

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA
RECHERCHE SCIENTIFIQUE

Université Larbi Ben M'hidi Oum El Bouaghi
Faculté des Sciences Exactes
Département de Mathématiques et informatique

No d'ordre:.....

Série:.....



Thèse

Présentée pour obtention du diplôme du Doctorat LMD

Thème :

**Analyse de l'effet d'utilisation du chaos
sur l'amélioration des performances des
méthodes numériques**

Option : Mathématiques appliquées

Présenté et soutenu par

Farida Derouiche

Membres du jury :

Président :	Zehrou Okba	Prof	Université de Oum El Bouaghi
Rapporteur :	Tayeb Hamaizia	M.C.A	Université de Constantine 1
Examineur :	Dalah Mohamed	Prof	Université de Constantine 1
Examineur :	Ouannas Adel	M.C.A	Université de Oum El Bouaghi

Année Universitaire 2020-2021

Remerciements

Mes remerciements les plus sincères vont à mon directeur de thèse Mr le docteur **Hamaizia Tayeb**, pour la grande qualité de son encadrement durant ce travail de thèse.

Toujours avec gentillesse, il a donné sans compter de son temps pour accompagner mes premiers pas dans le monde de la recherche, et c'est une chance d'avoir pu travailler avec un directeur sachant doser si justement patience, rigueur et exigence.

J'exprime ici ma profonde gratitude à Messieurs le professeur **Dalah Mohamed** et le docteur **Ouannas Adel** qui m'ont honoré d'accepter de faire partie du Jury. Ma reconnaissance va aussi à Mr le professeur **Zehrou Okba** pour avoir accepté de présider ce Jury.

DÉDICACES

A ma chère mère et mon cher père

A mon mari et ma fille

A mon frère et mes sœurs

A toute ma belle famille

Table des matières

Table des figures	4
Introduction générale	6
1 Systèmes dynamiques et chaos	9
1.1 Définitions et notations	9
1.2 Systèmes dynamiques discrets	10
1.2.1 Orbites ou trajectoires	10
1.2.2 Points fixes	11
1.2.3 Points périodiques et p-cycles	11
1.2.4 Étude de stabilité	12
1.2.5 Stabilité du point fixe	12
1.2.6 Stabilité des orbites périodiques.	13
1.2.7 Équivalence topologique des systèmes	13
1.2.8 Bifurcation	13
1.3 Chaos	17
1.3.1 Introduction	17
1.3.2 Systèmes dynamiques chaotiques	18
1.3.3 Les attracteurs	19
1.3.4 Détection du chaos	20
1.3.5 Routes vers le chaos	26
1.4 Exemples de Systèmes dynamiques chaotiques	28
1.4.1 Application logistique	28
1.4.2 L'attracteur de Hénon	32

1.4.3	L'attracteur de Lozi	37
2	Généralités sur l'optimisation	42
2.1	Introduction	42
2.2	Types d'optimum	43
2.3	Classification des problèmes et des approches	44
2.3.1	Problèmes avec ou sans contraintes	44
2.3.2	Approches déterministes et approches probabilistes	45
2.4	Méthodes d'optimisation globale	47
2.4.1	Méthode du gradient	47
2.4.2	Méthode du gradient conjugué	48
2.4.3	Méthode de Newton	50
2.4.4	Méthode de quasi newton	50
2.4.5	Méthode de relaxation	54
2.4.6	La méthode des initialisations multiples (multistart)	54
2.4.7	La méthode de tunnel (tunneling)	55
2.4.8	Méthodes de regroupement (clustering)	56
2.5	Méthodes méta-heuristiques	56
2.5.1	Méthodes de recherche locale	57
2.5.2	Les méthodes évolutives	58
2.5.3	Optimisation par essaims particulières	59
2.5.4	Optimisation par Colonie d'abeilles	62
3	Proposition d'une nouvelle approche pour les algorithmes d'optimisation chaotiques	65
3.1	Introduction	65
3.2	Stratégie de la recherche chaotique.	66
3.2.1	Génération de séquences chaotiques.	67
3.2.2	Mécanisme de l'algorithme BCOA	68
3.3	Proposition d'une nouvelle approche EBCOA	70
3.3.1	Description de la méthode EBCOA	70

3.3.2	Contrôle de la taille des pas (rôle du β)	72
3.4	Simulation numériques	73
3.4.1	Résultats pour la fonction de Rastrigin	76
3.4.2	Résultats pour la fonction de Creiwank	77
3.5	Applications de la nouvelle technique	78
3.5.1	La fonction de Rosenbroc	78
3.5.2	La fonction de Easom	79
3.5.3	La fonction de Matyas	81
3.6	Conclusion	82
	Conclusion générale et perspectives	83
	Bibliographie	84

Table des figures

1.1	Le point fixe $x^* = 0$ du système $f(x) = ax^3 - bx$. $a = 0.5, b = 0.7$ et $x_0 = 0.6$	11
1.2	attracteur de Lorenz	20
1.3	Exposant de Lyapounov du système logistique en fonction du paramètre μ .	22
1.4	Cascade de doublements de période	28
1.5	Orbite des itérations de la fonction f avec $\lambda = 0.7$ et $x_0 = 0.1$	29
1.6	Orbite des itérations de la fonction f avec $\lambda = 0.7$ et $x_0 = 0.2$	30
1.7	Orbite des itérations de la fonction f pour $\lambda = 0.8$ (à gauche), orbite des itérations de la fonction g pour $\lambda = 0.8$ (à droite)	30
1.8	Orbite des itérations de la fonction f pour $\lambda = 0.87$ (à gauche) , orbite des itérations de la fonction h pour $\lambda = 0.87$ (à droite)	31
1.9	Diagramme de bifurcation de Hénon	34
1.10	Attracteur de Hénon pour $a = 1.4$, $b = 0.3$	35
1.11	Illustration de la permanence de la structure de l'attracteur de Hénon à différentes échelles.	36
1.12	Les exposants de Lyapunov en fonction du temps pour l'application de Hénon.	37
1.13	Attracteur de Lozi pour $a = 1.7$ $b = 0.6$	40
1.14	Diagramme de bifurcation de Lozi	40
1.15	Les exposants de Lyapunov en fonction du temps pour l'application de Lozi	41
2.1	Exemple d'un problème d'optimisation : fonction de Schwefel	44
2.2	Déplacement d'une particule.	60
3.1	Attracteur et séries temporelles de la suite de Lozi	68

3.2	Schéma de la méthode EBCOA	72
3.3	fonction de Rastrigin à deux variables	74
3.4	fonction de Griewank à deux variables	75
3.5	fonction de Rosenbrock	78
3.6	fonction de Easom	80
3.7	fonction de Matyas	81

Introduction générale

Durant ces dernières années, de nombreuses études ont été menées dans le domaine de l'optimisation comme le montre le nombre important de publications sur ce thème dans les revues du domaine. Aujourd'hui, l'optimisation s'applique à tous les domaines de la science et même à notre vie quotidienne. Chacun cherche souvent à mieux gérer son temps, son argent, minimiser certaines consommations, ce sont autant de problèmes d'optimisation.

Les problèmes d'optimisation globale présentent plusieurs difficultés liées aux besoins de l'utilisateur (recherche d'une solution globale, fiabilité et précision de la solution, diversité des problèmes traités, temps de calculs raisonnable,), aux caractéristiques du problème d'optimisation (non linéarité, dérivées difficilement accessibles,..) et au temps de calculs importants. La résolution de telles difficultés a fait l'objet de nombreux travaux en utilisant diverses méthodes d'optimisation.

Les méthodes d'optimisation globale peuvent être subdivisées en deux classes : les méthodes déterministes et les méthodes non déterministes (stochastiques). Les méthodes déterministes s'appuient sur le calcul qui peut être fait d'une direction de recherche, généralement liée à la dérivée de certains résultats par rapport aux paramètres de conception de dispositif. Elles ne sont réellement utilisables que dans le cas restreint où la solution cherchée est réputée proche d'une solution connue, point de départ de cette recherche. C'est pour lever cette difficulté, que nous avons choisi de nous intéresser au développement des méthodes stochastiques et d'étudier leur application.

Les méthodes d'optimisation stochastique s'appuient sur des mécanismes de transition probabilistes et aléatoires. Cette caractéristique indique que plusieurs exécutions successives de ces méthodes peuvent conduire à des résultats différents pour une même configuration initiale d'un problème d'optimisation.

Les trois méthodes stochastiques les plus prometteuses : algorithmes génétiques, recuit simulé et recherche taboue ont été implantées et testées. Ces techniques ont une grande capacité trouver l'optimum global du problème. Contrairement à la plupart des méthodes déterministe, elles ne nécessitent ni point de départ, ni à la connaissance du gradient de fonction objectif pour atteindre la solution optimale. Cependant, elles demandent un nombre important d'évaluations de la fonction objectif.

Pour résoudre ce problème, quelques contributions originales ont été apportées en utilisant la théorie du chaos. Les algorithmes que nous allons étudier font partie des méthodes stochastiques globales permettant de résoudre des problèmes généraux d'optimisation.

Par ailleurs, ils permettent de fournir le plus haut degré de certitude concernant l'optimalité globale des solutions trouvées.

Ces algorithmes s'appellent : Algorithmes d'optimisation Chaotiques(COA).

En effet plusieurs applications des systèmes chaotiques en optimisation globale ont fait leur apparition due à leur comportement aléatoire. L'intérêt d'utiliser des signaux chaotiques réside dans deux propriétés du chaos :

– Un signal chaotique est un signal à large spectre et permet de transmettre des signaux très variés.

– D'autre part, un signal chaotique est obtenu à partir d'un système déterministe. Il n'existe pas encore de définition du chaos valable pour un système dynamique quelconque. Néanmoins, la communauté scientifique s'accorde, dans le cas d'un système dynamique discret correspondant à une fonction d'un espace métrique dans lui même, à le considérer comme chaotique s'il vérifie les propriétés suivantes :

1. La propriété de sensibilité
2. La propriété de mélange
3. La densité des points périodiques

A partir de ces propriétés, nous proposons une nouvelle approche simple basée sur le système chaotique est un bon candidat pour l'optimisation.

Ce travail est consacré à l'application de la théorie du chaos à l'optimisation. Son objective a visé une contribution au développement d'une méthode d'optimisation globale basée sur les séquences chaotiques.

Cette thèse comprend trois chapitres couvrant deux concepts principaux : « Les systèmes dynamiques chaotiques » et « Les méthodes d'optimisation globale ».

Le chapitre I se compose de plusieurs sections qu'elles contiennent l'ensemble des outils mathématiques, numériques nécessaires à une bonne compréhension du travail réalisé par la suite.

Le chapitre II dresse un aperçu non exhaustif des méthodes d'optimisation en accordant une importance particulière aux méthodes utilisées dans nos travaux de recherche. Il aborde aussi les définitions générales des méthodes d'optimisation qui se divisent en deux volets déterministes et non déterministes.

Le chapitre III présente un algorithme d'optimisation chaotique efficace basé sur le l'application logistique. Cet algorithme appelé BCOA. La deuxième partie de ce chapitre est une tentative de développer une technique d'optimisation chaotique. Dans cette approche nous avons amélioré cet algorithme en optimisant localement pendant quelques itérations le résultat global. Cette recherche globale localement moyennée est appelée EB-COA [Enhanced BCOA]. Ce chapitre a fait l'objet d'une publication :

F. Derouiche , T. Hamaizia. An Enhanced Bi-Directional Chaotic Optimization Algorithm Nonlinear Dynamics and Systems Theory, 20 (4) (2020) 365–373.

Chapitre 1

Systèmes dynamiques et chaos

1.1 Définitions et notations

Définition 1

On défénit un système dynamique par un triplet (X, T, f) constitué de l'espace d'états X , du domaine temporel T et d'une application de transition d'état : $f : X \times T \rightarrow X$ qui permet de définir à partir d'un vecteur de conditions initiales l'état du système à tout instant.

Représentations mathématiques des systèmes dynamiques

Un système dynamique décrit par une fonction mathématique présente deux types de variables : dynamiques et statiques, les variables dynamiques sont les quantités fondamentales qui changent avec le temps, les variables statiques, encore appelés paramètres du système, sont fixes.

Dans le cas où la composante “temps” est continue le système dynamique est présenté par un système d'équations différentielles de la forme :

$$\frac{dx}{dt} = f(x, t, p) \text{ où } x \in \mathbb{R}^n \text{ et } p \in \mathbb{R}^r. \quad (1.1)$$

Dans le cas où le temps est discret le système dynamique est présenté par une application itérative :

$$x_{k+1} = f(x_k, p), x \in \mathbb{R}^n \text{ et } p \in \mathbb{R}^r, k = 1, 2, 3, \dots, \quad (1.2)$$

où p un paramètre.

Lorsque le temps t ou l'indice k apparaissent explicitement dans les relations (1.1) et (1.2) le système est dit non-autonome.

En général, c'est un inconvénient majeur pour la résolution numérique et il est préférable de s'en affranchir.

Par un changement de variables approprié, on peut transformer un système non autonome avec $X \in \mathbb{R}^n$ en système autonome avec $X \in \mathbb{R}^{n+1}$.

1.2 Systèmes dynamiques discrets

Définition 2

Soit $f : D \rightarrow D, D \subseteq \mathbb{R}^n$ une application continue (ou une transformation), f^k désigne la $k^{\text{ième}}$ itérée de f , c'est-à-dire :

$$f^0(x) = x, f^1(x) = f(x), f^2(x) = f(f(x)), \dots, f^k(x) = f(f^{k-1}(x)), \forall k = 0, 1, 2, \dots$$

Dans la pratique $x_0, x_1 = f(x_0), x_2 = f^2(x_0), \dots$ représentent les valeurs d'une certaine quantité au temps $0, 1, 2, \dots$

Ainsi la valeur de la quantité au temps $k + 1$ est fonction de sa valeur au temps k .

L'application f est appelée un système dynamique discret.

1.2.1 Orbites ou trajectoires

Définition 3

L'orbite positive de x par le système dynamique f est définie par :

$$O_f^+ = \{f^k(x), k \in \mathbb{N}\}$$

Si f est bijectif, on définit l'orbite de x par :

$$O_f = \{f^k(x), k \in \mathbb{Z}\}$$

Ainsi que l'orbite négative :

$$O_f^- = \{f^{-k}(x), k \in \mathbb{N}\}$$

1.2.2 Points fixes

La notion de points d'équilibre est centrale dans l'étude de la dynamique de tout système physique. Dans de nombreuses applications en biologie, économie, physique, ingénierie, etc., il est souhaitable que tous les états (solutions) d'un système donné tendent à son état d'équilibre (point d'équilibre). C'est le sujet d'étude de la théorie de la stabilité, un sujet d'une grande importance pour les scientifiques et les ingénieurs.

Définition 4

Un point fixe d'une application $x_{k+1} = f(x_k, p)$ est un point invariant par f , c'est-à-dire un point x^* de l'espace des phases vérifiant l'équation $f(x^*) = x^*$.

Géométriquement, le point fixe est une intersection de la courbe de la fonction $y = f(x)$ avec la bissectrice $y = x$.

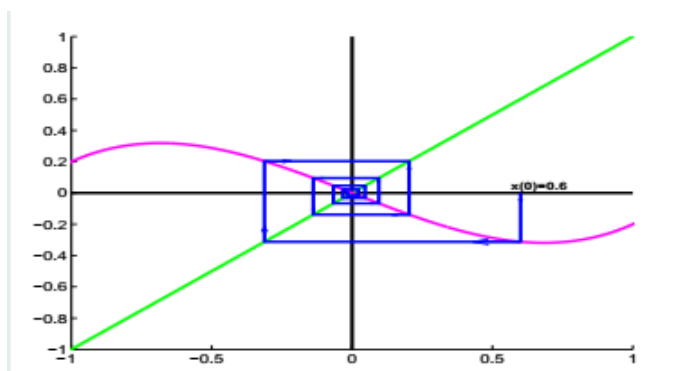


FIGURE 1.1 – Le point fixe $x^* = 0$ du système $f(x) = ax^3 - bx$. $a = 0.5$, $b = 0.7$ et $x_0 = 0.6$

1.2.3 Points périodiques et p-cycles

S'il existe $n \geq 1$, tel que $f^n(x) = x$, on dit que x est un point périodique.

La période d'un point périodique x est le plus petit entier $n \geq 1$ tel que :

$$f^n(x) = x.$$

Un ensemble $\{x_0, x_1, \dots, x_{p-1}\}$ forme un cycle d'ordre p (ou une orbite périodique d'ordre p , ou encore un p -cycle), si :

$$f(x_i) = x_{i+1} \text{ pour } i = 1, 2, \dots, p-1 \text{ et } f(x_{p-1}) = x_0$$

.

Autrement dit chaque point d'un cycle d'ordre p est un point fixe pour f^p , où $f^p(x_i) = x_i$ pour $i = 0, 1, 2, \dots, p-1$, et n'est pas un point fixe pour f^k si $k < p$.

1.2.4 Étude de stabilité

1.2.5 Stabilité du point fixe

Définition 5

Un point fixe x^* s'appelle stable si $\forall \varepsilon > 0$ il existe un $\delta > 0$ tel que si $|x_0 - x^*| < \delta$ alors pour tout $n > 0$ $|f^n(x_0) - x^*| < \varepsilon$

Autrement dit, toutes les orbites qui commencent près du point x^* restent dans un voisinage de ce point.

La stabilité du point x^* est déterminée par la pente $m = f'(x^*)$ de la tangente au point x^* à la courbe représentant f .

Nous avons les quatre cas suivants :

Définition 6

1. Le point x^* est attractif (ou stable) si $|m| < 1$.
2. Le point x^* est répulsif (ou instable) si $|m| > 1$.
3. Le point x^* est indifférent si $|m| = 1$:
4. Le point x^* est super attractif (ou super stable) si $m = 0$.

m s'appelle le multiplicateur de f au point x . En dimension n , pour décider si un point x^* est attractif ou non, il faut calculer les valeurs propres de la matrice jacobienne $Df(x) = J(x)$. Plus précisément si toutes les valeurs propres de $Df(x) = J(x)$ sont à l'intérieur du disque unité, x^* est stable. Si au moins une de ces valeurs propres a un

module plus grand que un, x^* est instable.

1.2.6 Stabilité des orbites périodiques.

Définition 7

Soit x_p un point périodique de période r d'un système dynamique discret défini par une fonction $f(x)$. L'orbite périodique correspondante $O(x_p)$ s'appelle stable (instable) si chacun de ses points est un point fixe stable (instable) de l'application f^r .

Exemple 1 Soit $f(x) = 1 - x$. Tout point $x \neq 0.5$ est un point périodique de période 2. Or l'application $f^2(x) = 1 - (1 - x) \equiv x$ est l'application identité. Tous les points sont des points fixes stables de cette application. Donc pour tout point $x \neq 0.5$ l'orbite correspondante $O(x) = \{x, 1 - x\}$ est une orbite périodique stable.

1.2.7 Équivalence topologique des systèmes

Définition 8 Soient (D, f) et (E, g) deux systèmes dynamiques. On dit qu'ils sont topologiquement conjugués s'il existe un homéomorphisme (une application continue et bijective) $h : D \rightarrow E$ tel que $hof = goh$

Théorème 1

Soient (D, f) et (E, g) deux systèmes dynamiques. Supposons qu'ils sont topologiquement conjugués par un homéomorphisme $h : D \rightarrow E$ Alors

1. L'application $h^{-1} : E \rightarrow D$ vérifie aussi la définition et assure donc l'équivalence topologique entre les systèmes (D, f) et (E, g) .
2. $hof^n = g^n oh$, pour tout $n \in \mathbb{N}$.
3. Si $x_p \in D$ est un point périodique de f de période fondamentale p alors $h(x_p) \in E$ est un point périodique de g de période fondamentale p .

1.2.8 Bifurcation

La bifurcation signifie un changement qualitatif de la dynamique du système, qui résulte du changement d'un des paramètres du système. Par exemple : déstabilisation

d'un équilibre stable, apparition ou disparition d'un cycle ou d'un attracteur,....etc

La valeur du paramètre pour laquelle la bifurcation se produit est nommée le point de bifurcation.

De nombreux types de bifurcation sont inventoriés dans [19, 12], Dans cette section on présentera trois types de bifurcations qui sont : la bifurcation noeud-col , bifurcation fourche et la bifurcation de Neimark-Sacker .

1.Bifurcation de type noeud-col

La bifurcation $\lambda = +1$ correspond à la situation où l'une des deux valeurs propres est égale à +1.

Ce type de bifurcation donne naissance à deux cycles d'ordre k en même temps, l'un est attractif et l'autre est instable. Cette situation peut être représentée par :

$\emptyset \leftrightarrow$ cycle attractif + cycle répulsif

Où \emptyset signifie absence de cycle.

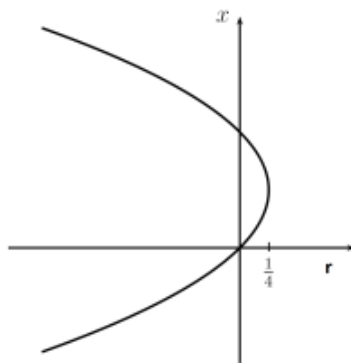
Exemple 2 *Considérons la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ avec $f_r(x) = x^2 + r$ où r est un paramètre réel .*

Les points fixes sont solutions de $x = x^2 + r$ qui sont :

$$x_1 = \frac{1}{2} + \frac{\sqrt{(1-4r)}}{2}, \text{ et } x_2 = \frac{1}{2} - \frac{\sqrt{(1-4r)}}{2}.$$

qui existent seulement pour $r \leq \frac{1}{4}$ (et coincident si $r = \frac{1}{4}$)

On représente ci-dessous l'ensemble $D = \{(r, x) \in \mathbb{R}^2 \mid x \text{ point fixe de } f\}$, c'est ici la parabole $x^2 - x + r = 0$.

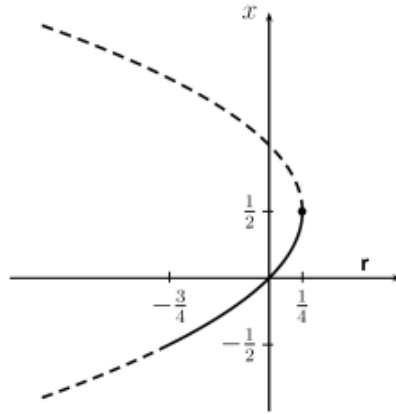


Etudions la stabilité de x_1 et x_2 en fonction de r .

On a $f'_r(x) = 2x$ d'où $f'_r(x_1) = 2x_1 = 1 + \sqrt{1 - 4r}$ et $f'_r(x_2) = 2x_2 = 1 - \sqrt{1 - 4r}$

pour $r < \frac{1}{4}$, x_1 est répulsif, et x_2 est attractif si et seulement si :

$|1 - \sqrt{1 - 4r}| < 1$ c'est-à-dire ($-\frac{3}{4} < r < \frac{1}{4}$) Afin de compléter la figure précédente , nous représentons en pointillés la partie de l'ensemble D qui correspond a des points fixes instables, ce qui donne :



On obtient ainsi une première portion de ce qui s'appelle le diagramme de bifurcation de f_r .

La situation, concernant les point fixes est décrite par :

$$\left\{ \begin{array}{l} r = \frac{1}{4} \quad f_r \text{ a un seul point fixe.} \\ \forall r > \frac{1}{4} \quad f_r \text{ n'a pas de points fixes réels.} \\ \forall r < \frac{1}{4} \quad f_r \text{ a deux points fixes réels : l'un est instable et l'autre est stable pour } r > -\frac{3}{4} \end{array} \right.$$

C'est ce que l'on appelle une bifurcation noeud-col, $r = \frac{1}{4}$ est appelée valeur de bifurcation et le point $(\frac{1}{4}, \frac{1}{2})$ est appelé point de bifurcation.

2. Bifurcation fourche

Cette bifurcation a lieu lorsqu'une des deux valeurs propres est égales à -1 . Un cycle d'ordre k qui subie cette bifurcation va changer de nature et crée un cycle d'ordre $2k$ de la même nature. C'est-à-dire, un point fixe stable d'ordre 1, par exemple, devient instable en même temps que l'apparition d'un cycle d'ordre 2 stable. Cette situation peut être représentée par :

Cycle attractif d'ordre $k \leftrightarrow$ cycle répulsif d'ordre k + cycle attractif d'ordre $2k$

Exemple 3 Considérons la fonction $f_r(x) = rx - x^3$. Ici, nous nous limiterons aux

valeurs de $r > 0$.

Recherche des Points fixes

Les points fixes de f_r sont les solutions de $rx - x^3 = x$. Ainsi 0 est un point fixe de f_r $\forall r > 0$ et, pour $r > 1$ deux points fixes vont apparaître $x_1 = \sqrt{r-1}$ et $x_2 = -\sqrt{r-1}$.

Stabilité des points fixes

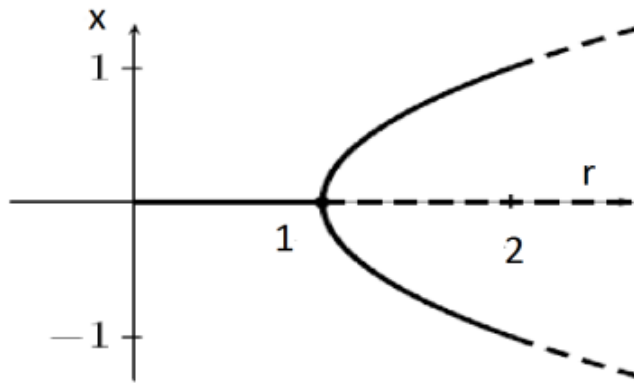
Nous avons $f'_r(x) = r - 3x^2$. Ainsi $f'_r(0) = r$ et $f'_r(x_1) = f'_r(x_2) = 3 - 2r$. Il s'ensuit donc que 0 est attractif si $0 < r < 1$ et répulsif si $r > 1$, alors que x_1 et x_2 sont attractif pour $1 < r < 2$ et répulsif pour $r > 2$.

Traçons maintenant l'ensemble

$$D = \{(x, r) | r > 0 \text{ et } x \text{ un point fixe de } f_r\}$$

où encore :

$$D = \{(x, r) | r > 0\} \cup \{(x, r) | x^2 = r - 1, \forall r \geq 1\}$$



3. Bifurcation de Neimark-Sacker

La bifurcation de Neimark-Sacker est définie pour un système de dimension plus grand que 1, cette bifurcation est caractérisé par la naissance (quand le paramètre de bifurcation μ dépasse le point de bifurcation μ_0) d'une courbe invariante fermée à partir d'un point fixe, lorsque le point fixe change de stabilité via une paire de valeurs propres complexes avec un module unitaire. La bifurcation peut être super-critique ou sous-critique, donnant lieu à une courbe invariante fermée stable ou instable, respectivement.

1.3 Chaos

1.3.1 Introduction

En 1963 le météorologue Edward Lorenz expérimentait une méthode lui permettant de prévoir les phénomènes météorologiques. C'est par pur hasard qu'il observa qu'une modification minime des données initiales pouvait changer de manière considérable ses résultats. Lorenz venait de découvrir le phénomène de sensibilité aux conditions initiales. Les systèmes répondant à cette propriété seront à partir de 1975 dénommés : systèmes chaotiques. C'est donc au cours des années soixante dix que la théorie du chaos a pris son essor. Cependant, les travaux de certains scientifiques menés bien avant cette découverte vont être très utiles à la compréhension de la dynamique chaotique. En effet, vers la fin du XIXe siècle le mathématicien, physicien et philosophe français Henri Poincaré avait déjà mis en évidence le phénomène de sensibilité aux conditions initiales lors de l'étude astronomique du problème des trois corps. On trouve dans le Calcul des Probabilités de Henri Poincaré l'affirmation suivante :

« Une cause très petite, qui nous échappe, détermine un effet considérable que nous ne pouvons pas ne pas voir, et alors nous disons que cet effet est dû au hasard. Si nous connaissions exactement les lois de la nature et la situation de l'univers à l'instant initial, nous pourrions prédire exactement la situation de ce même univers à un instant ultérieur. Mais, lors même que les lois naturelles n'auraient plus de secret pour nous, nous ne pourrions connaître la situation qu'approximativement. Si cela nous permet de prévoir la situation ultérieure avec la même approximation, c'est tout ce qu'il nous faut, nous disons que le phénomène a été prévu, qu'il est régi par des lois; mais il n'en est pas toujours ainsi, il peut arriver que de petites différences dans les conditions initiales en engendrent de très grandes dans les phénomènes finaux; une petite erreur sur les premières produirait une erreur énorme sur les derniers. La prédiction devient impossible et nous avons le phénomène fortuit. » Cette citation définit parfaitement le chaos en tant que sensibilité aux conditions initiales mais aussi le déterminisme qui réside dans le fait que si une condition initiale est parfaitement déterminée alors l'évolution du système l'est aussi. Le déterminisme traduit l'unicité de la solution pour l'équation différentielle d'un système

donné, c'est le théorème de Cauchy. Toujours au XIXe siècle, le mathématicien Russe Alexandre Lyapunov effectue des recherches sur la stabilité du mouvement. Il introduit l'idée de mesurer l'écart entre deux trajectoires ayant des conditions initiales voisines, lorsque cet écart évolue exponentiellement on parle de sensibilité aux conditions initiales. Les travaux de Lyapunov, d'abord tombés dans l'oubli, seront plus tard très précieux pour étudier certains aspects de la théorie du chaos. Les travaux des prédécesseurs de Lorenz ont donc été très importants pour la compréhension du chaos déterministe, mais il faut souligner que ce qui va permettre aux scientifiques une compréhension plus accrue des systèmes chaotiques c'est l'ordinateur. En effet, les équations différentielles régissant un système chaotique sont nécessairement non linéaires et, sans ordinateur, leur résolution est en général impossible.

1.3.2 Systèmes dynamiques chaotiques

On trouve dans la littérature plusieurs définitions mathématiques du chaos, mais jusqu'à présent, il n'existe aucune définition mathématique universelle du chaos. Avant de donner une définition du chaos, due à R.L Devaney [40], quelques définitions de base sont nécessaires.

Définition 9

$f : J \rightarrow J$ est dite topologiquement transitive si pour toute paire d'ensembles ouverts $U, V \subset J$ il existe $k > 0$ tel que $f^k(U) \cap V \neq \emptyset$.

Définition 10

$f : J \rightarrow J$ a une sensibilité aux conditions initiales s'il existe $\delta > 0$ tel que, pour tout $x \in J$ et tout voisinage N_x de x , il existe $y \in N_x$ et $n > 0$ tels que $|f^n(x) - f^n(y)| > \delta$

Définition 11

Supposons que X un ensemble et Y un sous-ensemble de X ($Y \subset X$), on dit que Y est dense dans X si, pour tout $x \in X$, il existe $y \in Y$ arbitrairement proche de x , autrement dit Y est dense dans X si pour tout $x \in X$ on peut trouver une suite $\{y_n\}_{n \in \mathbb{N}}$ de Y qui converge vers x .

On est maintenant en position d'énoncer la définition du chaos au sens de Devaney [40].

Définition 12

Soit un sous-ensemble V de X , la fonction $f : X \rightarrow X$ est dite chaotique sur V si :

1. La fonction f possède une sensibilité aux conditions initiales.
2. La fonction f est topologiquement transitive.
3. L'ensembles des points périodiques de la fonction f sont denses dans X .

1.3.3 Les attracteurs

Un système chaotique dissipatif possède (au moins) un attracteur d'un type particulier appelé attracteur étrange. Géométriquement, un tel attracteur peut être décrit comme le résultat d'une opération d'étirement et de repliement d'un cycle de l'espace des phases, répétée un nombre infini de fois. La "longueur" de l'attracteur est infinie, bien qu'il soit contenu dans un espace fini. Alors on peut donner cette définition :

Définition 13

Soit $U \supset \mathbb{R}^q$ un ouvert et soit l'application $f : U \rightarrow U$. Un ensemble fermé et borné $A \subset U$ est un attracteur si $f(A) = A$ et s'il existe $r > 0$ tel que $d(x_0, A) < r$ implique que $\lim_{n \rightarrow +\infty} d(x_n, A) = 0$

Les attracteurs jouent un rôle fondamental dans l'étude du comportement à long terme d'un système dynamique.

Certain de ces attracteurs ont une géométrie très compliquée au point que certains d'entre eux ont été appelés attracteurs étranges [4]. Les attracteurs peuvent être classés dans trois catégories [5] :

1. Point fixe (un seul point),
2. Orbite périodique (un ensemble fini des points).
3. Un attracteur chaotique (attracteur étrange) est tout autre type d'attracteurs.

Définition 14 [50] *Le bassin d'attraction B d'un attracteur A est l'ensemble des conditions initiales (l'ensemble des tous états initiaux des orbites) est à long-temps un comportement approche vers A*

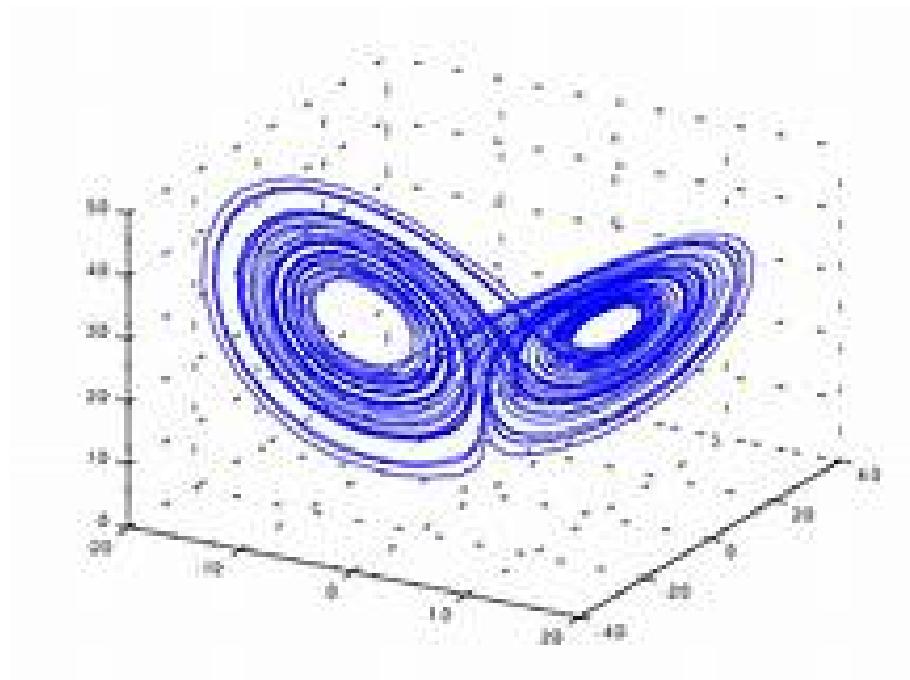


FIGURE 1.2 – attracteur de Lorenz

1.3.4 Détection du chaos

Il existe plusieurs méthodes qui permettent de déterminer si des systèmes non linéaires sont ou non chaotiques. Elles ne sont généralement pas très nombreuses, ni réparties sur un temps suffisamment long à l'échelle du système étudié. On a choisi de mettre en oeuvre deux des méthodes les plus couramment utilisées qui, d'ailleurs, sont complémentaires : la dimension fractale et les exposants de Lyapunov.

1. Les exposants de Lyapunov

Alexandre Lyapunov a développé une quantité permettant de mesurer la divergence des trajectoires qui sont voisines au départ, cette quantité est appelée “ exposant de Lyapunov ” qui est souvent utilisé pour déterminer si un système est chaotique ou non.

Cas d'une application discrète unidimensionnelle.

Soit une application discrète f de \mathbb{R} dans \mathbb{R} qui applique x_n sur x_{n+1} : Choisissons deux conditions initiales très proches, soit x_0 et $x_0 + \epsilon$ et regardons comment se comportent les trajectoires qui en sont issues. Supposons qu'elles s'écartent en moyenne à un rythme exponentielle. On pourra trouver un réel tel que après n itérations on a :

$$|f^n(x_0 + \epsilon) - f^n(x_0)| \approx \epsilon \exp(n\mu)$$

d'où

$$\ln \left| \frac{f^n(x_0 + \epsilon) - f^n(x_0)}{\epsilon} \right| \approx n\mu, \forall \epsilon > 0$$

En faisant tendre ϵ vers zéro, on trouve :

$$\mu \approx \ln \left| \frac{df^n(x_0)}{dx_0} \right|.$$

Finalement, en faisant tendre μ vers l'infini et utilisant la règle de dérivation en chaîne :

$$\frac{d}{dx} f^n(x_0) = \prod_{i=0}^{n-1} f'(x_i)$$

on obtient :

$$\mu = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln |f'(x_i)|,$$

μ est appelé exposant de Lyapunov.

Généralement, on peut distinguer trois cas des orbites suivant le signe de l'exposant de Lyapunov :

1. Si $\mu < 0$, l'orbite est attractive vers un point fixe ou une orbite périodique stable, il caractérise les systèmes dissipatifs. Ce type de système exhibe une stabilité asymptotique, plus l'exposant est négatif, plus la stabilité est grande. Les points fixes et les points périodiques super-stables ont un exposant de Lyapunov μ qui tend vers $-\infty$.
2. Si $\mu = 0$, l'orbite est un point fixe neutre. Un système physique avec un tel exposant est dit conservatif. Dans cette situation, les orbites gardent une séparation constante.
3. Si $\mu > 0$, l'orbite est instable et chaotique. Tous les points voisins doivent être visités, ces points sont dits instables. Pour un système discret, on a un ensemble de points sans aucun rapport de liaison.

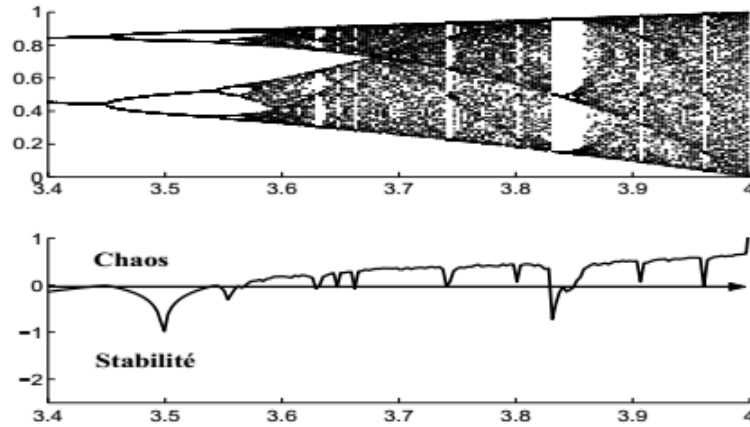


FIGURE 1.3 – Exposant de Lyapounov du système logistique en fonction du paramètre μ

Cas d'une application discrète multidimensionnelle

Soit f une application discrète de \mathbb{R}^m dans \mathbb{R}^m :

$$x_{n+1} = f(x_n)$$

Un système m -dimensionnel possède m exposants de Lyapounov, chacun d'entre eux mesure le taux de divergence suivant un des axes du système, de sorte qu'en moyenne un hyper-volume initial V_0 évolue selon une loi de type :

$$V = V_0 e^{(\mu_1 + \mu_2 + \dots + \mu_m)n}.$$

Pour avoir du chaos, il est nécessaire qu'au moins un μ_i soit positif, pour avoir étirement selon au moins un axe. Mais il faut aussi que la somme des μ_i soit négative. Puisque, dans le cas contraire, le volume initial finirait par remplir tout l'espace dans lequel il est immergé et on n'aurait plus un attracteur de faible dimension, ce qui signifie qu'on n'aura pas du chaos déterministe. Tout d'abord nous devons calculer les μ_i . Comme précédemment, nous nous intéressons à :

$$f^n(x_0 + \epsilon) - f^n(x_0).$$

Posons

$$x'_0 = x_0 + \epsilon.$$

On a le développement en série limité d'ordre 1 de $f_n(x_0)$ au voisinage de x'_0 suivant :

$$\begin{aligned}x_n - x'_n &\approx \frac{\partial f^n(x_0)}{\partial x_0}(x_0 - x'_0) \\ &\approx J(x_0)J(x_1)\dots J(x_n)(x_0 - x'_0) \\ &\approx \prod_{i=1}^n J(x_i)(x_0 - x'_0).\end{aligned}$$

On note

$$\prod_{i=1}^n J(x_i) \text{ par } J^n(x_0),$$

ainsi

$$x_n - x'_n \approx J^n(x_0)(x_0 - x'_0),$$

$J^n(x_0)$ dénote la matrice jacobienne de f^n au point x_0 : Il s'agit d'une matrice carrée $m \times m$, si elle est diagonalisable, alors il existe une matrice inversible P_n telle que :

$$D_m^t = P_n^{-1} J^n P_n,$$

D_m^t est une matrice diagonale des valeurs propre

$$\lambda_i(f^n(x_0)), i = 1, \dots, m \text{ de } J^n.$$

On définit alors les m exposants de Lyapunov de la manière suivante :

$$\mu_i = \lim_{n \rightarrow +\infty} \frac{1}{n} \ln |\lambda_i(f^n(x_0))|, i = 1, \dots, m.$$

Pour le point d'équilibre x^* , la formule devient :

$$\mu_i = \ln |\lambda_i(x^*)|, i = 1, \dots, m.$$

2. Dimension fractale

Cette méthode correspond à une mesure de la dimension de l'attracteur reconstruit du système étudié, nous pouvons calculer la dimension de l'attracteur du système étudié et ainsi déterminer si oui ou non il est construit de manière fractale. Si, à l'issue du calcul, nous obtenons une valeur positive non entière, cela signifie que le système possède un attracteur étrange. Plusieurs dimensions ont été proposées, on cite quelques une : il ya la dimension de Kolmogorov, dimension de Corrélacion et dimension de Lyapunov, il ya une différence légère entre chaqu'une de ces dimension, mais elles caractérisent, toutes, l'attracteur étrange avec sa dimension fractale et satisfont les propriétés suivantes :[51]

1. $A \subset B \Rightarrow d(A) \leq d(B)$
2. $A = \emptyset \Rightarrow d(A) = 0$
3. $d(A \times B) = d(A) + d(B)$

Plusieurs définition de dimension fractale ont été proposées depuis le debut de ce siècle. Certaines ont un intérêt purement théorique, pour une étude de ces définitions nous renvoyons à [7, 2, 15].

-Dimension de Corrélacion

Le principe de cette méthode repose sur les corrélacions entre les points de la série temporelle étudiée. Il s'agit de compter combien de cubes de côté de longueur R il faut pour contenir tous les points de l'espace de dimension N . Les données sont regroupées en multiples de p et on compte le nombre de cubes nécessaires pour prendre en compte tous les points de l'espace. On commence par utiliser une toute petite boîte puis on l'agrandi afin que tout les points soient contenus dans une seule et même grande boîte. Notons $C(R)$ le nombre de boîtes requis pour chaque dimension N : On peut caractériser la dimension de corrélacion à l'aide d'une fonction appropriée, une telle fonction est défini par :

$$C(R) = \lim_{N \rightarrow +\infty} \frac{1}{N^2} \vartheta \quad (1.3)$$

avec ϑ représente le nombre de paires i, j dont la distance $|X_i - X_j| < R$.

On peut réécrire la fonction (1.3) comme :

$$C(R) = \lim_{N \rightarrow +\infty} \frac{1}{N^2} \sum H(R - |X_i - X_j|),$$

avec H la fonction de Heaviside. La dimension de corrélation est alors définie par :

$$\vartheta = \lim_{R \rightarrow +\infty} \frac{\log(C(R))}{\log(R)}.$$

La distribution sur l'attracteur étant relativement facile à estimer, l'aspect fractal des attracteurs est souvent caractérisé simplement par cette dimension de corrélation.

-Dimension de Lyapunov

Soient $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ les exposants de Lyapunov d'un attracteur d'un système dynamique pour un espace de phase de dimension plus grand que 2.

Soit k , $1 \leq k \leq n - 1$, le plus grand entier naturel tel que :

$$\sum_{i=1}^k \lambda_i \geq 0 \text{ et } \sum_{i=1}^{k+1} \lambda_i < 0.$$

La dimension de Lyapunov défini par Kaplan et Yorke est [11][13] :

$$D_L = k + \frac{\sum_{i=1}^k \lambda_i}{\lambda_{k+1}}$$

-Dimension de capacité (Kolmogorov)

Soit X un ensemble de points de l'attracteur, on recouvre X par un nombre minimal $N(\varepsilon)$ d'hypercube de coté ε .

- Si X est un carré de coté L , il peut être recouvert par $N(\varepsilon) = (\frac{L}{\varepsilon})^2$ petit carrés de cotés ε .
- Dans le cas général on a :

$$N(\varepsilon) = \left(\frac{L}{\varepsilon}\right)^d,$$

$$d = \frac{\ln N(\varepsilon)}{\ln L - \ln \varepsilon} \text{ quand } \varepsilon \rightarrow 0, \quad \ln L \ll \ln \varepsilon$$

Définition 15

La dimension de Kolmogorov ou de capacité est définie par :

$$d_c = - \lim_{\varepsilon \rightarrow 0} \ln \frac{N(\varepsilon)}{\ln \varepsilon}$$

1.3.5 Routes vers le chaos

On ne sait pas à l'heure actuelle dans quelles conditions un système va devenir chaotique. Cependant il existe plusieurs types d'évolution possibles d'un système dynamique régulier vers le chaos. Supposons que la dynamique étudiée dépende d'un paramètre de contrôle [36][23].

Lorsqu'on varie ce paramètre, le système peut passer d'un état stationnaire à un état périodique, puis au-delà d'un certain seuil, suivre un scénario de transition et devenir chaotique. Il existe plusieurs scénarios qui décrivent le passage du point fixe au chaos. D'une manière générale, l'évolution du point fixe vers le chaos n'est pas progressive mais marquée par des changements discontinus appelés bifurcations. Une bifurcation marque le passage soudain d'un régime dynamique à un autre, qualitativement différent. Tous ces scénarios ont été prédits par la théorie et observés dans de nombreuses expériences. En physique, c'est notamment la convection thermique de Rayleigh-Bénard, dans laquelle une couche de fluide située entre deux plaques horizontales est soumise à un gradient de température vertical, qui a servi à l'origine de système modèle pour l'étude du chaos. Depuis, le chaos a été mis en évidence dans bien d'autres domaines. Nous allons en exposer brièvement trois types d'évolution possibles.

-Par doublement de période

Ce scénario de transition vers le chaos est sans doute le plus connu. Par augmentation du paramètre de contrôle de l'expérience, la fréquence du régime périodique double, puis est multipliée par 4, par 8, par 16 ... etc. Les doublements étant de plus en plus rapprochés, on tend vers un point d'accumulation auquel on obtiendrait hypothétiquement une fréquence infinie. C'est à ce moment que le système devient chaotique. Il a été étudié en particulier en dynamique de populations par RMay sur l'application logistique, $x_{n+1} = ax_n(1 - x_n)$ Selon la valeur du paramètre a , la suite converge soit vers un point

fixe nul ou pas. Dès que a est plus grand que 3 le système bifurque, c'est à dire qu'il oscille entre 2 valeurs autour du point fixe. On parle de cycle attracteur de période 2. En continuant à augmenter a , ces 2 attracteurs s'écartent du point fixe jusqu'à ce qu'une nouvelle bifurcation ait lieu. Chaque point se dédouble et on obtient un cycle attracteur de période 4. On dit qu'il y a doublement de période. C'est à partir de cet exemple que Feigenbaum pressentit l'existence d'une forme d'universalité dans cette transition vers le chaos sous forme de cascade de doublement de période.

-Par Intermittences

Ce scénario via les intermittences se caractérise par l'apparition erratique de bouffées chaotiques dans un système qui oscille de manière régulière. Le système conserve pendant un certain laps de temps un régime périodique ou pratiquement périodique, c'est à dire une certaine "régularité", et il se déstabilise, brutalement, pour donner lieu à une sorte d'explosion chaotique. Il se stabilise de nouveau ensuite, pour donner lieu à une nouvelle "bouffée" plus tard. On a constaté que la fréquence et la durée des phases chaotiques avaient tendance à s'accroître plus on s'éloignait de la valeur critique de la contrainte ayant conduit à leur apparition. L'intermittence suppose en particulier que le cycle limite (correspondant à l'état périodique d'où est issu ce phénomène de transition) bifurque de façon sous-critique et qu'il n'y ait pas d'attracteur à proximité. C'est ce que l'on observe dans le système de Rössler.

-Quasi-périodicité

Le scénario via la quasi-périodicité a été mis en évidence par les travaux théoriques de Ruelle et Takens (1971) illustré par exemple sur le modèle de Lorenz (1963). Ce scénario a été confirmé par de nombreuses expériences dont les plus célèbres se trouvent en thermo-hydrodynamique - convection de Rayleigh-Bénard dans une petite boîte - et en chimie - réaction de Bélousov-Zabotinsky - entre autres. Cette route vers le chaos résulte de la "concurrence" de différentes fréquences dans le système dynamique. Dans un système à comportement périodique à une seule fréquence, si nous changeons un paramètre alors il apparaît une deuxième fréquence. Si le rapport entre les deux fréquences est rationnelle le comportement est périodique. Mais, si le rapport est irrationnel, le comportement est quasi périodique. Dans ce cas, les trajectoires couvrent la surface d'un tore. Alors, on

change de nouveau le paramètre et il apparaît une troisième fréquence, et ainsi de suite jusqu'au chaos. n existe aussi des systèmes qui passent directement de deux fréquences au chaos.

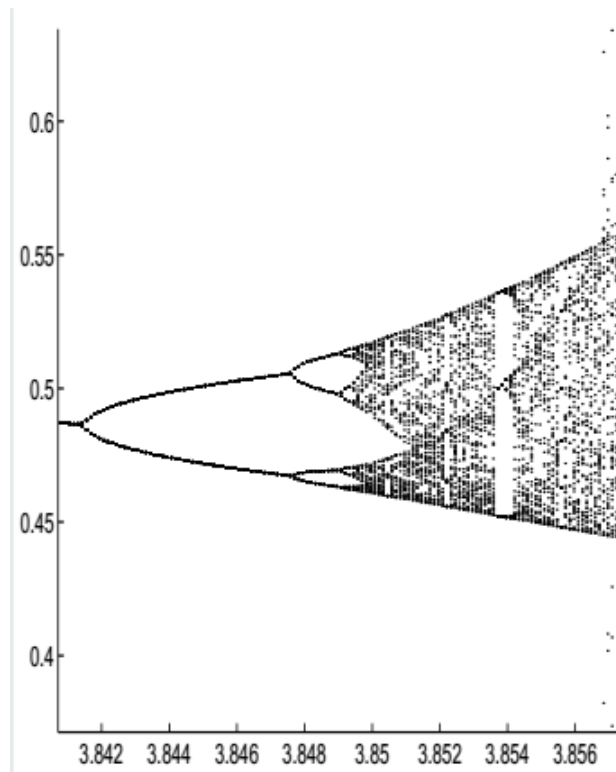


FIGURE 1.4 – Cascade de doublements de période

1.4 Exemples de Systèmes dynamiques chaotiques

1.4.1 Application logistique

Considérons l'application f qui est définie de $[0, 1]$ dans lui-même par l'itération suivante :

$$f(x_n) = 4\lambda x_n(1 - x_n), x_n \in [0, 1]$$

où $n = 0, 1, 2, \dots$ dénote le temps discret, x_n l'unique variable dynamique, et $0 \leq \lambda \leq 1$ un paramètre.

f est la fonction logistique, elle s'annule pour $x = 0$ et $x = 1$, et sa dérivée s'annule

pour $x = \frac{1}{2}$ donc atteint le maximum à $x = \frac{1}{2}$ et $f(\frac{1}{2}) = \lambda$.

Les points fixes de f sont les solutions de l'équation :

$$x = 4\lambda x(1 - x), \lambda > 0 \text{ d'où } x_1 = 0 \text{ et } x_2 = 1 - \frac{1}{4\lambda}$$

On voit bien sur la figure 1.5 que les points d'équilibre de la suite considérée correspondent aux intersections de la courbe d'équations $y = f$ et $y = x$.

La dynamique de cette application présente un comportement très différent selon la valeur du paramètre λ

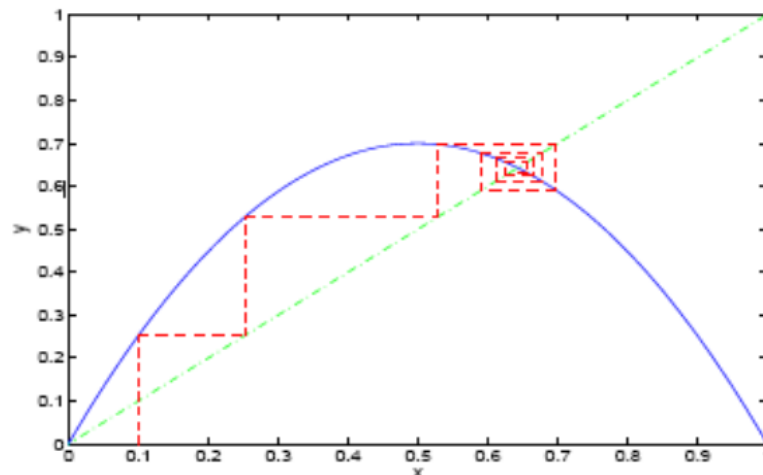


FIGURE 1.5 – Orbite des itérations de la fonction f avec $\lambda = 0.7$ et $x_0 = 0.1$

Valeurs de λ comprises entre 0 et 1

Les domaines de stabilité sont alors donnés par :

(x_i est stable si $|f'(x_i)| < 1$)

$0 < \lambda < 0.25$ pour x_1 .

$0.25 \leq \lambda \leq \lambda_1$ avec $\lambda_1 = 0.75$ pour x_2

Il existe alors, pour $0 < \lambda \leq \lambda_1$, un unique point fixe stable, en modifiant la condition initiale x_0 , la suite converge toujours mais la vitesse de convergence est différente.

pour $\lambda_1 = 0.75$, $f'(x_2) = -1$ donc $x = \frac{2}{3}$ est un point de bifurcation.

Valeurs de λ comprises entre λ_1 et λ_2

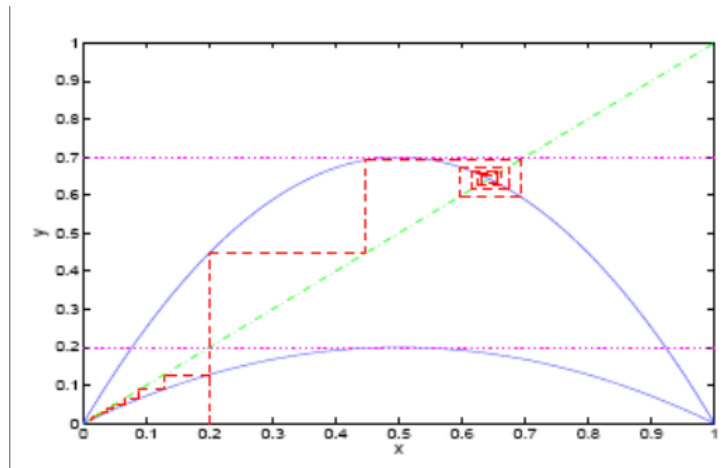


FIGURE 1.6 – Orbite des itérations de la fonction f avec $\lambda = 0.7$ et $x_0 = 0.2$

On relève figure 1.7 que pour des valeurs de λ comprises entre $\lambda_1 = 0.75$ et $\lambda_2 = 0.86237$, deux points de convergence x_3 et x_4 prennent naissance autour de x_2 , et vérifient :

$$x_3 = f(x_4) \text{ et } x_4 = f(x_3)$$

Ces points ne sont donc pas des points fixes de f mais de $g = f \circ f$, où \circ désigne la composition, ils forment un attracteur d'ordre 2 (cycle).

pour $\lambda_2 = 0.86237$ le cycle d'ordre 2 perd sa stabilité et donne lieu a un cycle d'ordre 4.

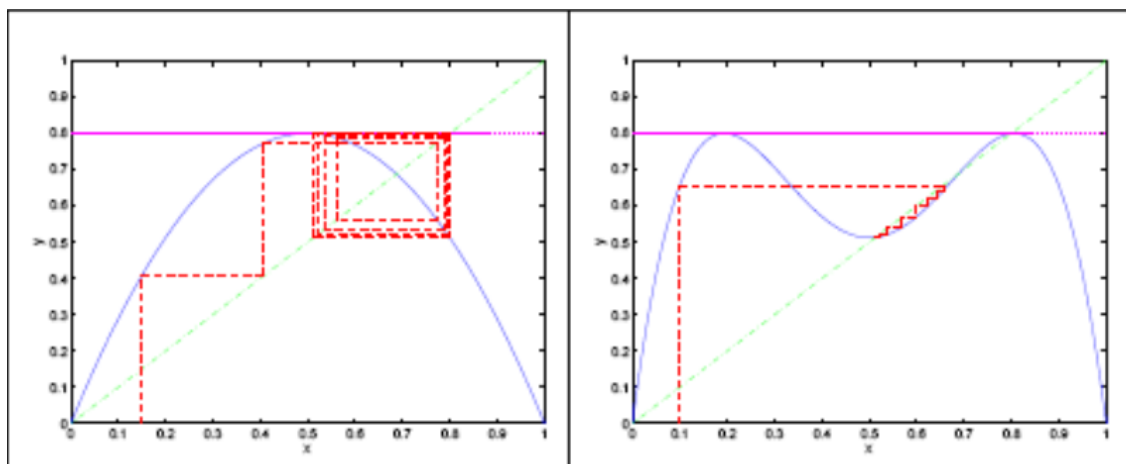


FIGURE 1.7 – Orbite des itérations de la fonction f pour $\lambda = 0.8$ (à gauche), orbite des itérations de la fonction g pour $\lambda = 0.8$ (à droite)

Valeurs de λ comprises entre λ_2 et λ_3

De même, pour des valeurs de λ comprises entre $\lambda_2 = 0.86237$ et $\lambda = 0.87$, la pente de f devient supérieure à 1, et la suite prend alors quatre valeurs différentes, qui sont :

$$x_5 = f(x_4), x_6 = f(x_5), x_7 = f(x_6), x_8 = f(x_7)$$

Ces points sont des points fixes de la fonction $h = g \circ g$.

Il y a ainsi quadruplement de période.

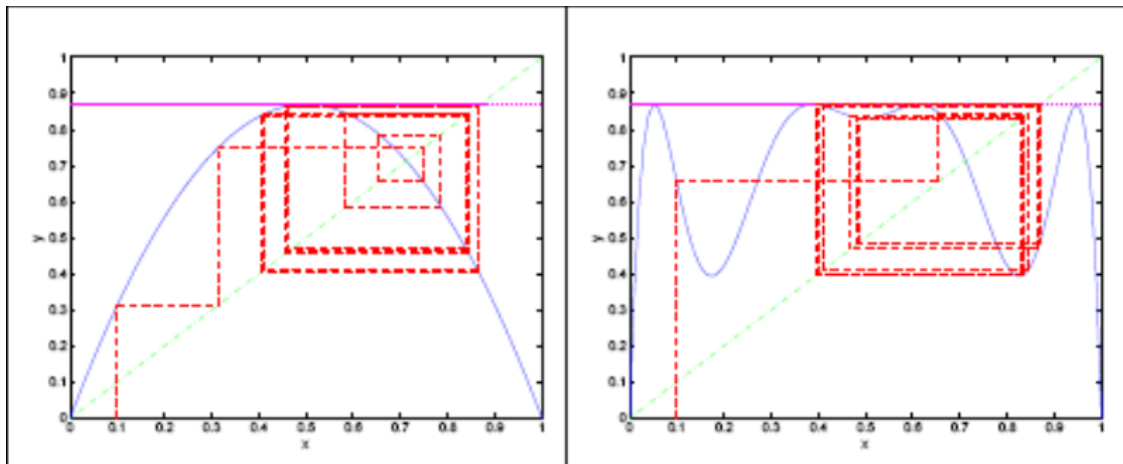


FIGURE 1.8 – Orbite des itérations de la fonction f pour $\lambda = 0.87$ (à gauche) , orbite des itérations de la fonction h pour $\lambda = 0.87$ (à droite)

On assiste ainsi à toute une série de doublement de période, pour des valeurs du paramètre de plus en plus rapprochées, ce qu'on appelle une “cascade sous-harmonique”. Cette cascade se produit jusqu'à atteindre une valeur limite du paramètre de bifurcation $\lambda = \lambda_c \simeq 0.892489418$. au-delà de laquelle le comportement devient chaotique.

La longueur des plages de paramètres correspondant à un comportement donné ($\lambda_j - \lambda_{j-1}$) diminue au fur et à mesure des bifurcations de la manière suivante :

$$\lim_{n \rightarrow +\infty} \frac{\lambda_n - \lambda_{n-1}}{\lambda_{n+1} - \lambda_n} = C_F = 4.66920166609102\dots$$

où C_F est la constante de Feigenbaum.

Diagramme de bifurcation

Il est intéressant de visualiser ces différents comportements sur un diagramme de bifurcation. On trace tous les points obtenus en fonction de la valeur du paramètre de bifurcation λ correspondante. Le nombre de points différents représentés sur une même droite verticale donne donc ainsi le facteur par lequel est multipliée la période initiale. On y retrouve bien les valeurs des seuils de bifurcation λ_i .

La constante de Feigenbaum

La découverte de cette constante est due entièrement au mathématicien Mitchell J. Feigenbaum, qui l'a calculée à l'aide d'une simple calculette vers 1975. Il avait observé que les bifurcations de l'application quadratique convergeaient vers leur limite d'une façon régulière. Il a donc été conduit à étudier la suite $\mu_k = \lambda_{k+1} - \lambda_k$ et il a remarqué qu'elle se conduisait presque comme une suite géométrique car le rapport $\delta_k = \frac{\mu_k}{\mu_{k+1}}$ n'est pas constant, mais δ_k tend vers une limite C_F . Avec les valeurs de λ_k que l'on a, cela donne : $\delta_1 = 4.7514, \delta_2 = 4.6562, \delta_3 = 4.6682, \delta_4 = 4.6687\dots$

1.4.2 L'attracteur de Hénon

L'attracteur de Hénon est défini par :

$$\begin{cases} X_{n+1} = Y_n + 1 - aX_n^2 \\ Y_{n+1} = bX_n \end{cases} \quad (1.4)$$

a et b étant deux paramètres réels, où la valeur de la constante a contrôle la non-linéarité de l'itération, et celle de b traduit le rôle de la dissipation.

Les valeurs habituellement utilisées pour a, b sont :

$a = 1.4$ et $b = 0.3$, ce sont les valeurs adoptées dans la suite.

Partant d'un point du plan de coordonnées (X_0, Y_0) on peut calculer les coordonnées (X_1, Y_1) du point suivant, et ainsi de suite.

L'application de Hénon est inversible, son inverse est :

$$f^{-1}(x, y) = (b^{-1}y, x - 1 + ab^{-2}y^2). \quad (1.5)$$

La matrice jacobienne ici a pour expression :

$$J = \begin{pmatrix} -2ax & 1 \\ b & 0 \end{pmatrix} \quad (1.6)$$

Le déterminant de la matrice jacobienne est égale à $|J| = -b$

Ainsi, les aires sont multipliées à chaque itération par un facteur de $|b|$, il y a donc contraction des aires si $|b| < 1$.

Considérons l'application de Hénon

$$f(x, y) = (1 + y - ax, bx) \quad (1.7)$$

avec $0 < b < 1$.

Cette application a deux points fixes

$$\begin{aligned} x_1 &= \frac{b - 1 + \sqrt{(1 - b)^2 + 4a}}{2a}, & y_1 &= bx_1 \\ x_2 &= \frac{b - 1 - \sqrt{(1 - b)^2 + 4a}}{2a}, & y_2 &= bx_2 \end{aligned} \quad (1.8)$$

trouvés en faisant la définition des points d'équilibres : $f(x, y) = (1 + y - ax, bx) = (x, y)$.

Le déterminant $(b - 1)^2 + 4a$ est négative si $a < a_0 = -\frac{(b-1)^2}{4} = -0.1225$, dans ce cas il n'y a pas des points fixes.

Stabilité des points fixes

La matrice jacobienne J a pour valeurs propres, les valeurs données par :

$$\lambda_{1,2} = -ax \pm \sqrt{a^2x^2 + b} \quad (1.9)$$

Si l'on calcule les valeurs absolues des valeurs propres, on constate que la plus petite des valeurs propres est toujours inférieure à 1, tandis que la plus grand est inférieure, égale ou supérieur à 1 suivant que $|x|$ inférieure, égale ou supérieur à $(1 - b)/2a$, on en déduit que le point fixe (x_2, y_2) est un point selle.

L'autre point fixe est stable si $a < 3(1 - b)^2/4 = 0.3675$.

Si $a = 3(1 - b)^2/4$, on a $\lambda_1(x_1, y_1) = b$ et $\lambda_2(x_2, y_2) = -1$.

Diagramme de bifurcation

La construction de diagramme de bifurcation est faite en faisant varier le paramètre a de 0 à 2 avec un pas de 0.0005, b est égale à 0.3.

Le diagramme obtenu est représenté par la figure 1.9

Ce diagramme est de type de bifurcation de doublement de période, l'attracteur de Hénon contient deux points fixes.

La partie stable se situe dans l'intervalle $[0, 0.3675]$.

Un 2-cycle stable commence à $a = 0.3675$ suivi d'un 4-cycle stable à $a = 0.9$ et ainsi de suite.

La période continue de doubler jusqu'à une valeur déterminée où le trajectoire commence à prendre une forme particulière.

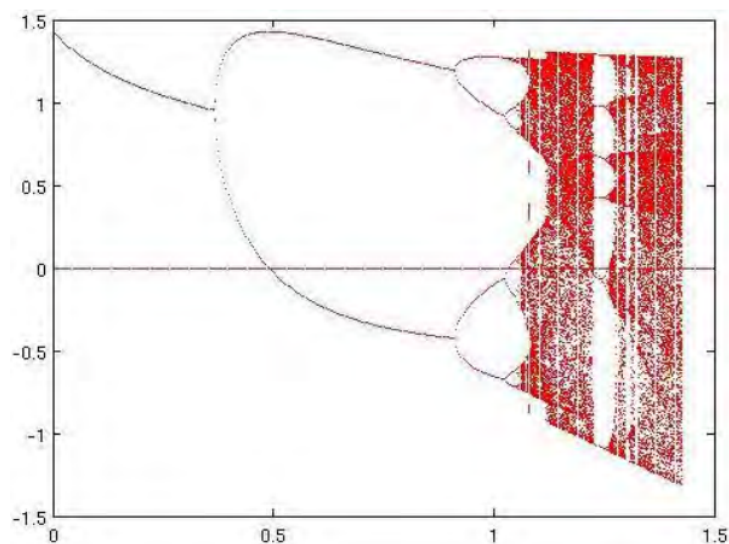


FIGURE 1.9 – Diagramme de bifurcation de Hénon

Pour $a = 1.4$, on ne distingue plus les cycle, le système est chaotique.

L'attracteur de Hénon pour $a = 1.4$, $b = 0.3$ est représenté dans la figure 1.10

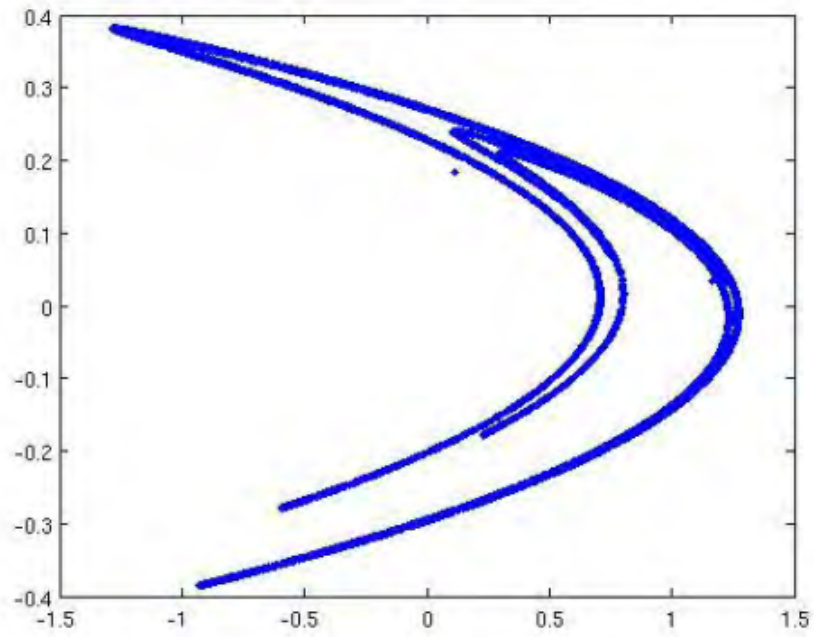


FIGURE 1.10 – Attracteur de Hénon pour $a = 1.4$, $b = 0.3$.

Pour déterminer la structure fractale de l'attracteur de Hénon, quelques graphes aident à préciser cette structure, représenté dans la figure 1.11

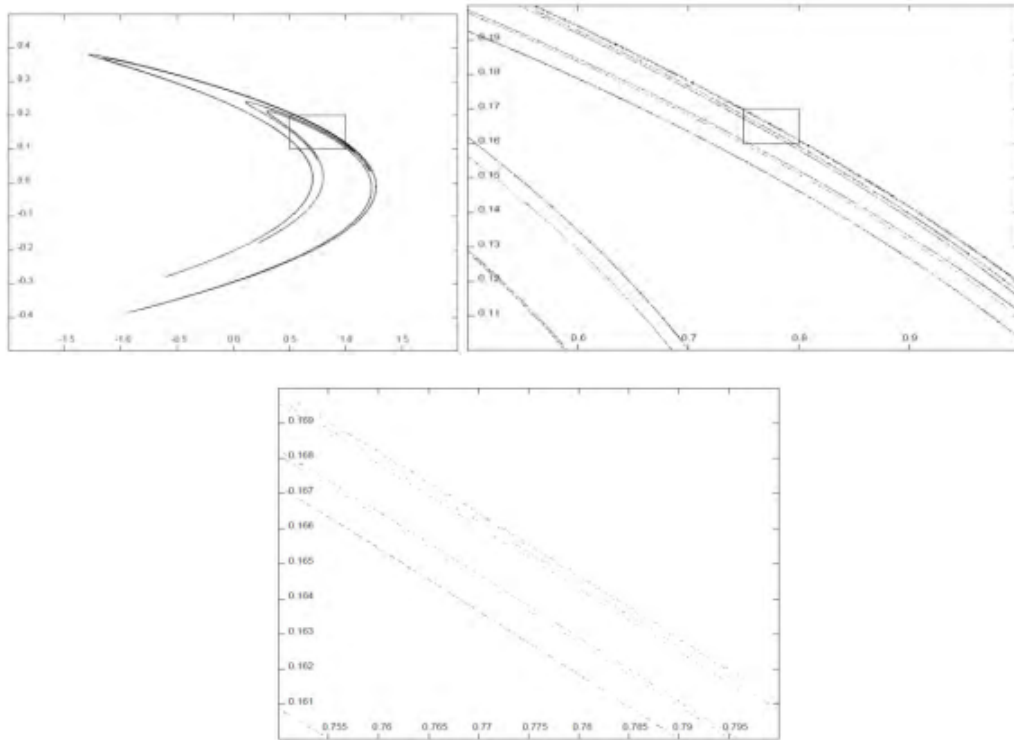


FIGURE 1.11 – Illustration de la permanence de la structure de l’attracteur de Hénon à différentes échelles.

Dans l’encadré de la figure 1.11 on note la présence de trois lignes parallèles, dont l’une paraît d’ailleurs plus épaisse. Un premier agrandissement relève que la ligne supérieure est, en réalité formé de trois lignes parallèle, de même un deuxième agrandissement fait apparaître que la ligne supérieure des trois lignes observées dans le premier agrandissement est elle aussi triple, et ainsi de suite. Autrement dit, la structure de l’attracteur se répète identiquement à elle-même aux échelles d’observation successives. Cette structure dont la permanence à différent échelle est caractéristique d’un objet fractale. Par ailleurs, la dimension de Hausdorff de l’attracteur de Hénon objet intermédiaire entre une ligne et une surface, est $D = 1.26$. On peut aussi calculer la dimension de l’attracteur de Hénon par la dimension de Lyapunov.

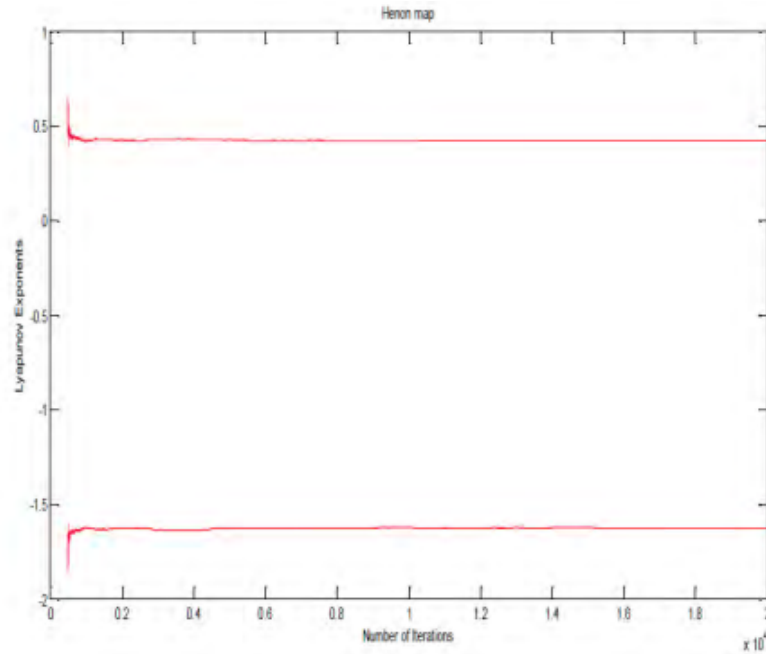


FIGURE 1.12 – Les exposants de Lyapunov en fonction du temps pour l’application de Hénon.

On a pour $a = 1.4$, $b = 0.3$, l’application de Hénon a deux exposants de Lyapunov $\lambda_1 = 0.42205$, $\lambda_2 = -1.626$ la dimension de Lyapunov par définition est égale à $D_L = 1.2596$.

1.4.3 L’attracteur de Lozi

Dans le but de simplifier l’attracteur de Hénon, René Lozi [8], propose l’application suivante :

$$\begin{aligned}
 L : \quad \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 & (1.10) \\
 (x, y) &\longmapsto (y + 1 - a|x|, bx)
 \end{aligned}$$

Propriétés de l’application de Lozi

1. La seule différence entre l’application de Hénon et de Lozi est que le terme non-linéaire x^2 de l’application de Hénon est remplacé par $|x|$ dans l’application de Lozi. Cette modification de l’application de Hénon est linéaire pour $x > 0$, $x < 0$.
2. L’application de Lozi n’est pas différentiable.

3. Si $a = 0$, L'application de Lozi est une application linéaire, donc on pose toujours $a \neq 0$.
4. L'application de Lozi est inversible, son inverse est :

$$f^{-1}(x, y) = (b^{-1}y, x - 1 + a/b|y|). \quad (1.11)$$

5. Le déterminant de la matrice Jacobienne est égale à $|J| = -b$, alors il y a contraction des aires pour $|b| < 1$.

Cette application possède deux points fixes hyperbolique définit par :

$$\begin{cases} P_1 = (\bar{x}_1, \bar{y}_1) = \left(\frac{1}{1+a-b}, \frac{b}{1+a-b}\right), & \text{si } b < a + 1 \\ P_2 = (\bar{x}_2, \bar{y}_2) = \left(\frac{1}{1-a-b}, \frac{b}{1-a-b}\right), & \text{si } b < -a + 1 \end{cases} \quad (1.12)$$

On peut facilement déterminer la stabilité locale de ces points par l'évaluation des valeurs propres de la matrice jacobienne :

$$Df(x) = \begin{pmatrix} -a \frac{\partial |x|}{\partial x} & 1 \\ b & 0 \end{pmatrix}. \quad (1.13)$$

Stabilité des points fixes

L'équation caractéristique de la matrice jacobienne est :

$$\begin{aligned} \lambda^2 + a\lambda - b & \text{ pour } P_1 \\ \lambda^2 - a\lambda - b & \text{ pour } P_2 \end{aligned} \quad (1.14)$$

Stabilité de P_1

Pour $b > \frac{-a^2}{4}$, les valeurs propres sont des réelles

Pour $b < \frac{-a^2}{4}$, les valeurs propres sont des complexes.

Elle sont de module inférieur à un si :

$b > -1$, $b < a + 1$ et $b < 1 - a$. Et le point fixe P_1 est stable.

Les valeurs propres sont de module supérieur à un si :

$b < -1$, $b < a + 1$ et $b < 1 - a$. Et le point fixe P_1 est instable.

Les valeurs propres λ_1 et λ_2 sont $|\lambda_1| < 1$ et $|\lambda_2| > 1$ si :

$b > a + 1$ et $b > 1 - a$. Et le point fixe P_1 est une point selle ou col.

Stabilité de P_2

L'équation caractéristique pour P_2 est défini par $\lambda^2 - a\lambda - b$ et donc :

$$\Delta = a^2 + 4b.$$

L'existence de P_2 est pour $b > -a + 1$, alors $\Delta = a^2 + 4b > 0$ et les valeurs propres sont toujours des réelles.

Elle sont de modules supérieurs à un si

$b > -a + 1$, $b > a + 1$. Et le point fixe P_2 est instable.

Les valeurs propres λ_1 et λ_2 sont $|\lambda_1| < 1$ et $|\lambda_2| > 1$ si :

$b > a + 1$ et $b > 1 - a$ et le point fixe P_2 est un point selle ou col. Plus de détails sont cités dans [30, 14, 31].

Le résumé de la stabilité des points fixes pour l'application de Lozi est donné dans le tableau (1.1)

Stabilité des poins fixes	Stabilité de P_1	Stabilité de P_2
$b < a + 1$ $b < -a + 1$ $b > -1$	Stable	N'existe pas
$b < a + 1$ $b < -a + 1$ $b < -1$	Instable	N'existe pas
$b > a + 1$ $b > -a + 1$	Selle	Selle
$b > -a + 1$ $b > a + 1$	N'existe pas	Instable

TABLE 1.1 – Variation de la stabilité de P_1 , P_2 par rapport à la variation du paramètres a et b

Misiurewicz dans [8] prouve l'existence d'un attracteur étrange de l'application de Lozi.

Pour les paramètres $a = 1.7$ et $b = 0.6$ la suite de points itérés par l'application de Lozi converge vers un attracteur étrange représenté par la figure 1.13

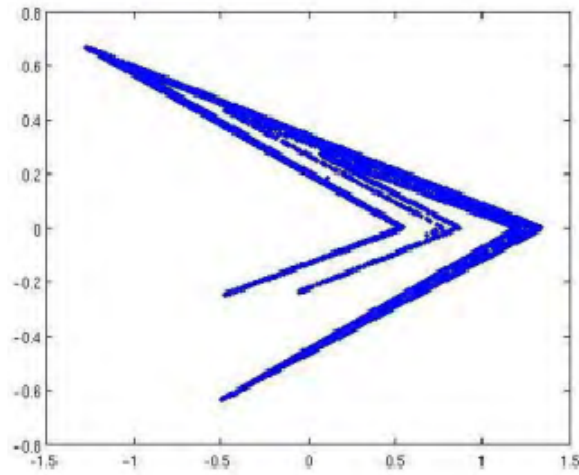


FIGURE 1.13 – Attracteur de Lozi pour $a = 1.7$ $b = 0.6$

Structure fractal

Pour déterminer la structure de l'attracteur de Lozi, on fait un agrandissement d'une région de l'attracteur et on observe que la structure se répète, voir la figure 1.14

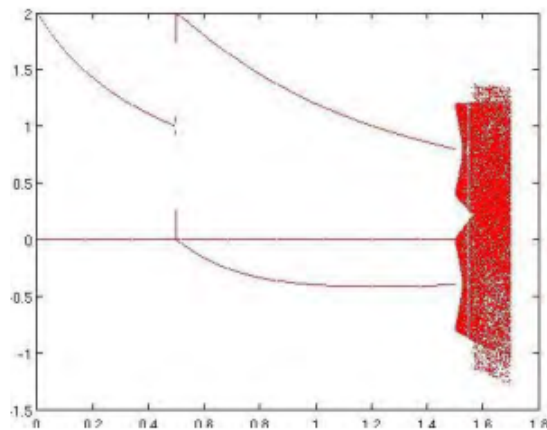


FIGURE 1.14 – Diagramme de bifurcation de Lozi

Contrairement au cas de l'application de Hénon, la route vers le chaos par l'application de Lozi n'est pas une bifurcation de doublement de période. Puisque la dérivée de l'application de Lozi n'est pas continue.

Exposants de Lyapunov pour l'application de Lozi

Pour $a = 1.7$, $b = 0.6$, l'application de Lozi a deux exposants de Lyapunov $\lambda_1 = 0.69314$ qui est positif, $\lambda_2 = -1.204$. On vérifie bien cependant que la somme de ces exposants est strictement négative.

La figure 1.15 présente la variation des exposants de Lyapunov en fonction du temps pour l'application de Lozi.

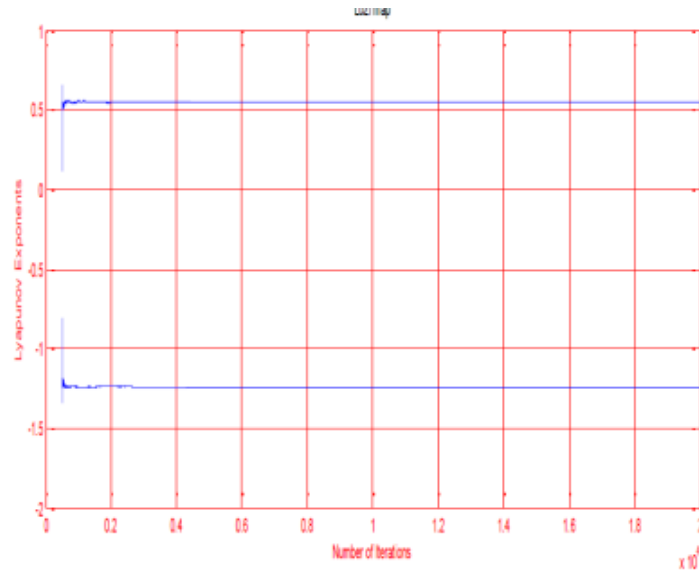


FIGURE 1.15 – Les exposants de Lyapunov en fonction du temps pour l'application de Lozi

Dimension de l'attracteur de Lozi

On a pour $a = 1.7$, $b = 0.6$, l'application de Lozi a deux exposants de Lyapunov $\lambda_1 = 0.69314$, $\lambda_2 = -1.204$, alors la dimension de Lyapunov est égale à $D_L = 1.5757$.

Chapitre 2

Généralités sur l'optimisation

2.1 Introduction

L'optimisation est une branche des mathématiques et de l'informatique en tant que disciplines, cherchant à modéliser, à analyser et à résoudre analytiquement ou numériquement les problèmes qui consistent à déterminer quelles sont la ou les solution(s) satisfaisant un objectif quantitatif tout en respectant d'éventuelles contraintes.

Dans la vie courante, nous sommes fréquemment confrontés à des problèmes d'optimisation plus ou moins complexes. Cela peut commencer au moment où l'on tente de ranger son bureau, de placer son mobilier, et aller jusqu'à un processus industriel, par exemple pour la planification des différentes tâches. Ces problèmes peuvent être exprimés sous la forme générale d'un problème d'optimisation.

D'un point de vue mathématique, l'optimisation consiste à rechercher le minimum ou le maximum d'une fonction avec ou sans contraintes. L'optimisation possède ses racines au 18ième siècle dans les travaux de :

- Taylor, Newton , Lagrange, qui ont élaboré les bases des développements limités.
- Cauchy ([1847]) fut le premier à mettre en œuvre une méthode d'optimisation, méthode du pas de descente, pour la résolution de problèmes sans contrainte.

Il n'existe pas de méthode unique permettant de résoudre efficacement tous les problèmes d'optimisation. C'est pourquoi de nombreuses méthodes d'optimisation ont été développées pour résoudre différents types de problèmes d'optimisation.

Un problème d'optimisation peut être énoncé comme suit :

Trouver $X = (x_1, x_2, \dots, x_n)$ qui minimise $f(X)$,

sous réserve des contraintes

$$g_j(X) \leq 0, \quad j = 1, 2, \dots, m$$

$$l_j(X) = 0, \quad j = 1, 2, \dots, p$$

où X est un vecteur à n dimension appelée vecteur de conception, $f(X)$ est appelée fonction objective et $g_j(X)$ et $l_j(X)$ sont respectivement contraintes d'inégalité et d'égalité. Certains problèmes d'optimisation n'imposent aucune contrainte et peuvent être définis comme suit :

Trouver $X = (x_1, x_2, \dots, x_n)$ qui minimise $f(X)$

Ce genre de problèmes sont appelés problèmes d'optimisation sans contrainte.

2.2 Types d'optimum

On veut résoudre $\min_{(x_1, x_2, \dots, x_n) \in C} f(x_1, x_2, \dots, x_n)$ ou $\max_{(x_1, x_2, \dots, x_n) \in C} f(x_1, x_2, \dots, x_n)$ i.e., on cherche v valeur optimale et x_0 tel que $f(x_0) = v$.

Définition 16

Soit $x_0 \in C$.

1. On dit que x_0 est un minimum local si

$$\exists V \in \vartheta(x_0) \text{ tel que } \forall x \in V, f(x) \geq f(x_0)$$

2. On dit que x_0 est un minimum local strict si

$$\exists V \in \vartheta(x_0) \text{ tel que } \forall x \in V, f(x) > f(x_0)$$

3. On dit que x_0 est un minimum global si

$$\forall x \in C, f(x) \geq f(x_0)$$

4. On dit que x_0 est un minimum global strict si

$$\forall x \in C, f(x) > f(x_0)$$

On définit de la même façon un maximum local, maximum local strict, maximum global, maximum global strict, en renversant les inégalités.

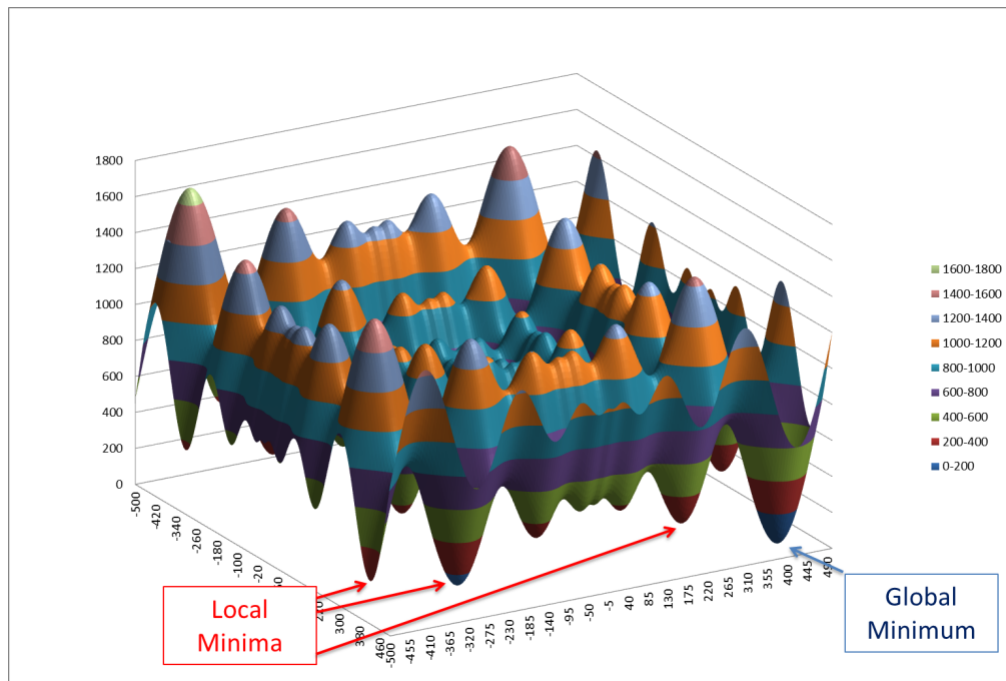


FIGURE 2.1 – Exemple d'un problème d'optimisation : fonction de Schwefel

2.3 Classification des problèmes et des approches

2.3.1 Problèmes avec ou sans contraintes

Soit le problème d'optimisation :

$$\min_{(x_1, x_2, \dots, x_n) \in C} f(x_1, x_2, \dots, x_n) \quad (P)$$

Une classification naturelle des problèmes d'optimisation globale serait de distinguer les problèmes sans contraintes (Psc) et ceux avec contraintes (Pc). Les problèmes sans contraintes, ce sont ceux dont l'ensemble faisable est $C = \mathbb{R}^n$. Cependant, sont aussi considérés essentiellement sans contraintes les problèmes dont les contraintes spécifient seulement la région d'intérêt pour l'exploration globale mais ne sont pas des contraintes importantes, comme par exemple le cas où C est un pavé de \mathbb{R}^n , i.e, $C = \prod_{i=1}^n [a_i, b_i]$ avec $(a_1, a_2, \dots, a_n) \in \mathbb{R}^n$ et $(b_1, b_2, \dots, b_n) \in \mathbb{R}^n$.

Si $C \neq \mathbb{R}^n$, le problème p est dit problème de minimisation avec contraintes.

Dans le cas $C = \mathbb{R}^n$, et pour que le problème (P) admette un minimum global, il faut imposer une hypothèse supplémentaire sur f dite condition de croissance à l'infini, connue sous le nom de la coercivité, c'est-à-dire $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$ où $\|x\|$ désigne la norme euclidienne de x dans \mathbb{R}^n . Notons que cette hypothèse implique l'existence d'un ensemble robuste compact $C' \subset \mathbb{R}^n$, tel que C' contient tous les minimiseurs globaux de f , et la valeur de $f(x)$ quand x est sur la frontière de C' est plus grande que la valeur de $f(x)$ pour tout point x à l'intérieur de C' . Rappelons qu'un sous ensemble fermé C' est dit robuste s'il est la fermeture d'un ensemble ouvert $U \neq \emptyset$ de \mathbb{R}^n , c'est-à-dire $C' = \overline{U}$.

Numériquement, il est plus aisé de résoudre les problèmes sans contraintes et cela, quelle que soit la nature de la méthode. En effet, le domaine (des variables) de recherche n'étant pas restreint, on a moins de soucis à se faire que dans les cas des problèmes avec contraintes où il faut tout mettre en œuvre pour s'assurer à chaque itération que l'on reste dans le domaine faisable ou sinon s'imposer un moyen d'y revenir à la limite du domaine.

2.3.2 Approches déterministes et approches probabilistes

Différentes classifications des méthodes d'optimisation globale sont trouvées dans la littérature. Une classification, basée sur [22], est présentée dans cette section. Elle n'a pas pour objectif d'être complète, ni de montrer tous les détails des méthodes, mais plutôt de mettre en évidence les caractéristiques de chaque classe.

-Approches déterministes

Dans ce type de méthodes, l'aléatoire n'intervient pas, c'est-à-dire que pour résoudre un problème, l'algorithme se comportera toujours de la même façon et donnera toujours

la même réponse. Ces algorithmes peuvent se classer en fonction du type de problèmes pouvant être résolus : les programmes linéaires, les problèmes convexes, quadratiques, polynomiaux ou plus généraux.

Ces techniques ont généralement l'avantage de ne pas nécessiter de point de départ. Mais avant tout elles fournissent une réponse déterminante sur la qualité des solutions trouvées : l'optimum est-il local ou global? Quel est le degré de certitude? etc. Cette précision a une importance significative, car il est souvent beaucoup moins coûteux de trouver une solution que de prouver qu'il s'agit bien de l'optimum global.

Pour ces méthodes, l'exploration de l'espace des solutions se fait grâce à des procédures de recherche qui sont élaborées à partir de la constante de Lipschitz ou celle de Hölder, des dérivées ou d'autres informations locales et globales concernant la fonction objectif. La théorie mathématique de l'optimisation globale est assez développée et possède de nombreuses applications importantes. Récemment, une série de travaux de Floudas, Horst et Tuy [22], Strongin et Sergeyev [21], Pardalos [24] et Thoai [25], Dans le cas où la fonction objectif est donnée sous forme de "boîte noire" le problème d'optimisation est particulièrement difficile. Les modèles déterministes ne traitent pas adéquatement les informations disponibles sur la fonction objectif.

-Approches stochastiques

Les approches stochastiques peuvent souvent faire face à ce genre de problèmes plus facilement et plus efficacement que les algorithmes déterministes. Les algorithmes stochastiques explorent l'espace des solutions grâce en partie à des procédures de transitions aléatoires. Ainsi, plusieurs exécutions successives de ces algorithmes, pourront conduire à des résultats différents (pour un même point initial). L'avantage des méthodes stochastiques est leur simplicité et leur pertinence pour les problèmes où les évaluations de la fonction objectif sont corrompues par un bruit aléatoire, ainsi que leur robustesse à l'égard de la croissance de la dimension. L'inconvénient majeur de ces méthodes est qu'elles peuvent diverger, et passer plusieurs fois à côté de la solution. L'obtention de l'optimum global n'est pas garantie, il est seulement repéré avec une probabilité proche de 1.

De nombreux algorithmes, où l'aléatoire et l'argument statistique sont impliqués ont

été proposés heuristiquement. Ces algorithmes sont basés sur des analogies avec les processus naturels. Les exemples les plus connus de tels algorithmes sont l'optimisation évolutive, le recuit simulé, l'algorithme génétique et la recherche tabou. Les algorithmes heuristiques d'optimisation globale sont très populaires dans les applications. Le recuit simulé, comme étant méthode importante, a été intensivement étudié par plusieurs chercheurs.

2.4 Méthodes d'optimisation globale

2.4.1 Méthode du gradient

La méthode (ou algorithme) du Gradient fait partie d'une classe plus grande de méthodes numériques appelées **méthodes de descente** [28], l'idée de ces méthodes est : On veut minimiser une fonction f . Pour cela on se donne un point de départ arbitraire x_0 . Pour construire l'itéré suivant x_1 il faut penser qu'on veut se rapprocher du minimum de f , on veut donc que $f(x_1) < f(x_0)$. On cherche alors x_1 sous la forme $x_1 = x_0 + \rho_1 d_1$ où d_1 est un vecteur non nul de \mathbb{R}^n et ρ_1 un réel strictement positif. En pratique donc, on cherche d_1 et ρ_1 pour que $f(x_0 + \rho_1 d_1) < f(x_0)$. On ne peut pas toujours trouver d_1 . Quand d_1 existe on dit que c'est une direction de descente et ρ_1 est le pas de descente. La direction et le pas de descente peuvent être fixes ou changer à chaque itération. Le schéma général d'une méthode de descente est le suivant :

$$\begin{cases} x_0 \in \mathbb{R}^n \text{ donné} \\ x_{k+1} = x_k + \rho_k d_k, d_k \in \mathbb{R}^n - \{0\}, \rho_k \in \mathbb{R}^{+*}, \end{cases} \quad (2.1)$$

où ρ_k et d_k sont choisis de telle sorte que $f(x_k + \rho_k d_k) \leq f(x_k)$.

Une idée naturelle pour trouver une direction de descente est de faire un développement de Taylor (formel) à l'ordre 2 de la fonction f entre deux itérés x_k et $x_{k+1} = x_k + \rho_k d_k$:

$$f(x_k + \rho_k d_k) = f(x_k) + \rho_k (\nabla f(x_k), d_k) + o(\rho_k d_k).$$

Comme on veut $f(x_k + \rho_k d_k) < f(x_k)$, on peut choisir en première approximation $d_k = -\nabla f(x_k)$. La méthode ainsi obtenue s'appelle l'algorithme du Gradient. Le pas ρ_k

est choisi constant ou variable.

- **Méthode du gradient à pas constant**

On utilise le plus souvent la méthode du gradient à pas constant ($\rho_k \equiv \rho$ constant). Toutefois, on peut faire varier le pas à chaque itération : on obtient alors la méthode du gradient à pas variable.

- **Méthode du gradient à pas optimal**

La méthode du gradient à pas optimal est un algorithme de gradient dans lequel la taille de pas ρ_k est choisie pour atteindre le maximum de diminution de la fonction objectif à chaque pas individuel. Plus précisément, un ρ_k est choisi pour minimiser $f(x_k - \rho \nabla f(x_k))$. En autre mots dit :

$$\rho_k = \arg \min_{\alpha \geq 0} f(x_k - \alpha \nabla f(x_k))$$

Pour plus de détail voir [38, 52].

2.4.2 Méthode du gradient conjugué

Les méthodes du gradient conjugué sont utilisées pour résoudre les problèmes d'optimisation non linéaires sans contraintes spécialement les problèmes de grandes tailles. On l'utilise aussi pour résoudre les grands systèmes linéaires.

Elles reposent sur le concept des directions conjuguées parce que les gradients successifs sont orthogonaux entre eux et aux directions précédentes. L'idée initiale était de trouver une suite de directions de descente permettant de résoudre le problème :

$$\min\{f(x) : x \in \mathbb{R}^n\},$$

où f est régulière (continûment différentiable)

Le principe général d'une méthode à directions conjuguées

Définition 17 Soit A une matrice symétrique $n \times n$, définie positive. On dit que deux vecteurs x et y de \mathbb{R}^n sont A -conjugués (ou conjugués par rapport à A) s'ils vérifient :

$$x^T A y = 0$$

Description de la méthode

Soit $\{d_0, d_1, \dots, d_n\}$ une famille de vecteurs A-conjugués. On appelle alors méthode de directions conjuguées toute méthode itérative appliquée à une fonction quadratique strictement convexe de n variables : $q(x) = \frac{1}{2}x^T A x + b^T x + c$, avec $x \in \mathbb{R}^n$ et $A \in M_{n \times n}$ est symétrique et définie positive, $b \in \mathbb{R}^n$ et $c \in \mathbb{R}$, conduisant à l'optimum en n étapes au plus. Cette méthode est de la forme suivante :

$$\begin{cases} x_0 \text{ donné ,} \\ x_{k+1} = x_k + \alpha_k d_k. \end{cases}$$

où α_k est optimal et d_1, d_2, \dots, d_n possédant la propriété d'être mutuellement conjuguées par rapport à la fonction quadratique. Si l'on note $g_k = \nabla q(x_k)$, la méthode se construit comme suit :

Calcul de α_k

Comme α_k minimise q dans la direction d_k , on a, $\forall k$:

$$q'(\alpha_k) = d_k^T \nabla q(x_{k+1}) = 0,$$

$$d_k^T \nabla q(x_{k+1}) = d_k^T (A x_{k+1} + b) = 0.$$

Soit :

$$d_k^T A (x_k + \alpha_k d_k) + d_k^T b = 0,$$

d'où l'on tire :

$$\alpha_k = \frac{-d_k^T (A x_k + b)}{d_k^T A d_k}.$$

Comment construire les directions A-conjuguées ?

Des directions A-conjuguées d_0, \dots, d_n peuvent être générées à partir d'un ensemble de vecteurs linéairement indépendants ζ_0, \dots, ζ_k en utilisant la procédure dite de Gram-Schmidt, de telle sorte que pour tout i entre 0 et k , le sous- espace généré par d_0, \dots, d_i soit égale au sous-espace généré par ζ_0, \dots, ζ_i . Alors d_{i+1} est construite comme suit :

$$d_{i+1} = \zeta_{i+1} + \sum_{m=0}^i \varphi_{(i+1)m} d_m.$$

Nous pouvons noter que si d_{i+1} est construite d'une telle manière, elle est effectivement linéairement indépendante avec d_0, \dots, d_i . En effet, le sous-espace généré par les directions d_0, \dots, d_i est le même que le sous-espace généré par les directions ζ_0, \dots, ζ_i , et ζ_{i+1} est linéairement indépendant de ζ_0, \dots, ζ_i . ζ_{i+1} ne fait donc pas partie du sous-espace généré par les combinaisons linéaires de la forme $\sum_{m=0}^i \varphi_{(i+1)m} d_m$, de sorte que d_{i+1} n'en fait pas partie non plus et est donc linéairement indépendante des d_0, \dots, d_i . Les coefficients $\varphi_{(i+1)m}$, eux sont choisis de manière à assurer la A-conjugaison des d_0, \dots, d_{i+1} .

2.4.3 Méthode de Newton

La méthode de Newton n'est pas a proprement parlé une méthode d'optimisation, c'est en réalité une méthode utilisée pour résoudre des équations non linéaires de la forme $F(x) = 0$ où F est une fonction de \mathbb{R}^n dans \mathbb{R}^n [17, 18]

Description de la méthode

Considérons le problème d'optimisation sans contraintes (P)

$$\left\{ \begin{array}{l} (p) : \min_{x \in \mathbb{R}^n} f(x) \\ f(x) \simeq f(x_k) + (x - x_k)^t \nabla f(x_k) + \frac{1}{2}(x - x_k)^t F(x_k)(x - x_k) = q(x), \end{array} \right.$$

$F(x_k)$ est la matrice hessienne. Soit x_{k+1} l'optimum de q , alors il vérifie $\nabla q(x_{k+1}) = 0$, d'où :

$$0 = \nabla q(x_{k+1}) = \nabla f(x_k) + F(x_k)(x - x_k).$$

Si la matrice $F(x_k)$ est définie positive, alors :

$$x_{k+1} = x_k - F(x_k)^{-1} \nabla f(x_k).$$

2.4.4 Méthode de quasi newton

Une méthode de quasi Newton est une méthode de type :

$$\left\{ \begin{array}{l} x_{k+1} = x_k + \lambda_k d_k, \\ d_k = -B_k g_k, \end{array} \right. \quad (2.2)$$

ou bien

$$\begin{cases} x_{k+1} = x_k + \lambda_k d_k, \\ d_k = -S_k^{-1} g_k, \end{cases} \quad (2.3)$$

où B_k (respectivement S_k) est une matrice destinées à approcher l'inverse du Hessien de f (respectivement le Hessien) de f en x_k . Le problème posé est : quelle stratégie à adopter pour faire cette approximation ? On peut par exemple poser $B_1 = I$, mais comment ensuite mettre à jour l'approximation B_k au cours des itérations ?

L'idée est la suivante :

Prenons $f \in C^2(\mathbb{R}^n)$, et faisons un développement de Taylor de $\nabla f(x)$ au voisinage de x_k ,

$$\begin{aligned} \nabla f(x) &= \nabla f(x_k) + H(x_k)(x - x_k) + o(\|x - x_k\|) \\ &\simeq \nabla f(x_k) + H(x_k)(x - x_k), \end{aligned} \quad (2.4)$$

ce qui implique

$$[H(x_k)]^{-1} [\nabla f(x) - \nabla f(x_k)] \simeq x - x_k.$$

Les approximations sont exactes si f est quadratique. En particulier avec $x = x_{k+1}$ et si B_k était une bonne approximation de $[H(x_k)]^{-1}$ alors

$$B_k [g(x_{k+1}) - g(x_k)] \simeq x_{k+1} - x_k.$$

On peut imposer que B_{k+1} satisfait cette équation exactement d'où

$$B_{k+1} [g(x_{k+1}) - g(x_k)] = x_{k+1} - x_k.$$

Formule de mise à jour de l'approximation du Hessien

Le principe de la mise à jour consiste à une itération donnée de l'algorithme

$$\begin{cases} x_{k+1} = x_k + \lambda_k d_k, \\ d_k = -B_k g_k, \end{cases}$$

à appliquer une formule de type

$$B_{k+1} = B_k + \Delta_k, \quad (2.5)$$

avec Δ_k symétrique, assurant la relation de quasi Newton. ainsi que B_{k+1} définie positive, sous l'hypothèse que B_k est définie positive. La formule (2.5) permet d'utiliser les nouvelles informations obtenues lors de l'étape k de l'algorithme, c'est-à-dire essentiellement le gradient $g_{k+1} = \nabla f(x_{k+1})$ au point x_{k+1} , obtenu par une recherche linéaire (exacte où approchée) dans la direction d_k . Il existe différentes formules de type (2.2). Suivant que Δ_k est de rang un ou deux, on perlera de correction de rang un ou de rang deux.

Méthode de correction de rang un

Etend donné que $[H(x_k)]^{-1}$ est symétrique, la formule de mise à jour de l'approximation du Hessien B_k est la suivante :

$$B_{k+1} = B_k + \alpha_k u_k u_k^T, u_k \in \mathbb{R}^n,$$

donc la condition de quasi Newton s'écrit comme suit :

$$s_k = (B_k + \alpha_k u_k u_k^T) y_k,$$

ou encore

$$s_k - B_k y_k = \alpha_k u_k u_k^T y_k.$$

D'où l'on déduit que u_k est proportionnel à $s_k - B_k y_k$, avec un facteur qui peut être pris en compte dans α_k . Un choix évident pour vérifier cette dernière équation est de prendre $u_k = s_k - B_k y_k$ et α_k tel que $\alpha_k (u_k^T y_k) = 1$, on obtient :

$$B_{k+1} = B_k + \frac{(s_k - B_k y_k) (s_k - B_k y_k)^T}{(s_k - B_k y_k)^T y_k}.$$

Méthode de Davidon Fletcher Powell (DFP)

Cette méthode a été proposée par Davidon en 1959 et développée plus tard en 1963 par Fletcher [38, 52]. La formule de mise à jour de DFP est une formule de correction de rang deux. De façon plus précise construisons B_{k+1} en fonction de B_k de la forme :

$$B_{k+1} = B_k + A_k + \Delta_k,$$

avec Δ_k et A_k deux matrices de rang un tel que

$$A_k = a_k u_k u_k^T \cdot \Delta_k = b_k v_k v_k^T,$$

a_k, b_k sont des constantes, u_k, v_k sont deux vecteurs de \mathbb{R}^n . B_{k+1} doit satisfaire la condition quasi Newton c'est-à-dire

$$x_{k+1} - x_k = B_{k+1} [g_{k+1} - g_k].$$

Si on pose par suite

$$s_k = x_{k+1} - x_k, y_k = g_{k+1} - g_k,$$

donc

$$\begin{aligned} s_k &= B_{k+1} y_k \\ &= (B_k + a_k u_k u_k^T + b_k v_k v_k^T) y_k, \end{aligned} \tag{2.6}$$

par suite

$$a_k u_k u_k^T y_k + b_k v_k v_k^T y_k = s_k - B_k y_k.$$

Un choix évident pour satisfaire cette équation est de prendre

$$u_k = s_k, v_k = B_k y_k, a_k u_k^T y_k = 1, b_k v_k^T y_k = -1,$$

d'où

$$B_{k+1} = B_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{B_k y_k^T y_k B_k}{y_k^T B_k y_k}.$$

Méthode de Broyden, Fletcher, Goldfarb et Shanno (BFGS)

La formule de mise à jour de Broyden, Fletcher, Goldfarb et Shanno est une formule de correction de rang deux, qui s'obtient à partir de la formule DFP en intervertissant les rôles de s_k et y_k . La formule obtenue permet de mettre à jour une approximation B_k de Hessien lui même et non de son inverse comme dans le cas de la méthode DFP. On exigera que posée dans les mêmes propriétés, à savoir B_{k+1} reste définie positive si B_k l'est et bien sur l'équation d'approximation de quasi Newton doit être vérifiée, c'est-à-dire :

$$B_{k+1}s_k = y_k.$$

On obtient donc

$$B_{k+1} = B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}.$$

2.4.5 Méthode de relaxation

On cherche à minimiser $f : \mathbb{R}^n \rightarrow \mathbb{R}$, posons $X = (x_1, \dots, x_n)$. Le principe de la méthode est le suivant :

étant donné un itéré X^k de coordonnées (x_1^k, \dots, x_n^k) , on fixe toutes les composantes sauf la première et on minimise sur la première :

$$\min f(x, x_2^k, x_3^k, \dots, x_n^k) \quad x \in \mathbb{R}.$$

On obtient ainsi la première coordonnée de l'itéré suivant X^{k+1} que l'on note x_1^{k+1} .

On recommence ensuite en fixant la première coordonnée à x_1^{k+1} et les $n - 2$ dernières comme précédemment. On minimise sur la deuxième coordonnée et ainsi de suite.

2.4.6 La méthode des initialisations multiples (multistart)

La technique Multistart est historiquement la première à être largement utilisée en optimisation globale [26]. Sa popularité est due à son importance pratique puisqu'elle utilise des méthodes locales qui étaient déjà parfaitement développées dans les années soixante et soixante dix. La méthode consiste à faire des recherches multiples des minima locaux en

commençant en différents points initiaux. Ces points sont générés d'une manière aléatoire parmi les éléments d'une grille uniforme. Pour s'assurer de l'obtention du minimum global on doit prendre le nombre de points initiaux beaucoup plus grand que le nombre des minimiseurs locaux (qui est généralement inconnu). La partie essentielle des efforts de calcul est consacrée donc à l'obtention des minima locaux d'une manière répétitive. Si les minimiseurs locaux sont éloignés l'un de l'autre, la version de base de Multistart est souvent modifiée par l'une des deux façons suivantes :

La première version consiste à appliquer une procédure de descente LS (local search) au voisinage de chaque minimiseur local. Cette approche n'est possible que si l'on est capable de choisir les voisinages qui sont des domaines d'attraction des minimiseurs locaux correspondants. Ceci est très difficile, voire impossible dans la majorité des cas.

La deuxième version est appelée méthode des points candidats. Elle consiste à faire des descentes simultanées à partir des points initiaux, en joignant les points voisins (les points voisins sont remplacés par celui qui a la plus petite valeur de la fonction objectif). Ces méthodes sont heuristiques, leur efficacité dépend des paramètres et des procédures auxiliaires utilisées dans l'algorithme [9]

2.4.7 La méthode de tunnel (tunneling)

L'algorithme tunneling est l'une des méthodes de pénalisation. La méthode de percement du tunnel (tunneling), présentée par Levy et Coll [16], se décompose en deux phases :

- Dans la première phase on utilise une technique de descente locale permettant de trouver un minimiseur local x^* de la fonction objectif f .
- Dans la deuxième phase (la phase de tunneling) on construit une fonction auxiliaire $T(x)$ appelée fonction de tunneling (appelée encore fonction pénalité ou fonction Filled). Cette fonction admet un maximum (peut être local) au point x^* possédant des dérivées premières continues (sauf, peut être en x^*) et dépend de f , de x^* et d'un nombre fini de paramètres choisis automatiquement par l'algorithme.

On cherche alors un minimiseur y^* de $T(x)$ dans l'ensemble :

$\mathbb{E}(x^*) = \{x \in D, f(x) \leq f(x^*)\} \setminus \{x^*\}$, en appliquant une minimisation locale qui com-

mence en un point x_0 proche du point record x^* .

Ensuite, on retourne à la minimisation de f en appliquant une descente qui commence cette fois-ci au point y^* , le minimiseur local obtenu est un nouveau point record. L'itération précédente peut être répétée. On arrête la recherche lorsque le même point $y^* \in E(x^*)$ est obtenu plusieurs fois par les itérations précédentes.

2.4.8 Méthodes de regroupement (clustering)

Ces méthodes s'appliquent aux fonctions ayant un nombre de minima locaux relativement petit (inférieur à 50) [26, 58, 47]. L'idée principale consiste à générer suivant la loi de probabilité uniforme, un échantillon de points relativement important et former des sous-ensembles constitués de points mutuellement proches (clusters) et appliquer une méthode de descente locale à partir d'un ou de plusieurs points de chaque groupe. Deux procédures pour créer de tels groupes à partir d'un échantillon initial sont utilisés, la première appelée méthode de réduction [3] retient uniquement une fraction γ de l'échantillon, constituée de points pour lesquels f a les meilleures valeurs. La deuxième appelée méthode de concentration [9] transforme l'échantillon en appliquant une ou plusieurs itérations de descente locale à partir de chaque point d'échantillon.

La procédure de base pour identifier les clusters des points qui résultent de l'échantillon obtenu par l'une des deux méthodes est toujours la même. Les clusters sont formés au cours des itérations, en commençant à partir d'un point foyer, qui peut être un point non clusterisé et ayant une petite valeur de f ou un minimiseur local trouvé en appliquant LS en ce point. Des points sont alors ajoutés au cluster à travers la règle de groupement. Un désavantage de ces méthodes est qu'elles sont peu performantes pour les fonctions ayant de nombreux minima. Dans ce cas, en effet, un grand nombre d'échantillons est nécessaire.

2.5 Méthodes méta-heuristiques

Une métaheuristique est un algorithme d'optimisation visant à résoudre des problèmes d'optimisation difficile (souvent issus des domaines de la recherche opérationnelle, de l'ingénierie ou de l'intelligence artificielle) pour lesquels on ne connaît pas de méthode

classique plus efficace.

Les métaheuristiques sont généralement des algorithmes stochastiques itératifs, qui progressent vers un optimum global, c'est-à-dire l'extremum global d'une fonction, par échantillonnage d'une fonction objectif. Elles se comportent comme des algorithmes de recherche, tentant d'apprendre les caractéristiques d'un problème afin d'en trouver une approximation de la meilleure solution (d'une manière proche des algorithmes d'approximation)

Il existe un grand nombre de métaheuristiques différentes, allant de la simple recherche locale à des algorithmes complexes de recherche globale. Ces méthodes utilisent cependant un haut niveau d'abstraction, leur permettant d'être adaptées à une large gamme de problèmes différents.

2.5.1 Méthodes de recherche locale

Le principe des méthodes de recherche locale (ou méthodes d'amélioration itérative) est inspirée des méthodes d'optimisation continue. Ces méthodes consistent à déterminer itérativement la solution d'une fonction continue en utilisant des outils comme les dérivées partielles ou les gradients, suivant que la fonction soit ou non dérivable.

La description de ces méthodes à partir d'une solution de départ X_0 , à engendrer une suite (finie) de solutions (X_n) , déterminées de proche en proche, c'est-à-dire itérativement (X_{i+1} étant déterminée à partir de X_i). Le choix de la solution X_{i+1} se fait dans un ensemble "localement proche" de la solution X_i et de manière à avoir une solution X_{i+1} "plus intéressante" au sens de la recherche locale. Les deux expressions mises entre guillemets dans la phrase précédente se formalisent à partir de la notion de voisinage et de recherche dans un voisinage.

La construction essentielle de ces méthodes reposent en effet sur la détermination d'un bon voisinage. Ensuite, suivant la façon de choisir une solution dans le voisinage, on obtient différentes méthodes de recherche locale : méthode tabou, descente "pure", descente stochastique, recuit simulé,...

2.5.2 Les méthodes évolutives

Le terme “algorithmes évolutifs” englobe une autre classe assez large de métaheuristiques. Ces algorithmes sont basés sur le principe du processus d’évolution naturelle. Les algorithmes évolutifs doivent leur nom à l’analogie entre leur déroulement et le mécanisme de sélection naturelle et de croisement des individus d’une population vivante sexuée [27, 1, 6, 33].

Un algorithme évolutif typique est composé de trois éléments essentiels :

1. Une population constituée de plusieurs individus représentant des solutions potentielles (configurations) du problème donné.
2. Un mécanisme d’évaluation de l’adaptation de chaque individu de la population à l’égard de son environnement extérieur.
3. Un mécanisme d’évolution composé d’opérateurs permettant d’éliminer certains individus et de produire de nouveaux individus à partir des individus sélectionnés.

Du point de vue opérationnel, un algorithme évolutif débute avec une population initiale souvent générée aléatoirement et répète ensuite un cycle d’évolution suivant les principes en 3 étapes séquentielles :

- Evaluation : mesurer l’adaptation (la qualité) de chaque individu de la population.
- sélectionner une partie des individus.
- Reproduction : produire de nouveaux individus par des recombinaisons d’individus sélectionnés.

Ce processus se termine quand la condition d’arrêt est vérifiée, par exemple, quand un nombre de cycles (générations) ou quand un nombre d’évaluations est atteint ou quand des solutions suffisamment bonnes sont trouvées. Si l’on s’imagine que le processus suit le principe d’évolution naturelle, la qualité des individus de la population doit s’améliorer au fur et à mesure du processus.

Parmi les composantes d’un algorithme évolutif, la population et la fonction d’adaptation correspondent respectivement à la notion de configuration et à la fonction d’évaluation dans la recherche locale. La notion de mécanisme d’évolution est proche de celle du mécanisme de parcours du voisinage de la recherche locale mais les opérateurs sont sen-

siblement différents. En effet, un algorithme évolutif comporte un ensemble d'opérateurs tels que la sélection, la mutation et éventuellement le croisement. La sélection a pour objectif de choisir les individus qui vont pouvoir survivre ou/et se reproduire pour transmettre leurs caractéristiques à la génération suivante. La sélection se base généralement sur le principe de conservation des individus les mieux adaptés et d'élimination des moins adaptés.

Le croisement ou recombinaison cherche à combiner les caractéristiques des individus parents pour créer des individus enfants avec de nouvelles potentialités dans la génération future.

La mutation effectue de légères modifications de certains individus.

2.5.3 Optimisation par essais particulières

L'optimisation par essaim particulaire "Particle Swarm Optimization (PSO)" est une méthode basée sur la reproduction d'un comportement social et la collaboration entre des individus souvent appelés particules. Ce comportement social s'appuie sur l'analyse de l'environnement et du voisinage et constitue alors une méthode de recherche de l'optimum par l'observation des tendances des particules voisines. Chaque particule cherche à optimiser ses chances en suivant la tendance de son voisinage qu'elle compose avec sa propre expérience. L'algorithme est généralement initialisé de façon aléatoire et les particules (solutions candidates) sont placées au hasard dans l'espace de recherche de la fonction objectif. Le concept principal de la PSO est que les solutions potentielles se dirigent vers les meilleures solutions. Les particules évaluent de manière itérative l'aptitude des solutions candidates et se souviennent de l'endroit où elles ont eu leur meilleure valeur de la fonction objectif. Chaque particule dispose d'une mémoire concernant sa meilleure solution visitée ainsi que la capacité de communiquer avec les particules de son entourage. À chaque itération k , les particules se déplacent en tenant compte de leur meilleure position, mais aussi de la meilleure position de ses voisines. L'objectif est de modifier leur trajectoires pour qu'elles se rapprochent le plus possible de l'optimum. L'optimum est obtenu par une procédure itérative sur la base des processus de mouvement et de l'intel-

ligence dans un système évolutif. La PSO met l'accent sur la coopération plutôt que sur la compétition et il n'y a pas de sélection (au moins dans la version de base). L'idée étant qu'une particule aujourd'hui encore pauvre mérite d'être préservée, peut-être justement parce que c'est elle qui saura trouver la meilleure solution, précisément parce qu'elle agit hors de l'espace prometteur à priori. Selon certains résultats, cette technique ne souffre pas des problèmes de pièges rencontrés par les méthodes évolutionnaires.

Un essaim de particules, qui sont des solutions potentielles au problème d'optimisation, survole l'espace faisable à la recherche de l'optimum global. Le déplacement d'une particule est influencé par les trois composantes suivantes :

1. Une composante d'inertie : la particule tend à suivre sa direction courante de déplacement.
2. Une composante cognitive : la particule tend à se diriger vers le meilleur site par lequel elle est déjà passée.
3. Une composante sociale : la particule tend à se fier à l'expérience de ses congénères et, ainsi, à se diriger vers le meilleur site déjà atteint par ses voisins.

La stratégie de déplacement d'une particule est illustrée dans Figure 2.2.

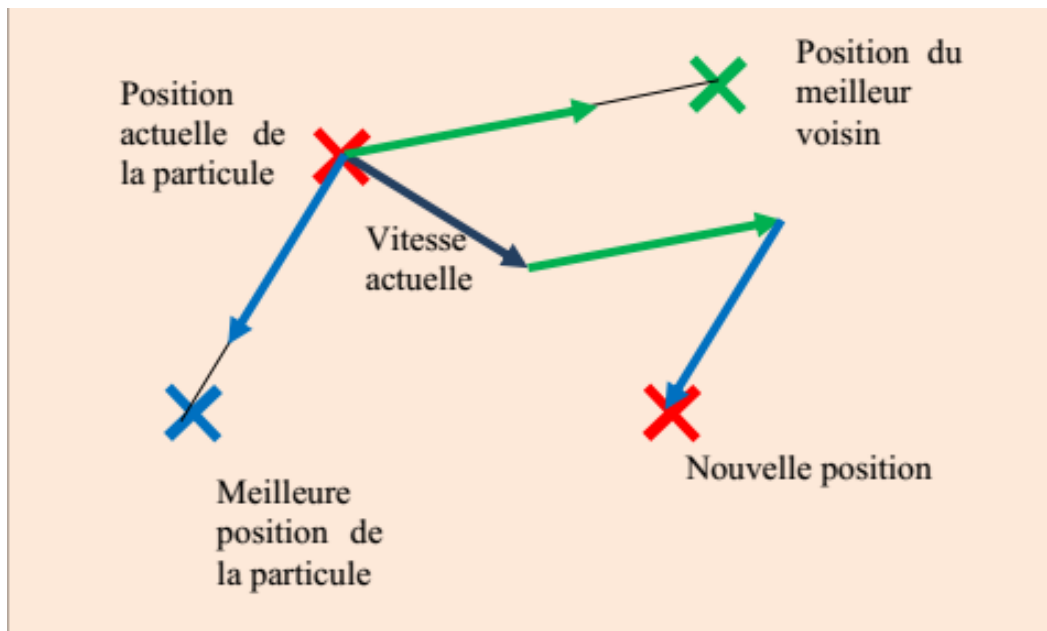


FIGURE 2.2 – Déplacement d'une particule.

Dans l'espace de recherche $D = \prod_{i=1}^n [a_i, b_i]$, une particule j de l'espace est représentée par sa position $x_j = (x_j^1, x_j^2, \dots, x_j^n)$ et son vecteur vitesse $v_j = (v_j^1, v_j^2, \dots, v_j^n)$. La qualité

de sa position est déterminée par la valeur de la fonction objectif en ce point. Cette particule garde en mémoire la meilleure position par laquelle elle est déjà passée, que l'on note $p_j = (p_j^1, p_j^2, \dots, p_j^n)$. La meilleure position atteinte par ses particules voisines est notée $g_j = (g_j^1, g_j^2, \dots, g_j^n)$. A l'itération k , le vecteur vitesse est calculé à partir de la formule suivante :

$$v_{j,k+1}^i = wv_{j,k}^i + c_1\alpha_i(k)(p_{j,k}^i - x_{j,k}^i) + c_2\beta_i(k)(g_{i,k}^i - x_{j,k}^i), i \in \{1, 2, \dots, n\}$$

où $w > 0$ est un paramètre, appelé coefficient d'inertie ; c_1 et c_2 sont deux constantes positives, appelées coefficients d'accélération et $\alpha_i(k)$ et $\beta_i(k)$ sont deux nombres aléatoires tirés uniformément de $[0, 1]$, à chaque itération k et pour chaque composante.

Il est à noter que le terme "vitesse" est ici abusif, car le vecteur v_j ne constitue pas une vitesse proprement dite. Cependant, pour respecter l'analogie avec le monde animal, la littérature a préféré utiliser ce terme.

Dans l'équation précédente, $v_{j,k}^i$ correspond à la composante d'inertie du déplacement. Le paramètre w contrôle l'influence de la direction de déplacement sur le déplacement futur. L'expression $c_1\alpha_i(k)(p_{j,k}^i - x_{j,k}^i)$ correspond à la composante cognitive du déplacement. Le paramètre c_1 contrôle le comportement cognitif de la particule. L'expression $c_2\beta_i(k)(g_{i,k}^i - x_{j,k}^i)$ correspond à la composante sociale du déplacement. Le paramètre c_2 contrôle l'aptitude sociale de la particule.

La combinaison des paramètres w , c_1 et c_2 permet de régler l'équilibre entre les phases de diversification et d'intensification du processus de recherche. Clerc et Kennedy [39] ont proposé l'utilisation d'un facteur pour prévenir la divergence de l'essaim. L'équation devient alors :

$$v_{j,k+1}^i = \psi(v_{j,k}^i + c_1\alpha_i(k)(p_{j,k}^i - x_{j,k}^i) + c_2\beta_i(k)(g_{i,k}^i - x_{j,k}^i), i \in \{1, 2, \dots, n\},$$

avec

$$\psi = \frac{2}{\phi - 2 + \sqrt{\phi^2 - 4\phi}}, \phi = c_1 + c_2 > 4;$$

la position à l'itération $k + 1$ de la particule j devient alors

$$x_{j,k+1}^i = x_{j,k}^i + v_{j,k+1}^i, i \in \{1, 2, \dots, n\}.$$

Si pour un certain i ; $x_{j,k+1}^i \notin [a_i, b_i]$ alors on pose :

$$\begin{cases} x_{j,k+1}^i = a_i & \text{si } x_{j,k+1}^i < a_i \\ x_{j,k+1}^i = b_i & \text{si } x_{j,k+1}^i > b_i. \end{cases} \quad (2.7)$$

2.5.4 Optimisation par Colonie d'abeilles

Les abeilles sont des insectes sociaux. Elles sont obligées de vivre en colonie très organisée, formée d'ouvrières, de faux-bourdon et d'une seule reine, et où chacune a un travail bien précis à faire. Les abeilles se nourrissent essentiellement de pollen et de miel. Elles vont butiner les fleurs pour prendre le nectar.

Les abeilles adultes (âgées de 20 à 40 jours) deviennent habituellement des butineuses. Les abeilles butineuses jouent en général l'un des trois rôles suivants : butineuses actives, butineuses éclaireuses et butineuses inactives.

Une colonie d'abeilles mellifères peut s'étendre sur de longues distances (plus de 10 km) et dans plusieurs directions simultanément pour exploiter un grand nombre de sources de nourriture. En principe, les abeilles visitent plus de parcelles de fleurs contenant beaucoup de nectar ou de pollen qui peuvent être récoltés avec moins d'effort, alors que les parcelles de fleurs contenant moins de nectar ou de pollen doivent recevoir moins d'abeilles [34]. Dans le processus de recherche, l'emplacement de la source de nourriture représente la solution possible au problème, et la quantité du nectar et de pollen de cette source correspond à une valeur objective dite fitness.

Le processus de recherche de nourriture commence dans une colonie par l'envoi des butineuses éclaireuses aux différentes sources de nourriture. Les butineuses éclaireuses se déplacent au hasard d'un champ de fleurs à un autre.

Quand les butineuses éclaireuses (qui ont trouvé une source de nourriture et qui ont évalué (fitness) au-dessus d'un certain seuil de qualité (mesuré à la combinaison de cer-

tains constituants, tels que la teneur en sucre)) retournent dans la ruche, déposent leur nectar ou pollen et commencent à danser. Cette danse mystérieuse est essentielle à la communication entre les abeilles et contient trois informations concernant un source de nourriture : la direction dans laquelle il sera trouvé, sa distance par rapport à la ruche et son indice de qualité [34]. Ces informations aide la colonie à envoyer ses abeilles à cette source avec précision, sans utiliser de guides. Après avoir dansé, les danseuses (c'est-à-dire les butineuses éclaireuses) retournent au source de nourriture avec des abeilles qui attendaient à l'intérieur de la ruche. Plus d'abeilles sont envoyées dans des sources plus prometteuses. Cela permet à la colonie de rassembler les aliments rapidement et efficacement. Lors de la récolte dans une source, les abeilles surveillent le niveau de nourriture. Cela est nécessaire pour choisir le type de la prochaine danse à leur retour dans la ruche.

Algorithme d'abeilles

L'algorithme des colonies d'abeilles [45] est une méthode basée sur la population pour trouver une solution optimale aux problèmes de recherche. Elle est inspirée du comportement des abeilles dans la nature.

L'algorithme nécessite la définition d'un certain nombre de paramètres :[45, 53]

1. n : nombre de butineuses éclaireuses,
2. m : nombre de sources sélectionnées sur n sources visitées,
3. e : nombre de meilleurs sources sur m champs sélectionnés,
4. nep : nombre d'abeilles recrutés pour les e meilleurs sources,
5. nsp : nombre d'abeilles recrutées pour les autres sources sélectionnées,
6. ngh : taille initiale des sources,
7. critère d'arrêt.

À l'étape 4, les abeilles qui ont les plus grand fitness sont choisies comme « abeilles sélectionnées » et les sources visités par celles-ci sont choisies pour une recherche au voisinage. Ensuite, aux étapes 5 et 6, l'algorithme effectue des recherches au voisinage des sources sélectionnées, affecter plus d'abeilles pour rechercher près de e meilleurs sources. À l'étape 7, les abeilles restantes de la population sont affectées.

L'algorithme d'abeille peut être illustré comme suit :

Étape 1 :Initialiser la population avec des solutions aléatoires.

Étape 2 : Évaluer la fitness de la population.

Étape 3 : Dès que (critère d'arrêt non rempli) Formation d'une nouvelle population.

Étape 4 : Sélectionnez les sources pour une recherche au voisinage.

Étape 5 : Recrutez des abeilles pour les sources sélectionnés (plus d'abeilles pour les meilleurs sources) et évaluer les fitness.

Étape 6 : Sélectionnez l'abeille correspond à $\max(\text{fitness})$ de chaque source.

Étape 7 : Affecter aux abeilles restantes une recherche aléatoire et évaluer leurs fitness.

Étape 8 : Fin While.

Chapitre 3

Proposition d'une nouvelle approche pour les algorithmes d'optimisation chaotiques

3.1 Introduction

L'identification de l'optimum global d'une fonction multimodale reste encore une tâche très difficile, malgré la puissance de la technologie de l'information d'aujourd'hui et les multitudes de techniques de recherche qui existent. Dans l'optimisation des problèmes non convexes, les algorithmes d'optimisation chaotiques (COAs) en deux phases possèdent une grande efficacité, et procèdent à une bonne exploration de l'espace de recherche par rapport aux autres méthodes probabiliste d'optimisation [46, 57]. Cependant, ces algorithmes présentent une faible exploitation des informations données par ces échantillons. Cette faiblesse engendre une perte de temps considérable, ce qui laisse l'algorithme converger vers l'optimum global lentement. Les algorithmes proposés par Ying Song [56], L. S. Coelho [54], et W.S. Jiang [29, 32] sont basés sur une recherche chaotique globale suivi d'une recherche local utilisant l'application Logistique et sont basés sur un partitionnement uniforme de l'espace de recherche de manière à ce que leur exploitations restent limités par ce mécanisme statique de partitionnement. Bien que Ying Song utilise des approches bidirectionnelles pour déterminer la prometteuse région la plus susceptible de contenir l'optimum global, l'efficacité de la méthode reste encore limitée.

Dans ce chapitre, nous proposons une nouvelle approche pour les algorithmes d'optimisations chaotiques. Dans cette approche appelée EBCOA, on introduit un mécanisme des phases imbriquées, une phase globale ponctuée des phases de recherche locale bidirectionnel basé sur un générateurs de nombres chaotiques bidimensionnel.

L'idée principale que nous développons dans ce chapitre, est de construire un algorithme basé sur la recherche chaotique autour des meilleurs individus d'une population associée à une sous recherche locale. La combinaison de la recherche locale avec la recherche globale peut assurer un temps de convergence très intéressant.

L'utilisation de cette approche améliore l'efficacité de l'algorithme de recherche puisque l'exploitation des renseignements est maximisée. L'utilisation d'une sous-phase de recherche locale bidirectionnelle est basée sur une division non uniforme (Gaussienne) de l'espace de recherche, ce qui laisse supposer que les sous régions n'ont pas les mêmes dimensions.

3.2 Stratégie de la recherche chaotique.

Parmi les stratégies de recherche couramment utilisées dans l'optimisation globale, la stratégie de recherche chaotique s'inspire du phénomène de chaos dans la nature. Le chaos peut être généré par un système dynamique non-linaire de structure assez simple. En effet, une simple équation de récurrence peut produire des dynamiques chaotiques assez complexes et riches [37, 41]. Les signaux chaotiques possèdent de nombreuses caractéristiques distinctes, telles que : ‘

- Haute sensibilité aux conditions initiales et aux paramètres du système.
- Auto-et inter corrélations proches de celles des signaux aléatoires.
- Pseudo-aléatoire et imprévisible à long terme.
- Déterministe.
- Ergodique, c.-à-d, la plupart des orbites conduisent à la même distribution.

L'idée d'utiliser des systèmes chaotiques au lieu des processus aléatoires a été remarquée dans plusieurs domaines. Un de ces champs est la théorie de l'optimisation. Dans les algorithmes d'optimisation aléatoire le rôle du hasard peut être joué par une dynamique

chaotique. Des études expérimentales ont affirmé que les avantages de l'utilisation des modèles chaotiques au lieu des modèles aléatoires sont souvent évidents mais il n'est pas encore mathématiquement prouvé [29, 32, 54, 55, 57].

3.2.1 Génération de séquences chaotiques.

Dans la littérature, il y a beaucoup de séquences et générateurs chaotiques mono et multidimensionnels qui sont utilisés dans diverses applications telles que : processus d'algorithmes évolutifs, sources pseudo- aléatoires, communication et sécurité de l'information, etc. . .

Rappelons, que le chaos peut être généré par tout système dynamique non-linéaire. En effet, des simples équations de récurrence sont capables de générer des dynamiques chaotiques riches, si les paramètres de contrôle sont bien positionnés .

Les séquences chaotiques sont utilisées dans différentes applications liées à l'optimisation mathématiques. Les séquences chaotiques sont des candidats possibles comme générateurs de nombres pseudo aléatoires et peuvent remplacer les générateurs pseudo-aléatoires traditionnels.

En effet, les séquences chaotiques possèdent des caractéristiques très intéressantes pour explorer l'espace de recherche et pour le processus d'algorithmes évolutifs telles que : haute sensibilité extrême aux conditions initiales et aux paramètres du système , auto et inter-corrélation similaires aux signaux pseudo-aléatoires, ergodicité, etc.

Les variables chaotiques peuvent passer par tous les états dans certaines plages selon leur propre régularité sans répétition [46, 42]. Plusieurs générateurs à base de modèle chaotiques ont été développés dans la littérature depuis 1990. Le générateur chaotique proposé dans cette thèse, est totalement numérique et est constitué à partir du modèle chaotiques de Lozi [10] donnée par :

$$\begin{cases} y_1(k) = 1 - a|(y_1(k-1))| + by(k-1) \\ y(k) = y_1(k-1) \end{cases} \quad (3.1)$$

$$z(k) = \frac{y(k) - \alpha}{\alpha' - \alpha}, \quad (3.2)$$

où k est le nombre d'itération. Dans ce travail, les valeurs de y sont normalisées dans l'intervalle $[0, 1]$ pour chaque variable dans l'espace à 2 dimensions du problème d'optimisation.

Par conséquent, $y_1 \in [-0.6417; 0.6716]$ et $(\alpha; \alpha') = (-0.6418; 0.6716)$.

Les paramètres utilisés dans cette étude sont $a = 1.7$ et $b = 0.5$, voir figure 3.1, ces valeurs sont illustrées dans [54].

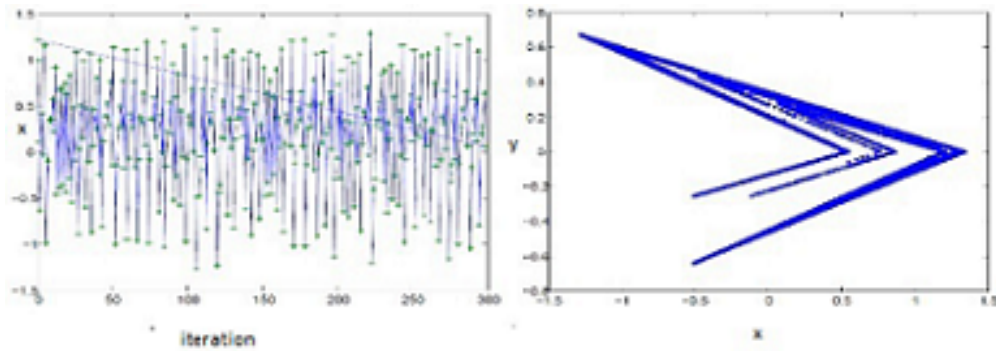


FIGURE 3.1 – Attracteur et séries temporelles de la suite de Lozi

3.2.2 Mécanisme de l'algorithme BCOA

Dans cette section, nous rappelons brièvement le BCOA introduit par Ying Song [56]. Beaucoup de stratégies chaotiques d'optimisation globale se composent de deux phases : La phase globale et la phase locale phase. Lors de la phase globale, des points chaotiques sont tirés du domaine des recherches X selon une certaine distribution, souvent uniforme. Ensuite, la fonction objectif est évaluée dans ces points. Pendant la phase locale, les points d'échantillonnage sont manipulés au moyen d'une recherche locale pour obtenir un minimum global candidat.

Considérons le problème d'optimisation pour une fonction non linéaire :

$$\begin{cases} \min f(X), & X = [x_1, x_2, x_3, \dots, x_n], \\ L_i \leq x_i \leq U_i. \end{cases}$$

Les variables chaotiques sont :

$$Z^{(k+1)} = g(Z^{(k)}),$$

où $Z^{(k)}$ sont des variables chaotiques générés par l'équation chaotique.

La méthode BCOA peut être décrit comme suit :

Etape 01 : (ainsi appelée la phase globale). On définit un générateur de séquences chaotiques basées sur une application chaotique (la suite logistique) puis on calcule les fonctions objectifs aux points générés, et on choisit le point où la fonction atteint son optimum comme l'optimum réel .

Les variables chaotiques sont donnés par :

$$X^k = c + d.Z^{(k)}, \quad (3.3)$$

où c et d sont des vecteurs constants constitués de n éléments respectivement $c_i = L_i$ et $d_i = U_i - L_i$

Etape 02 : (ainsi appelée la phase locale) . L'optimum réel est supposé proche de l'optimum global après certaines itérations, et il est considéré comme le centre avec un peu de perturbation chaotique et l'optimum global est obtenu grâce à la recherche fine. Répétez les deux étapes ci-dessus jusqu'à ce qu'un critère de convergence spécifié soit satisfait, et l'optimum global est obtenu. L'approche de la phase locale est la suivante :

$$X = \bar{X} + \beta\bar{X}(0.5 - Z), \quad (3.4)$$

\bar{X} est l'optimum approché, β est le paramètre de la deuxième étape. on a

$$-0.5\beta \leq \beta(0.5 - Z) \leq 0.5\beta \quad \text{si} \quad \beta \geq 0, \quad (3.5)$$

$$0.5\beta \leq \beta(0.5 - Z) \leq -0.5\beta \quad \text{si} \quad \beta \leq 0, \quad (3.6)$$

la recherche se fait donc des deux côtés de la solution.

Algorithme BCOA

Step 1 : Initialization. Set $k = 0$ and $k' = 0$, Set the initial best objective function $\bar{f} = +\infty$.

Step 2 : algorithm of chaotic global search.

Begin

While $k \leq \max K$ do

$k = k + 1$

First carrier wave using (3.3)

If $f(X^{(k)}) \leq \bar{f}$ then

$\bar{X} = X^{(k)}$ and $\bar{f} = f(X^{(k)})$

End If

End While

End

Step 3 : Algorithm of local search Begin

While $k' \leq \max K'$ and (CT is not satisfied) do

$k' = k' + 1$

Second carrier wave using (3.4)

If $f(x^{(k')}) \leq \bar{f}$ then

$\bar{X} = X^{(k')}$ and $\bar{f} = f(X^{(k')})$

End If

End While

End

3.3 Proposition d'une nouvelle approche EBCOA

3.3.1 Description de la méthode EBCOA

Dans cette section, nous proposons une nouvelle stratégie d'optimisation chaotique qui est une amélioration de la méthode BCOA [56] (bidirectionel Chaotic Optimization Algorithm) qui repose sur une recherche globale suivie d'une recherche locale et peut donner des résultats insatisfaisants.

Nous avons fait une modification dans la phase globale de la recherche. On a fait quelques étapes de recherches locales chaotiques autour de chaque point obtenu par la série chaotique, c-à-d en optimisant localement pendant quelques itérations le résultat global. Cette nouvelle recherche est appelée EBCOA. Donc EBCOA a améliorée la méthode BCOA par une recherche globale localement moyennée.

Dans de nombreuses stratégies de recherche chaotique, l'application Logistique [56, 29, 32] est utilisée pour générer une séquence chaotique, nous remplaçons la suite uniforme logistique par la suite gaussienne de Lozi pour accélérer le taux de convergence.

EBCOA peut être illustrée comme suit :

Entrées :

M_g : le nombre maximal d'itérations de la recherche globale chaotiques

M_l : le nombre maximum d'itérations de la recherche locale chaotiques.

M_{gl} : le nombre maximal d'itérations de la 1er recherche globale-locale.

$M_g \times M_{gl} + M_l$: le critère d'arrêt de la méthode d'optimisation chaotique.

β : le pas dans la recherche locale chaotique.

Sorties :

\bar{X} : la meilleure solution d'exécution en cours de recherche chaotiques.

\bar{f} : la meilleure fonction objectif.

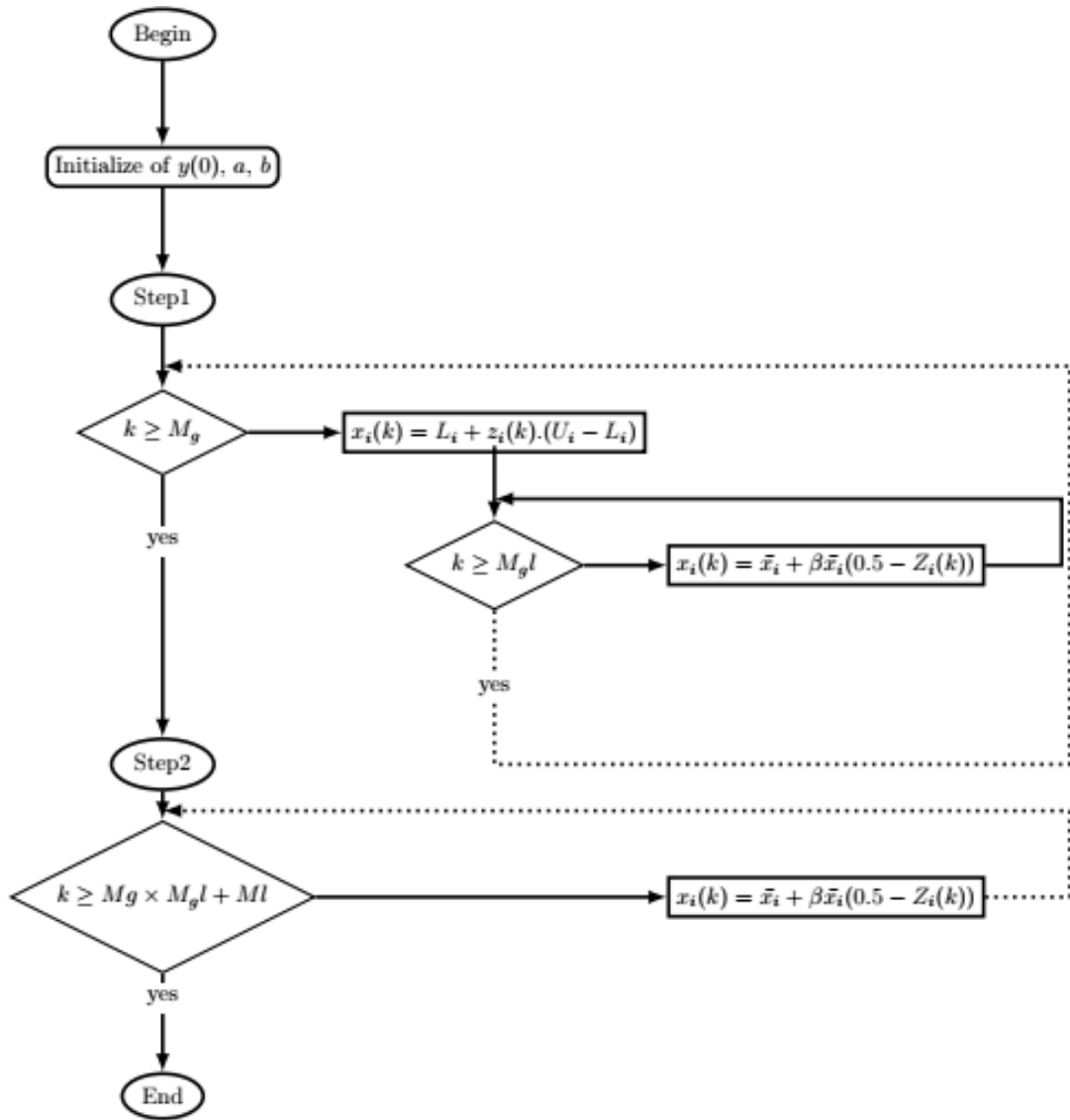


FIGURE 3.2 – Schéma de la méthode EBCOA

3.3.2 Contrôle de la taille des pas (rôle du β)

Au cours (Lors) de la recherche locale chaotique le pas β est un paramètre très important dans le comportement de convergence de la méthode d'optimisation qui permet de :

- Régler le rayon autour de \bar{X} .
- La taille du pas β est utilisée pour la génération d'une nouvelle solution.
- Utiliser pour contrôler l'impact de \bar{X}_k sur la génération \bar{X}_{k+1} .
- Un petit β sert à réaliser l'exploitation pour affiner les résultats par recherche locale.

- tandis qu'un grand β tend à faciliter une exploration globale de l'espace de recherche.
- Un petit β sert à bien exploiter \bar{X}
- Un grand β facilite l'exploration de l'espace de recherche. Une valeur convenable pour le pas β fournit l'équilibre entre les capacités d'exploration globales et locales et par conséquent une réduction du nombre d'itérations nécessaires pour trouver la solution optimale.

3.4 Simulation numériques

En mathématiques appliquées, les fonctions de test [48][43] sont utiles pour évaluer les caractéristiques des algorithmes d'optimisation. Pour tester notre approche, nous utilisons deux fonctions de référence non linéaires (La fonction Griewank et La fonction Rastrigin

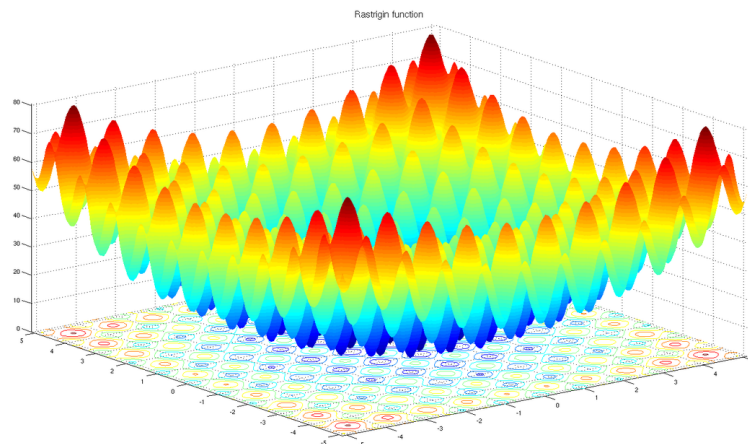
Dans notre étude, nous surmontons cette limitation en utilisant un certain nombre de dimensions 2 et en comparaison avec d'autres algorithmes d'optimisation heuristique.

La fonction de Rastrigin

La fonction de Rastrigin est une fonction mathématique souvent utilisée pour évaluer la performance d'algorithmes d'optimisation. Elle présente des pièges intéressants, sous la forme de ses nombreux minima et maxima locaux. Elle a été proposée, en 1974, par Rastrigin en deux dimensions et a été généralisée par Mühlenbein et D. Schomisch et J. Born.

Sa définition, en dimension n, est : $f(x) = 10n + \sum_{i=1}^n [x_i^2 - 10\cos(2\pi x_i)]$.

Son minimum global se trouve à l'origine, où sa valeur est nulle.



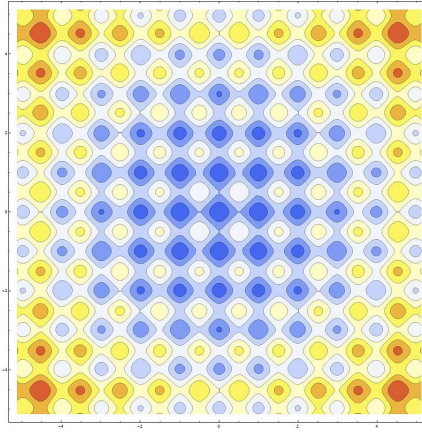
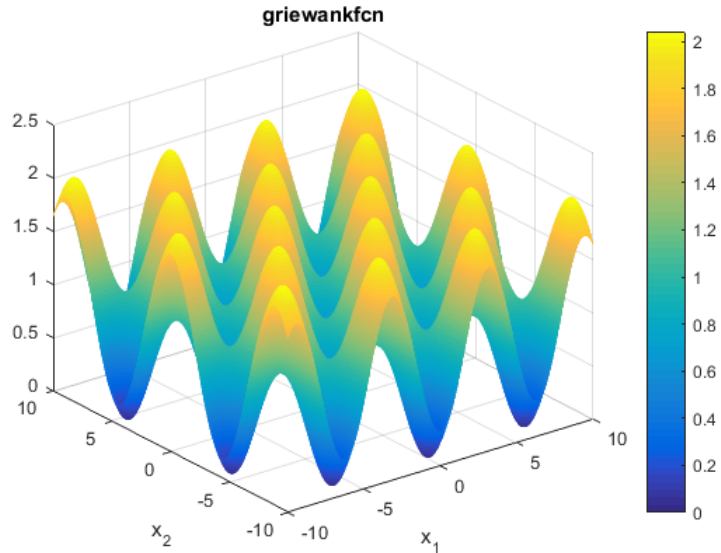


FIGURE 3.3 – fonction de Rastrigin à deux variables

La fonction de Griewank

La fonction Griewank est une fonction largement utilisée pour tester la convergence des fonctions d'optimisation. sa formule est donnée par : $f(x) = 1 + \frac{1}{4000} \sum_{i=1}^n [x_i^2 - \prod_{i=1}^n \cos(\frac{x_i}{\sqrt{i}})]$.

Son minimum global se trouve à l'origine, où sa valeur est nulle.



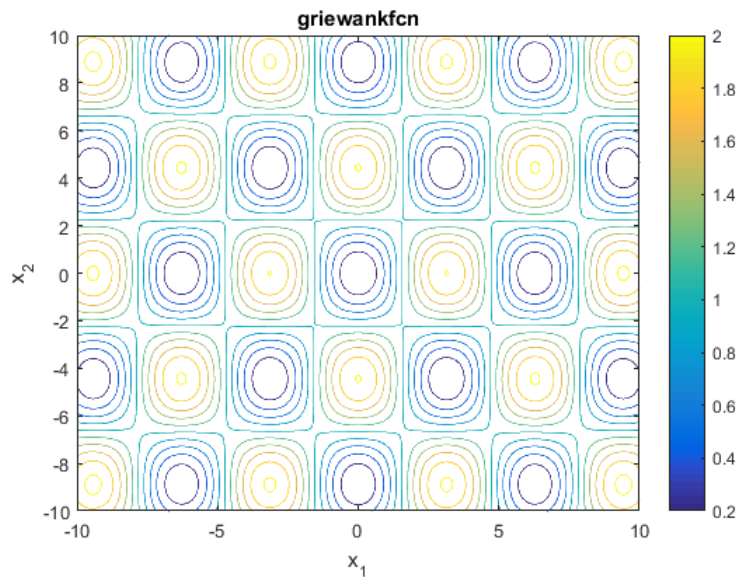


FIGURE 3.4 – fonction de Griewank à deux variables

Dans la suite, nous comparons les résultats obtenus par BCOA avec EBCOA en dimension 2, ces résultats montrent que la méthode EBCOA est meilleure que BCOA .

Dans chaque étude de cas, 50 essais indépendants ont été effectués pour chacune des méthodes EBCOA. Dans les cas testés pour comparer les problèmes, le nombre maximum d'itérations $maxK$ et $maxK'$ était de 10000 et 10000 itérations.

3.4.1 Résultats pour la fonction de Rastrigin

		BCOA	EBCOA
K'	β	optimum	optimum
1001	700	4.2752e-6	0
1001	500	4.7997e-9	0
1001	400	1.1219e-11	0
2405	200	3.5527e-15	0
1023	0.1	3.9080e-14	3.90798505 e -14
6965	0.01	4.3343e-13	4.192202141 e -13
$maxK'$	1e -3	2.6392e-5	5.419204974544 e - 6
$maxK'$	1e -4	4.7111e-4	8.3677132572291 e-5
$maxK'$	-(1 e -3)	2.8008e-5	2.1552183152806 e-5
$maxK'$	-(1 e -4)	4.7392e-4	5.0020093164866 e -5

TABLE 3.1 – L'optimum de la fonction de Rastrigin pour $n = 2$ avec différents β .

		BCOA	EBCOA
optimum	β	K'	K'
0	100	1206	10
	10	399	10
	1	368	10
	-100	1918	10
	-10	421	10

TABLE 3.2 – Nombre d'itérations avec différents β .

A partir du tableaux 3.1, pour $\beta \geq 200$, EBCOA peut trouver l'optimum réel 0.

Pour $|\beta| \in [1e - 4, 200] \cup] - 120, -(1e - 4)[$. l'optimum est amélioré .

A partir du tableau 3.2 por $\beta \in]0.1, 100]$, EBCOA peut également trouver l'optimum réel 0 mais avec nombres d'itérations moins que BCOA.

La valeur optimale et la vitesse de convergence sont meilleures que celles du COA [32]

et de ses améliorations, telles que MSCOA[35], COA-BFGS [49] et d'autres algorithmes évolutifs (comme le GA , PSO et ses améliorations) [44, 43].

3.4.2 Résultats pour la fonction de Creiwank

		BCOA	EBCOA
K'	β	optimum	optimum
802	11.12	0	0
346	11.10	0.2533	0
904	-9.93	0	0
372	-9.91	0.2516	0

TABLE 3.3 – L'optimum de la fonction de Griewank pour $n = 2$ avec différents β .

		BCOA	EBCOA
optimum	β	K'	K'
0	11.09	802	10
	10	550	10
	1.60	347	10
	-9.90	904	10
	-10	489	10
	-1.33	369	10

TABLE 3.4 – Nombre d'itérations avec différents β .

A partir du tableaux 3.3 et 3.4 , pour $\beta \geq 11.09$ et $\beta \leq -9.90$, EBCOA peut trouver toujours l'optimum réel 0.

Pour $\beta \in [1.60, 11.09] \cup [9.90, -1.33]$, EBCOA peut également trouver l'optimum réel 0 mais avec nombres d'itérations moins que BCOA.

La valeur optimale et la vitesse de convergence sont meilleures que celles du COA [32] et de ses améliorations, telles que MSCOA[35], COA-BFGS [49] et d'autres algorithmes évolutifs (comme le GA , PSO et ses améliorations) [44, 43].

3.5 Applications de la nouvelle technique

Afin de tester l'efficacité de la nouvelle approche proposée, et afin d'enrichir notre étude, nous allons l'appliquer à certaines fonctions test dont la recherche de minimum est difficile en choisissant différentes valeurs du nombre d'itérations et de pas β .

3.5.1 La fonction de Rosenbrock

La première fonction est la fonction de Rosenbrock à deux variables définie par :

$$f(x, y) = (1 - x)^2 + 100(y - x^2)^2$$

son minimum global est $f(1, 1) = 0$

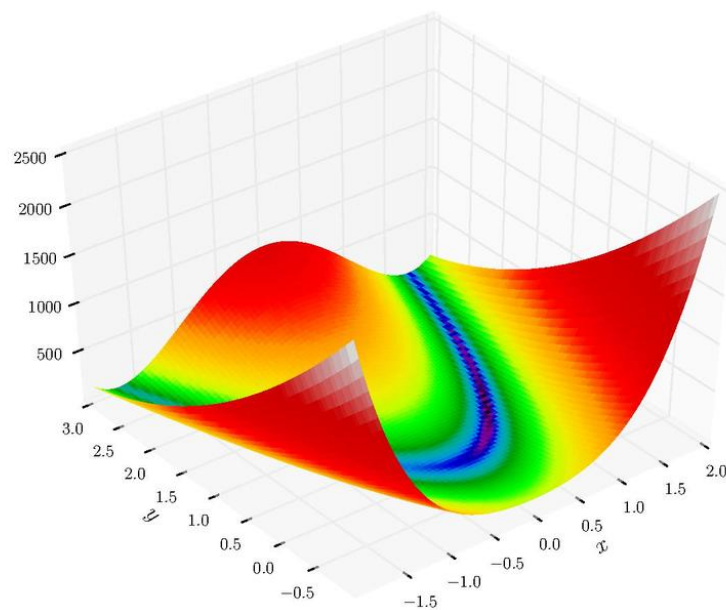


FIGURE 3.5 – fonction de Rosenbrock

M_g	M_{gl}	Ml	β	optimum
300	5	1000	700	0.000002606934759
100	50	600	500	0.235708453474101
150	6	100	10	0.000000028787292
326	4	1023	0.1	0.117853786900490
220	4	10000	0.01	0.000000000001901
220	4	10000	0.001	0
326	15	10000	-0.01	0.000000000035251
150	20	50	-0.001	0.000000000000010
400	20	900	-(1e-3)	0.000017329974426
300	5	6000	-(1e-4)	0
300	5	6000	-(1e-5)	0
300	8	1000	-10	0.000005743887348
400	5	1000	-100	0.000002349918484836
326	4	669	-200	0.004326082833209

TABLE 3.5 – Résultats d'optimisation de f_1

3.5.2 La fonction de Easom

La deuxième fonction est la fonction de Easom à deux variables définie par :

$$f(x, y) = -\cos(x)\cos(y)\exp(-(x - \pi)^2 - (y - \pi)^2).$$

son minimum global est $f(\pi, \pi) = -1$

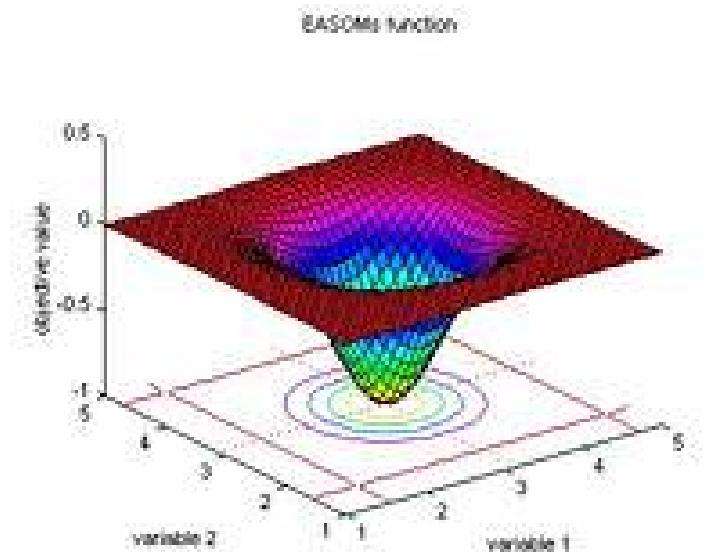


FIGURE 3.6 – fonction de Easom

M_g	M_{gl}	M_l	β	optimum
300	5	1000	700	-0.999664903059320
100	50	600	500	-0.998097535456517
150	6	100	10	-0.999881529010626
326	4	1023	0.1	-0.999999995240409
220	4	10000	0.01	-0.999999999955353
326	15	10000	-0.01	-0.99999999999181
150	20	50	-0.001	-1
400	20	900	-0.0001	-0.999999999999999
400	20	9000	-0.00001	-0.999998379519861
300	8	1000	-10	-0.999863011968368
400	5	1000	-100	-0.999822730309138
326	4	669	-200	-0.999963429422523

TABLE 3.6 – Résultats d'optimisation de f_2

3.5.3 La fonction de Matyas

La troisième fonction est la fonction de Rastrigin à deux variables définie par :

$$f(x, y) = 0.26(x^2 + y^2) - 0.48xy$$

son minimum global est $f(0, 0) = 0$

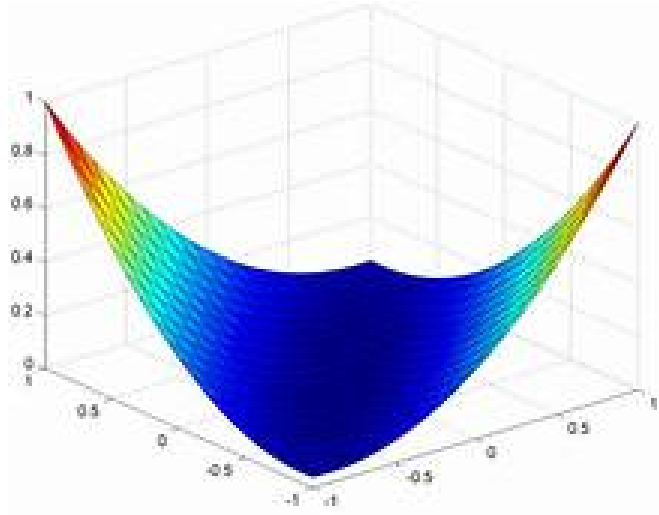


FIGURE 3.7 – fonction de Matyas

M_g	M_{gl}	Ml	β	optimum
300	5	1000	700	0.000000000387252
100	50	600	500	0
150	6	100	10	0
326	4	1023	0.1	0
220	4	10000	0.01	0.000000000003406e-9
326	15	10000	-0.01	0.000000000224091e-7
150	20	50	-0.001	0.000012192911768
400	20	900	-0.0001	0.000013953094222
300	5	6000	-0.00001	0.000025567240575
300	8	1000	-10	0.000010463822737
400	5	1000	-100	0
326	4	669	-200	0

TABLE 3.7 – Résultats d'optimisation de f_3

3.6 Conclusion

Dans cette thèse on a essayé de mettre certains concepts de la théorie du chaos à la disposition de l'optimisation globale en particulier l'attracteur de Lozi.

L'objectif de cette étude a visé une contribution au développement de méthodes d'optimisation globale basée sur les systèmes chaotiques. Cette méthode appelée EBCOA qui améliore BCOA et repose sur une modification dans la phase globale suivie d'une recherche locale.

L'étude permet de conclure que l'utilisation des algorithmes chaotiques avec l'approche proposée, représente une technique d'optimisation globale efficace.

Conclusion générale et perspectives

Récemment et en raison de plusieurs caractéristiques dynamiques importantes du chaos, à savoir : la sensibilité aux conditions initiales, l'ergodicité, la stochastique et la dimension fractale, de nombreux chercheurs se sont concentrés sur le développement d'algorithmes hybrides en combinant des algorithmes heuristiques avec les techniques de recherche chaotiques tout en évitant de rester bloqués dans les optimums locaux.

Ce travail est consacré à l'application de la théorie du Chaos à l'optimisation. Son objective a visé une contribution au développement d'une méthode d'optimisation globale basée sur les séquences chaotiques.

Cette nouvelle technique d'optimisation chaotique appelée EBCOA, basé sur la méthode des phases imbriquées, une phase globale ponctuée des phases de recherche locale bidirectionnel.

L'utilisation des algorithmes chaotiques avec l'approche proposée représente une technique d'optimisation globale très efficace. Les résultats des applications faites ont donné des résultats satisfaisants surtout en ce qui concerne le temps nécessaire pour atteindre un voisinage très étroit de l'optimum global et aussi la qualité de la solution obtenue.

Cependant, et comme perspectives des recherches supplémentaires sont nécessaires pour gagner en confiance et mieux comprendre la méthodologie proposée. Il est nécessaire d'évaluer l'algorithme proposé pour un grand nombre de fonctions de test à haute dimension ; analyser le comportement de notre approche lorsqu'elle est appliquée aux problèmes du monde réel et prouver l'algorithme avec d'autres types de cartes chaotiques de haute dimension.

Bibliographie

- [1] Holland, J. H. Outline for a logical theory of adaptive systems. Journal of the Association for Computing Machinery, 3, pp.297-314, 1962.
- [2] KAHANE, J.P. et SALEM, R. : Ensembles parfaits et séries trigonométriques, Hermann, paris(1963)
- [3] Becker, R. W., Lago, G. V. (1970). A global optimization algorithm. In Proceedings of the 8th Allerton Conference on Circuits and Systems Theory (pp. 3-12).
- [4] D. Ruelle and F. Takens, On the Nature of Turbulence, Commun. math. Phys. 20(1971) 167–192.
- [5] R. Abraham, L. Gardini & C. Mira, Chaos in discrete dynamical systems. Springer, 1997.
- [6] Holland, J. H. Genetic Algorithms and the Optimal Allocation of Trials, SIAM Journal on Computing, vol. 2, p. 88-105. 1973.
- [7] MANDELBORT, B.B. F ORME, Hasard et Dimension, Flammarion, Paris (1975).
- [8] Lozi R., (1978), “Un Attracteur étrange du Type Attracteur de Hénon ”. Journal de Physique. Colloque C5 ; Supplément au no8, Tome 39,pp.9-10.
- [9] Törn, A. A. (1978). A program for global optimization, multistart with clustering (MSC). Abo Akademi, datacentralen
- [10] R. Lozi, Un attracteur du type attracteur de Hénon, J. Phys. 39(C5), 1978, 9-10.
- [11] Kaplan, JL & Yorke, JA and Peitgen, HO and Walter, HO. “ Lecture Notes in Mathematics” Kapitel Chaotic behaviour of multidimensional difference equations, vol. 730, pp . 204, 1979. 51

- [12] GUCKENHEIMER, J. et HOLMES, P. (1983) *Nonlinear Oscillators, Dynamical Systems, and Bifurcations of Vector Fields*, volume 42 de *Applied Mathematical Sciences*. Springer verlag édition.
- [13] Frederickson, Paul & Kaplan, James L and Yorke, Ellen D and Yorke, James A. “The Liapunov dimension of strange attractors,” *Journal of Differential Equations*, vol. 49 no . 2, pp 185–207. 1983.51
- [14] Sawyer A. “The dynamics of piecewise linear mapping of the plane”, Ph.D. thesis, Boston, Michigan, USA ; 1984.
- [15] FALCONER : *The geometry of fractal sets*, Cambridge Univ.Press (1985).
- [16] A. V. Levy & A. Montalvo, The tunneling algorithm for the global minimization of functions, *SIAM J. Sci. Statist. Comput.*, 6, No. 1, 15-29, (1985)
- [17] R. Fletcher (1987), *Practical methods of optimization*, John Wiley and Sons, Chichester. of optimization, John Wiley and Sons, Chichester.
- [18] R. Fletcher, *Practical Methods of Optimization vol. 1 : Unconstrained Optimization* , John Wiley and Sons, New York, 1987.
- [19] BERGE, P. POMEAU, Y. et VIDAL, C. (1988). *L'ordre dans le chaos : vers une approche déterministe de la turbulence*. Hermann edition
- [20] O. Galor, *Discrete dynamical systems and chaos*. Longman Scientific, 1992.
- [21] R. G. Strongin & Ya. D. Sergeyev, Global multidimensional optimization on parallel computer, *Parallel Computing*, vol. 18, pp. 1259-1273. (1992)
- [22] R. Horst & H. Tuy, *Global Optimization, Deterministic Approach*, Springer-Verlag, Berlin. (1993).
- [23] KAPLAN, D. et GLASS, L. *Finite-difference equations. Understanding nonlinear dynamics*. New-York : Springer-Verlag, 1995, 1-53.
- [24] R. Horst & P. M. Pardalos, *Handbook of Global Optimization*, Kluwer Academic Publishers, Dordrecht. (1995).
- [25] R. Horst, P. M. Pardalos & N. V. Thoai, *Introduction to Global Optimization*, Kluwer Academic Publishers, Dordrecht. (1995)

- [26] Horst, R., Pardalos, P. M. (Eds.). (1995). Handbook of global optimization (Vol. 2). Springer Science and Business Media..
- [27] Back, T. Evolutionary Algorithms in Theory and Practice. Oxford Univ. Press (1996).
- [28] JF. Bonnans, J-C. Gilbert, C. Lemaréchal, C. Sagastizàbal, Optimisation Numérique, Aspects théoriques et pratiques, Springer M and A 27, 1997.
- [29] B. Li, W.S. Jiang, Chaos optimization method and its application. Journal of Control Theory and Application 14 (4) (1997) 613-615.
- [30] Ishii Y, Sands D. “Monoyonicity of the Lozi family near the tent map”. Comm Math Phys1998;198 :397-406.
- [31] Cao Y, Liu Z. “Orientation-preserving Lozi map”. Chaos, Soliton Fractals 1998; 9(11) :1857-63.
- [32] B. Li, W.S. Jiang, Optimizing complex function by chaos search. Cybernetics and Systems 29 (4) (1998) 409-419. Oxford, UK, 2003.
- [33] Michalewicz, Z. Genetic algorithms + Data structures = evolution programs. Springer, 1999.
- [34] E. Bonabeau, M. Dorigo and G. Theraulaz, Intelligence : from Natural to Artificial Systems. Oxford University Press, New York, 1999.
- [35] T. Zhang, H. W. Wang, and Z. C. Wang, Mutative scale chaos optimization algorithm and its application. Control and Decision, vol. 14, no. 3,pp. 285-288, 1999.
- [36] DANG-VU, H. et DELCARTE, C. Bifurcations et Chaos. Paris : Ellipses, 2000.
- [37] S. H. Strogatz, Nonlinear dynamics and chaos, Massachussets : Perseus Publishing, 2000.
- [38] E.K.P. Chong, S.H. Zak, An Introduction to Optimization. Wiley, 2001.
- [39] Clerc, M., Kennedy, J. (2002). The particle swarm-explosion, stability, and convergence in a multidimensional complex space. Evolutionary Computation, IEEE Transactions on, 6(1), 58-73.
- [40] R.L. Devaney, An introduction to chaotic dynamical systems. In Adissionwisley, Redwood City, CA 37, 2003.

- [41] J.C. Sprott, *Chaos and Times-Series Analysis*, Oxford University Press, Oxford, UK, 2003.
- [42] R. Caponetto, L. Fortuna, S. Fazzino and M. G. Xibilia, Chaotic Sequences to Improve the Performance of Evolutionary Algorithms, *IEEE Transactions on Evolutionary Computation*, Vol. 7, n 3, June 2003, 289-304.
- [43] S. Janson, and M. Middendorf, A hierarchical particle swarm optimizer and its adaptive variant. *IEEE Trans. Syst., Man, Cybern. - Part B : Cybernetics*, vol. 35, no. 6, pp. 1272-1282, 2005.
- [44] B. Liu, L. Wang, Y. H. Jin, F. Tang, and D. X. Huang, Improved particle swarm optimization combined with chaos. *Chaos, Soli. Fract.*, vol. 25, no. 5, pp. 1261-1271, 2005.
- [45] D.T. Pham, E. Koc, A. Ghanbarzadeh, S. Otri, S. Rahim, M. Zaidi, The Bees Algorithm—a novel tool for complex optimisation problems, *Proceedings of the Second International Virtual Conference on Intelligent Production Machines and Systems*. (2006) 454–461.
- [46] Li LX, Yang YX, Peng H, Wang XD. An optimization method inspired by chaotic behavior. *Int J Bifurcat Chaos* 2006 ; 16 : 2351-2364.
- [47] Zhigljavsky, A., Zilinskas, A. (2007). *Stochastic global optimization (Vol. 9)*. Springer Science and Business Media.
- [48] Y. Song, Z. Q. Chen, and Z. Z. Yuan, New chaotic PSO-based neural network predictive control for nonlinear process. *IEEE Trans. Neural Networks* vol. 18, no. 2, pp. 595-600, 2007.
- [49] D. X. Yang, G. Li, G. D. Cheng, On the efficiency of chaos optimization algorithms for global optimization. *Chaos, Soli. Fract*, vol. 34, no. 4, pp. 1366-1375, 2007.
- [50] E. Zeraoulia & J. C. Sprott, On the robustness of chaos in dynamical systems :Theories and applications. *Front. Phys. China*, 3 ;pp. 195 204 ; (2008)
- [51] O. Galor, *Discrete dynamical systems and chaos*. Longman Scientific, 1992.
- [52] S.S. Rao, *Engineering Optimization Theory and Practice*. Wiley, 2009.

- [53] D.T. Pham, M. Castellani, The Bees Algorithm : modelling foraging behaviour to solve continuous optimization problems, Proceeding of Institute Mechanical Engineering, C : Journal of Mechanical Engineering and Science. 223(2009)
- [54] L. S. Coelho. Tuning of PID controller for an automatic regulator voltage system using chaotic optimization approach. Chaos, Solitons and Fractals, 39, 2009, 1504-1514.
- [55] H. Shayeghi, S. Jalilzadeh, H.A. Shayanfar, A. Safari, Robust PSS Design using Chaotic Optimization Algorithm for a Multimachine Power System. ECTI-CON 2009, Pattaya, Thailand, May 2009, pp. 40 - 43. 2007.
- [56] Ying Song. A Bi-directional Chaos Optimization Algorithm. 2010 Sixth International Conference on Natural Computation (ICNC 2010).
- [57] T. Hamaizia, R. Lozi, An improved chaotic optimization algorithm using a new global locally averaged strategy, Journal of Nonlinear Systems and Applications 3 (2) (2012) 58-63.
- [58] Zabinsky, Z. B. (2013). Stochastic adaptive search for global optimization (Vol. 72). Springer Science and Business Media.
- [59] F. Derouiche , T. Hamaizia. An Enhanced Bi-Directional Chaotic Optimization Algorithm Nonlinear Dynamics and Systems Theory, 20 (4) (2020) 365–373.

Résumé

Basé sur la propriété ergodique, le chaos est adopté pour enrichir le comportement de recherche et empêcher les solutions d'être piégées dans l'optimum local dans les problèmes d'optimisation.

Cette thèse se concentre sur l'exploration des effets des suites chaotiques et donne des conseils pour améliorer l'algorithme d'optimisation chaotique bidirectionnelle dans la résolution de problèmes d'optimisation. En proposant un nouvel algorithme EBCOA, nous avons amélioré BCOA en faisant quelques modifications dans l'étape globale de recherche, nous avons affiné la solution finale en utilisant une deuxième méthode bidirectionnelle de recherche locale. L'étude présentée nous permet de conclure que la méthode proposée est rapide et converge vers le meilleur optimum.

Mots clés : Algorithme d'optimisation chaotique, Chaos, optimisation globale.

Abstract

Based on the ergodic property, chaos is adopted to enrich the search behavior and prevent solutions from being trapped in the local optimum in optimization problems.

This thesis focus on exploring the effects of chaotic maps and giving guidance for improving Bi-directional chaotic optimization algorithm in solving optimization problems. Through proposing a new algorithm EBCOA, we have improved BCOA doing some modification in the global step of research, we refined the final solution using a second bi-directional method of local search. The presented study allows us to conclude that the proposed method is fast and converges to a good optimum.

Keywords : Chaos optimization algorithm, Chaos, *global optimization*.

ملخص

استناداً إلى خاصية الأرجوديك، تم اعتماد الفوضى لإثراء سلوك البحث ومنع الحلول من الوقوع في الوضع المحلي الأمثل في مشكلات التحسين.

تركز هذه الأطروحة على استكشاف تأثيرات التسلسلات الفوضوية وتقدم نصائح لتحسين خوارزمية تحسين الفوضى ثنائية الاتجاه في حل مشكلات التحسين. من خلال اقتراح خوارزمية EBCOA جديدة، قمنا بتحسين BCOA من خلال إجراء بعض التعديلات في خطوة البحث العامة، قمنا بتحسين الحل النهائي باستخدام طريقة بحث محلية ثنائية الاتجاه ثانية.

تتيح لنا الدراسة المقدمة أن نستنتج أن الطريقة المقترحة سريعة وتتقارب نحو الحل الأفضل.

الكلمات المفتاحية: خوارزمية التحسين الفوضوي، فوضى، تحسين عام.