \varphi'(x) dx = -[\varphi]_{-\infty}^{+\infty} = 0,$$

et donc  $(T_{\mathbb{1}} * \delta') * W = 0$ . Or, d'un autre côté, on a  $\delta' * W = \delta$  donc  $T_{\mathbb{1}} * (\delta' * W) = T_{\mathbb{1}}$ . Ceci s'explique par le fait que  $T_{\mathbb{1}} * W$  n'est pas définie.

### Convolution et Dirac

1. Soit  $T$  une distribution. On a  $\delta * T = T * \delta = T$ . La distribution de Dirac est donc l'élément neutre du produit de convolution.
2. Soit  $a > 0$ . La translatée  $T_a$  d'une distribution  $T$  est égale au produit de convolution de  $T$  par la distribution de Dirac  $\delta_a$  au point  $a$  :  $T_a = T * \delta_a = \delta_a * T$ .
3. Les dérivations d'une distribution  $T$  s'obtiennent par convolution par les dérivées des distributions de Dirac : pour tout  $m \in \mathbb{N}$ ,  $T^{(m)} = \delta^{(m)} * T = T * \delta^{(m)}$ .

### Dérivée d'un produit de convolution

D'après le point 3. ci-dessus, on a  $(S * T)' = \delta' * (S * T) = \delta' * S * T = S' * T$ . De même  $(S * T)' = (S * T) * \delta' = S * T * \delta' = S * T'$ . On obtient donc le résultat suivant :

**Lemme 2.2.1.** *Pour dériver un produit de convolution, il suffit de dériver l'un des facteurs, i.e., pour tout  $S$  et  $T$  dans  $\mathcal{D}'$ , on a :*

$$(S * T)' = S' * T = S * T'.$$

## 2.3 Algèbre de convolution et résolution d'équations différentielles

### 2.3.1 Définition

**Définition 2.3.1.** On appelle algèbre de convolution tout espace vectoriel de distributions contenant  $\delta$  et sur lequel on peut définir le produit de convolution d'un nombre fini quelconque de distributions.

**Exemples :** les espaces  $\mathcal{D}'_+$  (resp.  $\mathcal{D}'_-$ ) des distributions à support contenu dans  $\{x \geq 0\}$  (resp.  $\{x \leq 0\}$ ) aussi appelé espace des distributions à support borné à droite (resp. borné à gauche), l'espace  $\mathcal{E}'$  des distributions à support compact.

### 2.3.2 Calcul algébrique

Ces algèbres permettent de résoudre des équations du type

$$A * X = B,$$

où  $A$  et  $B$  sont des distributions connues et  $X$  une distribution inconnue. On est amené à chercher s'il existe, dans l'algèbre de convolution considérée, une distribution notée  $A^{*-1}$  et appelée *inverse de convolution de  $A$*  (ou parfois *fonction de Green de  $A$* ) telle que

$$A * A^{*-1} = A^{*-1} * A = \delta.$$

Si  $A^{*-1}$  existe, alors il est unique et la solution de l'équation  $A * X = B$  dans l'algèbre de convolution considérée est donnée par

$$X = A^{*-1} * B.$$

**Exemple :** L'oscillateur harmonique

Un oscillateur harmonique est caractérisé par une équation de convolution du type

$$(\delta'' + \omega^2 \delta) * X(t) = B(t),$$

où  $B(t)$  est une distribution caractérisant le signal extérieur. On peut montrer (voir TD 2) que

$$(\delta'' + \omega^2 \delta) * \frac{1}{\omega} W(t) \sin(\omega t) = \delta.$$

Or la distribution  $\frac{1}{\omega} W(t) \sin(\omega t)$  appartient à  $\mathcal{D}'^+$  donc pour tout  $B \in \mathcal{D}'^+$ , l'équation admet l'unique solution

$$X(t) = \frac{1}{\omega} W(t) \sin(\omega t) * B(t).$$

On montre aussi que

$$(\delta'' + \omega^2 \delta) * -\frac{1}{\omega} W(-t) \sin(\omega t) = \delta.$$

donc si  $B \in \mathcal{D}'^-$ , l'équation admet l'unique solution

$$X(t) = -\frac{1}{\omega} W(-t) \sin(\omega t) * B(t).$$

On voit donc que l'inverse de convolution et donc la solution de l'équation dépend de l'algèbre de convolution dans laquelle on travaille. Ici l'équation n'admettra pas de solutions dans  $\mathcal{E}'$ .

### 2.3.3 Résolution d'une équation différentielle avec conditions initiales

Problème de Cauchy = Équation différentielle + **Conditions initiales**.

Par exemple pour le premier ordre, on considère une équation différentielle

$$\dot{u} + \alpha u = 0, \tag{2.1}$$

et on cherche une solution  $u : \mathbb{R}^+ \rightarrow \mathbb{R}$  telle que  $u(0) = u_0$ . Il est possible d'utiliser les distributions pour résoudre ce genre d'équation en cherchant non plus une fonction solution mais une distribution solution de la forme  $U = W(t) u(t)$ . Si  $u(t)$  est solution (2.1), alors  $\dot{U} = W \dot{u} + u_0 \delta$  et donc  $U$  vérifie l'équation différentielle  $\dot{U} + \alpha U = u_0 \delta^1$  qui se réécrit sous la forme d'une équation de convolution :

$$(\delta' + \alpha \delta) * U = u_0 \delta.$$

Il ne nous reste donc plus qu'à trouver dans  $\mathcal{D}'^+$  l'inverse de convolution de  $\delta' + \alpha \delta$ . On peut vérifier que ce dernier est donné par  $W(t) \exp(-\alpha t)$ . D'où la solution

$$U = W(t) \exp(-\alpha t) * u_0 \delta = W(t) u_0 \exp(-\alpha t).$$

#### Autre exemple : Équation de la chaleur

L'équation de la chaleur (ou équation de diffusion) à une dimension est donnée par

$$\left( \frac{\partial}{\partial t} - \alpha \frac{\partial^2}{\partial x^2} \right) u(x, t) = 0,$$

où  $u(x, t)$  représente par exemple la température d'une barre au point  $x$  et au temps  $t$ . On cherche une solution correspondant à une répartition initiale de température  $u(x, 0)$  donnée. On cherche donc une distribution  $U(x, t) = W(t) u(x, t)$  vérifiant l'équation différentielle. La dérivation au sens des distributions conduit à

$$\left( \frac{\partial}{\partial t} - \alpha \frac{\partial^2}{\partial x^2} \right) U(x, t) = u(x, 0) \delta(t).$$

---

<sup>1</sup> $\dot{U} + \alpha U = W \dot{u} + u_0 \delta + \alpha (W u) = W (\dot{u} + \alpha u) + u_0 \delta = u_0 \delta.$

En effet, on a

$$\begin{aligned}
 \left( \frac{\partial}{\partial t} - \alpha \frac{\partial^2}{\partial x^2} \right) U(x, t) &= \frac{\partial}{\partial t} (W(t) u(x, t)) - \alpha \frac{\partial^2}{\partial x^2} (W(t) u(x, t)) \\
 &= \delta(t) u(x, t) + W(t) \frac{\partial}{\partial t} (u(x, t)) - \alpha W(t) \frac{\partial^2}{\partial x^2} (u(x, t)) \\
 &= \delta(t) u(x, 0) + W(t) \left( \frac{\partial}{\partial t} (u(x, t)) - \alpha \frac{\partial^2}{\partial x^2} (u(x, t)) \right) \\
 &= \delta(t) u(x, 0).
 \end{aligned}$$

En écrivant cette équation comme une équation de convolution, on obtient alors :

$$\left( \frac{\partial}{\partial t} \delta(x, t) - \alpha \frac{\partial^2}{\partial x^2} \delta(x, t) \right) * U(x, t) = u(x, 0) \delta(t).$$

On peut alors vérifier que la fonction de Green (l'inverse de convolution) de  $\frac{\partial}{\partial t} \delta(x, t) - \alpha \frac{\partial^2}{\partial x^2} \delta(x, t)$  est donnée par

$$\frac{W(t)}{2\sqrt{\alpha\pi t}} \exp\left(-\frac{x^2}{4\alpha t}\right),$$

de sorte que la solution cherchée s'écrit finalement

$$U(x, t) = u(x, 0) \delta(t) * \frac{W(t)}{2\sqrt{\alpha\pi t}} \exp\left(-\frac{x^2}{4\alpha t}\right),$$

ou encore

$$U(x, t) = \frac{W(t)}{2\sqrt{\alpha\pi t}} \int_{-\infty}^{\infty} u(\xi, 0) \exp\left(-\frac{(x-\xi)^2}{4\alpha t}\right) d\xi.$$

## 2.4 Interprétation physique de la convolution

### 2.4.1 Systèmes décrits par un opérateur de convolution

Nous allons maintenant voir que beaucoup de systèmes physiques (*e.g.*, systèmes de mesures) peuvent être représentés par des opérateurs de convolution.

**Définition 2.4.1.** Soit ( $S$ ) un système physique décrit par un opérateur qui à un signal d'entrée  $E(t)$  (ou excitation) fait correspondre un signal de sortie (ou réponse)  $S(t)$ . Soit  $R$  l'opérateur tel que  $S(t) = R(E(t))$ . Le système est dit linéaire si  $R$  est linéaire. Il est dit continu si  $R$  est continu c'est-à-dire si des excitations peu différentes conduisent à des réponses peu différentes. Un système est dit invariant par translation si l'opérateur  $R$  commute avec la translation c'est-à-dire si le diagramme suivant commute :

$$\begin{array}{ccc}
 E(t) & \longrightarrow & S(t) = R(E(t)) \\
 \downarrow & & \downarrow \\
 E(t-a) & \longrightarrow & S(t-a) = R(E(t-a))
 \end{array}$$

La plupart des systèmes physiques sont invariants par translation dans le temps c'est-à-dire que si l'on retarde l'entrée de  $\tau$ , alors la sortie est aussi retardée de  $\tau$ .

**Définition 2.4.2.** *On dit qu'un système physique est décrit par un opérateur de convolution s'il existe une distribution  $T$  caractéristique du système telle que  $R(E(t)) = E(t) * T(t)$ .*

**Définition 2.4.3.** *La distribution  $T$  telle que  $R(E(t)) = E(t) * T(t)$  est appelée réponse impulsionnelle ou percussionnelle car elle correspond à la réponse d'une excitation élémentaire  $E = \delta$ .*

**Théorème 2.4.1.** *Un système physique peut être décrit par un opérateur de convolution si et seulement si il est linéaire, continu et invariant par translation.*

**Exemples :** Les filtres linéaires en électronique (systèmes formés de résistances, selfs, capacités, amplificateurs sans saturation). Les systèmes formés de masses, de ressorts et d'amortisseurs en mécanique.

## 2.4.2 Système causal

**Définition 2.4.4.** *Un système décrit par un opérateur de distribution temporel (c'est-à-dire à une variable qui représente le temps) est dit causal si sa réponse impulsionnelle  $T(t)$  est nulle pour  $t < 0$  (c'est-à-dire que  $\langle T, \varphi \rangle = 0$  dès que  $\varphi$  est nulle pour  $t \geq 0$ ).*

Dans le cas d'un système causal, l'effet d'un signal ne peut précéder sa cause c'est-à-dire que la réponse  $S(\tau)$  d'un tel système à un temps  $\tau$  ne dépend que des valeurs du signal  $E(t)$  pour  $t \leq \tau$ .

## 2.4.3 Réponse à une excitation exponentielle

**Théorème et Définition 2.4.1** (Régularisation d'une distribution singulière). *Soit  $\Psi$  une fonction indéfiniment dérivable et  $T$  une distribution. On note  $\Psi * T$  l'application*

$$\Psi * T : x \mapsto \langle T(t), \Psi(x - t) \rangle .$$

*Cette application est indéfiniment dérivable et vérifie  $T_\Psi * T = T_{\Psi * T}$ . On l'appelle la régularisée de  $T$  par  $\Psi$ .*

*Démonstration.*

$$\begin{aligned} \forall \varphi \in \mathcal{D}(\mathbb{R}^2), \langle T_\Psi * T, \varphi \rangle &= \langle T(t) \otimes T_\psi(x), \varphi(x + t) \rangle \\ &= \langle T(t), \langle T_\Psi(x), \varphi(x + t) \rangle \rangle \\ &= \langle T(t), \int \Psi(x) \varphi(x + t) dx \rangle \\ &= \langle T(t), \int \Psi(x - t) \varphi(x) dx \rangle \\ &= \langle T(t), \langle T_\varphi(x), \Psi(x - t) \rangle \rangle \\ &= \langle T(t) \otimes T_\varphi(x), \Psi(x - t) \rangle \\ &= \langle T_\varphi(x), \langle T(t), \Psi(x - t) \rangle \rangle \\ &= \int \varphi(x) \langle T(t), \Psi(x - t) \rangle dx \\ &= \langle T_{\Psi * T}, \varphi \rangle . \end{aligned}$$

□

La réponse impulsionnelle d'un système décrit par un opérateur de convolution joue un rôle important. Un autre type de réponse joue un rôle important : celle aux signaux exponentiels. Considérons donc un système décrit par un opérateur de convolution de réponse impulsionnelle  $T$  c'est-à-dire

$$S = E * T,$$

et examinons la réponse à un signal

$$E(t) = \exp(2i\pi\nu t),$$

correspondant à une oscillation harmonique de fréquence  $\nu$ . La réponse sera alors

$$S = T * \exp(2i\pi\nu t),$$

et comme la fonction exponentielle est indéfiniment dérivable, il vient

$$S(t) = \langle T(\tau), \exp(2i\pi\nu(t-\tau)) \rangle = \langle T(\tau), \exp(-2i\pi\nu\tau) \rangle \exp(2i\pi\nu t).$$

En posant  $\hat{T}(\nu) = \langle T(\tau), \exp(-2i\pi\nu\tau) \rangle$ , on obtient

$$S(t) = \hat{T}(\nu) \exp(2i\pi\nu t).$$

La fonction  $\hat{T}(\nu)$  s'appelle la *transformée de Fourier* de la distribution  $T$ . Le système transforme donc l'excitation  $\exp(2i\pi\nu t)$  en  $\hat{T}(\nu) \exp(2i\pi\nu t)$ .

Plus généralement la réponse à une excitation  $E(t) = \exp(pt)$  où  $p$  est un scalaire complexe quelconque est

$$S(t) = \langle T(\tau), \exp(-p\tau) \rangle \exp(pt).$$

La fonction de  $p$  donnée par  $\langle T(\tau), \exp(-p\tau) \rangle$  est appelée *transformée de Laplace* de la distribution  $T$ .

**Théorème 2.4.2.** *Les fonctions exponentielles sont des fonctions propres pour les opérateurs de convolution. Les valeurs propres correspondantes sont données par les transformées de Fourier (ou de Laplace) de la réponse impulsionnelle.*

On voit ici le rôle important que jouent les transformées de Fourier et de Laplace que l'on va étudier dans les chapitres suivants. Un moyen de calculer  $T * E$  sera de décomposer  $E$  en combinaison linéaires de fonctions propres

$$E(t) = \int \hat{E}(\nu) \exp(2i\pi\nu t) d\nu.$$

On verra que  $\hat{E}(\nu)$  est justement la transformée de Fourier de  $E$ . En effet, par linéarité, la réponse  $S(t)$  sera alors donnée par

$$S(t) = \int \hat{E}(\nu) \hat{T}(\nu) \exp(2i\pi\nu t) d\nu.$$

La transformée de Fourier  $\hat{T}(\nu)$  est appelée *fonction de transfert* du système.



# Chapitre 3

## La Transformation de Fourier

### 3.1 Transformée de Fourier des fonctions

#### 3.1.1 Définition et existence

**Définition 3.1.1.** Soit  $f : \mathbb{R} \rightarrow \mathbb{R}$  ou  $\mathbb{C}$  une fonction de la variable réelle à valeurs réelles ou complexes. On appelle transformée de Fourier (ou spectre) de  $f$ , si elle existe, la fonction  $\hat{f} : \mathbb{R} \rightarrow \mathbb{C}$  définie par

$$\hat{f}(\nu) = \int_{-\infty}^{+\infty} f(x) \exp(-2i\pi\nu x) dx.$$

On écrira symboliquement

$$\hat{f} = \mathcal{F}[f] \quad \text{ou} \quad \hat{f}(\nu) = \mathcal{F}[f(x)].$$

L'intégrale et donc la transformée de Fourier n'existe pas toujours, par exemple la fonction  $x \mapsto x^2$  n'admet pas de transformée de Fourier car l'intégrale

$$\int_{-\infty}^{+\infty} x^2 \exp(-2i\pi\nu x) dx$$

n'existe pour aucune valeur de  $\nu$ .

Si des conditions d'existence de la transformée de Fourier d'une fonction sont difficiles à écrire, on a en revanche la condition suffisante suivante :

**Théorème 3.1.1.** Toute fonction intégrable possède une transformée de Fourier qui est une fonction continue, bornée et tendant vers 0 lorsque  $|\nu|$  tend vers l'infini.

**Exemples :**

1.  $\mathcal{F}[\Pi(x)] = \frac{\sin(\pi\nu)}{\pi\nu}$  ( $= 1$  si  $\nu = 0$ );
2.  $\mathcal{F}[\exp(-\pi x^2)] = \exp(-\pi\nu^2)$ ;
3.  $\mathcal{F}[\exp(-a|x|)] = \frac{2a}{a^2 + 4\pi^2\nu^2}$ .

### 3.1.2 Inversion

Soit  $f$  une fonction intégrable admettant une transformée de Fourier  $\hat{f}$  elle-même intégrable. Alors, en tout point  $x$  où  $f$  est continue, on a :

$$f(x) = \int_{-\infty}^{+\infty} \hat{f}(\nu) \exp(2i\pi\nu x) d\nu.$$

Cette transformation est appelée *transformée de Fourier inverse*. On écrira symboliquement  $f(x) = \mathcal{F}^{-1}[\hat{f}(\nu)] = \mathcal{F}^{-1}[\hat{f}(\nu)]$  si  $f$  est continue en  $x$  et de manière générale, si  $f$  est continue :  $f = \mathcal{F}[\hat{f}] = \mathcal{F}^{-1}[f]$ .

Si  $f$  est continue par morceaux, on peut donc obtenir  $f(x)$  à partir de  $\hat{f}(\nu)$  presque partout. Si  $f$  n'est pas continue en  $x$ , on a plus généralement :

$$\int_{-\infty}^{+\infty} \hat{f}(\nu) \exp(2i\pi\nu x) d\nu = \frac{1}{2} (f(x^+) + f(x^-)).$$

### 3.1.3 Transformée de Fourier en sinus et cosinus

Soit  $f$  une fonction de la variable réelle à valeurs réelles ou complexes. Il est connu que  $f$  peut se décomposer en somme d'une fonction paire  $p$  et d'une fonction impaire  $q$  :

$$\forall x \in \mathbb{R}, f(x) = p(x) + q(x),$$

avec

$$p(x) = \frac{1}{2} (f(x) + f(-x)), \quad q(x) = \frac{1}{2} (f(x) - f(-x)).$$

On a alors

$$\hat{f}(\nu) = \int_{-\infty}^{+\infty} (p(x) + q(x)) (\cos(2\pi\nu x) - i \sin(2\pi\nu x)) dx,$$

d'où

$$\hat{f}(\nu) = 2 \int_0^{+\infty} p(x) \cos(2\pi\nu x) dx - 2i \int_0^{+\infty} q(x) \sin(2\pi\nu x) dx.$$

On écrit alors

$$\mathcal{F}[f(x)] = \mathcal{F}_{\cos}[p(x)] - i \mathcal{F}_{\sin}[q(x)],$$

où  $\mathcal{F}_{\cos}$  et  $\mathcal{F}_{\sin}$  sont les *transformées de Fourier* respectivement *en cosinus et sinus* définies par :

$$\mathcal{F}_{\cos}[f(x)] = 2 \int_0^{+\infty} f(x) \cos(2\pi\nu x) dx, \quad \mathcal{F}_{\sin}[f(x)] = 2 \int_0^{+\infty} f(x) \sin(2\pi\nu x) dx.$$

Lorsque la fonction  $f$  est à valeurs complexes, il faut décomposer  $p$  et  $q$  en parties réelles et imaginaires. On obtient alors la correspondance :

$$\begin{array}{rcccccccc}
 f(x) & = & \text{partie réelle paire} & + & \text{imag. paire} & + & \text{réelle imp.} & + & \text{imag. imp.} \\
 & & \downarrow & & \downarrow & & & & \searrow \swarrow \\
 \hat{f}(\nu) & = & \text{partie réelle paire} & + & \text{imag. paire} & + & \text{réelle imp.} & + & \text{imag. imp.}
 \end{array}$$

Finalement, on a

$f(x)$	$\longrightarrow$	$\hat{f}(\nu)$
paire	$\longrightarrow$	paire
impaire	$\longrightarrow$	impaire
réelle	$\longrightarrow$	hermitienne ( $\hat{f}(\nu) = \overline{\hat{f}(-\nu)}$ )
imaginaire	$\longrightarrow$	antihermitienne ( $\hat{f}(\nu) = -\overline{\hat{f}(-\nu)}$ )
réelle paire	$\longrightarrow$	réelle paire
réelle impaire	$\longrightarrow$	imaginaire impaire
imaginaire paire	$\longrightarrow$	imaginaire paire
imaginaire impaire	$\longrightarrow$	réelle impaire

### 3.1.4 Propriétés

**Linéarité :**

$$\begin{aligned}
 \mathcal{F}[\lambda f(x) + \mu g(x)] &= \int (\lambda f(x) + \mu g(x)) \exp(-2i\pi\nu x) dx \\
 &= \lambda \int f(x) \exp(-2i\pi\nu x) dx + \mu \int g(x) \exp(-2i\pi\nu x) dx \\
 &= \lambda \mathcal{F}[f(x)] + \mu \mathcal{F}[g(x)]
 \end{aligned}$$

**Transposition :**

$$\begin{aligned}
 \mathcal{F}[f(-x)] &= \int f(-x) \exp(-2i\pi\nu x) dx \\
 &= \int f(y) \exp(2i\pi\nu y) dy \\
 &= \hat{f}(-\nu)
 \end{aligned}$$

**Conjugaison :**

$$\begin{aligned}
 \mathcal{F}[\overline{f(x)}] &= \int \overline{f(x)} \exp(-2i\pi\nu x) dx \\
 &= \overline{\int f(x) \exp(2i\pi\nu x) dx} \\
 &= \overline{\hat{f}(-\nu)}
 \end{aligned}$$

**Changement d'échelle :**

$$\begin{aligned}
 \mathcal{F}[f(ax)] &= \int f(ax) \exp(-2i\pi\nu x) dx \\
 &= \int f(y) \exp\left(\frac{-2i\pi\nu y}{a}\right) \frac{1}{|a|} dy \\
 &= \frac{1}{|a|} \hat{f}\left(\frac{\nu}{a}\right)
 \end{aligned}$$

En d'autres termes, une dilatation dans le monde réel entraîne une compression dans le monde de Fourier et inversement.

**Translation :**

$$\begin{aligned}
 \mathcal{F}[f(x-a)] &= \int f(x-a) \exp(-2i\pi\nu x) dx \\
 &= \int f(y) \exp(-2i\pi\nu(y+a)) dy \\
 &= \exp(-2i\pi\nu a) \int f(y) \exp(-2i\pi\nu y) dy \\
 &= \exp(-2i\pi\nu a) \hat{f}(\nu).
 \end{aligned}$$

En d'autres termes, une translation dans le monde réel correspond à un déphasage (proportionnel à la fréquence  $\nu$ ) dans le monde de Fourier.

**Modulation :**

$$\begin{aligned}
 \mathcal{F}[\exp(2i\pi\nu_0 x) f(x)] &= \int \exp(2i\pi\nu_0 x) f(x) \exp(-2i\pi\nu x) dx \\
 &= \int f(x) \exp(-2i\pi(\nu - \nu_0) x) dx \\
 &= \hat{f}(\nu - \nu_0).
 \end{aligned}$$

Moduler la fonction  $f$  par une exponentielle imaginaire revient à traduire sa transformée de Fourier.

### 3.1.5 Dérivation

**Par rapport à  $x$  :**

Supposons  $f$  sommable, dérivable et à dérivée sommable. Par intégration par partie, il vient alors

$$\begin{aligned}
 \mathcal{F}[f'(x)] &= \int f'(x) \exp(-2i\pi\nu x) dx \\
 &= [f(x) \exp(-2i\pi\nu x)]_{-\infty}^{+\infty} + 2i\pi\nu \int f(x) \exp(-2i\pi\nu x) dx \\
 &= 2i\pi\nu \hat{f}(\nu).
 \end{aligned}$$

Plus généralement, on obtient

$$\mathcal{F}[f^{(m)}(x)] = (2i\pi\nu)^m \hat{f}(\nu).$$

De cette formule on tire (en prenant les modules)

$$|2\pi\nu|^m |\hat{f}(\nu)| \leq \int |f^{(m)}(x)| dx,$$

et on conclut que plus  $f$  est dérivable, à dérivées sommables, plus  $\hat{f}$  décroît rapidement à l'infini. En effet, si  $f$  est  $m$  fois dérivable et à dérivée  $m$ -ième sommable,  $\hat{f}$  décroît au moins en  $1/\nu^m$ .

**Par rapport à  $\nu$  :**

$$\begin{aligned}
 \frac{\partial}{\partial\nu} \hat{f}(\nu) &= \int f(x) \frac{\partial}{\partial\nu} \exp(-2i\pi\nu x) dx \\
 &= \int (-2i\pi x) f(x) \exp(-2i\pi\nu x) dx \\
 &= \mathcal{F}[(-2i\pi x) f(x)].
 \end{aligned}$$

De manière générale, on obtient

$$\hat{f}^{(m)}(\nu) = \mathcal{F}[(-2i\pi x)^m f(x)].$$

Ce résultat conduit aussi à une majoration

$$|\hat{f}^{(m)}(\nu)| \leq \int |2\pi x|^m |f(x)| dx.$$

et donc : plus  $f$  décroît à l'infini, plus  $\hat{f}$  est dérivable (avec ses dérivées bornées). En effet, si  $f$  décroît en  $1/x^m$  à l'infini, alors,  $\hat{f}$  est  $m$  fois dérivable (car  $|2\pi x|^m |f(x)|$  est alors sommable) et sa dérivée  $m$ -ième est bornée.

### 3.1.6 Transformée de Fourier et convolution

Supposons  $f$  et  $g$  sommables telles que  $f * g$  existe. On a alors

$$\mathcal{F}[f * g] = \int \exp(-2i\pi\nu x) \int f(t) g(x-t) dx dt.$$

Le théorème de Fubini donne alors

$$\mathcal{F}[f * g] = \int f(t) dt \int g(x-t) \exp(-2i\pi\nu x) dx,$$

et après le changement de variable  $y = x - t$  dans la seconde intégrale, il vient

$$\mathcal{F}[f * g] = \int f(t) \exp(-2i\pi\nu t) dt \int g(y) \exp(-2i\pi\nu y) dy,$$

d'où

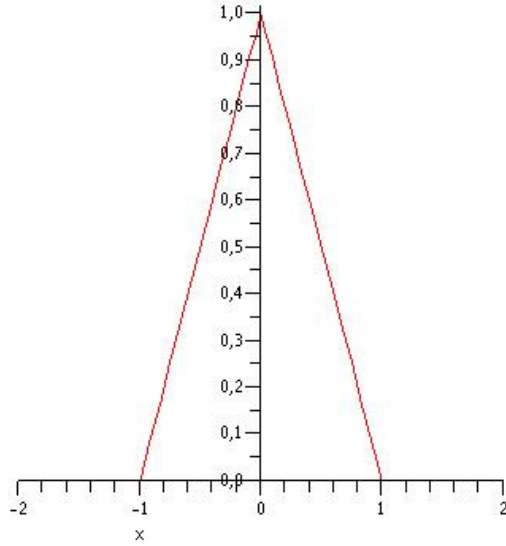
$$\mathcal{F}[f * g] = \mathcal{F}[f] \mathcal{F}[g].$$

**Théorème 3.1.2.** *La transformée de Fourier du produit de convolution de deux fonctions est le produit ordinaire des transformées de Fourier des deux fonctions.*

**Exemple :** On se propose de calculer la transformée de Fourier de la fonction  $\Lambda$  définie par :

$$\Lambda(x) = \begin{cases} 1+x & \text{pour } -1 \leq x \leq 0 \\ 1-x & \text{pour } 0 \leq x \leq 1 \\ 0 & \text{pour } |x| \geq 1, \end{cases}$$

ayant le graphe suivant :



On commence par montrer que  $\Lambda(x) = (\Pi * \Pi)(x)$  et on en déduit que

$$\mathcal{F}[\Lambda(x)] = \mathcal{F}[\Pi(x)]^2 = \left( \frac{\sin(\pi \nu)}{\pi \nu} \right)^2.$$

Inversement, on montre que l'on a le résultat suivant :

**Théorème 3.1.3.** *Lorsque ces expressions sont définies, on a*

$$\mathcal{F}[f g] = \mathcal{F}[f] * \mathcal{F}[g].$$

### 3.1.7 Formule de Parseval-Plancherel

La relation suivante a été établie par Parseval et généralisée par Plancherel aux transformées de Fourier :

**Théorème 3.1.4.** *Soient  $f$  et  $g$  deux fonctions de carré sommable. On a alors :*

$$\int f(x) \overline{g(x)} dx = \int \hat{f}(\nu) \overline{\hat{g}(\nu)} d\nu.$$

*Un cas particulier important est le cas  $f = g$  c'est-à-dire*

$$\int |f(x)|^2 dx = \int |\hat{f}(\nu)|^2 d\nu.$$

*Démonstration.* On a

$$\int f(x) \overline{g(x)} dx = \mathcal{F}[f \bar{g}]|_{\nu=0} = [\hat{f}(\nu) * \overline{\hat{g}(-\nu)}]|_{\nu=0} = \left[ \int \hat{f}(t) \overline{\hat{g}(t - \nu)} dt \right]_{\nu=0} = \int \hat{f}(t) \overline{\hat{g}(t)} dt.$$

□

En physique, si  $f$  est une onde ou une vibration et si la variable  $x$  est temporelle ( $x = t$ ), alors  $\int |f(x)|^2 dx$  peut représenter la puissance (ou l'énergie) totale dans le domaine temporel et  $\int |\hat{f}(\nu)|^2 d\nu$  représente la puissance totale dans le domaine fréquentiel.

**Remarque :** Transformée de Fourier des fonctions de carré sommable

Il existe des fonctions non intégrables mais dont le carré l'est (par exemple la fonction sinus cardinal  $x \mapsto \text{sinc}(x) = \frac{\sin(x)}{x}$  prolongée par continuité en 0 par  $\text{sinc}(0) = 1$ ). De telles fonctions apparaissent fréquemment en physique (fonction d'onde d'une particule en mécanique quantique, en électricité ou traitement du signal où  $\int |f(t)|^2 dt$  représente l'énergie totale d'un signal temporel  $t \mapsto f(t)$ ). Il existe un moyen (que nous ne traiterons pas ici) d'étendre la transformée de Fourier aux fonctions non sommables mais de carré sommable.

## 3.2 Transformée de Fourier des distributions

Nous allons maintenant définir une notion de transformée de Fourier pour les distributions. L'intérêt est de :

1. Pouvoir définir la transformée de Fourier des distributions comme  $\delta, \text{III}, \dots$ ,
2. Espérer pouvoir étendre la transformée de Fourier des fonctions sommables (et de carré sommable) à des fonctions intervenant tout le temps en physique et n'étant ni sommables ni de carré sommable comme  $H$ .

### 3.2.1 Définition

Comme d'habitude on va tout d'abord essayer de définir la transformée de Fourier pour les distributions régulières. Soit donc  $f$  une fonction intégrable qui définit une distribution régulière  $T_f$  et intéressons nous à la distribution régulière associée à  $\hat{f}$  : on a :

$$\forall \varphi \in \mathcal{D}, \langle T_{\hat{f}}, \varphi \rangle = \int \hat{f}(t) \varphi(t) dt = \int \left( \int f(x) \exp(-2i\pi xt) dx \right) \varphi(t) dt,$$

et en utilisant le théorème de Fubini pour intervertir les deux intégrales :

$$\forall \varphi \in \mathcal{D}, \langle T_{\hat{f}}, \varphi \rangle = \int f(x) \left( \int \exp(-2i\pi xt) \varphi(t) dt \right) dx = \int f(x) \hat{\varphi}(x) dx = \langle T_f, \hat{\varphi} \rangle .$$

Le problème ici est que si  $\varphi$  appartient à  $\mathcal{D}$ , il n'y a aucune raison pour que sa transformée de Fourier  $\hat{\varphi}$  appartienne à  $\mathcal{D}$  et donc  $\langle T_f, \hat{\varphi} \rangle$  n'a en général pas de sens. Pour obtenir une définition satisfaisante de la transformée de Fourier des distributions, on doit donc se placer sur un espace plus grand que  $\mathcal{D}$ .

### 3.2.2 Espace $\mathcal{S}$ et transformée de Fourier

**Définition 3.2.1.** Une fonction est dite à décroissance rapide si pour tout  $k$  dans  $\mathbb{N}$ ,  $\lim_{x \rightarrow \infty} |x^k f(x)| = 0$ . Une telle fonction décroît plus vite que toutes puissance de  $1/|x|$  à l'infini. On note  $\mathcal{S}$  l'ensemble des fonctions de  $\mathbb{R}$  dans  $\mathbb{R}$  ou  $\mathbb{C}$  qui sont indéfiniment dérivables et à décroissance rapide ainsi que toutes leurs dérivées.

Intéressons nous maintenant à la transformée de Fourier de telles fonctions. Soit  $\varphi$  une fonction de  $\mathcal{S}$ . On a :

$$\forall m \in \mathbb{N}, \hat{\varphi}^{(m)}(\nu) = \int (-2i\pi x)^m \exp(-2i\pi \nu x) \varphi(x) dx,$$

et on en déduit que  $\hat{\varphi}$  ainsi que toutes ses dérivées sont aussi à décroissances rapides (voir aussi le paragraphe 3.1.5). D'une manière générale, on a le résultat suivant :

**Théorème 3.2.1.** La transformation de Fourier est une application linéaire (et continue) de  $\mathcal{S}$  dans  $\mathcal{S}$ .

### 3.2.3 Transformée de Fourier des distributions tempérées

**Définition 3.2.2.** On appelle distribution tempérée toute fonctionnelle linéaire et continue sur l'espace de fonctions  $\mathcal{S}$ . Les distributions tempérées forment un sous-espace de  $\mathcal{D}'$  noté  $\mathcal{S}'$ .

Dans la pratique, la plupart des distributions sont tempérées ( $\delta_a$ ,  $\text{vp} \frac{1}{x}$ ). Cependant, en général, si  $f$  est une fonction localement sommable, alors  $T_f$  n'est pas tempérée.

**Théorème 3.2.2** (Caractérisation des distributions tempérées). Pour qu'une forme linéaire continue  $T$  sur  $\mathcal{S}$  soit tempérée, il faut et il suffit qu'il existe  $A > 0$  et  $p \in \mathbb{N}^+$  tels que, pour tout  $\varphi \in \mathcal{S}$ , on ait :

$$|\langle T, \varphi \rangle| \leq A \|\varphi\|_p,$$

où

$$\|\varphi\|_p = \int |\varphi(t)|^p dt.$$

**Proposition 3.2.1.** L'ensemble des distributions tempérées contient toutes les distributions à support bornée comme les Dirac, les dérivées des Dirac et les distributions régulières associées aux fonctions à croissance lente comme les polynômes, les fonctions périodiques localement sommables.

**Exemple :** la distribution régulière  $T_{\text{exp}}$  n'est pas tempérée puisque la fonction exponentielle croît trop rapidement à l'infini.

**Théorème et Définition 3.2.1.** Toute distribution tempérée  $T$  admet une transformée de Fourier, notée  $\mathcal{F}[T]$  ou  $\hat{T}$ , qui est également une distribution tempérée. Elle est définie par :

$$\forall \varphi \in \mathcal{S}, \langle \hat{T}, \varphi \rangle = \langle T, \hat{\varphi} \rangle.$$



**Exemples :**

- $\mathcal{F}[T_{\mathbb{1}}] = \delta.$

$$\langle \hat{T}_{\mathbb{1}}, \varphi \rangle = \langle T_{\mathbb{1}}, \hat{\varphi} \rangle = \int_{-\infty}^{+\infty} \hat{\varphi}(\nu) d\nu = \varphi(0) = \langle \delta, \varphi \rangle,$$

(cf. formule de la transformée de Fourier inverse)

- $\mathcal{F}[\delta] = T_{\mathbb{1}}.$

$$\langle \hat{\delta}, \varphi \rangle = \langle \delta, \hat{\varphi} \rangle = \hat{\varphi}(0) = \int_{-\infty}^{+\infty} \varphi(x) dx = \langle T_{\mathbb{1}}, \varphi \rangle,$$

- $\mathcal{F}[T_{\exp(2i\pi\nu_0 x)}] = \delta(\nu - \nu_0),$

- $\mathcal{F}[T_{\cos(2\pi\nu_0 x)}] = \frac{1}{2} (\delta(\nu - \nu_0) + \delta(\nu + \nu_0)),$

- $\mathcal{F}[T_{\sin(2\pi\nu_0 x)}] = \frac{1}{2i} (\delta(\nu - \nu_0) - \delta(\nu + \nu_0)).$

### 3.2.4 Propriétés

On peut définir une transformée de Fourier inverse comme pour les fonctions : on a

$$\forall \varphi \in \mathcal{S}, \langle \mathcal{F}^{-1}\mathcal{F}[T], \varphi \rangle = \langle \mathcal{F}[T], \mathcal{F}^{-1}[\varphi] \rangle = \langle T, \mathcal{F}\mathcal{F}^{-1}[\varphi] \rangle = \langle T, \varphi \rangle .$$

On peut montrer, de la même façon que pour la transformée de Fourier des fonctions, que l'on a les propriétés suivantes :

**Théorème 3.2.3.** *Soit  $T$  une distribution tempérée. On a alors :*

- $\mathcal{F}[T^{(m)}] = (2i\pi\nu)^m \mathcal{F}[T];$

- $\mathcal{F}[T(x - a)] = \exp(-2i\pi\nu a) \mathcal{F}[T];$

- $\mathcal{F}[T(ax)] = \frac{1}{|a|} \hat{T}\left(\frac{\nu}{a}\right);$

- $\mathcal{F}[\exp(2i\pi a x) T] = \hat{T}(\nu - a).$

**Théorème 3.2.4.** *Si  $T$  est une distribution tempérée à support borné, alors sa transformée de Fourier  $\mathcal{F}[T]$  est une distribution régulière associée à une fonction indéfiniment dérivable.*

### 3.2.5 Transformée de Fourier de la distribution peigne de Dirac

On rappelle que la distribution peigne de Dirac notée  $\text{III}$  est égale à  $\sum_{n=-\infty}^{+\infty} \delta(x - n)$ . Le théorème suivant n'est pas du tout évident à démontrer mais est très important pour les applications comme, par exemple, l'échantillonnage.

**Théorème 3.2.5.** *La distribution peigne de Dirac est une distribution tempérée. Elle admet donc une transformée de Fourier au sens des distributions. Sa transformée de Fourier est la distribution peigne de Dirac elle-même :  $\mathcal{F}[\text{III}(x)] = \text{III}(\nu)$  ou encore*

$$\mathcal{F}\left[\sum_{n=-\infty}^{+\infty} \delta(x - n)\right] = \sum_{n=-\infty}^{+\infty} \delta(\nu - n).$$

De plus, on a :

$$\mathcal{F}\left[\sum_{n=-\infty}^{+\infty} \delta(x - nT)\right] = \frac{1}{T} \sum_{n=-\infty}^{+\infty} \delta(\nu - \frac{n}{T}).$$

En utilisant la définition de  $\text{III}$  et la linéarité de la transformée de Fourier, on peut alors écrire :

$$\mathcal{F}\left[\sum_{n=-\infty}^{+\infty} \delta(x - n)\right] = \sum_{n=-\infty}^{+\infty} \mathcal{F}[\delta(x - n)] = \sum_{n=-\infty}^{+\infty} \exp(-2i\pi n\nu) T_{\mathbb{1}}.$$

D'où

$$\sum_{n=-\infty}^{+\infty} \delta(\nu - n) = \sum_{n=-\infty}^{+\infty} \exp(2i\pi n\nu) T_{\mathbb{1}}.$$

Grâce à ces formules, on peut obtenir le résultat très intéressant suivant :

**Théorème 3.2.6** (Formule sommatoire de Poisson). *Soit  $f$  une fonction continue admettant une transformée de Fourier. Lorsque ces sommes ont un sens, on a :*

$$\sum_{n=-\infty}^{+\infty} f(x - nT) = \frac{1}{T} \sum_{n=-\infty}^{+\infty} \hat{f}\left(\frac{n}{T}\right) \exp\left(\frac{2i\pi x n}{T}\right).$$

Dans le cas particulier  $x = 0$  et  $T = 1$ , la formule de Poisson s'écrit :

$$\sum_{n=-\infty}^{+\infty} f(n) = \sum_{n=-\infty}^{+\infty} \hat{f}(n).$$

## 3.3 Séries de Fourier et Échantillonnage

### 3.3.1 Transformée de Fourier des fonctions périodiques

#### Développement en série de Fourier d'une fonction

Soit  $f$  une fonction périodique de période  $T$ , i.e.,  $f(x) = f(x + nT)$  pour tout  $n \in \mathbb{N}$ . En notant

$$V_0 = \frac{1}{T},$$

on peut écrire le développement de  $f(x)$  en série de Fourier sous la forme

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{+\infty} (a_n \cos(2\pi n V_0 x) + b_n \sin(2\pi n V_0 x)),$$

avec

$$a_n = 2V_0 \int_{-T/2}^{T/2} f(t) \cos(2\pi n V_0 t) dt, \quad b_n = 2V_0 \int_{-T/2}^{T/2} f(t) \sin(2\pi n V_0 t) dt.$$

En transformant les cosinus et sinus en exponentielles complexes, on peut encore écrire

$$f(x) = \sum_{n=-\infty}^{+\infty} c_n \exp(2i\pi n V_0 x),$$

où

$$c_n = \frac{1}{2}(a_n - ib_n) = V_0 \int_{-T/2}^{T/2} f(t) \exp(-2i\pi n V_0 t) dt.$$

**Théorème 3.3.1** (formule de Parseval). *La série  $\sum_{n=-\infty}^{+\infty} |c_n|^2$  est convergente et on a*

$$\sum_{n=-\infty}^{+\infty} |c_n|^2 = \frac{1}{T} \int_{-T/2}^{T/2} |f(x)|^2 dx = \|f\|^2,$$

ce qui peut aussi s'écrire

$$\|f\|^2 = \frac{a_0^2}{4} + \frac{1}{2} \sum_{n \geq 1} (a_n^2 + b_n^2).$$

## Distributions et fonctions périodiques

**Théorème 3.3.2.** *Si  $F$  est une distribution périodique de période  $T$ , alors il existe une distribution  $F_0$  dont le support a une longueur inférieure à  $T$  et telle que*

$$F = F_0 * \sum_{n=-\infty}^{\infty} \delta(t - nT).$$

Sa transformée de Fourier est alors un peigne de Dirac modulé dont les Dirac sont en  $\frac{n}{T}$  avec  $n \in \mathbb{Z}$  :

$$\hat{F}(\nu) = \sum_{n=-\infty}^{+\infty} c_n \delta(\nu - \frac{n}{T}), \quad c_n = \frac{1}{T} \hat{F}_0(\frac{n}{T}).$$

*Démonstration.* On a  $\hat{F}(\nu) = \hat{F}_0(\nu) \mathcal{F}[\sum_{n=-\infty}^{\infty} \delta(t - nT)]$ . D'après le théorème 3.2.5, on obtient donc  $\hat{F}(\nu) = \hat{F}_0(\nu) \frac{1}{T} \sum_{n=-\infty}^{\infty} \delta(\nu - \frac{n}{T})$ . Or d'après le théorème 3.2.4,  $\hat{F}_0$  est une distribution régulière associée à une fonction indéfiniment dérivable de sorte que  $\hat{F}_0(\nu) \delta(\nu - \frac{n}{T}) = \hat{F}_0(\frac{n}{T}) \delta(\nu - \frac{n}{T})$  d'où le résultat.  $\square$

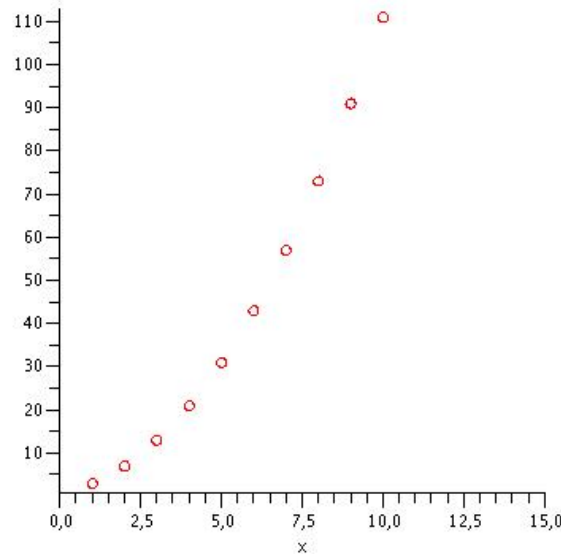
**Théorème 3.3.3.** Si  $f$  est une fonction périodique de période  $T$  et si l'on note  $c_n$  les coefficients dans son développement en série de Fourier complexe, alors on a

$$\mathcal{F}[T_f] = \sum_{n=-\infty}^{+\infty} c_n \delta(\nu - \frac{n}{T}).$$

*Démonstration.* Si  $f_0$  représente  $f$  sur une période, alors  $f(t) = \sum_{n=-\infty}^{+\infty} f_0(t - nT)$ . D'où  $T_f = \sum_{n=-\infty}^{+\infty} T_{f_0} * \delta(t - nT)$ . On a donc  $\mathcal{F}[T_f] = \mathcal{F}[T_{f_0}] \frac{1}{T} \sum_{n=-\infty}^{+\infty} \delta(\nu - \frac{n}{T})$ . Or  $f_0$  est localement sommable donc  $\mathcal{F}[T_{f_0}] = T_{\hat{f}_0}$  et  $\hat{f}_0(\nu) = \int_{-T/2}^{T/2} f_0(t) \exp(-2i\pi\nu t) dt$  et on retrouve bien le coefficient  $c_n = \frac{1}{T} \int_{-T/2}^{T/2} f_0(t) \exp(-2i\pi \frac{\nu}{T} t) dt$  du développement en série de Fourier complexe de  $f$  (voir aussi la preuve du théorème précédent).  $\square$

### 3.3.2 Échantillonnage

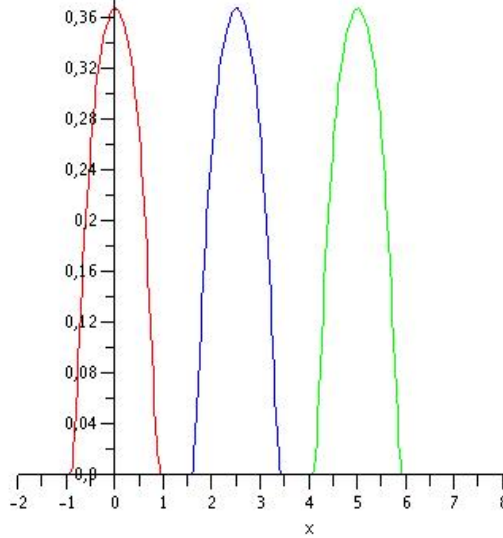
L'expérimentateur est souvent confronté au problème suivant : il ne dispose que d'une suite de mesures de la valeur d'une fonction en certains points.



Supposons que les  $x_i$  soient équidistants ( $x_j - x_{j-1} = c$  où  $c$  est une constante) et suffisamment rapprochés. Peut-on déterminer  $f$ ? La réponse est oui à condition que la transformée de Fourier de  $f$  soit à support borné et, dans ce cas, on va déterminer l'intervalle maximal entre deux valeurs de  $x$  permettant de reconstruire  $f$ .

Considérons le produit  $\text{III}(\frac{n}{T}) f(x)$ . En appliquant la transformée de Fourier, on obtient  $\mathcal{F}[\text{III}(\frac{n}{T}) f(x)] = T \text{III}(T\nu) * \hat{f}(\nu) = \sum_{n=-\infty}^{+\infty} \delta(x - \frac{n}{T}) * \hat{f}(\nu) = \sum_{n=-\infty}^{+\infty} \hat{f}(\nu - \frac{n}{T})$ .

Plaçons nous dans le cas où  $\hat{f}$  est à support borné compris dans  $[-V_0, V_0]$ . Pour  $T$  donné, on prélève un échantillon  $(f(nT), n \in \mathbb{N})$ . Si  $\frac{1}{T} > 2V_0$  c'est-à-dire  $T < \frac{1}{2V_0}$ , alors on obtient une série de fonctions disjointes.



On peut alors multiplier par la fonction porte prenant la valeur 1 pour  $|\nu| < V_0$  et 0 ailleurs, *i.e.*,  $\Pi(\frac{\nu}{2V_0})$  pour obtenir  $\hat{f}(\nu)$ .

**Théorème 3.3.4** (Théorème d'échantillonnage). *Une fonction réelle ayant une transformée de Fourier dont le support est contenu dans l'intervalle  $[-V_0, V_0]$  est entièrement déterminée par ses valeurs aux points  $x = nT$  pour  $n \in \mathbb{N}$  et  $T < \frac{1}{2V_0}$ .*

Ce théorème aussi connu sous le nom de théorème de Nyquist-Shannon signifie que pour pouvoir transformer un signal d'une forme discrète à une forme continue, sa fréquence d'échantillonnage doit être supérieure ou égale au double de la fréquence maximale contenue dans ce signal (application : conversion analogique-numérique des signaux).

On doit maintenant reconstruire  $f$  à partir des valeurs des  $f(nT)$ . On appelle ce processus *l'interpolation*. Si  $\hat{f}(\nu)$  est nulle pour  $|\nu| \geq V_0$ , alors  $\hat{f}(\nu) = (T\Pi(T\nu) * \hat{f}(\nu))\Pi(\frac{\nu}{2V_0})$ , pour  $T < \frac{1}{2V_0}$ . D'où en prenant la transformée de Fourier inverse  $f(x) = \Pi(\frac{x}{T}) f(x) * \frac{\sin(2\pi V_0 x)}{\pi x}$  et au final on trouve :

$$f(x) = \sum_{n=-\infty}^{+\infty} T f(nT) \frac{\sin(2\pi V_0(x - nT))}{\pi(x - nT)},$$

que l'on peut aussi écrire

$$f(x) = \sum_{n=-\infty}^{+\infty} f\left(\frac{n}{2V_0}\right) \frac{\sin(2\pi V_0 x - n\pi)}{2\pi V_0 x - n\pi}.$$



# Chapitre 4

## La Transformation de Laplace

La transformation de Laplace est une sorte de généralisation de la transformation de Fourier qui permet entre autre d'éviter d'utiliser les distributions lorsqu'une fonction n'admet pas de transformée de Fourier.

### 4.1 Transformée de Laplace des fonctions

#### 4.1.1 Définition

**Définition 4.1.1** (Transformée de Laplace unilatérale). Soit  $f$  une fonction défini sur  $\mathbb{R}_+$  et à valeurs réelles ou complexes :  $f : \mathbb{R}_+ \rightarrow \mathbb{R}$  ou  $\mathbb{C}$ . Lorsque  $f$  est localement sommable, on définit sa transformée de Laplace notée  $\mathcal{L}[f(t)]$  ou  $L(s)$  avec  $s \in \mathbb{C}$  par

$$\mathcal{L}[f(t)] = L(s) = \int_0^{+\infty} f(t) \exp(-s t) dt.$$

**Remarque :** Transformée de Laplace bilatérale.

Lorsque  $f : \mathbb{R} \rightarrow \mathbb{R}$  ou  $\mathbb{C}$  est sommable, on définit sa transformée de Laplace par

$$\mathcal{L}[f(t)] = L(s) = \int_{-\infty}^{+\infty} f(t) \exp(-s t) dt.$$

#### 4.1.2 Lien entre transformées de Laplace et de Fourier

Soit  $f$  une fonction défini sur  $\mathbb{R}_+$  et supposons que  $f$  admette une transformée de Fourier  $\hat{f}$ . On a alors

$$F(i\omega) = \int_{-\infty}^{+\infty} f(t) \exp(-i\omega t) dt = \int_{-\infty}^{+\infty} f(t) \exp(-2i\pi(\frac{\omega}{2\pi})t) dt = \hat{f}(\frac{\omega}{2\pi}).$$

La transformée de Laplace peut donc se voir comme une extension de la transformée de Fourier.  $F(x + i\omega)$  est la transformée de Fourier de  $t \mapsto f(t) \exp(-xt)$  prise en  $\omega/(2\pi)$ .

### 4.1.3 Domaine de définition, abscisse de sommabilité

Soit  $f$  une fonction localement sommable et soit  $s = x + i\omega \in \mathbb{C}$ . La fonction  $t \mapsto f(t) \exp(-st)$  est intégrable si et seulement si  $t \mapsto f(t) \exp(-xt)$  est intégrable.

**Définition 4.1.2.** On appelle abscisse de sommabilité de la fonction  $f$  et on note  $\alpha$  la borne inférieure de tous les  $x$  pour lesquels il y a sommabilité :

$$\alpha = \inf\{x \in \mathbb{R}; t \mapsto |f(t)| \exp(-xt) \text{ est sommable}\}$$

La transformée de Laplace  $F$  de  $f$  est donc défini pour tout  $s = x + i\omega \in \mathbb{C}$  avec  $x > \alpha$  où  $\alpha$  est l'abscisse de sommabilité de  $f$ . Dans certains cas,  $F$  est aussi définie pour  $x = \alpha$ .

**Exemple :** on considère la fonction de Heaviside  $H$ . Pour  $x \in \mathbb{R}$ , on a

$$\int_0^{+\infty} H(t) \exp(-xt) dt = \int_0^{+\infty} \exp(-xt) dt = -\frac{1}{x} [\exp(-xt)]_0^{+\infty}$$

L'abscisse de sommabilité de  $H$  est donc  $\alpha = 0$  et  $\mathcal{L}[H(t)]$  est défini pour tout complexe  $s$  ayant une partie réelle strictement positive : on a alors  $\mathcal{L}[H(t)] = \frac{1}{s}$ .

### 4.1.4 Formule d'inversion

Soit  $f$  une fonction d'abscisse de sommabilité  $\alpha$  et notons  $F$  sa transformée de Laplace. On a alors, pour  $x > \alpha$ ,

$$F(x + 2i\pi\nu) = \int_{-\infty}^{+\infty} H(t) f(t) \exp(-xt) \exp(-2i\pi\nu t) dt.$$

D'où

$$F(x + 2i\pi\nu) = \mathcal{F}[H(t) f(t) \exp(-xt)],$$

et en appliquant la transformée de Fourier inverse, il vient qu'en tout point  $t$  où  $f$  est continue :

$$H(t) f(t) \exp(-xt) = \int_{-\infty}^{+\infty} F(x + 2i\pi\nu) \exp(2i\pi\nu t) d\nu.$$

Ceci entraîne

$$H(t) f(t) = \int_{-\infty}^{+\infty} F(x + 2i\pi\nu) \exp(xt) \exp(2i\pi\nu t) d\nu.$$

On obtient donc

$$H(t) f(t) = \frac{1}{2i\pi} \int_{D_x} F(s) \exp(st) ds,$$

où  $s = x + 2i\pi\nu$  et  $D_x = \{x + i\omega; \omega \in \mathbb{R}\}$  est appelé *contour de Bromwich*.



**Théorème 4.1.1** (Inversion de la transformée de Laplace). *Soient  $f$  une fonction localement sommable et  $F$  sa transformée de Laplace. Si l'on note  $\alpha$  l'abscisse de sommabilité de  $f$ , on a la formule d'inversion suivante (valable en tout point de continuité de  $f$ ) :*

$$f(t) = \frac{1}{2i\pi} \int_{x_0-i\infty}^{x_0+i\infty} F(s) \exp(st) ds,$$

avec  $x_0 > \alpha$  quelconque.

### 4.1.5 Propriétés et Exemples

On montre assez facilement en utilisant des intégrations par parties les égalités suivantes :

- $\mathcal{L}[f'(t)] = -f(0) + s \mathcal{L}[f(t)]$  ;
- $\mathcal{L}[f''(t)] = -f'(0) - s f(0) + s^2 \mathcal{L}[f(t)]$  ;
- $\mathcal{L}[f^{(n)}(t)] = -f^{(n-1)}(0) - s f^{(n-2)}(0) - \dots - s^{n-1} f(0) + s^n \mathcal{L}[f(t)]$  ;
- $\mathcal{L}[\int_0^t f(u)du] = \frac{1}{s} \mathcal{L}[f(t)]$  ;
- $\mathcal{L}[f(t - T)] = \exp(-sT) \mathcal{L}[f(t)]$  ;
- $\mathcal{L}[f * g] = \mathcal{L}[f] \mathcal{L}[g]$  ;
- $\mathcal{L}[H(t)] = \frac{1}{s}$  ;
- $\mathcal{L}[H(t) \frac{t^n}{n!}] = \frac{1}{s^{n+1}}$  ;
- $\mathcal{L}[H(t) \exp(-at)] = \frac{1}{s+a}$  ;
- $\mathcal{L}[H(t) \cos(\omega t)] = \frac{s}{s^2 + \omega^2}$  ;
- $\mathcal{L}[H(t) \sin(\omega t)] = \frac{\omega}{\omega^2 + s^2}$ .

## 4.2 Transformée de Laplace des distributions

**Définition 4.2.1.** *Soit  $T$  une distribution à support bornée à gauche. S'il existe  $\alpha \in \mathbb{R}$  telle que, pour tout  $x > \alpha$ , la distribution  $\exp(-xt)T(t)$  soit tempérée, alors on peut définir la transformée de Laplace de  $T$  par l'application*

$$s \mapsto \mathcal{L}[T(t)] = \langle T(t), \exp(-st) \rangle .$$

On voit donc que la transformée de Laplace d'une distribution n'est pas une distribution mais une fonction qui à un complexe  $s$  associe le complexe  $\langle T(t), \exp(-st) \rangle$ .

### 4.2.1 Lien entre transformée de Laplace et de Fourier

Soit  $f$  une fonction localement sommable et  $T_f$  la distribution régulière associée. Soit  $\mathcal{L}[T_f]$  la transformée de Laplace de  $T_f$  et notons  $\alpha$  son abscisse de sommabilité. On a alors :

1. Si  $\alpha > 0$ , alors  $T_f$  n'est pas tempérée et n'admet donc pas de transformée de Fourier.
2. Si  $\alpha < 0$ , alors  $\hat{T}_f(\nu) = \mathcal{L}[T_f](2i\pi\nu)$ .
3. Si  $\alpha = 0$ , alors

$$\mathcal{L}[T_f](s) = \mathcal{L}(T_g)(s) + \sum_{n \in I} \frac{\lambda_n}{(s - i\omega_n)^{m_n}}$$

$$\hat{T}_f(\nu) = \text{Pf } \mathcal{L}(T_g)(2i\pi\nu) + \sum_{n \in I} \frac{(2i\pi)^{m_n-1}}{2(m_n-1)!} \lambda_n \delta^{(m_n-1)}(\nu - \nu_n),$$

où Pf est une distribution appelée *partie fractionnaire*.

### 4.2.2 Exemples

On montre assez facilement que l'on a :

- $\mathcal{L}[\delta(t)] = \langle \delta(t), \exp(-st) \rangle = 1$  ;
- $\mathcal{L}[\delta'(t)] = s$  ;
- $\mathcal{L}[\text{III}_+(t)] = \frac{1}{1 - \exp(-s)}$  ;
- $\mathcal{L}(W(t) * T(t)) = \frac{\mathcal{L}[T(t)]}{s}$ .

## 4.3 Application à la résolution d'équations de convolution

La transformation de Laplace peut être utile pour calculer des inverses de convolution et donc résoudre des équations de convolution.

On considère une équation de convolution

$$a * x = b,$$

où l'inconnue est la distribution  $x$  et  $a$  et  $b$  sont deux distributions données. On a déjà vu que la solution  $x$  est donnée par

$$x = a^{*-1} * b,$$

où  $a^{*-1}$  est l'inverse de convolution de  $a$  que l'on doit déterminer. Soit  $A(s)$  (resp.  $B(s)$ ) la transformée de Laplace de  $a(t)$  (resp.  $b(t)$ ). On peut alors montrer la proposition suivante :

**Proposition 4.3.1.** Si le module du complexe  $\frac{1}{A(s)}$  est majoré par un polynôme (en la variable  $s$ ), alors  $\frac{1}{A(s)}$  est la transformée de Laplace de l'inverse de convolution  $a^{*-1}$  de  $a$ .

On en déduit donc que si le module du complexe  $\frac{B(s)}{A(s)}$  est majoré par un polynôme (en la variable  $s$ ), alors  $\frac{B(s)}{A(s)}$  est la transformée de Laplace de la solution  $x$  de l'équation de convolution.

**Exemple :** Considérons l'équation de convolution

$$D \delta(t) * x(t) = b(t),$$

où  $D$  est un opérateur différentiel à coefficients constants. La transformée de Laplace de  $D \delta(t)$  est alors un polynôme  $P$  de la variable  $s$ . D'après ce qui précède, on sait alors que la transformée de Laplace de l'inverse de convolution  $(D \delta(t))^{*-1}$  de  $D \delta(t)$  est donné par  $\frac{1}{P(s)}$ , fraction rationnelle que l'on peut décomposer en éléments simples pour obtenir

$$\frac{1}{P(s)} = \sum_k \frac{a_k}{(s - s_k)^{\alpha_k}}.$$

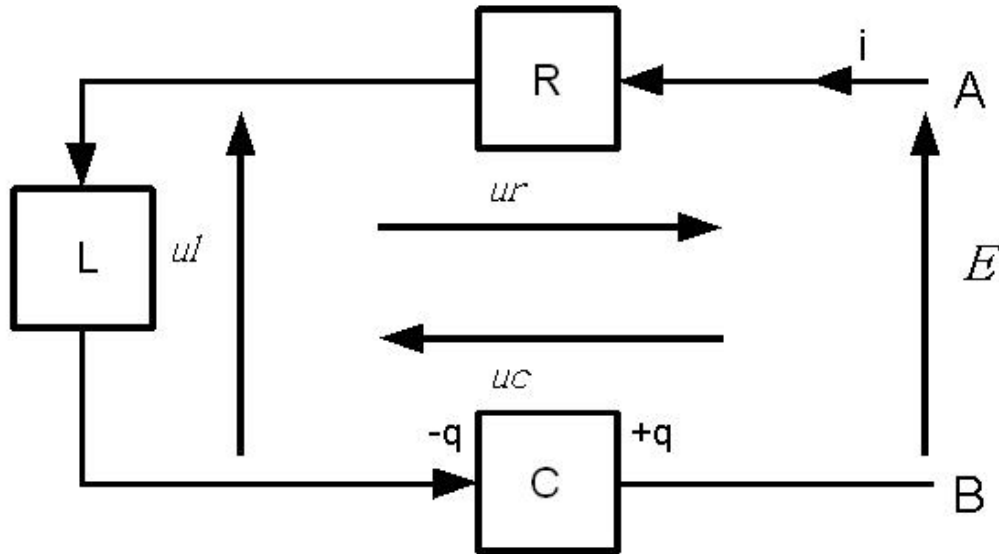
Finalement on en déduit (par Laplace inverse)

$$(D \delta(t))^{*-1} = W(t) \sum_k a_k \exp(-s_k t) \frac{t^{\alpha_k - 1}}{(\alpha_k - 1)!}.$$

## 4.4 Utilisation de la transformée de Laplace en physique

### 4.4.1 Calcul des fonctions de transfert en électronique

On considère un circuit RLC.



L'équation différentielle régissant un tel circuit est alors donnée par

$$E(t) = L \frac{d i(t)}{dt} + \frac{1}{C} \int_0^t i(u) du + R i(t).$$

Dans l'espace  $\mathcal{D}'_+$  des distributions à support borné à droite, cette équation s'écrit alors

$$E(t) = (L \delta' + \frac{1}{C} W + R \delta) * i(t).$$

En prenant la transformée de Laplace, il vient alors

$$E(s) = (L s + \frac{1}{C s} + R) I(s),$$

d'où

$$Z(s) := \frac{E(s)}{I(s)} = L s + \frac{1}{C s} + R.$$

Ce rapport de la tension à l'intensité en régime exponentiel est appelé *fonction de transfert* du circuit en régime exponentiel.

#### 4.4.2 En mécanique

On considère une cabine en translation le long de l'axe  $Oz$  d'un référentiel galiléen  $Oxyz$  et une masse  $m$  suspendue à son plafond par l'intermédiaire :

- d'un ressort de constante de raideur  $k$ ,
- d'un amortisseur de coefficient  $a$ .

Le principe fondamental de la dynamique nous permet alors d'écrire

$$m \ddot{x} = -a \dot{x} - k x - m \ddot{u}.$$

En prenant  $\ddot{u} = a H(t)$  et en appliquant la transformée de Laplace, il vient alors

$$m (-x'(0) - s x(0) + s^2 X(s)) + a (-x(0) + s X(s)) + k X(s) = -m \frac{a}{s}.$$

D'où avec les conditions initiales  $x(0) = x'(0) = 0$ ,

$$X(s) = \frac{-m a}{s (m s^2 + a s + k)} = \frac{-a}{s (s^2 + \frac{a}{m} s + \frac{k}{m})}.$$

On pose alors

$$w_0^2 = \frac{k}{m}, \quad \epsilon_1 = \frac{a}{2 \sqrt{m k}},$$

pour obtenir

$$X(s) = \frac{-a}{s(s^2 + 2\epsilon_1\omega_0 s + \omega_0^2)}.$$

Le discriminant du trinôme  $s^2 + 2\epsilon_1\omega_0 s + \omega_0^2$  est égal à  $4\omega_0^2(\epsilon_1^2 - 1)$  donc, en supposant,  $\epsilon_1 > 1$  on a  $s^2 + 2\epsilon_1\omega_0 s + \omega_0^2 = (s - s_1)(s - s_2)$  avec  $s_1 = -\omega_0\epsilon_1 + \omega_0\sqrt{\epsilon_1^2 - 1}$  et  $s_2 = -\omega_0\epsilon_1 - \omega_0\sqrt{\epsilon_1^2 - 1}$ . On peut alors décomposer la fraction rationnelle  $X(s)$  en éléments simples pour l'écrire sous la forme

$$X(s) = \frac{-a}{s(s - s_1)(s - s_2)} = \frac{A}{s} + \frac{B}{s - s_1} + \frac{C}{s - s_2},$$

avec

$$A = \frac{-a}{s_1 s_2} = \frac{-a}{\omega_0^2}, \quad B = \frac{-a}{s_1(s_1 - s_2)} = \frac{-a}{2\omega_0\sqrt{\epsilon_1^2 - 1}s_1}, \quad C = \frac{-a}{s_2(s_2 - s_1)} = \frac{a}{2\omega_0\sqrt{\epsilon_1^2 - 1}s_2}.$$

Finalement, en prenant la transformée de Laplace inverse, il vient

$$x(t) = \frac{-a}{\omega_0^2} H(t) + \frac{-a \exp(s_1 t)}{2\omega_0\sqrt{\epsilon_1^2 - 1}s_1} H(t) + \frac{a \exp(s_2 t)}{2\omega_0\sqrt{\epsilon_1^2 - 1}s_2} H(t).$$



# Chapitre 5

## Introduction à l'Optimisation

L'objectif de ce chapitre est d'introduire l'optimisation (sous contraintes) en partant du calcul différentiel. Nous rappellerons tout d'abord des notions indispensables de calcul différentiel (*e.g.*, dérivées premières et secondes). Ensuite, nous donnerons des conditions nécessaires et suffisantes d'extremum relatif (équation d'Euler, multiplicateurs de Lagrange, ...) pour les fonctions réelles. Enfin, nous utiliserons ceci pour étudier des problèmes (simples) d'optimisation sous contraintes (*e.g.*, d'égalité) dans une partie de  $\mathbb{R}^n$ .

### 5.1 Rappels et compléments de calcul différentiel

#### 5.1.1 Dérivées premières

Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  et  $(e_1, \dots, e_n)$  la base canonique de  $\mathbb{R}^n$ , *i.e.*,  $e_i$  est le vecteur ayant un 1 en  $i$ -ème position et 0 ailleurs.

**Définition 5.1.1.** Soit  $a = (a_1, \dots, a_n)$ ,  $d = (d_1, \dots, d_n) \in \mathbb{R}^n$ . On appelle dérivée directionnelle de  $f$  en  $a$  dans la direction  $d$  en on note  $\frac{\partial f}{\partial d}(a)$  la limite

$$\frac{\partial f}{\partial d}(a) = \lim_{h \rightarrow 0} \frac{f(a + h d) - f(a)}{h} = \lim_{h \rightarrow 0} \frac{f(a_1 + h d_1, \dots, a_n + h d_n) - f(a_1, \dots, a_n)}{h},$$

si elle existe. L'application  $\frac{\partial f}{\partial d} : a \mapsto \frac{\partial f}{\partial d}(a)$  est une fonction de  $\mathbb{R}^n$  dans  $\mathbb{R}$  que l'on appelle dérivée directionnelle de  $f$  dans la direction  $d$ .

**Exemples :**

- La pente d'un relief dans une direction donnée est la dérivée directionnelle dans cette direction.
- Soit  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $(x, y) \mapsto x^3 y - y$  et soit  $d = (\frac{1}{2}, \frac{\sqrt{3}}{2}) \in \mathbb{R}^2$ . Un calcul direct montre que :

$$\frac{\partial f}{\partial d}(x, y) = \frac{3}{2} x^2 y + \frac{\sqrt{3}}{2} (x^3 - 1).$$

- En physique, la dérivée normale (resp. tangentielle) en un point d'une courbe ou d'une surface est la dérivée directionnelle dans la direction normale (resp. tangente) à la courbe ou la surface.

**Définition 5.1.2.** On appelle dérivée partielle de  $f$  par rapport à la  $i$ -ème variable et on note  $\frac{\partial f}{\partial x_i}$  la dérivée directionnelle de  $f$  dans la direction  $e_i$ .

En pratique, on calcule  $\frac{\partial f}{\partial x_i}$  comme une dérivée classique en supposant  $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$  constants et en dérivant par rapport à  $x_i$ .

**Définition 5.1.3.** Soit  $a \in \mathbb{R}^n$ . On appelle gradient de  $f$  au point  $a$  et on note  $\text{grad } f(a)$  ou  $\nabla f(a)$  le vecteur

$$\text{grad } f(a) = \nabla f(a) = \left( \frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right).$$

**Proposition 5.1.1.** Soit  $d \in \mathbb{R}^n$ . Si  $\nabla f(a)$  existe, alors  $\frac{\partial f}{\partial d}(a) = \nabla f(a) \cdot d$ , où  $\cdot$  désigne le produit scalaire de  $\mathbb{R}^n$ , i.e.,  $\forall x, y \in \mathbb{R}^n, x \cdot y = \sum_{i=1}^n x_i y_i$ .

**Exemple :** Soit  $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^3 y - y, d = (\frac{1}{2}, \frac{\sqrt{3}}{2}) \in \mathbb{R}^2$  et  $a = (0, 0)$ . On a  $\frac{\partial f}{\partial x}(x, y) = 3x^2 y, \frac{\partial f}{\partial y}(x, y) = x^3 - 1$ , d'où  $\nabla f(a) = (0, -1)$ . D'où

$$\frac{\partial f}{\partial d}(a) = \nabla f(a) \cdot d = (0, -1) \cdot \left( \frac{1}{2}, \frac{\sqrt{3}}{2} \right) = -\frac{\sqrt{3}}{2}.$$

**Définition 5.1.4.** Soit  $a \in \mathbb{R}^n$ . On dit que  $f$  est (partiellement) dérivable en  $a$  si toutes ses dérivées partielles en  $a$  existent. On dit que  $f$  est continûment (partiellement) dérivable en  $a$  si toutes ses dérivées partielles existent dans un voisinage de  $a$  et sont continues en  $a$  : dans ce cas on dit aussi que  $f$  est de classe  $\mathcal{C}^1$  en  $a$ .

**Définition 5.1.5.** Soit  $f$  de classe  $\mathcal{C}^1$  en  $a$ . On appelle différentielle de  $f$  en  $a$  et on note  $Df[a]$  l'application linéaire définie par :

$$Df[a] : \begin{cases} \mathbb{R}^n & \rightarrow \mathbb{R} \\ h & \mapsto \nabla f(a) \cdot h \end{cases}$$

**Exemple :** pour l'exemple précédent, on a

$$Df[(x, y)](dx, dy) = 3x^2 y dx + (x^3 - 1) dy.$$

**Proposition 5.1.2.** Avec les notations précédentes, on a :

$$f(a + h) = f(a) + Df[a](h) + o(\|h\|),$$

c'est-à-dire

$$\lim_{h \rightarrow 0} \frac{f(a + h) - f(a) - Df[a](h)}{\|h\|} = 0.$$



**Définition 5.1.6.** Avec les notations précédentes, on dit que  $f$  est différentiable en  $a$  s'il existe une application linéaire, notée  $Df[a]$ , telle que, pour tout  $h \in \mathbb{R}^n$ ,  $f(a+h) = f(a) + Df[a](h) + o(\|h\|)$ .

La définition précédente s'applique d'une manière générale. En particulier si  $f$  est  $\mathcal{C}^1$ , alors elle est différentiable mais il peut aussi exister des applications non  $\mathcal{C}^1$  qui sont différentiables.

**Définition 5.1.7.** Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ . On note  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$  les applications composantes. Alors, on dit que  $f$  est différentiable en  $a \in \mathbb{R}^n$  si et seulement si, pour tout  $i \in \{1, \dots, p\}$ ,  $f_i$  est différentiable en  $a$ . De plus :

$$Df[a](h) = (Df_1[a](h), \dots, Df_p[a](h)).$$

Puisque  $Df_i[a](h) = \nabla f_i(a) \cdot h$ , on a

$$\begin{pmatrix} Df_1[a](h) \\ \vdots \\ Df_p[a](h) \end{pmatrix} = \begin{pmatrix} \partial_1 f_1(a) & \partial_2 f_1(a) & \dots & \partial_n f_1(a) \\ \partial_1 f_2(a) & \partial_2 f_2(a) & \dots & \partial_n f_2(a) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 f_p(a) & \partial_2 f_p(a) & \dots & \partial_n f_p(a) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix},$$

avec la notation

$$\partial_i f_j(a) = \frac{\partial f_j}{\partial x_i}(a).$$

La matrice précédente s'appelle *matrice jacobienne de  $f$  en  $a$*  et se note  $J_f(a)$ . Dans le cas où  $n = p$ , son déterminant s'appelle le *jacobien de  $f$  en  $a$* .

## 5.1.2 Dérivation de fonctions composées

**Théorème 5.1.1** (Dérivée d'une application composée). Soient  $f : E_1 \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$  et  $g : E_2 \subset \mathbb{R}^p \rightarrow \mathbb{R}$  deux applications. On suppose que  $f(E_1) \subset E_2$  de sorte que la composée  $g \circ f$  soit bien définie. Soit  $a \in E_1$ . On suppose que  $f$  est différentiable en  $a$  et que  $g$  est différentiable en  $f(a)$ . Alors  $g \circ f$  est différentiable en  $a$  et

$$D(g \circ f)[a] = Dg[f(a)] \circ Df[a],$$

ou encore, matriciellement,

$$J_{g \circ f}(a) = J_g(f(a)) J_f(a).$$

Une application utile de ce résultat est le changement de variables. Soit  $\Phi : (x_1, \dots, x_n) \mapsto (y_1, \dots, y_n)$  un changement de variables. Soit  $f$  une fonction de classe  $\mathcal{C}^1$  et  $g = f \circ \Phi$ , i.e.,  $g(x_1, \dots, x_n) = f(y_1, \dots, y_n)$ . On a alors la *formule des dérivées totales* :

$$\forall i = 1, \dots, n, \quad \frac{\partial g}{\partial x_i}(x_1, \dots, x_n) = \sum_{k=1}^n \frac{\partial f}{\partial y_k}(y_1, \dots, y_n) \frac{\partial y_k}{\partial x_i}.$$

**Exemple** : considérons le changement de variables  $\Phi : (r, \theta) \mapsto (x = r \cos(\theta), y = r \sin(\theta))$ ,  $f : (x, y) \mapsto x^2 + y^2$  et  $g = f \circ \Phi$ . On a alors

$$\begin{aligned} \frac{\partial g}{\partial r}(r, \theta) &= \frac{\partial f}{\partial x}(x, y) \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}(x, y) \frac{\partial y}{\partial r} \\ &= 2x \cos(\theta) + 2y \sin(\theta) \\ &= 2r \cos^2(\theta) + 2r \sin^2(\theta) \\ &= 2r, \end{aligned}$$

et

$$\begin{aligned} \frac{\partial g}{\partial \theta}(r, \theta) &= \frac{\partial f}{\partial x}(x, y) \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}(x, y) \frac{\partial y}{\partial \theta} \\ &= 2x(-r \sin(\theta)) + 2y(r \cos(\theta)) \\ &= 2(r \cos(\theta))(-r \sin(\theta)) + 2(r \sin(\theta))(r \cos(\theta)) \\ &= 0. \end{aligned}$$

Ceci est cohérent avec le fait que  $g(r, \theta) = r^2$ .

### 5.1.3 Dérivées d'ordre supérieurs

Soit  $f : E \subset \mathbb{R}^n \rightarrow \mathbb{R}$ .

**Définition 5.1.8.** Soit  $\{i_1, \dots, i_k\} \in \{1, \dots, n\}^k$ . On dit que  $f$  admet une dérivée partielle d'ordre  $k$  en  $a \in \mathbb{R}^n$  par rapport à  $\{i_1, \dots, i_k\}$  si  $\frac{\partial f}{\partial x_{i_1}}, \frac{\partial}{\partial x_{i_2}}(\frac{\partial f}{\partial x_{i_1}}), \dots, \frac{\partial}{\partial x_{i_{k-1}}}(\dots)$  existent dans un voisinage de  $a$  et si  $\frac{\partial}{\partial x_{i_k}}(\dots)(a)$  existe. On note alors  $\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}(a)$ .

**Définition 5.1.9.** On dit que  $f$  est de classe  $\mathcal{C}^k$  sur  $E$  si toutes ses dérivées partielles d'ordre inférieur ou égale à  $k$  existent et sont continues sur  $E$ .

**Théorème 5.1.2** (Schwarz). Si  $f$  est  $\mathcal{C}^1$  sur  $E$  et si  $\frac{\partial^2 f}{\partial x_i \partial x_j}$  et  $\frac{\partial^2 f}{\partial x_j \partial x_i}$  existent et sont continues en  $a \in E$ , alors

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

D'une manière générale, lorsque tout est bien défini, l'ordre de dérivation dans le calcul des dérivées partielles n'a pas d'importance.

## 5.2 Extrema d'une fonction de plusieurs variables réelles

### 5.2.1 Définitions et premiers résultats

Soit  $f : E \subset \mathbb{R}^n \rightarrow \mathbb{R}$ .

**Définition 5.2.1.** On dit que  $f$  admet un minimum local en  $a \in E$  s'il existe un voisinage  $\mathcal{V}$  de  $a$  tel que,  $\forall x \in \mathcal{V}$ ,  $f(x) \geq f(a)$ . On dit que  $f$  admet un minimum global en  $a \in E$  si  $\forall x \in E$ ,  $f(x) \geq f(a)$ .

On définit de même la notion de maximum local ou global et on parle de minimum (ou maximum) *strict* lorsque l'inégalité est stricte. On parle d'*extremum* pour désigner indifféremment un maximum ou un minimum.

**Définition 5.2.2.** Soit  $a \in E$ . On dit que  $a$  est un point critique de  $f$  si les dérivées partielles premières de  $f$  en  $a$  existent et sont nulles ou, autrement dit, si le gradient de  $f$  en  $a$  est nul, i.e.,  $\nabla f(a) = 0$ .

**Théorème 5.2.1.** On suppose que les dérivées partielles de  $f$  existent en tout point  $a$  n'appartenant pas au bord de  $E$ . Une condition nécessaire pour que  $f$  admette un extremum en  $a$  est que  $a$  soit un point critique de  $f$ .

**Théorème 5.2.2.** Si  $E$  est un ensemble fermé et borné et que  $f$  est continue sur  $E$ , alors  $f$  admet au moins un minimum et un maximum sur  $E$ .

**Théorème 5.2.3.** Les extrema d'une fonction  $\mathcal{C}^1$  sur un ensemble fermé et borné sont soit des points critiques, soit des points du bord.

Dans le cas d'une fonction de deux variables réelles, on a le résultat plus précis suivant :

**Théorème 5.2.4.** Soit  $f : E \subset \mathbb{R}^2 \rightarrow \mathbb{R}$  que l'on suppose de classe  $\mathcal{C}^2$  au voisinage d'un point critique  $(x_0, y_0) \in E$ . On pose

$$r = \frac{\partial^2 f}{\partial x^2}(x_0, y_0), \quad s = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0), \quad t = \frac{\partial^2 f}{\partial y^2}(x_0, y_0),$$

et on note  $u = s^2 - rt$ . On a alors les quatre cas suivants :

- Si  $u < 0$  et  $r < 0$ , alors  $f$  admet un maximum local stricte en  $(x_0, y_0)$ ,
- Si  $u < 0$  et  $r > 0$ , alors  $f$  admet un minimum local stricte en  $(x_0, y_0)$ ,
- Si  $u > 0$ , alors  $f$  admet un "point selle" en  $(x_0, y_0)$ ,
- Si  $u = 0$ , alors on ne sait pas a priori la nature du point critique et une étude plus précise doit être menée.

**Application :** méthode d'approximation des moindres carrées

Une question souvent rencontrée est celle de la modélisation du lien entre deux variables  $X$  et  $Y$ . En pratique, on dispose d'un échantillon de  $n$  mesures  $\{(x_1, y_1), \dots, (x_n, y_n)\}$ . Si le nuage de points semble suivre une certaine allure, on peut essayer de modéliser la relation entre  $X$  et  $Y$  sous la forme

$$Y = \sum_{k=1}^K a_k f_k(X) + \epsilon,$$

où les  $f_i$  sont des fonctions élémentaires (e.g., puissances, logarithmes, exponentielles, sinus, cosinus, ...) devinée à partir de l'allure de la courbe et où  $\epsilon$  est l'erreur entre le modèle et

la réalité. On dit que l'on *explique*  $Y$  par  $X$  ou encore que  $Y$  est la *variable expliquée* et  $X$  la *variable explicative*. On doit donc avoir :

$$\forall i \in \{1, \dots, n\}, \quad y_i = \sum_{k=1}^K a_k f_k(x_i) + \epsilon_i.$$

On définit l'*erreur globale entre le modèle et la réalité* :

$$E(a_1, \dots, a_K) = \sum_{i=1}^n \epsilon_i^2 = \sum_{i=1}^n \left( y_i - \sum_{k=1}^K a_k f_k(x_i) \right)^2.$$

La *méthode des moindres carrés* consiste alors à déterminer les  $a_k$  qui minimisent cette erreur. D'après les résultats précédents, nous sommes donc amenés à résoudre le système linéaire de  $K$  équations à  $K$  inconnues suivant :

$$\begin{cases} \frac{\partial E}{\partial a_1}(a_1, \dots, a_K) = -2 \sum_{i=1}^n f_1(x_i) (y_i - a_1 f_1(x_i) - \dots - a_K f_K(x_i)) = 0, \\ \vdots \\ \frac{\partial E}{\partial a_K}(a_1, \dots, a_K) = -2 \sum_{i=1}^n f_K(x_i) (y_i - a_1 f_1(x_i) - \dots - a_K f_K(x_i)) = 0, \end{cases}$$

qui peut s'écrire :

$$\begin{cases} a_1 \sum_{i=1}^n f_1(x_i)^2 & + & a_2 \sum_{i=1}^n f_1(x_i) f_2(x_i) & + & \dots & + & a_K \sum_{i=1}^n f_1(x_i) f_K(x_i) & = & 0, \\ \vdots & & \vdots & & \vdots & & \vdots & & \\ a_1 \sum_{i=1}^n f_K(x_i) f_1(x_i) & + & a_2 \sum_{i=1}^n f_K(x_i) f_2(x_i) & + & \dots & + & a_K \sum_{i=1}^n f_K(x_i)^2 & = & 0. \end{cases}$$

Lorsque les valeurs optimales  $\hat{a}_1, \dots, \hat{a}_K$  sont déterminées (en résolvant le système linéaire précédent), l'erreur globale  $E(\hat{a}_1, \dots, \hat{a}_K)$  est appelée *erreur résiduelle* et  $\sqrt{\frac{1}{n} E(\hat{a}_1, \dots, \hat{a}_K)}$  est appelée *écart-type résiduel*. Si l'on souhaite comparer deux modèles de régression  $Y = \sum_{k=1}^K a_k f_k(X)$  et  $Y = \sum_{k=1}^L b_k g_k(X)$ , on comparera alors leurs écarts-types résiduels. Notons enfin que dans le cas particulier d'un modèle de régression linéaire  $Y = aX + b$ , la résolution du système linéaire précédent conduit à

$$\hat{a} = \frac{\frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y}}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}, \quad \bar{b} = \bar{y} - a \bar{x}, \quad \text{avec} \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

## 5.2.2 Extrema relatifs

Les résultats suivants sont énoncés dans un cadre général. En pratique, le cas qui nous intéressera est le cas où  $\Omega = \mathbb{R}^n$  et où, pour une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f'$  représente le gradient  $\nabla f$  de  $f$ .

**Définition 5.2.3.** Soit maintenant  $f : \Omega \rightarrow \mathbb{R}$  une fonction définie sur un espace topologique  $\Omega$  et soit  $U$  une partie de  $\Omega$ . On dit que  $f$  admet en un point  $u \in U$  un minimum (ou maximum ou extremum) relatif par rapport à  $U$  si la restriction de  $f$  à  $U$  (muni de la topologie induite) admet un minimum (ou maximum ou extremum) relatif en  $u$ . On parle d'extremum relatif lié lorsque  $\Omega$  est un produit  $V_1 \times V_2$  d'espaces vectoriels normés et que  $U$  est de la forme

$$U = \{(v_1, v_2) \in \Omega; \varphi(v_1, v_2) = 0\},$$

où

$$\varphi : \Omega \subset V_1 \times V_2 \rightarrow V_2.$$

**Théorème 5.2.5** (Condition nécessaire d'extremum relatif lié). Soit  $\Omega$  un ouvert d'un produit  $V_1 \times V_2$  d'espaces vectoriels normés, l'espace  $V_1$  étant complet. Soit  $\varphi : \Omega \rightarrow V_2$  une fonction continûment dérivable sur  $\Omega$  et soit  $u = (u_1, u_2)$  un point de l'ensemble

$$U = \{(v_1, v_2) \in \Omega; \varphi(v_1, v_2) = 0\} \subset \Omega,$$

en lequel

$$\partial_2 \varphi(u_1, u_2) = \frac{\partial \varphi}{\partial v_2}(u_1, u_2) \in \text{Iso}(V_2),$$

où  $\text{Iso}(V_2)$  désigne l'ensemble des isomorphismes de  $V_2$ . Soit  $f : \Omega \rightarrow \mathbb{R}$  une fonction dérivable en  $u$ . Si la fonction  $f$  admet un extremum relatif par rapport à l'ensemble  $U$ , alors il existe un élément  $\Lambda(u) \in \mathcal{L}(V_2, \mathbb{R})$  tel que

$$f'(u) + \Lambda(u) \varphi'(u) = 0.$$

Dans la pratique, ce résultat s'emploie surtout dans la situation suivante : soient  $m$  et  $n$  deux entiers tels que  $1 \leq m < n$ . On considère les fonctions  $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$  et  $\varphi_i : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$  pour  $1 \leq i \leq m$ . On se pose alors le problème de trouver une condition nécessaire d'extremum relatif par rapport à l'ensemble

$$U = \{v \in \Omega; \varphi_i(v) = 0, 1 \leq i \leq m\}.$$

En pratique, cet ensemble  $U$  comprend les *contraintes* du problème. Le théorème précédent s'écrit alors :

**Corollaire 5.2.1** (Condition nécessaire d'extremum relatif lié). Soit  $\Omega$  un ouvert de  $\mathbb{R}^n$ , soit  $\varphi_i : \Omega \rightarrow \mathbb{R}$ ,  $1 \leq i \leq m$ , des fonctions continûment dérivables sur  $\Omega$  et soit  $u$  un point de l'ensemble

$$U = \{v \in \Omega; \varphi_i(v) = 0, 1 \leq i \leq m\} \subset \Omega,$$

en lequel les dérivées  $\varphi'_i(u) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ ,  $1 \leq i \leq m$  sont linéairement indépendantes. Soit  $f : \Omega \rightarrow \mathbb{R}$  une fonction dérivable en  $u$ . Si la fonction  $f$  admet en  $u$  un extremum relatif par rapport à l'ensemble  $U$ , alors il existe des nombres  $\lambda_i(u)$ ,  $1 \leq i \leq m$ , définis de façon unique tels que

$$f'(u) + \lambda_1(u) \varphi'_1(u) + \cdots + \lambda_m(u) \varphi'_m(u) = 0.$$

Les nombres  $\lambda_i(u)$ ,  $1 \leq i \leq m$ , du corollaire ci-dessus sont appelés *les multiplicateurs de Lagrange* associés à l'extremum lié  $u$ .

**Exemple :** Soient à trouver les extremum relatifs de la fonction  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $(x_1, x_2) \mapsto -x_2$  par rapport à l'ensemble  $U = \{(x_1, x_2) \in \mathbb{R}^2; \varphi(x_1, x_2) = x_1^2 + x_2^2 - 1 = 0\}$ . En notant  $\partial_1 = \frac{\partial}{\partial x_1}$  et  $\partial_2 = \frac{\partial}{\partial x_2}$ , la condition nécessaire du corollaire précédent s'écrit alors

$$\begin{cases} \partial_1 f(u_1, u_2) + \lambda \partial_1 \varphi(u_1, u_2) = 0, \\ \partial_2 f(u_1, u_2) + \lambda \partial_2 \varphi(u_1, u_2) = 0, \end{cases}$$

ce qui équivaut alors à

$$\begin{cases} 2\lambda u_1 = 0, \\ -1 + 2\lambda u_2 = 0. \end{cases}$$

Finalement les conditions nécessaires d'extremum relatif lié obtenues sont :

- $u = (u_1, u_2) = (0, 1)$  avec multiplicateur de Lagrange  $\lambda = 1/2$ , ou bien,
- $u = (u_1, u_2) = (0, -1)$  avec multiplicateur de Lagrange  $\lambda = -1/2$ .

Cette condition est seulement nécessaire et, en général, on ne peut donc pas conclure que les points  $u$  candidats trouvés sont effectivement des extrema. Pour obtenir une condition suffisante, on a besoin de prendre en compte la dérivée seconde.

### 5.2.3 Conditions nécessaires et suffisantes en utilisant la dérivée seconde

Cette sous-section est aussi écrite dans un cadre très général.

On note  $f''(a)(h, k) = \sum_{i,j} h_i k_j \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$ .

L'utilisation de la dérivée seconde permet tout d'abord d'obtenir la condition nécessaire suivante :

**Théorème 5.2.6** (Condition nécessaire de minimum relatif). *Soit  $\Omega$  un ouvert d'un espace vectoriel normé  $X$  et soit  $f : \Omega \subset X \rightarrow \mathbb{R}$  une fonction dérivable dans  $\Omega$ , deux fois dérivable en un point  $u \in \Omega$ . Si la fonction  $f$  admet un extremum relatif en  $u$ , alors*

$$f''(u)(x, x) \geq 0, \forall x \in X.$$

La réciproque de ce théorème est fautive. On a besoin d'une hypothèse plus forte au point  $u$  pour avoir une condition suffisante.

**Théorème 5.2.7** (Condition suffisante de minimum relatif). *Soit  $\Omega$  un ouvert d'un espace vectoriel normé  $X$ , soit  $u$  un point de  $\Omega$  et  $f : \Omega \subset X \rightarrow \mathbb{R}$  une fonction dérivable dans  $\Omega$  telle que  $f'(u) = 0$ .*

1. Si la fonction  $f$  est deux fois dérivable en  $u$  et s'il existe un nombre  $\alpha$  tel que

$$\alpha > 0, \text{ et } f''(u)(x, x) \geq \alpha \|x\|_X^2, \forall x \in X,$$

alors la fonction  $f$  admet un minimum relatif strict en  $u$ .

2. Si la fonction  $f$  est deux fois dérivable dans  $\Omega$  et s'il existe une boule  $\mathcal{B} \subset \Omega$  centrée en  $u$  telle que

$$f''(v)(x, x) \geq 0, \forall v \in \mathcal{B}, \forall x \in X,$$

alors la fonction  $f$  admet un minimum relatif en  $u$ .

## 5.3 Optimisation sous contrainte dans une partie de $\mathbb{R}^n$

**Exemple** : considérons la fonction  $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x_1, x_2) \mapsto x_1^2 + 2x_2^2 - 2x_1$  et intéressons nous au problème de trouver le minimum de  $f$  (c'est-à-dire le point où la valeur de  $f(x_1, x_2)$  est minimale). On s'aperçoit alors que  $f(x_1, x_2) = (x_1 - 1)^2 + 2x_2^2 - 1$  de sorte que l'on trouve simplement  $\min f(x_1, x_2) = (1, 0)$ . Considérons maintenant la contrainte  $g(x_1, x_2) = x_2 - x_1^2 = 0$ . On s'intéresse donc au problème de trouver le point  $(u_1, u_2)$  qui

1. minimise  $f(x_1, x_2)$ ;
2. satisfait  $g(u_1, u_2) = 0$ .

Le problème se ramène donc à minimiser la fonction  $f$  donnée par  $f(x_1) = (x_1 - 1)^2 + 2x_1^4 - 1$ . On a alors  $f'(x_1) = 2(x_1 - 1) + 8x_1^3 = 2(2x_1 - 1)(2x_1^2 + x_1 + 1)$  de sorte que  $f'(x_1) = 0 \Leftrightarrow x_1 = 1/2$ . On remarque que la solution trouvée  $(1/2, 1/4)$  est la projection de la solution  $(1, 0)$  du problème sans contrainte sur la courbe représentative de la contrainte.

### 5.3.1 Optimisation avec contrainte d'égalité

**Définition 5.3.1.** Soit  $\Omega$  un ouvert de  $\mathbb{R}^n$ . Soit  $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$  et soit  $\varphi : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ . On appelle problème d'optimisation avec contraintes d'égalité le problème

$$\text{Trouver } x^* = \min_{x \in \Omega} f(x) \text{ sous les contraintes } \varphi(x) = 0.$$

On déduit des théorèmes vu précédemment les théorèmes suivants :

**Théorème 5.3.1** (Condition nécessaire du 1<sup>er</sup> ordre : théorème de Lagrange). Soit  $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$  dérivable et soit  $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^m$  telle que les applications composantes  $\varphi_i$  soient dérivables. Soit

$$\mathcal{C} = \{x \in \Omega; \varphi(x) = 0\}.$$

Si  $x^*$  est une solution (locale) du problème

$$\text{Trouver } x^* = \min_{x \in \Omega} f(x) \text{ sous les contraintes } \varphi(x) = 0,$$

et si les  $\varphi'_i(x^*)$  sont linéairement indépendants alors :

$$\exists! (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m ; f'(x^*) + \lambda_1 \varphi'_1(x^*) + \dots + \lambda_m \varphi'_m(x^*) = 0.$$

**Définition 5.3.2.** On appelle fonction de Lagrange ou Lagrangien associé au problème

$$\text{Trouver } x^* = \min_{x \in \Omega} f(x) \text{ sous les contraintes } \varphi(x) = 0,$$

la fonction

$$L : (x, \lambda) \in \Omega \times \mathbb{R}^m \mapsto L(x, \lambda) := f(x) + \sum_{i=1}^m \lambda_i \varphi_i(x).$$

**Théorème 5.3.2** (Condition suffisante de minimalité). Soit  $x^* \in \mathcal{C}$  et soit  $\lambda \in \mathbb{R}^m$  tels que :

$$\begin{cases} \partial_1 L(x^*, \lambda) & = 0, \\ \partial_{11} L(x^*, \lambda)(d, d) & > 0, \end{cases}$$

pour tout  $d$  appartenant à l'ensemble

$$T(\mathcal{C}, x^*) := \{d \in \mathbb{R}^n ; \varphi'_i(x^*)d = 0, \forall 1 \leq i \leq m\}.$$

Alors  $x^*$  est une solution locale du problème

$$\text{Trouver } x^* = \min_{x \in \Omega} f(x) \text{ sous les contraintes } \varphi(x) = 0,$$

**Théorème 5.3.3** (Condition nécessaire du 2<sup>nd</sup> ordre). Soit  $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$  deux fois dérivable et soit  $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^m$  telle que les applications composantes  $\varphi_i$  soient deux fois dérivables. Soit

$$\mathcal{C} = \{x \in \Omega ; \varphi(x) = 0\}.$$

Si  $x^*$  est une solution (locale) du problème

$$\text{Trouver } x^* = \min_{x \in \Omega} f(x) \text{ sous les contraintes } \varphi(x) = 0,$$

et si les  $\varphi'_i(x^*)$  sont linéairement indépendants alors :

$$\exists! \lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m ; \begin{cases} \partial_1 L(x^*, \lambda) & = 0, \\ \partial_{11} L(x^*, \lambda)(x, x) & \geq 0, \end{cases}$$

pour tout  $d$  appartenant à

$$T(\mathcal{C}, x^*) := \{d \in \mathbb{R}^n ; \varphi'_i(x^*)d = 0, \forall 1 \leq i \leq m\}.$$



### 5.3.2 Optimisation sous d'autres contraintes

Les contraintes peuvent aussi s'exprimer à l'aide d'inégalités ou d'appartenance à un ensemble.

**Exemple :** Une usine fabrique deux produits  $P_1$  et  $P_2$  à l'aide de matières premières  $Q_1$ ,  $Q_2$  et  $Q_3$ . La fabrication d'une unité du produit  $P_1$  nécessite une unité de  $Q_1$ , deux unités de  $Q_2$  et quatre unités de  $Q_3$ . La fabrication d'une unité du produit  $P_2$  nécessite six unités de  $Q_1$ , deux unités de  $Q_2$  et une unité de  $Q_3$ . L'usine dispose de trente unités de  $Q_1$ , quinze unités de  $Q_2$  et vingt quatre unités de  $Q_3$ . Enfin la vente d'une unité de  $P_1$  rapporte un bénéfice de deux euros alors que celle d'une unité de  $P_2$  rapporte un bénéfice de un euro. L'objectif de l'usine étant la recherche d'un bénéfice maximal, comment doit-elle organiser sa production ? Ce problème s'écrit comme un problème d'optimisation : trouver  $\min(-2x_1 - x_2)$  sous les contraintes  $x_1 + 6x_2 \leq 30$ ,  $2x_1 + 2x_2 \leq 15$ ,  $4x_1 + x_2 \leq 24$ .

#### Inégalité

**Définition 5.3.3.** Soit  $\Omega$  un ouvert de  $\mathbb{R}^n$ . Soit  $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$  et soit  $\varphi : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ . On appelle problème d'optimisation avec contraintes d'inégalité le problème

Trouver  $x^* = \min_{x \in \Omega} f(x)$  sous les contraintes  $\varphi(x) \leq 0$ .

Soit  $\Omega$  un ouvert de  $\mathbb{R}^n$ . Soit  $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$  et soit  $\varphi : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ . On a alors

$$\forall x \in \Omega, \varphi(x) \leq 0 \iff \exists \alpha \in \mathbb{R}^m ; \varphi(x) + \alpha^2 = 0.$$

On considère alors la fonction de Lagrange

$$L : (x, \lambda, \alpha) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^m \mapsto L(x, \lambda, \alpha) := f(x) + \sum_{i=1}^m \lambda_i \varphi_i(x) + \alpha_i^2,$$

et on étudie le problème comme précédemment.

#### Appartenance à un ensemble

Les contraintes sont, dans ce cas, exprimées sous la forme  $x \in \mathcal{C}$ . On considère alors la fonction indicatrice  $\chi_{\mathcal{C}}$  de  $\mathcal{C}$  et on se ramène à un problème avec contraintes d'égalité de la forme  $\chi_{\mathcal{C}}(x) - 1 = 0$ .



# Bibliographie

- [1] Walter Appel Mathématiques pour la physique et les physiciens, H & K Éditions (2<sup>e</sup> édition), 2002
- [2] François Roddier Distributions et transformation de Fourier (à l'usage des physiciens et des ingénieurs) Ediscience, 1971