



# Analyse Non Standard et Représentation du Réel: Deux exemples en Automatique

Claude Lobry, Tewfik. Sari

► **To cite this version:**

Claude Lobry, Tewfik. Sari. Analyse Non Standard et Représentation du Réel: Deux exemples en Automatique. 2007.

**HAL Id: inria-00163365**

**<https://hal.inria.fr/inria-00163365>**

Submitted on 17 Jul 2007

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

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

# Analyse Non Standard et Représentation du Réal: Deux exemples en Automatique

C. Lobry<sup>1</sup>

T. Sari<sup>2</sup>

Juillet 2007

<sup>1</sup>**Université de Nice**, Département de Sciences de la Vie, Faculté des Sciences, Parc Valrose, 06000 Nice, et Equipe-Projet INRA-INRIA, **Modélisation Et Ressources en Eau** ; claude.lobry@inria.fr

<sup>2</sup>**Université de Haute Alsace**, Mathématiques, 2 rue des frères Lumière, F-68093 Mulhouse Cedex ; tewfik.sari@uha.fr

## Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Analyse Non Standard</b>	<b>3</b>
2.1	Les “grands” entiers . . . . .	4
2.2	Les infinitésimaux : Vocabulaire . . . . .	5
2.3	Les fonctions continues . . . . .	6
2.4	Regard sur quelques fonctions non standard . . . . .	7
2.4.1	Les fonctions trigonométriques rapidement oscillantes . . . . .	7
2.4.2	Les impulsions . . . . .	8
2.5	Commentaires historiques et bibliographiques . . . . .	10
<b>3</b>	<b>Le phénomène du “peaking”</b>	<b>10</b>
3.1	Stabilisation des systèmes en cascade . . . . .	11
3.2	Une vision externe du phénomène de “peaking” . . . . .	15
3.2.1	Stabilité . . . . .	15
3.2.2	Stabilité instantanée . . . . .	16
3.2.3	Stabilité des systèmes en cascade . . . . .	18
3.3	L’ANS et les équations différentielles . . . . .	20
<b>4</b>	<b>Théorie Nonstandard du bruit selon M. Fliess</b>	<b>21</b>
4.1	La théorie du Moiré de Harthong . . . . .	21
4.2	La théorie de la moyennisation de C. Reder . . . . .	25
4.2.1	Valeur apparente en un point . . . . .	25
4.2.2	Moyenne mobile . . . . .	27
4.3	Radically Elementary Probability Theory . . . . .	28
4.4	La théorie de la mesure de Cartier-Perrin . . . . .	30
4.5	La définition Non Standard du bruit de Michel Fliess . . . . .	32
<b>5</b>	<b>Dernières remarques</b>	<b>33</b>

## Avant propos

Ce texte a été rédigé par Tewfik Sari et moi même à partir des notes que j’avais préparées pour la conférence en l’honneur de Michel Fliess à laquelle j’avais eu le plaisir d’être invité. Voici l’introduction que j’avais écrite en vue de ma présentation orale :

*Il y a un peu plus de vingt ans Marc Diener et moi éditions un ouvrage collectif [13] dont j’ai repris le titre pour cette conférence car la question me semble toujours d’actualité. Dans les années dix neuf cent quatrevingt il était difficile d’évoquer l’Analyse Non Standard sans déclencher des passions : Il y avait les “contre” et les “pour” dont je faisais partie. Maintenant que la passion est moins forte il semble possible d’aborder cette question de façon un peu plus scientifique. C’est ce que je vais tenter de faire en analysant sur deux exemples choisis en*

automatique la capacité de l'Analyse Non Standard à traduire formellement des discours mathématiques informels.

*Le premier exemple que j'ai choisi est celui du phénomène de "peaking" et de son influence sur la stabilisation des systèmes; il est tiré d'un travail commun avec T. Sari [22, 24, 25] et je me sens assez à l'aise pour en parler. En revanche, je me sens bien moins à qualifié pour parler du second exemple puisqu'il touche au traitement du signal, discipline que je ne connais pas. J'ai cependant décidé d'en parler à cause de son actualité dans le cadre de ces journées en l'honneur de Michel Fliess*

*En effet, peu de temps avant le colloque j'ai découvert, avec un immense plaisir, qu'il a l'intention d'aborder des questions de traitement du signal en proposant une représentation du "bruit" par des fonctions "rapidement oscillantes", cette dernière expression étant définie de façon précise dans le cadre Non Standard. Comme je l'ai déjà dit, je ne connais rien au traitement du signal, donc je ne peux pas avoir un avis scientifiquement fondé sur l'avenir de cette idée. Ce qui ne m'empêche pourtant pas d'y croire et donc de vouloir lui faire de la publicité.*

*Qu'est ce donc qui me donne cette foi du charbonnier? C'est simplement le fait, maintes fois constaté au long de sa carrière remarquablement féconde, que chaque fois que Michel Fliess a eu une idée, cette dernière a fini par s'imposer. Je ne vois pas pourquoi il n'en serait pas de même encore une fois!*

Il y a très loin d'un texte préparé pour une conférence orale à un texte destiné à la publication. Le travail de mise en forme a été fait avec mon complice T. Sari. C'est pourquoi nous signons ensemble le texte qui suit.

Claude Lobry

## 1 Introduction

Quelques mots sur le titre : Pourquoi Analyse Non Standard (ANS) et Représentation du Réel? L'Analyse Non Standard est une façon de faire de l'analyse mathématique dans laquelle il est licite de dire d'un nombre réel qu'il est "infinitement-petit-fixé-une-fois-pour-toutes". Par "licite" nous voulons dire que du point de vue de la rigueur mathématique, selon les canons les plus modernes de la logique, on ne peut rien lui reprocher. Ainsi une démonstration, via l'ANS, d'un résultat "classique" peut-elle être transformée en une démonstration classique, ce qui a fait dire à certains que l'ANS n'apporte "rien de neuf". Ce n'est pas tout à fait exact. Ce que l'ANS apporte est une nouvelle manière de formaliser le discours mathématique informel dans un formalisme dont on peut démontrer qu'il n'est pas plus dangereux que le formalisme classique : Si il existait une contradiction dans les mathématiques formalisées dans l'ANS il en existerait une dans les mathématiques formalisées classiquement. Ainsi la question n'est elle pas de savoir si les mathématiques non standard sont rigoureuses mais de savoir si cette nouvelle manière de travailler est préférable à l'ancienne.

Le premier critère est celui de la fécondité : Combien de résultats nouveaux significatifs ont-ils été obtenus grâce à l'ANS? Combien de conjectures

anciennes démontrées ? Bien peu, pour ainsi dire pas du tout. Cela disqualifie-t-il définitivement la méthode ? Certainement pas. Il y a, à l'heure actuelle, au plus un mathématicien sur cent qui utilise l'ANS dans ses recherches. La disproportion est trop grande pour que des comparaisons globales aient un sens et, dans quelques domaines particuliers comme celui des équations différentielles ordinaires, le bilan est loin d'être nul.

Un second critère, moins décisif parce que moins objectif, est celui de l'élégance. L'élégance du style a toujours préoccupé les scientifiques, particulièrement les mathématiciens. Nous pensons que pour certaines questions le style ANS est plus élégant que le style classique et c'est sur ce terrain que nous nous plaçons dans cet article. Plus précisément nous essayerons de nous placer du point de vue d'un utilisateur non mathématicien, plus particulièrement de l'ingénieur automatique, pour qui les mathématiques sont un outil de représentation et de compréhension de phénomènes concrets.

En Automatique il s'agit d'observer un système en temps réel afin d'en déduire les lois de commande qui réaliseront l'objectif poursuivi. Pour cela on réalise un "modèle", c'est-à-dire une "représentation mathématique" du système réel qui nous intéresse. Nous pensons que, chaque fois que la réalité modélisée se présente avec des échelles d'observations différentes, une représentation à l'aide de l'ANS permet de simplifier considérablement le langage. Mais à quel prix direz-vous ? L'effort à consentir pour manipuler correctement les outils de l'ANS n'est-il pas disproportionné avec le bénéfice attendu ? La réponse est non. Au prix d'un effort minime, n'importe quel automatique peut se doter, en plus de ses outils traditionnels, d'un langage très adapté à l'expression de certaines idées. C'est du moins ce que nous essayons de démontrer sur deux exemples : La représentation des systèmes à deux échelles de temps, d'une part, la représentation du bruit, d'autre part. Notre objectif étant purement pédagogique aucun résultat proposé n'est vraiment original sauf, peut être, l'idée d'écrire systématiquement en caractères gras toutes les "notions externes" dans l'espoir de faciliter la lecture de ce formalisme. En tout état de cause, nous espérons que notre présentation poussera quelques lecteurs à se rendre compte par eux mêmes. En particulier ils pourront consulter l'ouvrage collectif [12] qui est en anglais, qui contient une grande partie de ce que nous exposons ici et surtout bien d'autres choses !

L'article est divisé en trois parties. La première présente, sans prérequis particulier, les outils d'ANS nécessaires ; les quelques exemples sont introductifs à des notions exposées dans les parties deux et trois. La seconde est consacrée au phénomène du "peaking", la troisième à la question du bruit.

## 2 Analyse Non Standard

L'Analyse Non Standard a été inventée par A. Robinson [32]. Nous nous plaçons dans le cadre de la théorie développée par E. Nelson [29]. Ce n'est pas la seule façon de faire mais c'est celle que nous connaissons !

Le système IST de Nelson [29] est un langage formel  $\mathcal{L}'$  qui contient le langage classique  $\mathcal{L}$  de la théorie des ensembles ZFC (pour Zermelo, Fraenkel +

axiome du choix). Il contient le prédicat (non défini dans  $\mathcal{L}$ ) qui s'écrit dans les formules  $st(x)$  et se prononce “ $x$  est standard”. Une série de trois axiomes I pour “Idéalisation”, S pour “Standardisation” et T pour “Transfert” régleme l'usage de “st”. Nelson démontre que la théorie IST est “conservative” et donc irréprochable du point de vue de la rigueur : si il existait une contradiction dans IST on la retrouverait automatiquement dans la théorie ZFC Il n'est pas question dans ce texte de détailler ces axiomes et leurs conséquences les plus immédiates mais il faut insister sur le “double langage” que permet le système de Nelson, dans la pratique qui a été largement répandue par G. Reeb et son école. Dans  $\mathcal{L}$  on distingue naturellement les formules qui ne contiennent pas le prédicat “st” et qui sont dites “internes” (sous entendu à la théorie des ensembles) des formules qui le contiennent qui sont dites “externes”. Bien entendu quand on pratique les mathématiques on ne s'exprime pas en formules mais dans une langue technique assez précise (au moins quand les choses sont bien faites) pour indiquer le cheminement qui permettrait la formalisation du texte. Ainsi, dès les premiers mois d'université, sait-on que la phrase “lorsque  $t$  tend vers 0 la fonction  $f$  tend vers  $l$ ” se formalise en :

$$\forall \epsilon > 0 \exists \eta > 0 \forall t \{ |t| \leq \eta \implies |f(t) - l| \leq \epsilon \} \quad (1)$$

Mais tout l'art du formalisme bien compris est de savoir faire allusion aux formules qui pourraient être écrites, tout en ne les utilisant que le moins possible.

Ainsi, dans le système de Nelson, il est essentiel, quand on utilise des expressions, de savoir si elles sont internes ou externes et de prendre garde que ce qui dans le système classique était une façon laxiste de parler peut éventuellement devenir tout à fait rigoureux en ce sens que le chemin vers la formalisation a été clairement indiqué. Seule la lecture d'un manuel de cours (par exemple : [10, 14, 27] et un peu d'expérience peuvent familiariser avec cette pratique. Toutefois il est assez facile d'en donner une idée comme nous allons le faire maintenant.

## 2.1 Les “grands” entiers

De l'axiome de transfert T il découle que 0 est standard et que le successeur d'un entier standard est standard. Donc 1 est standard, 2 est standard, etc... De l'axiome d'idéalisation I on déduit la formule <sup>1</sup> :

$$\exists \omega \in \mathbb{N} : \forall^{st} n \in \mathbb{N} \quad n \leq \omega \quad (2)$$

Cet entier  $\omega$  qui est plus grand que tous les entiers standard mérite le nom de “infiniment grand” ainsi que tous ceux qui comme lui seront plus grands que tous les entiers standard. Donc “infiniment grand” n'est pas une “manière de parler” devant conduire à une formalisation classique, mais un “résumé” d'une formule externe (la formule (2) est externe). Dans l'espoir de rendre les choses plus claires, dans cet exposé, nous utiliserons les caractères gras pour désigner

<sup>1</sup>On utilise l'abréviation :  $\forall^{st} x \ A(x)$  pour :  $\forall x \ \{st(x) \implies A(x)\}$

les formulations dans la langue naturelle qui ont une traduction formelle externe immédiate dans le système IST. Ainsi dire que  $\omega$  est un entier **infiniment grand** est rigoureusement équivalent à la formule externe

$$\forall^{st} n \in \mathbb{N} \quad n \leq \omega \quad (3)$$

et bien entendu nous lui donnons toute la connotation affective qui accompagne l'idée de "plus grand que toute chose". Le choix de cette terminologie peut être critiqué et a été critiqué. En effet dans le langage classique, que nous tenons à conserver, un ensemble est "fini" lorsqu'il ne peut pas être mis en bijection avec une partie propre de lui même, ce qui reste le cas pour le sous ensemble de  $\mathbb{N}$  :

$$[0, 1, 2, \dots, \omega]$$

même si  $\omega$  est **infiniment grand**. Ainsi, dans le système de Nelson, la phrase :

Soit  $\omega$  un entier **infiniment grand** : l'ensemble  $[0, 1, 2, \dots, \omega]$  est fini.

est parfaitement correcte.

C'est troublant, mais pas plus que pour l'étudiant qui découvre qu'une boule peut être à la fois ouverte et fermée. Il aurait peut être été sage de choisir une autre terminologie <sup>2</sup> mais la tradition à laquelle nous appartenons en a décidé ainsi et nous nous y conformons. Elle a l'avantage de souligner que ce qui est métaphysique et troublant ce n'est pas tant l'existence de "grands entiers" mais celle d'ensembles infini dans la théorie des ensembles. Ce qui nous renvoie aux discussions du début du vingtième siècle sur "l'infini actuel" et à la possibilité de fonder l'Analyse Non Standard dans une tradition intuitioniste, ce qui est un autre sujet.

Notons pour terminer ce fait important que des formules externes peuvent "désigner" des ensembles au sens naïf du terme mais qui ne sont pas des "vrais" ensembles de la théorie formelle des ensembles. Ainsi la formule externe :

$$n \in N \wedge st(n)$$

désigne l'ensemble naïf des entiers standards qui *n'est pas un ensemble* au sens formel. Ce fait est à l'origine d'un outil puissant dont nous reparlerons : les "principes de permanence".

## 2.2 Les infinitésimaux : Vocabulaire

Oublions de ce qui précède les considérations métaphysiques et passons à l'infiniment petit qui véhicule, nous semble-t-il, une charge émotionnelle moindre que l'infiniment grand. Un réel sera dit **infiniment grand** s'il est plus grand qu'un entier **infiniment grand** et un **infiniment petit** sera un réel nul ou, sinon, tel que son inverse soit **infiniment grand**. Deux nombres sont **infiniment proches** si leur différence est **infiniment petite**. Un nombre qui n'est

<sup>2</sup>Certaines traditions tiennent pour "hyper fini" on a aussi proposé **I-grand** pour "idéalement grand"

pas **infiniment grand** est **limité**, un nombre **limié** qui n'est pas **infiniment petit** est **appréciable**. On désigne également par  $hal(x)$  et on appelle **halo de x** l'ensemble (naïf) des réels qui sont **infiniment proches** de  $x$ .

Nous pouvons donc travailler avec des infinitésimaux en acte. Ainsi, dans le système IST il est légal de dire : Soit  $\epsilon$  un réel strictement positif fixé une fois pour toute. La somme de deux **infiniment petits** est bien un **infiniment petit** et le fameux paradoxe des infinitésimaux :

Si  $\epsilon$  est un infinitésimal,  $2\epsilon$  doit l'être aussi, et ainsi de suite ; soit  $n_0$  le dernier entier tel que  $n_0\epsilon$  soit infinitésimal ; la somme des deux infinitésimaux  $n_0\epsilon + \epsilon$  n'est pas un infinitésimal, ce qui pose un problème.

qui avait été contourné au dix neuvième siècle par la méthode du "passage-à-la-limite" l'est dans IST par l'interdiction qui est faite de considérer l'ensemble "naïf" des entiers tels que  $n\epsilon$  est **infiniment petit** comme un ensemble "formel" de la théorie des ensembles.

### 2.3 Les fonctions continues

Les infinitésimaux sont là pour faire de l'analyse, donc, en particulier, nous parler de la continuité. La définition suivante est une définition externe.

**Définition 2.1** (S-continuité) *Une fonction  $f$  est **S-continue** si  $x$  **infiniment proche** de  $y$  implique  $f(x)$  **infiniment proche** de  $f(y)$ , c'est à dire <sup>3</sup> :*

$$\forall x \forall y : x \approx y \implies f(x) \approx f(y)$$

Pour comprendre le sens de cette définition nous nous donnons un nombre réel  $a$  strictement positif et considérons la fonction constante par morceaux définie par :

$$x \in [ka, (k+1)a[ \implies f(x) = ka \quad (4)$$

Cette fonction n'est pas continue, comme chacun sait. Elle est standard si le paramètre  $a$  est standard, elle ne l'est pas sinon. Il est facile de voir que  $f$  est **S-continue** si le nombre  $a$  est infinitésimal. En effet on a clairement

$$|f(x) - f(y)| \leq |x - y| + a$$

et donc si  $a$  est **infiniment petit**  $f(x)$  est bien **infiniment proche** de  $f(y)$  dès que  $x$  et  $y$  le sont. En revanche, si  $a$  est standard cette fonction n'est pas **S-continue**. Maintenant nous considérons une fonction standard **S-continue**. Nous allons montrer qu'elle est uniformément continue.

– Montrons la formule :

$$\forall^{st} \epsilon \exists \eta \forall x \forall y \{ |x - y| \leq \eta \implies |f(x) - f(y)| \leq \epsilon \} \quad (5)$$

Il suffit de prendre  $\eta$  **infiniment petit** quelconque. Si  $|x - y| \leq \eta$ ,  $x$  et  $y$  sont **infiniment proches** donc  $f(x)$  et  $f(y)$  aussi et leur différence, **infiniment petite**, est inférieure au standard  $\epsilon$ .

<sup>3</sup>Nous introduisons la notation  $x \approx y$  pour dire que  $x$  est **infiniment proche** de  $y$ .



– Par transfert, la formule (5) est équivalente à la formule :

$$\forall \epsilon \exists \eta \forall x \forall y \{ |x - y| \leq \eta \implies |f(x) - f(y)| \leq \epsilon \} \quad (6)$$

que nous reconnaissons comme la définition de la continuité uniforme. L'axiome de transfert est applicable parce que  $f$ , qui est un paramètre dans la formule, est standard.

Nous comprenons que cette preuve est quelque peu obscure dans la mesure où nous n'avons jamais énoncé l'axiome de transfert ! Nous l'avons écrite juste pour donner une idée de ce genre de preuves. La première partie de la démonstration se comprend très facilement. La seconde partie n'est que du non sens formel : on applique l'axiome. Il n'y a rien à comprendre !

Réciproquement nous aurions pu montrer que si  $f$  est standard et uniformément continue elle est **S-continue**. Pour les fonctions standard la continuité uniforme equivaut donc à la **S-continuité**.

Nous commençons donc à voir fonctionner le jeu non standard. La même définition externe (ici **S-continue**) est équivalente à la définition classique (continue) lorsqu'elle est appliquée à un objet standard (une fonction standard) et est porteuse de son propre sens lorsqu'elle s'applique à un objet non standard, ici une fonction constante par morceau de pas infinitésimal  $a$  ; une telle fonction “semble” effectivement continue si son graphe est regardé de suffisamment loin. La fonction “semble continue”, d'où le **S** de **S-continue**.

## 2.4 Regard sur quelques fonctions non standard

Nous nous intéressons à quelques fonctions non standard particulières qui reviendront dans la suite. Pour “observer”<sup>4</sup> une fonction non standard il suffit de considérer une famille à un paramètre de fonctions et de fixer le paramètre à une valeur non standard, par exemple **infinitement grande**. Attention à ne pas confondre une fonction non standard *qui doit être une vraie fonction au sens classique, c'est à dire définie par une formule interne* avec une correspondance externe comme, par exemple,

$$f(x) = 1 \text{ si } x \text{ est } \mathbf{infinitement petit}, f(x) = 0 \text{ sinon,}$$

qui n'est pas une fonction au sens de ZFC

### 2.4.1 Les fonctions trigonométriques rapidement oscillantes

Soit la fonction :

$$t \rightarrow \sin(\omega t)$$

Si  $\omega$  est non **infinitement grand**, cette fonction est **S-continue**. En revanche, si  $\omega$  est **infinitement grand** le nombre  $\frac{\pi}{2\omega}$  est **infinitement proche** de 0 et  $\sin(\frac{\pi}{2})$

<sup>4</sup>Entendons nous sur “observer” ! Il a été reproché à l'ANS de ne pas être capable de montrer explicitement un réel infinitement petit. C'est parfaitement exact (mais cela ne prouve pas l'inutilité de l'ANS car à ce compte le nombre  $\pi$  serait bien inutile). Notre “observer” doit donc être pris dans un sens assez faible !

n'est pas **équivalent** à  $\sin(0)$ . Calculons l'intégrale :

$$I_\delta(x) = \frac{1}{\delta} \int_x^{x+\delta} \sin(\omega t) dt \quad (7)$$

Soit :

$$I_\delta(x) = \frac{-1}{\omega\delta} (\cos(\omega(x+\delta)) - \cos(\omega x)) \quad (8)$$

Nous voyons donc que si  $\omega$  est **infiniment grand** l'intégrale  $I_\delta(x)$  est toujours **infiniment petite** dès que  $\omega\delta$  est **infiniment grand**, donc en particulier pour tout  $\delta$  non **infiniment petit**. Dans les cours de physique élémentaires on nous explique que la moyenne mobile ci-dessus est un "filtre passe bas", c'est à dire que si l'on considère qu'une intensité de 0.01 est "imperceptible" (ou "équivalente à zéro") toutes les fréquences supérieures à  $\omega_0$  avec  $\omega_0 = \frac{2}{0.01\delta}$  sont coupées. Comme le seuil 0,01 peut varier suivant les circonstances un énoncé mathématique idéal, un peu universel serait le bienvenu. En voici un possible dans le système IST. Une moyenne mobile de largeur non **infiniment petite** filtre toutes les fréquences **infiniment grandes**

Mais reconnaissons tout de suite que cette affaire n'est pas encore très convaincante. Sans faire appel à l'Analyse Non Standard nous avons l'énoncé classique :

$$\omega \geq \frac{2}{\epsilon\delta} \implies I_\delta(x) \leq \epsilon$$

parfaitement clair et surtout bien plus précis. D'une manière générale l'ANS, comme toutes les formulations asymptotiques classiques, est inutile chaque fois que nous avons une majoration explicite. Mais c'est loin d'être toujours le cas.

#### 2.4.2 Les impulsions

Dans cette sous section nous nous penchons sur une classe de fonctions bien connue des physiciens, les fonctions de Dirac. Rappelons que pour les physiciens, avant la popularisation de la théorie des distributions, la "fonction de Dirac en 0" était la fonction nulle en dehors de 0, infinie en 0 et dont l'intégrale sur les réels vaut 1. Dans le cadre de l'ANS on dira :

**Définition 2.2** (Impulsion de Dirac) *On dit qu'une fonction intégrable  $f$  appartient à la classe des fonctions de Dirac en 0 si il existe un **infiniment petit**  $\delta > 0$  tel que :*

$$|x| \geq \delta \implies f(x) \approx 0$$

$$\int_{\mathbb{R}} f(x) dx \approx 1$$

On remarque que la "classe" des fonctions de Dirac, définie par une formule externe, n'est pas un ensemble de fonctions. On ne parlera donc pas de "classe d'équivalence". Une des plus célèbres parmi les fonctions de Dirac en 0 est la fonction de Gauss :

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

considérée pour une valeur **infiniment petite** de  $\sigma$ . Une autre fonction que nous utiliserons par la suite est la fonction définie par :

$$f(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \omega^2 t e^{-\omega t} & \text{si } x > 0 \end{cases}$$

qui pour  $\omega$  **infiniment grand** est aussi une impulsion de Dirac en 0. Ces fonctions non standard sont des analogues externes des distributions classiques ; on vérifie très facilement la :

**Proposition 2.3** *Si  $f$  est une fonction de la classe de Dirac en 0, pour toute fonction standard  $\psi$  continue à support compact on a :*

$$\int_{\mathbb{R}} f(x)\psi(x)dx \approx \psi(0)$$

En fait la classe des fonctions non standard contient tous les objets les plus fous dont un analyste peut rêver.

Cette façon de voir les choses permet de clarifier très facilement certaines questions. Par exemple celle du contrôle impulsionnel pour des systèmes non linéaires. On sait que si on considère le système non linéaire :

$$\frac{dx}{dt} = u_1 X^1(x) + u_2 X^2(x) \quad (9)$$

il est problématique de définir ce que sont ses solutions lorsque les entrées sont des distributions. On comprend immédiatement pourquoi en prenant les deux fonctions de Dirac :

$$u_1(t) = \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{1}{\epsilon} & \text{si } 0 < t < \epsilon \\ 0 & \text{si } \epsilon \leq t \end{cases}$$

$$u_2(t) = \begin{cases} 0 & \text{si } t < \epsilon \\ \frac{1}{\epsilon} & \text{si } \epsilon \leq t < 2\epsilon \\ 0 & \text{si } 2\epsilon \leq t \end{cases}$$

Pour  $\epsilon$  **infiniment petit** ces deux entrées sont des “gentilles fonctions”, (quoi que non standard) et rien n’empêche d’intégrer (9). Il est facile de voir que le résultat est, avec les notations usuelles sur les groupes à un paramètres engendrés par un champ de vecteurs :

$$X_1^2 \circ X_1^1(x_0) \quad (10)$$

Si les deux champs de vecteurs  $X^1$  et  $X^2$  ne commutent pas l’échange de  $u_1$  avec  $u_2$  change le résultat de l’intégration. C’est pour cette raison qu’il n’est pas possible de mettre dans le même sac toutes les fonctions de Dirac, sauf dans le cas commutatif. Cela dit il n’est pas interdit d’utiliser des fonctions de Dirac comme entrées, à condition de faire attention !

## 2.5 Commentaires historiques et bibliographiques

Faire une histoire de l'Analyse Non Standard nous entraînerait beaucoup trop loin. Donnons seulement quelques éléments d'histoire récente. Le livre de Robinson [32], considéré comme le créateur de l'ANS moderne, est publié en 1966. Une importante école mathématique travaille dans le formalisme développé dans ce livre. Il n'en est pas question ici en raison de notre ignorance.

Au début des années soixante-dix le mathématicien français G. Reeb, connu pour ses travaux en géométrie (le “feuilletage de Reeb”) découvre le livre de Robinson et n'a de cesse de faire partager son enthousiasme à tout son entourage. Sa position radicale, rapidement connue à travers le slogan : *Les entiers naïfs ne remplissent pas  $\mathbb{N}$* , ne facilite pas toujours le dialogue. Il crée cependant autour de lui à Strasbourg une école qui décidera d'adopter le point de vue de l'article IST de 1977 de Nelson [29]. Les premiers travaux de l'école de Reeb sont synthétisés en partie dans le livre de 1982 de R. Lutz et M. Goze [27] puis dans deux ouvrages collectifs : [13] puis [12] auxquels nous nous référerons beaucoup. La méthode avait été présentée devant une partie de la communauté française des mathématiques appliquées lors du colloque d'analyse numérique de 1981 [20] et du colloque de Belle Ile de 1983 [16].

L'existence d'une école d'analyse nonstandard en France a suscité une violente polémique dont on peut trouver un compte rendu (bien entendu tout à fait partisan !) dans le livre [21]. Des points de vue plus philosophiques, donc en principe moins passionnés, se trouvent dans les deux ouvrages collectifs [1, 33]. Sur des questions philosophiques soulevées par l'ANS (et autres questions) nous recommandons très vivement les sites de J. Harthong [<http://moire4.u-strasbg.fr/>] et G. Wallet [<http://perso.univ-lr.fr/gwallet/>].

Il existe beaucoup d'ouvrages de cours d'initiation à l'ANS. Nous renvoyons à [13, 12] pour une initiation rapide et à la bibliographie de [12] pour des ouvrages plus fouillés.

## 3 Le phénomène du “peaking”

En automatique on aime bien stabiliser les systèmes le plus rapidement possible. Si le système est linéaire on essaye de placer très à gauche dans le plan complexe les valeurs propres mais il est bien connu que ce n'est pas sans danger en présence de perturbations, même petites mais non linéaires : Loin d'être stabilisé le système peut “exploser à l'infini”. En 1989 M. Canalis et P. Yalo ont remarqué un article [19] de Kokotovic et Marino qui mettait en garde sur ce problème. Ils ont publié un court papier [8] qui abordait cette question et mettait en évidence une des raisons profondes de l'instabilité aux perturbations non linéaires : avant de tendre vers zéro, une trajectoire peut faire une excursion “au voisinage de l'infini”. De façon indépendante, Sussman et Kokotovic ont popularisé ce phénomène sous le nom de “peaking phenomenon” dans un article [34] qui en plus de décrire le phénomène donne des conditions pour s'en prémunir, ce qui évidemment est bien plus difficile que de simplement le constater comme

dans [8]. Nous allons voir comment l’ANS permet de parler de ce phénomène.

### 3.1 Stabilisation des systèmes en cascade

Rappelons ce vieux piège qu’il convient d’éviter. On s’intéresse aux systèmes :

$$\begin{cases} \frac{dx}{dt} = f(x, y) \\ \frac{dy}{dt} = g(y) \end{cases} \quad (11)$$

tels que le système :

$$\frac{dx}{dt} = f(x, 0) \quad (12)$$

soit globalement asymptotiquement stable en 0 (GAS en 0) ainsi que le système :

$$\frac{dy}{dt} = g(y) \quad (13)$$

Le raisonnement ci-dessous, un peu trop rapide comme nous allons le voir, conduit à conclure à la stabilité du système global :

*Puisque  $y(t)$  tend vers 0 le système :*

$$\frac{dx}{dt} = f(x, y(t)) \quad (14)$$

*tend vers le système autonome (12) qui est GAS en 0.*

Il n’en est rien comme le montre l’exemple bien connu :

$$\begin{cases} \frac{dx}{dt} = -x(1 - xy) \\ \frac{dy}{dt} = -y \end{cases} \quad (15)$$

On vérifie immédiatement que :

$$t \longrightarrow \begin{cases} x(t) = e^t \\ y(t) = 2e^{-t} \end{cases} \quad (16)$$

est une solution qui sépare le plan des phases comme indiqué sur la figure 1.

On peut penser que la taille du bassin d’attraction dépend de la vitesse avec laquelle  $y$  tend vers 0. Plus vite on tend vers 0 plus grand serait le bassin d’attraction. On peut le constater sur la famille de systèmes :

$$\begin{cases} \frac{dx}{dt} = -x(1 - xy) \\ \frac{dy}{dt} = -\gamma y \end{cases} \quad (17)$$

et faire tendre  $\gamma$  vers l’infini. Pour ce système la séparatrice est la trajectoire définie par :

$$t \longrightarrow \begin{cases} x(t) = e^{\gamma t} \\ y(t) = (\gamma + 1)e^{-\gamma t} \end{cases} \quad (18)$$

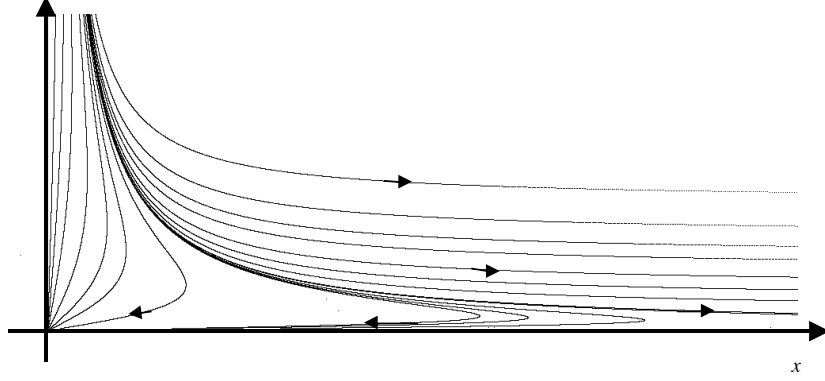


FIG. 1 – Portrait de phase de (15)

On voit que le bassin d’attraction de (17) grossit indéfiniment quand  $\gamma$  tend vers l’infini. Toutefois ceci n’est pas général à cause du phénomène de “peaking” que nous exposons maintenant. Considérons le système :

$$\begin{cases} \frac{dx}{dt} = -\frac{x^3}{2}(1+y_2) \\ \frac{dy_1}{dt} = y_2 \\ \frac{dy_2}{dt} = -\gamma^2 y_1 - 2\gamma y_2 \end{cases} \quad (19)$$

Le système linéaire en  $y_1, y_2$  ne dépend pas de  $x$ . Il agit comme un terme forçant sur l’équation en  $x$ . Les valeurs propres du système linéaire sont confondues et égales à  $-\gamma$ . Le système linéaire a ses solutions qui tendent vers 0 d’autant plus vite que  $\gamma$  est grand. On peut intégrer explicitement ce système d’équations et l’on obtient :

$$x(t, \gamma) = \frac{x_0}{\sqrt{1+x_0^2[t-y_{10}+(y_{10}(\gamma t+1)+y_{20}t)e^{-\gamma t}]}} \quad (20)$$

Prenons pour condition initiale  $y_{10} = 1$  et  $y_{20} = 0$ , il vient :

$$x(t, \gamma) = \frac{x_0}{\sqrt{1+x_0^2[t-1+(\gamma t+1)e^{-\gamma t}]}} \quad (21)$$

On voit sur cette formule que lorsque  $\gamma$  est grand, pour  $t$  petit et non nul, la quantité  $[t-1+(\gamma t+1)e^{-\gamma t}]$  est très proche de  $-1$  et donc pour  $x_0^2 > 1$  la quantité sous le radical est négative ce qui veut dire que la solution  $x(t)$  a “explosé à l’infini” avant. Sur la figure (2) on peut voir les solutions de condition initiale  $x_0 = 1, 4$ ,  $y_{10} = 1$  et  $y_{20} = 0$  pour des valeurs de plus en plus grandes de  $\gamma$ . Pour comprendre ce qui s’est passé il suffit de considérer  $y_1(t)$  et  $y_2(t)$  pour

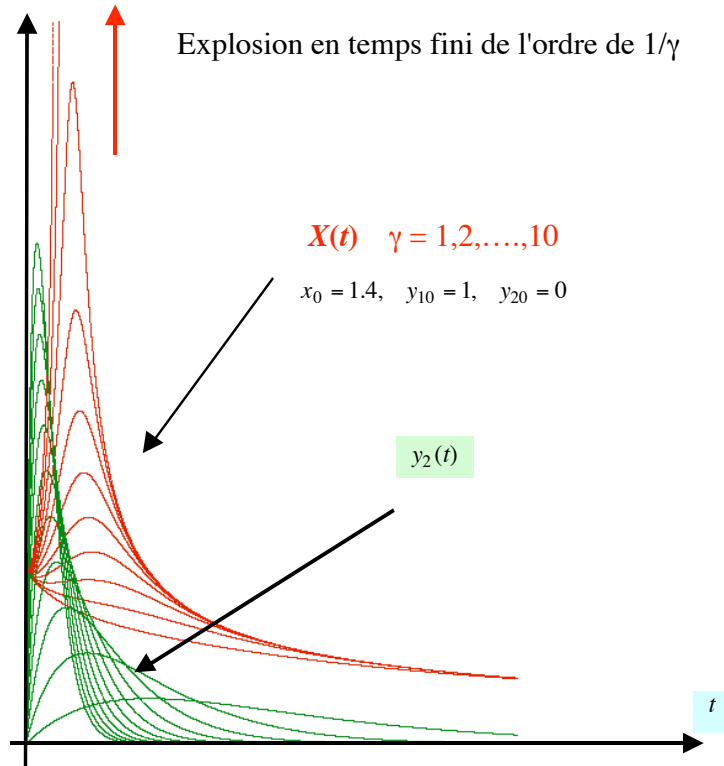


FIG. 2 – Le “peaking” dans l’équation (19)

les conditions initiales précédentes, soit :

$$t \longrightarrow \begin{cases} y_1(t) &= (\gamma t + 1)e^{-\gamma t} \\ y_2(t) &= -\gamma^2 t e^{-\gamma t} \end{cases} \quad (22)$$

La fonction  $y_2(t)$  qui rentre dans l’expression de la première équation :

$$\frac{dx}{dt} = -\frac{x^3}{2}(1 + y_2(t)) \quad (23)$$

est une fonction nulle pour  $t = 0$ , négative, dont le minimum est atteint pour  $t = \frac{1}{\gamma}$  et vaut  $-\frac{\gamma}{e}$ . Pour  $\gamma$  assez grand ce minimum est plus petit que  $-2$  et, donc, pendant un certain temps nous avons :

$$\frac{dx}{dt} > \frac{x^3}{2}$$

qui peut “exploser à l’infini”. Le fait que  $y_2(t)$  soit une entrée de (23) de plus en plus proche d’une impulsion et que le champ de (23) ne soit pas complet est

à l’origine de l’instabilité. Le “peaking” est donc, dans ce cas, le fait que quand on considère le système :

$$\begin{cases} \frac{dy_1}{dt} = & y_2 \\ \frac{dy_2}{dt} = & -\gamma^2 y_1 - 2\gamma y_2 \end{cases} \quad (24)$$

si nous voulons que, partant de la condition initiale  $y_1 = 0$  avec une vitesse nulle rejoindre  $(0, 0)$  en un temps de l’ordre de  $\frac{1}{\gamma}$  il faudra bien avoir, à un moment donné, une vitesse négative, en valeur absolue, de l’ordre de  $\gamma$ . Sur les figures 3

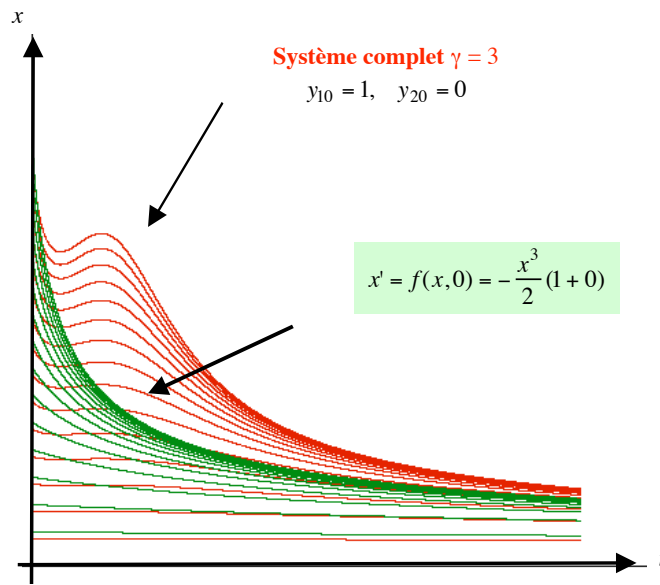


FIG. 3 – Le “peaking” dans l’équation (19)

et 4 nous représentons, pour deux valeurs successives de  $\gamma$  et pour des conditions initiales de plus en plus grandes en  $x_0$  en vert les solutions du système limite, en rouge les solutions du système complet. Tout ceci n’est pas très mystérieux mais assez compliqué à formaliser en vue d’une étude générale. On pourrait, par exemple, dire que la famille de systèmes linéaires :

$$\frac{dY}{dt} = A(\gamma)Y \quad (25)$$

“fait du peaking” quand  $\gamma$  tend vers l’infini si, d’une part, les parties réelles des valeurs propres de  $A(\gamma)$  tendent vers  $-\infty$  quand  $\gamma$  tend vers l’infini (pour dire que les solutions tendent de plus en plus vite vers 0) et, d’autre part, il existe pour tout  $\gamma$  au moins une condition initiale de norme 1 telle que le maximum de la trajectoire correspondante soit plus grand qu’une fonction de  $\gamma$  qui tend vers



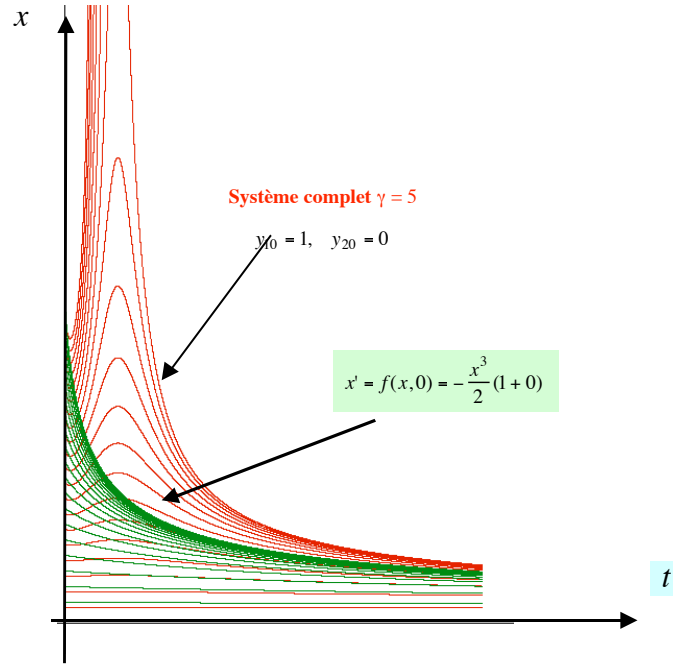


FIG. 4 – Le “peaking” dans l’équation (19)

l’infini avec  $\gamma$ . C’est compliqué, sans être insurmontable puisque c’est ce qui est fait dans ([34]). Notons cependant que le système (25) est linéaire. Comment faire lorsque ce n’est pas le cas? Voyons comment ces choses se disent à l’aide de l’ANS.

### 3.2 Une vision externe du phénomène de “peaking”

#### 3.2.1 Stabilité

**Définition 3.1** (S-Globalement Asymptotiquement Stable) *Le système différentiel :*

$$\frac{dx}{dt} = f(x)$$

*est S-Globalement Asymptotiquement Stable en 0 (S-GAS) si pour toute condition initiale  $x_0$  limitée et tout  $t$  infiniment grand  $x(t, x_0)$  est infiniment petit (où  $x(t, x_0)$  désigne la trajectoire de condition initiale  $x_0$ ).*

Cette définition externe de la stabilité asymptotique, ainsi désignée par **S-GAS** pour la distinguer de la stabilité asymptotique globale classique, est équivalente

pour les systèmes standard à la définition classique car on peut facilement démontrer le

**Théorème 3.2** *Si le système différentiel*

$$\frac{dx}{dt} = f(x)$$

*est standard il est **S-GAS** en 0 si et seulement si les deux propriétés suivantes sont satisfaites :*

- *Le système est stable en 0, c’est à dire : pour tout voisinage  $\mathcal{V}$  de 0 il existe un voisinage  $\mathcal{W}$  de 0 tel que pour toute condition initiale  $x_0$  dans  $\mathcal{W}$  la demi trajectoire positive issue de  $x_0$  est contenue dans  $\mathcal{V}$  ou, si l’on préfère avec des quantificateurs :*

$$\forall \mathcal{V} \exists \mathcal{W} : \forall x_0 \in \mathcal{W}, \forall t > 0 x(t, x_0) \in \mathcal{V}$$

- *Le point 0 est attractif, c’est à dire que pour toute condition initiale la trajectoire issue de cette condition initiale tend vers 0.*

Le théorème peut se paraphraser en : Un système standard est Globalement Asymptotiquement Stable si et seulement si il est **S-Globalement Asymptotiquement Stable**. Donc, grâce à ce théorème, nous nous retrouvons avec la **S-GAS** dans la même situation que pour la **S-continuité** : C’est une définition externe qui coïncide, lorsque le système différentiel est standard, avec la notion de stabilité asymptotique globale. On remarque aussi que la définition de **S-GAS** est spectaculairement plus compacte que celle de la stabilité asymptotique globale.

Sur l’exemple suivant nous voyons ce que peut vouloir dire **S-GAS** dans le cas d’un système qui n’est pas standard. Considérons le système défini en coordonnées polaires par les équations :

$$\begin{aligned} \frac{d\theta}{dt} &= 1 \\ \frac{d\rho}{dt} &= \rho(\epsilon - \rho) \end{aligned} \tag{26}$$

C’est un système qui, pour  $\epsilon$  **infinitement petit**, possède un cycle limite globalement asymptotiquement stable de rayon **infinitement petit**. Donc, pour  $t$  **infinitement grand**, la trajectoire se trouve **infinitement proche de 0**, mais à l’extérieur d’un cercle de rayon  $\epsilon$ . Donc le système n’est pas stable *stricto sensu*. Nous ne chercherons pas à démontrer le théorème (3.2) et nous renvoyons à [22] pour sa preuve qui est un exercice de routine.

### 3.2.2 Stabilité instantanée

Nous abordons maintenant la question de la mesure de la vitesse avec laquelle un système dynamique GAS en 0 tend vers 0. Ce n’est, à priori, pas très simple à formuler puisque, en principe, la trajectoire n’atteint jamais 0. En principe, le temps est toujours infini. Mais ceci est une vision mathématique “idéale”, dans la pratique, tout système réel finit effectivement par se trouver

en sa position d'équilibre, “à la précision des mesures près”. Le seul cas où il y a une réponse simple est le cas linéaire où, grâce à l'invariance par homothétie, on peut définir le “temps caractéristique” comme le temps nécessaire pour que la norme soit divisée par deux. Rien de tel n'est possible en non linéaire. Grâce aux formulations externes les choses peuvent être dites simplement.

**Définition 3.3** (S-Instantanément Globalement Asymptotiquement Stable(IGAS))

Le système

$$\frac{dx}{dt} = f(x)$$

est **S-Instantanément Globalement Asymptotiquement Stable en 0** (**S-IGAS en 0**) si pour toute condition initiale  $x_0$  **limitée** et tout  $t > 0$  **non infiniment petit** on a  $x(t, x_0)$  **infiniment petit**.

Le sens de cette définition est clair : toute condition initiale “raisonnable” est transférée instantanément presque en 0. Pour voir ce qu'est l'équivalent classique de cette définition externe il faut procéder ainsi. On commence par se donner une famille de systèmes dynamiques :

$$\frac{dx}{dt} = f(x, \gamma) \tag{27}$$

où  $\gamma$  est un réel positif.

**Définition 3.4** (semble-IGAS) *On dit que la familles de systèmes (27) “semble Instantanément Globalement Asymptotiquement Stable en 0 quand  $\gamma$  tend vers l'infini” (semble-IGAS en 0) si :*

$$\forall R > 0 \forall r > 0 \forall t_0 > 0 \exists \gamma_0 > 0$$

tel que

$$\forall \gamma > \gamma_0 \forall x_0 \leq R \forall t > t_0 \quad x(t, x_0) < r$$

Nous pouvons paraphraser cette définition par :

Pour tout ensemble borné de conditions initiales et toute boule  $B$  centrée en 0 de rayon arbitrairement petit, lorsque  $\gamma$  tend vers l'infini, les trajectoires pénètrent de plus en plus vite dans la boule  $B$ .

ce qui est (peut être?) un peu plus parlant. Alors nous pourrions démontrer le théorème suivant :

**Théorème 3.5** *On suppose que la famille (27) est standard. Alors la famille de systèmes (27) semble-IGAS en 0 quand  $\gamma$  tend vers l'infini si et seulement si pour tout  $\gamma$  **infiniment grand** le système (27) est **S-IGAS**.*

Donc, grace à ce théorème (ou d'autres du même type) tout ce que nous pourrions démontrer sur des *systèmes* non standard en terme de stabilité instantané aura une contrepartie classique sur des propriétés asymptotiques, lorsque le paramètre tend vers l'infini, de *familles de systèmes*.

### 3.2.3 Stabilité des systèmes en cascade

Dans le paragraphe introductif nous avons vu sur l'exemple (19) qu'une cause de non stabilité était le fait que des solutions “explosaient à l'infini” très rapidement quand  $\gamma$  augmente. Il suffit d'empêcher ce phénomène pour obtenir la stabilité. Soit le système, non nécessairement standard :

$$\begin{aligned} \frac{dx}{dt} &= f(x, y) \\ \frac{dy}{dt} &= g(y) \end{aligned} \quad (28)$$

**Définition 3.6** (Uniformément Infinitésimalement Borné) *Le système (28) est **Uniformément Infinitésimalement Borné (UIB)** si pour toute condition initiale **limitée**  $(x_0, y_0)$  et tout  $t$  **infinitésimal** la composante en  $x$  de la solution correspondante est **limitée**.*

Nous pouvons maintenant énoncer le

**Théorème 3.7** *Supposons que le système (28) soit **UIB**, que le sous système :*

$$\frac{dx}{dt} = f(x, 0)$$

*soit **S-GAS** en 0 et que sous système :*

$$\frac{dx}{dt} = g(y)$$

*soit **I-GAS** en 0, alors (28) est **S-GAS** en 0*

La démonstration est très simple. Nous commençons par montrer que pour toute condition initiale  $(x_0, y_0)$  **limitée** existe un réel  $t_1$  **appréciable** (non **infinitement petit**) tel que  $x(t_1, x_0, y_0)$  est **limité**. Pour montrer ce résultat nous utilisons une technique connue sous le nom “principe de permanence” qui remonte à Robinson.

**Lemme 3.8 (Robinson)** *Soit  $u_n$  une suite telle que pour tout  $n$  standard on ait  $u_n$  **infinitement petit**, alors il existe un  $\omega$  **infinitement grand** tel que  $u_\omega$  soit **infinitement petit**.*

**Preuve** On considère l'ensemble :

$$A = \{n \in \mathbb{N} : nu_n \leq 1\}$$

Il convient de noter que  $A$  est un véritable ensemble, au sens formel. L'implication suivante est évidente :

$$\forall n \in \mathbb{N} \quad \{st(n) \implies n \in A\}$$

puisque le produit d'un **standard** par un **infinitement petit** est **infinitement petit** et donc plus petit que 1. D'autre part un ensemble qui contient tous les standard de  $\mathbb{N}$  contient forcément au moins un

autre élément, donc un **infiniment grand**; soit  $\omega$  un tel élément. On a  $\omega u_\omega \leq 1$  ce qui entraîne que  $u_\omega$  est **infiniment petit** et prouve le lemme.

Il existe diverses variantes de ce lemme dont on peut trouver une synthèse dans [15].

Utilisons la même idée pour notre démonstration. Soit :

$$A = \{t \in \mathbb{R} : t \geq 0 \quad t x(t, x_0) \leq 1\}$$

L'ensemble  $A$  est un véritable ensemble (au sens formel) qui contient tous les réel positifs **infiniment petits** par définition de **UIB** (le produit d'un **infiniment petit** par un **limité** est **infiniment petit** donc plus petit que 1). L'ensemble  $A$  qui contient tous les **infiniment petits** contient donc au moins un non **infiniment petit** c'est à dire un **appréciable**; soit  $t_1$  un tel nombre. L'inégalité :

$$t_1 x(t_1, x_0) \leq 1$$

avec  $t_1$  non **infiniment petit** entraîne que  $x(t_1, x_0)$  est forcément **limité**, ce que nous voulions démontrer. Comme le sous système :

$$\frac{dx}{dt} = g(y)$$

est **I-GAS**, à partir de  $t \geq t_1$  nous savons que  $y(t, y_0)$  est **infiniment petit**. A partir de l'instant  $t_1$  nous intégrons le système :

$$\frac{dx}{dt} = f(x, y(t, y_0))$$

qui est une perturbation régulière du système

$$\frac{dx}{dt} = f(x, 0)$$

Notons  $x^0(t, t_1, x(t_1, x_0, y_0))$  la solution de ce dernier système issue du point  $x(t_1, x_0, y_0)$  à l'instant  $t_1$ . De la dépendance continue des solutions d'une équation différentielle par rapport aux paramètres découle immédiatement que pour  $t \geq t_1$  et **limité**<sup>5</sup> :

$$\|x^0(t, t_1, x(t_1, x_0, y_0)) - x(t, x_0, y_0)\|$$

est **infiniment petit**. Le Lemme de Robinson nous permet une fois de plus d'aller un peu plus loin et donc de dire qu'il existe un **infiniment grand**  $\omega$  tel que sur  $[t_1, \omega]$  :

$$\|x^0(t, t_1, x(t_1, x_0, y_0)) - x(t, x_0, y_0)\|$$

est **infiniment petit** et donc, puisque le système :

$$\frac{dx}{dt} = f(x, y(t, y_0))$$

---

<sup>5</sup>Le théorème de dépendance continue n'affirme que des convergences uniformes sur des intervalles de temps bornés, ce qui se traduit ici par des convergences pour des durées **limitées**

est **S-GAS** en 0, pour  $t$  **infiniment grand** plus petit que  $\omega$  on est assuré que  $x(t, x_0, y_0)$  est infiniment petit. Il reste à montrer que c’est vrai pour n’importe quel infiniment grand. Supposons que ce ne soit pas vrai. Alors il existe  $\omega_1 > \omega$  tel que  $x(\omega_1, x_0, y_0)$  ne soit pas infiniment petit et la différence  $\omega_1 - \omega$  est nécessairement **infiniment grande** (toujours par dépendance continue, partant d’un **infiniment petit** on le reste pendant des temps **limités**). L’ensemble :

$$B = \{t : t > \omega \text{ (} (t - \omega)x(t, x_0, y_0) > 1 \text{)}\}$$

est non vide (il contient  $\omega_1$ ) et on vérifie que sa borne inférieure  $I$  est telle que  $I - \omega$  est **limité**, ce qui est une contradiction et achève la démonstration.

Cette démonstration est parfaitement résumée par la figure 5.

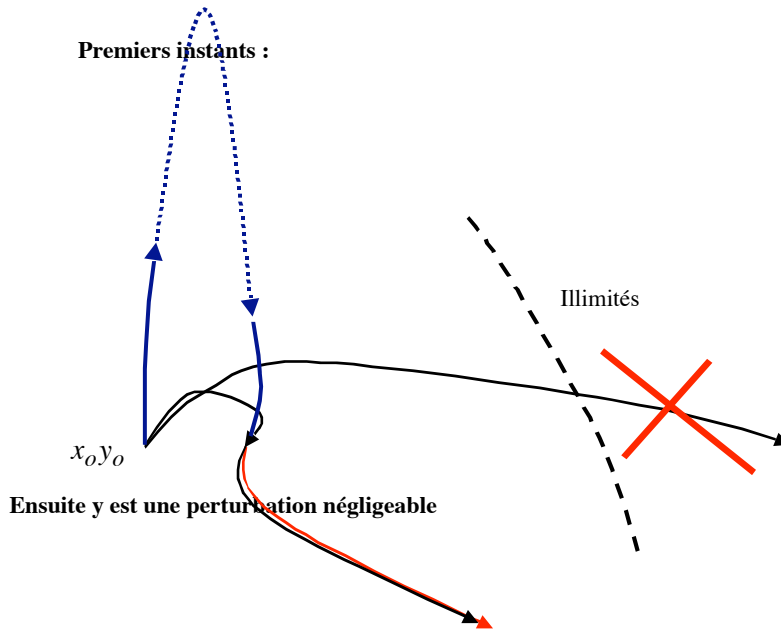


FIG. 5 – Illustration de la preuve du théorème (3.7)

Une petite remarque pour terminer. Le théorème 3.7 ne doit pas faire illusion. La condition suffisante **S-UIB** ne porte pas directement sur les données. Il est difficile de donner des conditions suffisantes portant sur les second membres pour qu’un système soit **S-UIB**. C’est ce qui est fait dans [34]. Nous n’avons pas cherché d’équivalents nonstandard des conditions qui y sont proposées.

### 3.3 L’ANS et les équations différentielles

G. Reeb a certainement été le premier à voir tout le bénéfice que l’ANS pouvait apporter à la *réduction* dans le domaine des équations différentielles

ordinaires où les arguments géométriques ne sont pas forcément simples à formaliser. C'est lui qui a poussé à la fin des années soixante dix quelques jeunes chercheurs de Strasbourg à s'intéresser à l'équation de Van der Pol via l'ANS. Dans une France mathématique qui commençait à peine à concevoir qu'il existe d'autres mathématiques pures que bourbachiques et appliquées que celle de la simulation numérique des E.D.P., pousser des mathématiciens à s'intéresser à une toute petite équation que seuls les électroniciens des écoles d'ingénieurs enseignaient, était résolument provocateur. Reeb ne doutait pas que sur ce vieux sujet un regard neuf ne manquerait pas d'être fécond. C'est ce qui s'est passé avec la découverte du phénomène "canard", c'est à dire de l'importance de certaines solutions instables dans la description du portait de phase de certaines familles à un paramètre d'équations différentielles [4]. Notre traitement de la question du "peaking" se réclame très clairement de la philosophie inaugurée dans cet article. L'ANS a fait de nombreuses autres intrusions dans la théorie des équations différentielles comme, par exemple, la méthode de stroboscopie, qui est la vision ANS des méthodes classiques de moyennisation [7], la théorie des équations différentielles à second membres discontinus [23] la théorie des "fleuves" [14], qui n'a pas d'équivalent classique, et la considération des champs lent-rapides complexes [6]. Nous renvoyons à [12] et sa bibliographie pour une vision assez complète du sujet.

## 4 Théorie Nonstandard du bruit selon M. Fließ

Dans une note récente [17] Michel Fließ s'appuie sur un résultat de P. Cartier et Y. Perrin [9] pour proposer une approche purement déterministe de la question du traitement du bruit en théorie du signal. Nous proposons dans cette partie une petite histoire des idées qui ont conduit au résultat de [9] utilisé dans [17]. Il s'agit essentiellement de quelques résultats des articles [9, 18, 30, 31]<sup>6</sup>. Si nous ne respectons pas la lettre de ces auteurs (en particulier certaines définitions et/ou résultats que nous leur attribuons ne figurent pas nécessairement dans les originaux sous la forme que nous leur donnons) nous espérons, en revanche, en respecter l'esprit.

### 4.1 La théorie du Moiré de Harthong

Passer d'une échelle microscopique à une échelle macroscopique en faisant une moyenne est le B.A. B.A. de l'art du physicien. L'école de G. Reeb, principalement à travers J. Harthong, a apporté une vision nonstandard à ces questions.

Le phénomène de *moiré*, ou simplement *moiré*, est le phénomène suivant. Sur la figure (6) : sont dessinés deux réseaux de fines lignes noires et blanches alternées. Si l'on superpose deux tels réseaux comme sur la figure (7) on voit apparaître des "bandes" foncées séparées par des bandes plus claires. On voit

<sup>6</sup>Comme il sera question notamment de théories de la mesure notons l'existence de très nombreux développements dans ce domaine connus sous le nom de "mesure de Loeb". Nous ne les évoquerons pas car ils se situent dans le formalisme de Robinson.

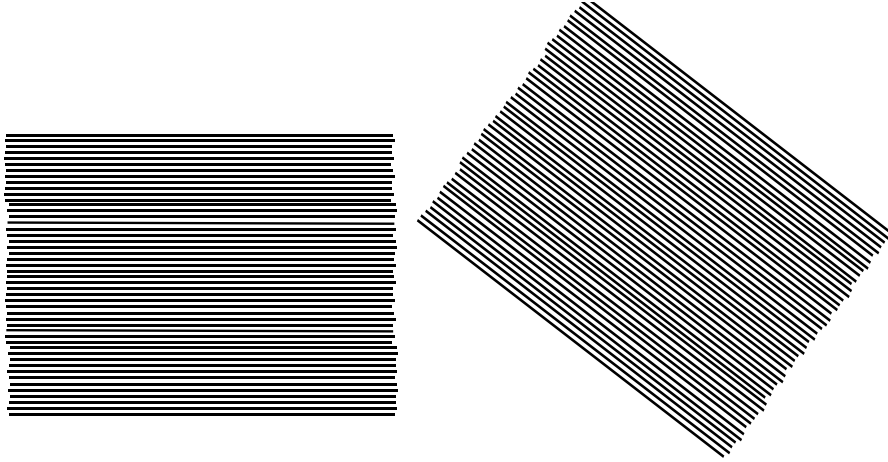


FIG. 6 – Deux réseaux “microscopiques” de droites

également que ces bandes sont d’autant plus larges que l’angle entre les deux réseaux est plus faible. Ce phénomène optique est très répandu. Il est visible notamment quand on regarde des rideaux par transparence, à la télévision quand la nature “pixelisée” de l’écran interfère avec certaines figures de l’image etc... Jacques Harthong a tout de suite vu comment tirer profit de l’ANS pour théoriser ce phénomène. Il s’agit de modéliser le fait que dans certaines régions de l’espace, localement, la superposition des lignes noires se fait sur des lignes blanches, résultant en un noir complet, alors que dans d’autres régions de l’espace une ligne noire se superpose avec une autre ligne noire, laissant claire la ligne voisine. Harthong modélise un “réseau de pas  $h$ ” de la manière suivante (voir la figure 8). On décide que le réseau de fines droites est à l’échelle infiniment petite. On se donne donc un réel  $h$  **infiniment petit** strictement positif qui sera le *pas* du réseau et une fonction périodique  $\psi$ , de période 1 qui vaut alternativement 0 ou 1 sur des intervalles de longueur  $\frac{1}{2}$ . La valeur 0 code pour “noir”, la valeur 1 pour “blanc”. (La fonction  $\psi$  mesure une “transmittance” : 0 si la lumière ne passe pas (noir), 1 si toute la lumière passe. La fonction :

$$(x, y) \rightarrow \psi\left(\frac{\lambda x - y}{h}\right)$$

vaut 1 si :

$$h(k - 0.5) \leq \lambda x - y < hk$$

c’est à dire sur des droites de largeur **infinitésimale**  $h$  et de pente  $\lambda$ . La fonction :

$$(x, y) \rightarrow \psi\left(\frac{-y}{h}\right) \psi\left(\frac{\lambda x - y}{h}\right)$$

vaut 0 où 1 et code pour la superposition d’un réseau horizontal avec un réseau de pente  $\lambda$ .



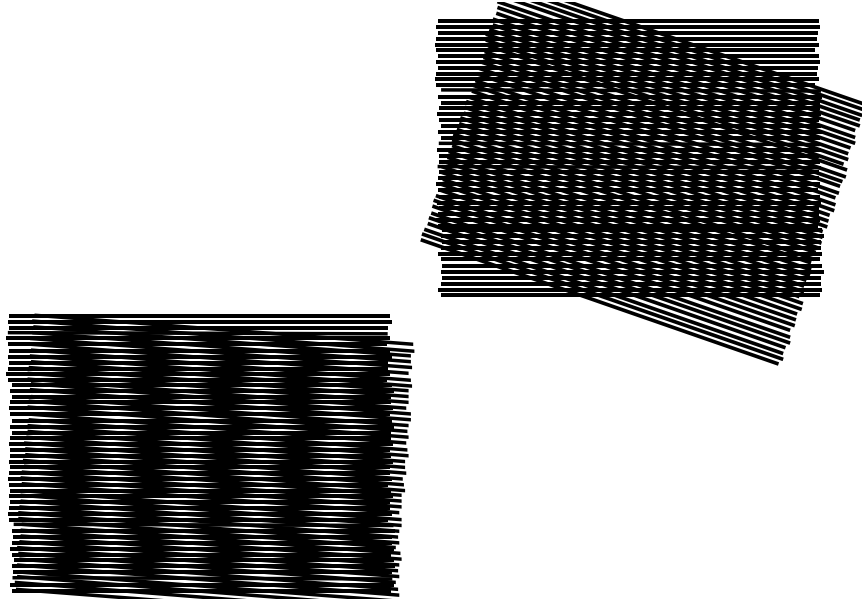


FIG. 7 – Les “bandes” de moiré

Pour représenter le phénomène de “moiré”, Harthong propose de considérer la fonction :

$$\mathcal{M}(x, y) = \frac{1}{\mu(\text{hal}(x, y))} \int \int_{\text{hal}(x, y)} \psi\left(\frac{-v}{h}\right) \psi\left(\frac{\lambda u - v}{h}\right) dudv$$

où  $\text{hal}(x, y)$  est le **halo** du point  $(x, y)$  et  $\mu$  désigne la mesure de Lebesgue de  $R^2$ . Le **halo** de  $(x, y)$  étant constitué des points du plan qui sont “très proches” de  $(x, y)$  on peut dire que c’est un domaine à la fois “petit” dans l’absolu, mais “grand” à l’échelle de  $h$ . En effet les points à distance de  $\sqrt{h}$  de  $(x, y)$  sont à une distance **infiniment grande** par rapport à  $h$  de  $(x, y)$ . La “moyenne sur le halo” de  $(x, y)$  du produit des transmittances représente donc bien localement (au point  $(x, y)$ ), à l’échelle macroscopique, l’effet du produit des transmittances des deux réseaux microscopiques.

Comme les figures ont montré que les bandes apparaissent pour les petites valeurs de la pente, nous allons prendre pour  $\lambda$  un nombre du même ordre de grandeur que  $h$ , c’est à dire  $\lambda = kh$  avec  $k$  **limité**. Remplaçons  $\lambda$  par cette valeur non pas dans  $\mathcal{M}(x, y)$  mais dans :

$$M_\epsilon(x, y) = \frac{1}{(2\epsilon)^2} \int_{v=y-\epsilon}^{v=y+\epsilon} \int_{u=x-\epsilon}^{u=x+\epsilon} \psi\left(\frac{-v}{h}\right) \psi\left(ku + \frac{-v}{h}\right) dudv.$$

On prend pour  $\epsilon$  un nombre **infiniment grand** par rapport à  $h$ . Le changement

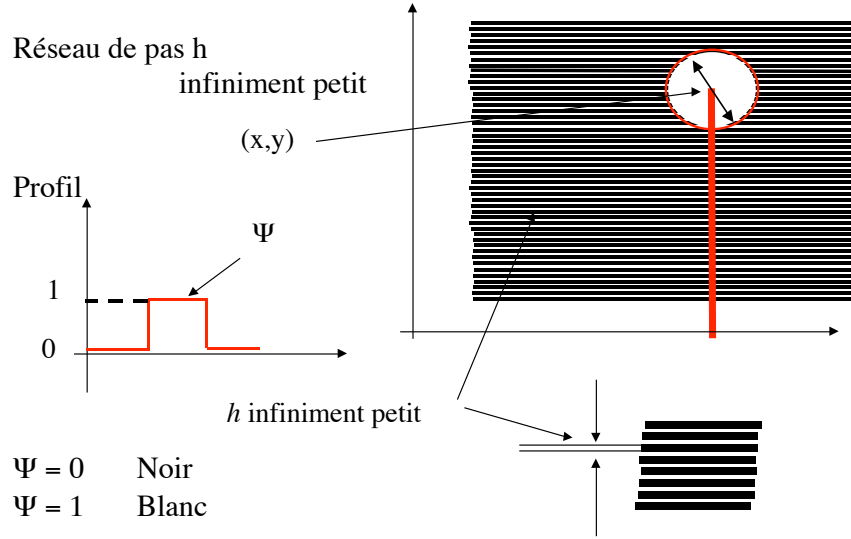


FIG. 8 – Modélisation

de variable  $\frac{-v}{h} = w$  donne :

$$\frac{1}{2\epsilon} \int_{v=y-\epsilon}^{v=y+\epsilon} \psi\left(\frac{-v}{h}\right) \psi\left(ku + \frac{-v}{h}\right) dv = \frac{h}{2\epsilon} \int_{w=-y/h-\epsilon/h}^{w=-y/h+\epsilon/h} \psi(w)\psi(ku+w)dw.$$

Lorsque  $\epsilon/h$  est **infiniment grand** cette dernière intégrale est la moyenne sur un très grand nombre de périodes de la fonction périodique  $w \mapsto \psi(w)\psi(ku+w)$ , et est donc **infiniment proche** de la moyenne de cette fonction sur une période, c'est à dire

$$\int_0^1 \psi(w)\psi(ku+w)dw.$$

De plus la fonction

$$u \rightarrow \int_0^1 \psi(w)\psi(ku+w)dw$$

est **S-continue**. Nous pouvons donc écrire

$$M_\epsilon(x,y) \approx \frac{1}{2\epsilon} \int_{u=x-\epsilon}^{u=x+\epsilon} \int_0^1 \psi(w)\psi(ku+w)dw du$$

et donc, pour  $\epsilon$  **infiniment petit** cette quantité est **infiniment proche** de :

$$\int_0^1 \psi(w)\psi(kx+w)dw$$

et finalement :

$$M_\epsilon(x,y) \approx \int_0^1 \psi(w)\psi\left(\frac{\lambda}{h}x+w\right)dw$$

On voit que  $M_\epsilon(x, y)$  est une quantité indépendante de  $y$ , ce qui explique, d'une part, que les bandes sont verticales, et que la période est d'autant plus grande que  $\lambda$  est petit, ce qui explique, d'autre part, que la largeur des bandes est d'autant plus grande que la pente est faible.

Nous avons donc une théorie<sup>7</sup> qui rend compte convenablement du phénomène optique présenté au début de ce paragraphe. Le principe en a été simple : faire une moyenne sur un ensemble à la fois assez grand à l'échelle du pas des réseaux considérés, mais assez petit pour refléter une propriété locale. On devine que cette théorie peut s'étendre à des réseaux non linéaires. C'est ce qui est fait dans [18].

## 4.2 La théorie de la moyennisation de C. Reder

Il n'a sans doute pas échappé au lecteur que, du point de vue mathématique, la formule de définition :

$$\mathcal{M}(x, y) = \frac{1}{\mu(\text{hal}(x, y))} \int \int_{\text{hal}(x, y)} \psi\left(\frac{-v}{h}\right) \psi\left(\frac{\lambda u - v}{h}\right) dudv$$

de la "moyenne sur le halo" pose un problème, même dans le cadre de l'ANS<sup>8</sup>. En effet "halo (x,y)" n'est pas un ensemble interne. Nous ne disposons donc pas d'une théorie de l'intégration valable pour de tels objets. Dans le calcul du paragraphe précédent nous nous avons résolu le problème en ne considérant pas tout le halo de  $(x, y)$  mais simplement un carré de côté  $2\epsilon$ . Mais cela manque un peu de "canonicité". La clarification de ces questions n'est pas un problème purement formel et a été entreprise par C. Reder dans [31] d'où nous extrayons la matière de ce paragraphe.

Alors que la question du moiré est essentiellement un problème à 2 dimensions C. Reder se place, pour commencer, dans la cas unidimensionnel qui est plus simple.

### 4.2.1 Valeur apparente en un point

On se place sur  $\mathbb{R}$  muni de sa mesure de Lebesgue. On considère une fonction  $f$  intégrable Lebesgue, non nécessairement standard, mais pour simplifier nous supposons que  $|f|$  est majorée par une constante **limitée**.

**Définition 4.1** (Observabilité en un point) *Soit  $x$  un réel **limité**. On dit que la fonction  $f$  est **observable en  $x$**  et que  $a$  est une valeur apparente de  $f$  en  $x$  si il existe un réel **infinitement petit**  $h_0 > 0$  tel que :*

$$\forall h_1 \forall h_2 \quad (h_1 \approx 0 \ h_2 \approx 0 \ h_1 \geq h_0 \ h_2 \geq h_0) \implies \frac{1}{h_1 + h_2} \int_{x-h_1}^{x+h_2} f(s) ds \approx a$$

<sup>7</sup>Ici nous prenons théorie au sens de "théorie physique", c'est à dire de système convaincant de représentation de la réalité, pas au sens de théorie mathématique, c'est à dire un texte irréprochable du point de vue des canons de la rigueur du moment.

<sup>8</sup>D'ailleurs Harthong s'est bien gardé dans son article d'écrire une telle formule!

Il est immédiat de constater que deux valeurs apparentes en un même point sont infiniment proches. Par abus de langage nous appelons “valeur apparente” (lorsqu’elle existe) et notons

$$\mathcal{F}(x)$$

la collection des valeurs apparentes de  $f$  au point  $x$ . Si cette collection était un ensemble formel nous pourrions invoquer l’axiome du choix et “choisir” une valeur. Dans le système IST cette possibilité de “singulariser” une valeur particulière dans une collection de nombres infiniment proches est offerte par l’axiome de “standardisation”. Nous pourrions l’utiliser (c’est d’ailleurs ce qui est fait dans [31]) mais cela nous conduirait à des développements formels qui ne nous semblent pas utiles dans cet article. Pour paraphraser ce qui vient d’être fait nous dirons que la “valeur apparente”, lorsqu’elle existe, est la “moyenne” sur un intervalle **infiniment petit** suffisamment grand.

Dans la définition (4.2) on fait la moyenne sur un intervalle qui est contenu dans  $hal(x)$  et qui contient  $x$  dans son intérieur. La définition suivante ne fait plus cette restriction.

**Définition 4.2** (Observabilité forte en un point) *Soit  $x$  un réel **limité**. On dit que la fonction  $f$  est **fortement observable en  $x$**  et que  $a$  est une valeur apparente, si il existe un réel **infiniment petit**  $h_0 > 0$  tel que :*

$$\forall h \forall y \quad (h \approx 0 \ h > h_0 \ |y - x| \approx 0) \implies \frac{1}{h} \int_y^{y+h} f(s) ds \approx a$$

- On constate facilement qu’une fonction  $f$  **S-continue** en  $x$  est observable en  $x$  et a pour valeur apparente  $f(x)$
- Une fonction  $f$  périodique de période  $T$  avec  $T$  **infinitésimal** est fortement observable en  $x$  et a pour valeur apparente n’importe quel  $a$  tel que :

$$a \approx \frac{1}{T} \int_0^T f(t) dt$$

Pour le voir il suffit de prendre  $h_0 = \sqrt{T}$  qui est **infiniment grand** par rapport à  $T$ . Soit  $h$  un **infiniment petit** plus grand que  $h_0$ . On peut écrire :

$$h = nT + r$$

avec  $n$  **infiniment grand** et  $r < T$ . La moyenne sur  $[y, y+h]$  se décompose en :

$$\frac{1}{h} \int_y^{y+h} f(t) dt = \frac{1}{nT + r} [n \int_y^{y+T} f(t) dt + \int_y^{y+r} f(t) dt]$$

d’où nous déduisons immédiatement ce qui vient d’être avancé.

- La fonction qui vaut 0 pour  $x$  négatif et 1 pour  $x$  positif et indéterminée en 0 n’a pas de valeur apparente en 0. En effet, la moyenne de cette fonction sur un intervalle qui contient 0 peut varier entre 0 et 1 selon que 0 est à une extrémité ou l’autre de l’intervalle.

- On montre dans [31] que la fonction  $f$  qui est nulle pour  $x \leq 0$  et qui vaut  $\sin(\frac{1}{x})$  pour  $x$  positif est observable en 0 mais n’y est pas fortement observable.

On peut se demander à quelles conditions une fonction aura souvent une valeur apparente. La réponse est que si  $f$  est intégrable elle possède une valeur apparente en presque tout point (voir [31]).

#### 4.2.2 Moyenne mobile

La définition purement externe de la valeur apparente ne permet pas de parler de la fonction qui à  $x$  associe, lorsqu’elle existe, une valeur apparente de  $f$  au point  $x$ . Le résultat de ce paragraphe va nous donner un moyen de résoudre ce problème.

Considérons la fonction (définie pour  $h$  positif non nul) :

$$x \rightarrow M_h(x) = \frac{1}{h} \int_x^{x+h} f(t) dt$$

c’est la “moyenne mobile” de  $f$  sur une fenêtre de largeur  $h$  au sens le plus élémentaire du terme. Depuis des temps immémoriaux, prendre la moyenne mobile est un procédé utilisé pour régulariser les fonctions. Y a-t-il un lien entre la moyenne mobile et la valeur apparente? Si nous revenons à la définition de la valeur apparente en  $x$  nous trouvons que la moyenne sur un intervalle **infinitésimal**  $h$  “assez grand” ne doit plus dépendre de  $h$ , mais, à priori, le “assez grand” dépend de  $x$ . En fait il n’en est rien, comme le montre le résultat de [31] :

**Théorème 4.3** *Il existe un réel **infinitement petit** strictement positif  $h_0$  tel que pour tout  $h$  **infinitement petit** supérieur à  $h_0$ , en tout point  $x$  où  $f$  est observable, la moyenne  $M_h(x)$  est une valeur apparente de  $f$ . De plus, en ce point, la fonction  $x \rightarrow M_h(x)$  est observable et une valeur apparente de  $M_h$  est une valeur apparente de  $f$ .*

Ce théorème n’est pas trivial et nous renvoyons à [31] pour sa preuve. En revanche ce qui s’en déduit est évident :

- Soit  $f$  intégrable. Il existe un **infinitésimal**  $h_0$  tel que pour  $h_1$  et  $h_2$  **infinitésimaux** plus grands que  $h_0$  on a  $M_{h_1}(x) \approx M_{h_2}(x)$  en tout point où  $f$  est observable.
- Soit  $h$  tel que dans le théorème 4.3. La fonction  $x \rightarrow M_h(x)$  est continue mais n’est pas **S-continue**.
- Soit  $I$  un intervalle en tout point duquel  $f$  est observable. Alors  $x \rightarrow M_h(x)$  est **S-continue** sur  $I$ .

Nous pouvons donc, à partir de maintenant, associer à une fonction  $f$  sa “régularisée” que nous notons  $\mathcal{M}_f$ . Nous entendons par là une quelconque des fonctions  $M_h$  pour  $h$  suffisamment grand, définies par le théorème 4.3. Ce procédé de régularisation par moyenne mobile est appelé régularisation par convolution dans [31]. D’un point de vue concret, si nous partons d’un signal

bruité, si nous traçons le graphe de  $M_h$  pour des valeurs de  $h$  croissantes à partir de 0 nous observons que le graphe devient indépendant de  $h$  à partir d'une certaine valeur de  $h$ , comme il est illustré sur la figure 9. Dans cette expérience

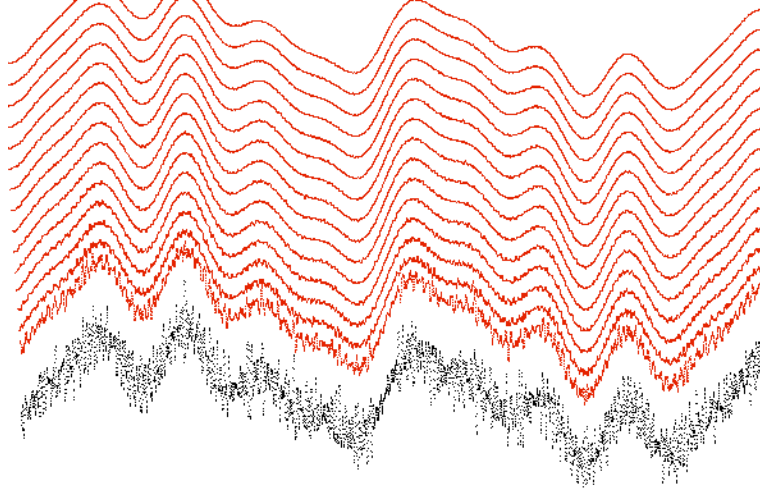


FIG. 9 – Evolution de  $M_h$

nous sommes partis d'un signal bruité (en noir). La largeur de la figure représente au total sept unités; les courbes en rouge sont les graphes (décalés vers le haut pour chaque nouvelle valeur de  $h$ ) des moyennes mobiles de fenêtre de largeur  $h$ , pour des valeurs de  $h$  variant de 0.02, par pas de 0.02, à 0.3. On peut considérer qu'à partir de  $h = 0.1$  la fonction  $M_h$  ne dépend plus de  $h$ .

### 4.3 Radically Elementary Probability Theory

Dans [30] (Livre téléchargeable en version anglaise, française et russe sur <http://www.math.princeton.edu/nelson/books/rept/rept.pdf>) E. Nelson propose une version nonstandard de la théorie du mouvement brownien, ce qui ne peut laisser indifférent quand on sait qu'il est aussi l'auteur de "Dynamical theories of brownian motion" [28] (Livre téléchargeable <http://www.math.princeton.edu/nelson/books/bmotion.pdf>), un très grand classique. Le principe en est le suivant. Tout le monde connaît la "marche de l'ivrogne". C'est le processus stochastique défini par :

$$x_{t+dt} = x_t + Z_t \sqrt{dt} \quad (29)$$

où  $t$  prend les valeurs discrètes  $0, dt, \dots, kdt, \dots$  et où  $Z_t$  est une suite de variables aléatoires indépendantes prenant les valeurs  $+1$  ou  $-1$  avec la probabilité  $\frac{1}{2}$ , ce que nous symbolisons dans l'écriture  $Z_t = \pm 1$  et conduit à réécrire (29) sous la forme :

$$x_{t+dt} = x_t \pm \sqrt{dt} \quad (30)$$

La longueur du pas  $\sqrt{dt}$  est choisie pour normaliser la variance de  $x_1$  à la valeur 1.

Le mouvement brownien physique est le processus dans lequel une particule de taille de l'ordre du micron subit chaque seconde un nombre incroyablement grand de chocs de la part des molécules du fluide dans lequel elle est plongée d'où l'invention du mouvement brownien mathématique, ou processus de Wiener, qui se veut l'idéalisation de cette situation. Comme il est courant en mathématiques classiques on idéalise le "petit" par le "passage au continu" ce qui conduit à l'invention du mouvement brownien mathématique qui est un *un processus en temps continu* dans lequel à *chaque instant* on tire à pile ou face la direction du parcours. Il y a une antinomie très forte entre l'idée de "continu mathématique", représenté par la droite réelle, et une "succession d'instant". C'est la raison pour laquelle le processus de Wiener est un objet abstrait si difficile à définir.

Nelson propose, lui, pour idéaliser le mouvement brownien physique, de considérer simplement le processus (30) avec  $dt$  **infiniment petit**. Le caractère indéterminé de  $dt$  donne toute sa canonicité au processus au moins dans toutes les assertions où  $dt$  ne figure pas explicitement comme par exemple :

**Théorème 4.4 (Nelson)** *Presque sûrement les trajectoires du mouvement brownien sont continues.*

On a certainement compris que dans ce théorème "continu" doit être pris au sens de **S-continu**. Le "presque sûrement" demande une explication. Supposons que nous travaillions sur l'intervalle de temps  $[0, 1]$  discrétisé en

$$0, dt, 2dt, \dots, kdt, \dots, Ndt = 1$$

L'espace probabilisé sur lequel nous travaillons est donc l'espace :

$$\{-1, +1\}^N$$

muni de la mesure de probabilité produit de la mesure uniforme sur l'ensemble  $\{-1, +1\}$ . Chaque élément de cet ensemble a une probabilité de  $\frac{1}{2^N}$  et il n'existe donc pas d'ensemble de mesure nulle à proprement parler. D'autre part, pour une trajectoire, être **S-continu** est une propriété externe, qui ne définit donc pas un ensemble. C'est pourquoi on dit qu'une propriété  $P$  (éventuellement externe) est **rare** si, pour tout standard  $\epsilon$ , il existe un ensemble  $A$  de mesure inférieure à  $\epsilon$  tel que si  $P(z)$  est vérifiée alors  $z$  appartient à  $A$ . Un évènement est **presque certain** si son contraire est un évènement **rare**.

La théorie de Nelson est *radicalement élémentaire* en ce sens qu'elle n'utilise qu'une version faible de la théorie IST plus intuitive et plus simple que la théorie complète, et une théorie (triviale) de l'intégration sur les ensembles finis. Le plus étonnant est que, dans une annexe qui, elle, n'est pas élémentaire (on utilise toute la force de IST), Nelson montre que la théorie élémentaire contient potentiellement tout résultat que la théorie continue pourrait produire. Il n'ésite pas à présenter son annexe par ces mots :

*Dans cette annexe, nous voulons montrer que les théorèmes de la théorie conventionnelle des probabilités peuvent être établis par leurs*

*analogues élémentaires grâce à des arguments que l'on qualifie habituellement de non-sens généralisé; aucun raisonnement probabiliste n'apparaîtra ici. On illustre ainsi le fait que la théorie non standard élémentaire des processus stochastiques peut être utilisée pour obtenir des résultats conventionnels. Inversement, cela indique que la machinerie sophistiquée de la théorie classique et les stratagèmes de l'Analyse Non Standard, nécessaires pour prouver l'équivalence entre les résultats élémentaires et leurs équivalents classiques, n'apportent rien d'important aux probabilités : la théorie élémentaire a le même contenu scientifique que la théorie classique. Cette annexe est en quelque sorte autodestructrice.*

Ce livre de Nelson, s'il n'a pas un rapport direct avec la moyennisation, le sujet qui nous intéresse ici, est certainement une source d'inspiration pour les auteurs qui souhaitent aborder la question de la théorie de la mesure d'un point de vue élémentaire avec les moyens de l'ANS.

#### 4.4 La théorie de la mesure de Cartier-Perrin

Il est donc tentant d'oublier les probabilités et de suivre la démarche de Nelson pour construire une théorie de la mesure élémentaire. Prenons sur l'intervalle  $[0, 1]$  la suite de points  $\{dt, 2dt, \dots, kdt, \dots, Ndt = 1\}$ , attribuons à chaque point la "masse"  $\frac{1}{N}$ . Appelons "mesure de dénombrement" d'une partie  $A$  de  $[0, 1]$  le nombre :

$$\delta(A) = \frac{1}{N} \sum_{k=1}^N \chi_A(kdt)$$

où  $\chi_A$  est la fonction indicatrice de l'ensemble  $A$ . La quantité  $\delta(A)$  a de bonnes propriétés comme l'additivité, l'invariance par translation mais possède le défaut inacceptable de donner la masse 1 à l'ensemble des rationnels de  $[0, 1]$ . Il est donc difficile d'envisager une théorie de la mesure sur une base aussi naïve.

Dans [9], P. Cartier et Y. Perrin proposent une "théorie de l'intégration sur des ensembles finis" qui tente d'être assez riche pour couvrir les besoins de l'analyse courante. Ils considèrent des ensembles finis (de cardinal **infiniment grand**)  $X = \{a_1, a_2, \dots, a_i, \dots, a_N\}$ ; à chaque élément  $a_i$  est associé un nombre positif ou nul : sa "masse"  $m_i$ . La mesure  $m(A)$  d'un ensemble  $A$  est la somme des masses des points qui lui appartiennent. Etant donnée une fonction  $f$  définie sur  $X$ , à valeur dans  $\mathbb{R}$  on peut toujours considérer la quantité :

$$\sum_{i=1}^N m_i f(a_i)$$

et la noter :

$$\sum_{i=1}^N m_i f(a_i) = \int_X f dm$$

La classe de toutes les fonctions étant trop vaste on se restreindra aux :



**Définition 4.5** (Fonction **S-intégrables**) *Une application  $f$  de  $X$  dans  $\mathbb{R}$  est dite **S-intégrable** si et seulement si :*

$$\int_x |f| dm$$

est **limité** et

$$\int_A f dm \approx 0$$

pour tout ensemble  $A$  **rare**.

où un ensemble  $A$  **rare**<sup>9</sup> est, comme dans la théorie de Nelson, un ensemble (éventuellement externe) tel que pour tout standard  $\alpha > 0$  il existe un ensemble interne  $B$  tel que  $A$  soit contenu dans  $B$  et  $m(B) \leq \alpha$ .

On munit ensuite l'ensemble  $X$  d'une distance  $d$ . On suppose que  $X$  est précompact, ce qui veut dire que le diamètre de  $X$  est **limité** et que pour tout nombre  $r$  non **infinitement petit** il existe un recouvrement de l'espace par un nombre **limité** de boules de rayon plus petit que  $r$ .

La définition de **S-continu** est la même que celle que nous avons déjà donnée et

**Définition 4.6** (Fonction **presque S-continue**) *Une fonction  $f$  de  $X$  dans  $\mathbb{R}$  est **presque S-continue** si et seulement si elle est **S-continue** sur le complémentaire d'un ensemble rare.*

Ceci permet de définir la notion de fonction "Lebesgue intégrable".

**Définition 4.7** (Fonction **L-intégrable**) *Une fonction  $f$  de  $X$  dans  $\mathbb{R}$  est **L-intégrable** si et seulement si elle est **S-intégrable** et **presque S-continue**.*

Ajoutons la définition de fonction "rapidement oscillante" :

**Définition 4.8** (Fonction **rapidement oscillante** (Cartier-Perrin)) *Une fonction  $h$  de  $X$  dans  $\mathbb{R}$  est **rapidement oscillante** si et seulement si elle est **S-intégrable** et si pour tout sous ensemble  $A$  **quarrable** de  $X$  on a :*

$$\int_A h dm \approx 0$$

où **quarrable** est défini par la :

**Définition 4.9** *Un ensemble  $A$  est **quarrable** si sa frontière est un ensemble **rare**.*

et nous pouvons maintenant énoncer le :

<sup>9</sup>Dans [2] qui est l'extension de la théorie des probabilité "radicalement élémentaire" de Nelson aux processus de diffusion, E. Benoit établit des propriétés non triviales de certains ensembles **rares**

**Théorème 4.10 (Théorème de décomposition (Cartier-Perrin))** *Soit  $f$  une fonction **S-intégrable** sur un presque intervalle. Alors on a :*

$$f = g + h$$

où  $g$  est **L-intégrable** et  $h$  **rapidement oscillante**. La décomposition est unique à des **infinitement petits** près.

Pour faire le lien avec la théorie de la moyennisation de J. Harthong et de C. Reder donnons une idée de la démonstration de ce résultat. L'idée est de définir une suite  $\mathcal{P}_n$  de partitions de  $X$  par des ensembles de diamètres de plus en plus petits et de calculer la moyenne de  $f$  sur chaque atome de la partition  $\mathcal{P}_n$  ce qui donne une suite de fonction  $f_n$  ; pour des valeurs convenables des indices la fonction  $f_n$  sera **presque S-continue** donc **intégrable Lebesgue**. Comme toutes les définitions en cause sont externes cette construction doit être effectuée avec soin.

#### 4.5 La définition Non Standard du bruit de Michel Fliess

Dans une note récente, Michel Fliess [17] propose de considérer un signal (bruité), non comme une fonction continue, mais pour ce qu'il est réellement, un signal échantillonné, c'est à dire une fonction  $f$  (les notations ne sont pas celles de [17]) définie sur un ensemble fini discret, mais idéalisé en un "presque intervalle".

Un "presque-intervalle" (pour suivre la terminologie de Nelson reprise par Cartier-Perrin) est un ensemble fini de points :

$$\{t_0, t_1, \dots, t_i, \dots, t_N\}$$

de l'intervalle  $[a, b]$  tels que  $a = t_0$ ,  $t_{i-1} \approx t_i$ ,  $T_N = b$  affectés des masses  $m_0 = 0$  et  $m_i = t_i - t_{i-1}$ ;  $i = 1, 2, \dots, N$  et muni de la distance induite par cette de  $\mathbb{R}$ . C'est donc un ensemble fini, précompact au sens du paragraphe précédent.

On admet alors que le signal bruité est une fonction **S - intégrable** et, d'après le théorème de décomposition de Cartier-Perrin [9], on a :

$$f = g + h$$

avec  $g$  **L-intégrable** et  $h$  **rapidement oscillante**.

Michel Fliess propose de définir le "bruit" comme la partie  $h$ , **rapidement oscillante**, du signal  $f$ .

Cette double décision de définir le bruit dans un cadre *discret* et *purement déterministe* est une décision forte qui va à l'encontre de la tradition mathématique actuelle qui traite le bruit comme un "bruit blanc" donc, en quelque sorte, comme une réalisation continue du "hasard pur". Nous ne discuterons pas de la pertinence de ce choix car, nous le répétons, notre compétence en théorie du signal est nulle. Nous pouvons simplement faire quelques remarques d'ordre mathématique.

- M. Fliess définit également le bruit pour des signaux à plusieurs dimensions.
- On aurait pu aussi bien choisir de traiter des signaux continus en se plaçant dans le cadre plus classique (mais nonstandard) de C. Reder en utilisant les résultat de [31].
- Comme nous l’avons fait dans le cas du peaking, proposons un équivalent classique de “rapidement oscillant”. On se donne une famille  $f_\gamma(t)$  d’applications de  $\mathbb{R}$  dans  $\mathbb{R}$ . On dira que  $f_\gamma$  est asymptotiquement (quand  $\gamma \rightarrow +\infty$ ) rapidement oscillante si et seulement si :

$$\forall \epsilon > 0 \forall M > 0 \exists \gamma_0 \forall \gamma \forall a \forall b$$

$$\{\gamma > \gamma_0 \mid b - a < M \Rightarrow \left| \int_a^b f_\gamma(t) dt \right| < \epsilon\}$$

qu’on comparera à : “Sur tout intervalle **limité** l’intégrale de  $f$  est **infiniment petite**”.

- Intéressons nous à des fonctions de période  $[0, 2\pi]$ . Que l’on adopte le point de vue relativement classique des fonctions (éventuellement non standard) intégrable ou celui plus radical où  $[0, 2\pi]$  est remplacé par un presque - intervalle (comme dans [11]) on peut décomposer la fonction en série de Fourier . Soit donc (cas classique) :

$$f(t) = \sum_{n=-\infty}^{+\infty} c_n e^{int}$$

Le théorème de décomposition de  $f$  en une partie régulière (fonction d’observation ou **L-intégrable**) et une partie **rapidement oscillante** suggère la possibilité d’une décomposition du genre :

$$\sum_{n=-\infty}^{+\infty} c_n e^{int} = \sum_{n \in \{\text{limités}\}} c_n e^{int} + \sum_{n \in \{\text{infiniment grands}\}} c_n e^{int}$$

où la première somme serait la partie régulière et la seconde la partie **rapidement oscillante**. Pour cela il faudrait déterminer les conditions sous lesquelles il est possible de donner un sens aux deux sommes (sur des ensembles d’indices externes) qui pour le moment n’en ont pas.

- Dans un travail très récent [3] E. Benoit aborde ce point de vue fréquentiel dans le cas non périodique via la transformée de Laplace. Il montre que si  $f$  est **rapidement oscillante** sa transformée de Laplace  $F(s)$  est **infiniment petite** dès lors que  $\Re e(s) > 0$  et  $\frac{\Im m(s)}{\Re e(s)}$  est **limité**.

## 5 Dernières remarques

Le lecteur un peu familier avec l’ANS de [29] aura remarqué que nous n’avons pas utilisé l’axiome de Standardisation. De quoi s’agit-t-il ? On sait qu’en théorie formelle des ensembles (dans ZFC) il existe un axiome qui s’énonce ainsi :

Soit  $P$  une formule (interne!)

$$\forall x \exists y \forall z \{z \in x \wedge P(z) \iff z \in y\}$$

qui dit que l'ensemble (intuitif) des  $z$  qui possèdent une propriété (interne) est un véritable ensemble, un ensemble au sens formel dans ZFC. Nous avons suffisamment insisté sur le fait que si  $P$  n'est pas une formule interne un tel ensemble formel n'existe pas nécessairement et que c'était là une des clefs de l'efficacité du langage non standard.

L'axiome de Standardisation dans IST est un pourvoyeur d'ensembles standard à partir de formules externes. Il s'énonce comme suit :

Soit  $P$  une formule (pas nécessairement interne)

$$\forall^{st} x \exists^{st} y \forall^{st} z \{z \in x \wedge P(z) \iff z \in y\}$$

En d'autres termes, pour toute propriété il existe un ensemble standard dont les éléments standard sont les standard qui vérifient la propriété et seulement eux.

C'est, en particulier, l'usage de cet axiome qui permet à Nelson de montrer dans l'annexe de [30] que sa théorie radicalement élémentaire des probabilités est équivalente à la théorie classique. Mais comme dit Nelson "Cette annexe est en quelque sorte auto destructrice". C'est pourquoi un point de vue nonstandard plus "radical" est de préconiser l'usage de théories nonstandard faibles comme Nelson dans [30], Callot dans [5] ou Lutz dans [26]. Dans de telles théories on ne cherche pas à distinguer des objets standard d'objets non standard en toute généralité. On se contente de le faire sur les réels ce qui est suffisant pour faire de l'analyse. Mais il n'est plus possible de comparer avec les résultats classiques. Ainsi "Une fonction standard est uniformément continue si et seulement si elle est **S-continue**" n'a plus de sens puisque "fonction standard" n'est pas défini. En gardant la notion d'objet standard, sans utiliser l'axiome qui l'accompagne, nous avons choisi de permettre de telles comparaisons. Mais tous ceux qui sont convaincus par l'Analyse Non Standard pensent que ces comparaisons sont *auto-destructrices*

## Références

- [1] H. Barreau and J. Harthong eds. *La mathématique non standard*. Editions du CNRS, Paris, 1989.
- [2] E. Benoit. Diffusions discrètes et mécanique stochastique. Technical report, Université de Nice, Laboratoire J. A. Dieudonné, 1989.
- [3] E. Benoit. Nonstandard averaging and signal processing. 2006. Slides at EMS 2006 congress; <http://www.dm.unipi.it/nsm2006/slides/benoit-s.pdf>.
- [4] E. Benoit, J-L. Callot, F. Diener, and M. Diener. Chasse au canard. *Collectedanea Mathematica*, 31(1-3) :37–119, 1981.

- [5] J-L. Callot. Trois leçons d'analyse infinitésimale. In J-M Salanskis and H. Sinaceur, editors, *Le Labyrinthe du Continu (Colloque de Cerisy)*, Lecture Notes in Control and Information Sciences, pages 369–381. Springer-Verlag, 1992.
- [6] J-L. Callot. Champs lent-rapides complexes à une dimension lente. *Ann. Scient. Ec. Norm. Sup*, 4(26) :149–176, 1993.
- [7] J-L. Callot and T. Sari. Stroboscopie infinitésimale et moyennisation dans les systèmes d'équations différentielles à solution rapidement oscillantes. In I.D. Landau, editor, *Outils et Modèles Mathématiques pour l'Automatique, l'Analyse de Systèmes et le Traitement du Signal*, pages 345–353. Edition du CNRS, Paris, 1983.
- [8] M. Canalis-Durand and P. Yalo. Robustesse des systèmes linéaires bouclés aux perturbations non linéaires. In A. Bensoussan and J.L. Lions, editors, *Analysis and Optimization of Systems*, volume 111 of *Lecture Notes in Control and Information Sciences*. Springer, 2005.
- [9] P. Cartier and Y. Perrin. Integration over finite sets. In F. Diener and M. Diener, editors, *Non Standard Analysis in Practice*, Universitext, pages 185–204. Springer-Verlag, Berlin, 1995.
- [10] A. Deledicq and M. Diener. *Leçons de calcul infinitésimal*. Armand Colin, Paris, 1989.
- [11] P. Delfini and C. Lobry. The vibrating string. In F. Diener and M. Diener, editors, *Non Standard Analysis in Practice*, Universitext, pages 51–70. Springer-Verlag, Berlin, 1995.
- [12] F. Diener and M. Diener eds. *Non Standard Analysis in Practice*. Springer-Verlag, Berlin, 1995.
- [13] M. Diener and C. Lobry eds. *Analyse Non Standard et représentation du réel*. Outils et Modèles Mathématiques pour l'Automatique et le Traitement du Signal. OPU, Alger and CNRS, Paris, 1985.
- [14] M. Diener and G. Reeb. Champs polynômiaux : nouvelles trajectoires remarquables. *Bull. Soc.Math. Belgique*, 38 :131–150, 1987.
- [15] M. Diener and I. van den Berg. Halos et galaxies. *C. R. Acad. Sci. Paris, Ser I*, 293 :385–388, 1981.
- [16] I. D. Landau ed. *Outils et Modèles Mathématiques pour l'Automatique, l'Analyse de Systèmes et le Traitement du Signal*. Edition du CNRS, Paris, 1983.
- [17] M. Fliess. Analyse non standard du bruit. *C. R. Acad. Sci. Paris, Ser I*, 342 :797–802, 2006.
- [18] J. Harthong. Le moiré. *Adv. in Appl. Math.*, 2(1) :24–75, 1981.
- [19] P.V. Kokotovic and R. Marino. On vanishing stability region in non-linear systems with high gain feedback. *IEEE. Trans. Autom. Control.*, AC-31(10), 1986.

- [20] C. Lobry. Mathématiques non standard mathématiques appliquées? Technical report, Université de Bordeaux I, Mathématiques, 22p., 1981. Conférence plénière, Colloque national d'analyse numérique, Aussois.
- [21] C. Lobry. *Et pourtant ils ne remplissent pas  $\mathbb{N}$ !* Editeur ALEAS, 15 quai Lassagne, Lyon, 1989.
- [22] C. Lobry and T. Sari. The peaking phenomenon and singular perturbations : an extension of tikhonov's theorem. Technical report, INRIA, 2001.
- [23] C. Lobry and T. Sari. *Equations différentielles à second membres discontinus*, pages 237–266. Ecole du CIMPA, Quelques Questions de Théorie du Contrôle. Hermann, 2004.
- [24] C. Lobry and T. Sari. *Singular perturbations in control theory*, pages 151–177. Ecole du CIMPA, Quelques Questions de Théorie du Contrôle. Hermann, 2004.
- [25] C. Lobry, T. Sari, and S. Touhami. On tykhonov's theorem for convergence of solutions of slow and fast systems. *Electronic Journal of Differential Equations*, 1998(19) :1–22, 1998.
- [26] R. Lutz. Rêveries infinitésimales. *Gazette des Mathématiciens*, 34 :79–87, 1987.
- [27] R. Lutz and M. Goze. *Nonstandard analysis : a practical guide with applications*. Lecture notes in maths vol 881, Springer, 1982.
- [28] E. Nelson. *Dynamical Theories of Brownian Motion*. Princeton University Press, 1967.
- [29] E. Nelson. Internal set theory. *Bull. Amer. Math. Soc.*, 83 :1165–1198, 1977.
- [30] E. Nelson. *Radically Elementary Probability Theory*. Princeton University Press, 1987.
- [31] C. Reder. Observation macroscopique de phénomènes microscopiques. In M. Diener and C. Lobry, editors, *Analyse Non Standard et représentation du réel*, Outils et Modèles Mathématiques pour l'Automatique et le Traitement du Signal, pages 1995–244. OPU, Alger and CNRS, Paris, 1985.
- [32] H. Robinson. *Non standard analysis*. Noth Holland, Amsterdam, 1966.
- [33] J-M. Salanskis and H. Sinaceur (Eds). *Le Labyrinthe du Continu*. Springer-Verlag France, Paris, 1992.
- [34] H.J. Sussmann and P.V. Kokotovic. The peaking phenomenon and the global stabilization of nonlinear systems. *IEEE. Trans. Autom. Control.*, 36(4) :224–440, 1991.