

THESE

Présentée devant

L'ECOLE CENTRALE DE LYON

Pour obtenir le grade de
DOCTEUR
(Arrêté du 30/03/1992)

Spécialité: Génie Electrique

Préparée au sein de

L'ECOLE DOCTORALE

ELECTRONIQUE, ELECTROTECHNIQUE, AUTOMATIQUE

DE LYON

par

Bruno SARENI

**METHODES D'OPTIMISATION MULTIMODALES
ASSOCIEES A LA MODELISATION NUMERIQUE
EN ELECTROMAGNETISME**

Soutenu le 20 janvier 1999 devant la commission d'examen :

JURY : MM

J. L. COULOMB	Professeur - ENSIEG - LEG	Président - Rapporteur
M. MOUSSAOUI	Professeur - ECL - MI	Rapporteur
A. PETROWSKI	Maître de conférences - INT	Examineur
X. ROBOAM	CR CNRS - INPT - LEEI	Examineur
S. RUSSENSCHUCK	Docteur - CERN - LHC	Examineur
A. NICOLAS	Professeur - ECL - CEGELY	Examineur
L. KRÄHENBÜHL	DR CNRS - ECL - CEGELY	Examineur

C'est dans l'inconnu que l'on définit notre existence

Gene Roddenberry

A mes parents, *Claude et Jacqueline*,

A ma sœur, *Anne Marie*,

A ma famille,

A tous mes amis.

AVANT-PROPOS

Les travaux de recherche présentés dans ce mémoire de thèse ont été effectués au sein du Centre de Génie Electrique de Lyon à l'Ecole Centrale de Lyon.

Je remercie tout particulièrement :

Monsieur *Jean-Louis Coulomb*, Professeur à l'Institut Polytechnique de Grenoble pour avoir accepté de rapporter ce travail et de présider ce jury.

Monsieur *Mohand Moussaoui*, Professeur au département de Mathématiques et d'Informatique de l'Ecole Centrale de Lyon pour l'intérêt qu'il a porté à ce travail en acceptant d'être rapporteur.

Monsieur *Stephan Russenschuck*, chercheur au CERN, pour le grand honneur qu'il me fait en participant à ce jury.

Monsieur *Xavier Roboam*, chargé de recherche CNRS au LEEI, pour s'être intéressé à ce travail et avoir accepté de siéger à ce jury.

Monsieur *Alain Petrowski*, Maître de Conférences à l'Institut National des Télécommunications, pour les nombreuses discussions et les échanges scientifiques que nous avons eus autour des algorithmes génétiques. Sa présence dans le jury m'honore.

Monsieur *Alain Nicolas*, Professeur à l'Ecole Centrale de Lyon et directeur du CEGELY, pour son accueil au sein du Laboratoire et son soutien inconditionnel au cours de ces quatre années. Je lui exprime toute ma reconnaissance pour la confiance qu'il m'a accordée en me confiant ce travail.

Monsieur *Laurent Krähenbühl*, Directeur de recherche CNRS au CEGELY, pour avoir accepté d'encadrer ce travail, pour ses conseils avisés et toujours pertinents, sa rigueur scientifique et ses grandes qualités humaines. Qu'il trouve à travers ce mémoire l'expression de ma gratitude et de mon amitié qui j'espère se poursuivra au-delà des ces années passées au Laboratoire.

Mes remerciements s'adressent aussi à :

Philippe Auriol, Professeur à l'Ecole Centrale de Lyon et directeur de la formation doctorale pour avoir accepté d'être mon tuteur pédagogique durant mes années de monitorat.

Guy Grellet, Professeur à l'Université Claude Bernard de Lyon, pour m'avoir communiqué sa passion pour le génie électrique lors de mes études universitaires.

Abderrahmane Bérroual, Maître de Conférences à l'Ecole Centrale de Lyon pour le soutien et les conseils qu'il m'a apportés au cours de ces années.

Henri Foch, Professeur à l'Institut National Polytechnique de Toulouse, pour son accueil au sein du LEEI lors de mon service national scientifique.

Philippe Ladoux, Maître de Conférences à l'Institut Universitaire de Formation des Maîtres de Toulouse, pour m'avoir fait découvrir, lors de mon passage au LEEI, la recherche en Electronique de Puissance (et le cassoulet toulousain!)

Charles Joubert, Maître de Conférences à l'Institut Universitaire de Formation des Maîtres de Montpellier pour l'aide qu'il m'a apportée lors de mes premières années passées au CEGELY.

Je ne saurais oublier l'ensemble du personnel du CEGELY, enseignants chercheurs, doctorants, personnels administratifs et techniques pour la chaleureuse ambiance qu'ils ont su créer.

J'associe à ces remerciements *Michel Berland*, Professeur agrégé à l'IUT de Génie Electrique de Lyon ainsi que l'ensemble des enseignants du département pour m'avoir accueilli pendant mon Monitorat au sein des Labos d'automatismes industriels et d'électrotechnique de l'IUT.

Je remercie également mes amis Toulousains du LEEI, *Fernando Iturriz*, l'argentin qui pensait que son équipe pourrait un jour gagner la coupe du monde de football en France et *Philippe Baudesson*, l'albigeois toujours souriant et joyeux qui a témoigné tant de sympathie à mon égard.

Merci aussi à mes amis trekkers, mes amis du club de Tennis et du Tarot : *Hervé, le "Gros" Daniel, Maurice, Céline, Valérie* et tous les autres qui m'ont permis de m'évader pendant mes loisirs.

Enfin, je dois ce travail à mes parents, *Claude* et *Jacqueline*, et à ma sœur, *Anne Marie*, qui m'ont toujours soutenu pendant mes études universitaires. Je les en remercie du fond du cœur.

RESUME

**METHODES D'OPTIMISATION MULTIMODALES ASSOCIEES A
LA MODELISATION NUMERIQUE EN ELECTROMAGNETISME**

L'essor de l'informatique et des techniques d'intelligence artificielle a conduit ces dernières années à un développement sans précédent des procédés d'optimisation automatique qui peuvent aujourd'hui prendre en compte des dizaines de paramètres de conception. En particulier, les méthodes évolutionnistes ont connu depuis le début des années soixante une croissance exponentielle et s'affirment peu à peu comme les techniques les plus robustes : d'une part, elles permettent de localiser l'optimum d'une fonction dans l'espace des paramètres sans avoir recours aux dérivées de la fonction par rapport à ces paramètres ; d'autre part elles ne se laissent pas piéger par un optimum local et réussissent le plus souvent à déterminer l'optimum global de la fonction considérée.

Cependant, la traduction d'un problème d'optimisation réel avec tous ses aspects (performance "pure" mais aussi sensibilité, facilité de fabrication, prix de revient, ...) sous forme d'une fonction à optimiser n'est pas toujours une chose simple. Dès lors, le concepteur apprécie lorsqu'il étudie l'un des aspects d'être conduit à plusieurs possibilités (plus ou moins parfaites suivant cet aspect là) plutôt qu'à une solution unique. Les méthodes génétiques multimodales ou méthodes de nichage offrent des perspectives intéressantes en permettant la localisation de solutions optimales multiples, aussi bien locales que globales.

Notre travail est centré sur la caractérisation de ces nouvelles techniques d'optimisation numériques. Chaque méthode a été testée de façon classique à partir de fonctions mathématiques ainsi que sur des problèmes d'électromagnétisme et sur un procédé très original de conception de formes optimales d'électrodes. Nous décrivons une nouvelle approche pour des systèmes 2D-plan ou axisymétriques, où la forme de l'électrode est identifiée à une ligne équipotentielle obtenue par optimisation du positionnement et de la valeur d'un certain nombre de charges fictives.

MOTS CLES

Conception automatique optimale, algorithmes génétiques, méthodes de nichage, optimisation de forme, électrostatique, électrode, charges équivalentes.

ABSTRACT

MULTIMODAL OPTIMIZATION METHODS RELATED TO NUMERICAL MODELING IN ELECTROMAGNETICS

In recent years, expansion of computer sciences and artificial intelligence led to a unprecedented development of optimization methods which are today capable of taking into account a great number of design variables. In particular, evolutionary algorithms follow an exponential growth since the beginning of the sixties and are progressively considered as the most robust optimization techniques. On the one hand, they converge to an optimum of a given objective function without using any partial derivatives of this function in relation to design variables. On the other hand, they are not trapped by local optima and often succeed in determining the global optimum.

However, translating a real optimization problem in a mathematical form, by taking into account several characteristics (efficiency, sensitivity, cost price,...), is not easy. Consequently, engineers appreciate when they study one aspect to get various "good" possibilities rather than a single solution. For that purpose, niching genetic algorithms offer interesting perspectives by locating multiple optimal solutions either local as global.

Our work is centered on the characterization of these new optimization techniques. Each niching method is investigated on mathematical test functions and on electromagnetic benchmarks. Comparisons are also carried out on an original process of automatic electrode shape design. We describe a new approach for 2D and axisymmetric devices where the electrode shape is identified from an equipotential obtained from a system of fictitious point charges.

KEYWORDS

Automated optimal design, genetic algorithms, niching methods, shape optimization, electrostatics, electrode, charge simulation method.

TABLE DES MATIERES

INTRODUCTION GÉNÉRALE	1
CHAPITRE I : MÉTHODES D'OPTIMISATION MATHÉMATIQUES	5
I.1. DÉMARCHE SYSTÉMATIQUE EN OPTIMISATION.....	5
I.1.1. DÉFINITION PRÉCISE DU CAHIER DES CHARGES	6
I.1.2. MISE EN ÉQUATION MATHÉMATIQUE.....	6
I.1.3. RÉOLUTION DU PROBLÈME D'OPTIMISATION.....	6
I.1.4. EXPLOITATION DE LA SOLUTION	7
I.2. FORMULATION D'UN PROBLÈME D'OPTIMISATION	7
I.3. RÉOLUTION D'UN PROBLÈME NON CONTRAINT.....	8
I.3.1. DIFFICULTÉ D'UN PROBLÈME NON CONTRAINT	8
I.3.2. CARACTÉRISTIQUES DES MÉTHODES DE RÉOLUTION	9
<i>I.3.2.1. Modèle de la boîte noire</i>	<i>9</i>
<i>I.3.2.2. Efficacité d'une méthode de résolution.....</i>	<i>10</i>
<i>I.3.2.3. Opérateurs de recherche fondamentaux.....</i>	<i>10</i>
<i>I.3.2.4. Spécificités liées au mode de recherche de l'optimum.....</i>	<i>11</i>
<i>I.3.2.5. Ordre d'une méthode de résolution</i>	<i>11</i>
I.3.3. ETAT DE L'ART DES MÉTHODES DE RÉOLUTION.....	11
<i>I.3.3.1. Méthodes d'optimisation unidimensionnelles.....</i>	<i>11</i>
<i>I.3.3.2. Méthodes d'optimisation multidimensionnelles.....</i>	<i>12</i>
<i>I.3.3.3. Méthodes d'optimisation stochastiques</i>	<i>14</i>
I.4. PRISE EN COMPTE DES CONTRAINTES DE CONCEPTION.....	18
I.4.1. MÉTHODES DE TRANSFORMATION	18
<i>I.4.1.1. Méthodes de pénalités intérieure et extérieure</i>	<i>18</i>
<i>I.4.1.2. Techniques de minimisation séquentielles sans contraintes.....</i>	<i>19</i>
<i>I.4.1.3. Pénalisation radicale des solutions</i>	<i>20</i>
<i>I.4.1.4. Continuité de la fonction de transformation.....</i>	<i>20</i>
<i>I.4.1.5. Mise à l'échelle des contraintes.....</i>	<i>21</i>
I.4.2. OPTIMISATION SÉQUENTIELLE DES CONTRAINTES.....	21
I.4.3. TRANSFORMATION D'UN PROBLÈME CONTRAINT EN UN PROBLÈME ÉQUIVALENT NON CONTRAINT À OBJECTIFS MULTIPLES.....	22

I.5.	OPTIMISATION A OBJECTIFS MULTIPLES	22
I.5.1.	FORMULATION D'UN PROBLÈME D'OPTIMISATION VECTORIEL	23
I.5.2.	OPTIMUM DE PARETO.....	23
I.5.3.	MÉTHODE DE PONDÉRATION	24
I.5.4.	MÉTHODE DE MARGLIN	25
I.5.5.	OPTIMISATION À OBJECTIFS MULTIPLES PAR LA LOGIQUE FLOUE	26
I.5.5.1.	<i>Notion de sous-ensemble flou</i>	<i>26</i>
I.5.5.2.	<i>Représentation des fonctions d'appartenance</i>	<i>27</i>
I.5.5.3.	<i>Opérations sur les sous-ensembles flous</i>	<i>29</i>
I.5.5.4.	<i>Résolution de problèmes à objectifs multiples par la logique floue</i>	<i>30</i>
I.6.	CONCLUSION	31
	CHAPITRE II : L'ALGORITHME GÉNÉTIQUE STANDARD	33
II.1.	GÉNÉRALITÉS.....	34
II.1.1.	FORMULATION DU PROBLÈME D'OPTIMISATION.....	34
II.1.2.	TERMINOLOGIE, DÉFINITIONS ET NOTATIONS EMPLOYÉES	35
II.1.2.1.	<i>Terminologie et définitions</i>	<i>35</i>
II.1.2.2.	<i>Notations.....</i>	<i>35</i>
II.1.3.	LE CODAGE DES PARAMÈTRES	36
II.1.3.1.	<i>Le codage binaire standard</i>	<i>36</i>
II.1.3.2.	<i>Le code Gray.....</i>	<i>37</i>
II.1.3.3.	<i>Le codage dynamique des paramètres</i>	<i>38</i>
II.1.4.	LES SCHÉMAS.....	39
II.1.4.1.	<i>Principe.....</i>	<i>39</i>
II.1.4.2.	<i>Caractéristiques d'un schéma.....</i>	<i>39</i>
II.1.4.3.	<i>Interprétation géométrique des schémas</i>	<i>40</i>
II.1.4.4.	<i>Localisation des schémas dans l'espace de recherche réel.....</i>	<i>40</i>
II.2.	STRUCTURE DE L'ALGORITHME GÉNÉTIQUE.....	42
II.3.	INITIALISATION DE LA POPULATION	43
II.4.	LA SÉLECTION DES PARENTS	43
II.4.1.	LES MÉTHODES DE SÉLECTION.....	43
II.4.1.1.	<i>La sélection proportionnelle RWS</i>	<i>43</i>
II.4.1.2.	<i>La sélection proportionnelle à reste stochastique</i>	<i>45</i>
II.4.1.3.	<i>La sélection stochastique universelle.....</i>	<i>46</i>
II.4.1.4.	<i>La sélection linéaire par rapport au rang</i>	<i>47</i>

II.4.1.5.	<i>La sélection uniforme par rapport au rang</i>	47
II.4.1.6.	<i>La sélection par tournoi</i>	47
II.4.2.	LES CARACTÉRISTIQUES DES MÉCANISMES DE SÉLECTION	48
II.4.2.1.	<i>La sélection dynamique et la sélection statique</i>	48
II.4.2.2.	<i>La sélection préservatrice et la sélection extinctrice</i>	48
II.4.2.3.	<i>La sélection élitiste</i>	49
II.4.2.4.	<i>L'intensité de la sélection</i>	49
II.4.3.	EFFET DE LA SÉLECTION	50
II.5.	LES OPÉRATEURS DE RECOMBINAISON GÉNÉTIQUE	51
II.5.1.	LE CROISEMENT	51
II.5.1.1.	<i>Le croisement à un site</i>	52
II.5.1.2.	<i>Le croisement multi-sites</i>	53
II.5.1.3.	<i>Le croisement uniforme</i>	54
II.5.1.4.	<i>Quel type de croisement appliquer ?</i>	55
II.5.2.	LA MUTATION	56
II.5.2.1.	<i>Principe</i>	56
II.5.2.2.	<i>Taux optimal de mutation et mutation adaptative</i>	57
II.5.3.	LE RÉORDONNANCEMENT DES GÈNES.....	58
II.6.	LA SÉLECTION DES INDIVIDUS D'UNE NOUVELLE GÉNÉRATION	58
II.6.1.	LA SÉLECTION PAR DESCENDANCE	59
II.6.2.	LES SÉLECTIONS ÉLITISTES	59
II.7.	LA CONVERGENCE DES ALGORITHMES GÉNÉTIQUES	60
II.7.1.	ANALYSE DE LA CONVERGENCE	60
II.7.1.1.	<i>Le théorème des schémas</i>	60
II.7.1.2.	<i>Parallélisme implicite</i>	61
II.7.2.	CRITÈRES DE CONVERGENCE.....	61
II.7.3.	TAUX DE RÉUSSITE.....	62
II.8.	HYBRIDATION DES ALGORITHMES	62
II.9.	ALGORITHMES GÉNÉTIQUES À CODAGE RÉEL	63
II.9.1.	CODAGE DES PARAMÈTRES	63
II.9.2.	LE CROISEMENT	63
II.9.3.	LA MUTATION	64
II.9.4.	L'ALGORITHME GÉNÉTIQUE RECOMBINATEUR (BGA).....	65
II.9.5.	PERFORMANCES COMPARÉES	66
II.10.	CONCLUSION	69

CHAPITRE III : ALGORITHMES GÉNÉTIQUES MULTIMODAUX ET MÉTHODES DE NICHAGE.....	71
III.1. MÉTHODES DE NICHAGE.....	71
III.2. CRITÈRES DE DISTANCE.....	73
III.2.1. DISTANCE GÉNOTYPIQUE	73
III.2.2. DISTANCES PHÉNOTYPIQUES	74
III.2.3. DISTANCE NON PARAMÉTRIQUE	75
III.3. LES MÉTHODES DE SURPEUPLEMENT.....	75
III.3.1. LE SURPEUPLEMENT STANDARD	76
III.3.2. LA PRÉSÉLECTION	76
III.3.3. LE SURPEUPLEMENT DÉTERMINISTE	76
III.3.4. LA SÉLECTION PAR TOURNOI RESTREINT.....	77
III.4. LES MÉTHODES DE PARTAGE.....	78
III.4.1. LE PARTAGE STANDARD	78
III.4.1.1. Principe du partage	78
III.4.1.2. Choix du rayon de niche	80
III.4.1.3. Taille de la population.....	81
III.4.1.4. Sélection et croisement avec la méthode de partage	82
III.4.1.5. L'union restrictive	83
III.4.2. LE PARTAGE DYNAMIQUE ET LE PARTAGE RAPIDE	85
III.4.3. LA MÉTHODE DES SOUS-POPULATIONS	85
III.4.4. LE PARTAGE SÉQUENTIEL	86
III.4.4.1. La recherche d'une solution optimale	86
III.4.4.2. Réajustement de la fonction objectif.....	86
III.4.4.3. Critère d'arrêt.....	87
III.4.4.4. Conséquences du partage séquentiel.....	87
III.5. LA MÉTHODE D'ÉCLAIRCISSEMENT.....	88
III.5.1. PRINCIPE	88
III.5.2. IMPLANTATION DE LA MÉTHODE D'ÉCLAIRCISSEMENT	88
III.5.3. ELITISME.....	89
III.5.4. COMPLEXITÉ	89
III.5.5. CAPACITÉ DES NICHES.....	90
III.6. CRITÈRES DE PERFORMANCE.....	90
III.6.1. NOMBRE TOTAL DE SOMMETS MAINTENUS	90

III.6.2. TAUX DE DÉTECTION DES SOMMETS	90
III.6.3. CRITÈRE DE DISTRIBUTION PROPORTIONNELLE.....	91
III.7. ALGORITHMES GÉNÉTIQUES MULTIMODAUX COMPARÉS.....	92
III.7.1. COMPARAISON DES MÉTHODES DE NICHAGE AVEC L'ALGORITHME GÉNÉTIQUE STANDARD	92
<i>III.7.1.1. Efficacité dans la recherche de l'optimum global</i>	92
<i>III.7.1.2. Comportement qualitatif sur des problèmes multimodaux</i>	94
III.7.2. RÉOLUTION DE PROBLÈMES MULTIMODAUX	96
III.7.3. ETUDE DE LA MÉTHODE DE PARTAGE STANDARD	97
<i>III.7.3.1. Influence du mode de sélection</i>	97
<i>III.7.3.2. Influence du mode de recombinaison</i>	98
III.7.4. COMPARAISONS DES PRINCIPALES MÉTHODES DE NICHAGE	99
III.7.5. DISCUSSION	104
III.8. CONCLUSION	106
CHAPITRE IV : APPLICATIONS DES ALGORITHMES GÉNÉTIQUES MULTI- MODAUX A DES PROBLÈMES D'ÉLECTROMAGNÉTISME.....	71
IV.1. OPTIMISATION D'UN COMPOSITE DIÉLECTRIQUE.....	107
IV.1.1. DÉFINITION DU PROBLÈME	107
IV.1.2. RÉOLUTION PAR DES MÉTHODES GÉNÉTIQUES	109
IV.2. OPTIMISATION DE LA FORME D'UN PÔLE D'INDUCTEUR	113
IV.2.1. DÉFINITION DU PROBLÈME	113
IV.2.2. COUPLAGE AVEC LA MÉTHODE DES ÉLÉMENTS FINIS	114
IV.2.3. INDUCTION CONSTANTE SUR L'ARC AB.....	116
IV.2.4. INDUCTION SINUSOÏDALE SUR L'ARC AB.....	122
IV.3. CONCLUSION	124
CHAPITRE V : OPTIMISATION DE FORME D'ÉLECTRODES	127
V.1. EQUATIONS DE L'ÉLECTROSTATIQUE.....	127
V.2. DISTRIBUTIONS DE CHARGES DANS LE VIDE	129
V.3. SIMULATION DE CONDUCTEURS PAR DES CHARGES FICTIVES	130
V.4. MÉTHODE DE GÉNÉRATION DE FORMES D'ÉLECTRODES OPTIMALES.....	131

V.4.1. INTRODUCTION.....	131
V.4.2. PRINCIPE DE LA MÉTHODE.....	132
V.4.3. OPTIMISATION DES CHARGES	134
V.4.4. CALCUL DES GRANDEURS ÉLECTRIQUES.....	134
V.4.5. RESPECT DU GABARIT GÉOMÉTRIQUE.....	134
V.4.6. MISE À L'ÉCHELLE DU POTENTIEL	135
V.4.6.1. <i>Normalisation intégrée</i>	136
V.4.6.2. <i>Normalisation optimisée</i>	136
V.4.6.3. <i>Normalisation fixe</i>	137
V.4.7. IDENTIFICATION ET EXPLORATION D'UNE ÉQUIPOTENTIELLE VALIDE.....	137
V.4.7.1. <i>Identification d'un point de l'équipotentielle</i>	137
V.4.7.2. <i>Exploration de l'ensemble des points de l'équipotentielle</i>	137
V.4.8. CRITÈRES D'OPTIMISATION ET CALCUL DE L'OBJECTIF	139
V.4.9. TRAITEMENT DES CONTRAINTES ET PROCÉDÉ DE PÉNALISATION	139
V.4.10. SPÉCIFICITÉ DU PROBLÈME D'OPTIMISATION	140
V.4.10.1. <i>Modes géométriques recherchés</i>	140
V.4.10.2. <i>Modes dus aux possibles permutations de charges</i>	140
V.4.10.3. <i>Modes dus à la mise à l'échelle du potentiel</i>	140
V.5. EXEMPLES D'APPLICATIONS	142
V.5.1. ELECTRODE SIMPLE.....	142
V.5.1.1. <i>Définition du problème</i>	142
V.5.1.2. <i>Résolution du problème par des algorithmes génétiques multimodaux.</i>	143
V.5.1.3. <i>Influence de la méthode de normalisation du potentiel</i>	145
V.5.1.4. <i>Influence de la méthode de pénalisation</i>	146
V.5.1.5. <i>Localisation des profils optimaux dans l'espace de recherche</i>	149
V.5.1.6. <i>Exemple de profils optimaux</i>	152
V.5.2. CONDENSATEUR PLAN À LAMES PROFILÉES	154
V.5.2.1. <i>Définition du problème</i>	154
V.5.2.2. <i>Résolution numérique</i>	155
V.5.2.3. <i>Utilisation de charges fixes</i>	158
V.5.2.4. <i>Adaptation du gabarit</i>	159
V.5.2.5. <i>Profil optimal</i>	161
V.5.3. ELECTRODE AVEC BLINDAGE	163
V.5.3.1. <i>Définition du problème</i>	163
V.5.3.2. <i>Résolution numérique</i>	163
V.5.3.3. <i>Profil optimal</i>	165
V.6. CONCLUSION	166

CONCLUSION GÉNÉRALE	169
ANNEXE I : VALIDATION DU GÉNÉRATEUR DE NOMBRES ALÉATOIRES....	171
ANNEXE II : FONCTIONS TESTS MATHÉMATIQUES	177
ANNEXE III : EFFICACITÉ DES OPÉRATEURS GÉNÉTIQUES	189
ANNEXE IV : COEFFICIENTS DE POTENTIEL ET DE CHAMP	199
ANNEXE V : MÉTHODE DES CHARGES ÉQUIVALENTES	203
BIBLIOGRAPHIE.....	211
GLOSSAIRE DES TERMES EMPLOYÉS	221

INTRODUCTION GENERALE

Depuis l'avènement de la modélisation numérique sur calculateurs, les techniques de C.A.O. ont connu un fort développement. Les méthodes d'éléments finis [Coulomb 81] et d'intégrales de frontières [Krähenbühl 83][Nicolas 83] ont permis ces dernières années de caractériser de façon précise le comportement des dispositifs électrotechniques. Nous sommes maintenant capables, dès la conception, de prédire avec exactitude les performances d'un système électromagnétique en fonction de ses caractéristiques géométriques et structurales. La C.A.O. prend donc naturellement deux orientations liées au développement en parallèle des techniques d'optimisation numériques. D'un côté, nous cherchons, à "automatiser" les procédés de conception (*conception automatique*), de l'autre à améliorer simultanément les performances des dispositifs (*conception optimisée*).

La part de la recherche dédiée à l'optimisation en génie électrique est devenue de plus en plus importante. En 1993, seulement 7% des travaux présentés lors de conférences internationales telles que COMPUMAG étaient consacrés à l'optimisation, contre 15% en 1997. Le nombre d'articles sur ce thème a doublé en quatre ans! Le Tableau 1 montre que les applications concernées dans le domaine du calcul de champs électromagnétiques, sont très variées.

Tableau 1. Exemples d'applications en génie électrique

<i>Application concernée</i>	<i>paramètres</i>	<i>objectif</i>	<i>Référence</i>
machine synchrone	forme de pôle	réduction des harmoniques de couple	[Gizolme 97]
capteur à courants de Foucault (C.N.D.)	bobinage	détection des défauts	[Thollon 95]
électrodes et connecteur H.T.	forme géométrique	réduction du champ électrique	[Vasconcelos 94]
aimants supraconducteurs	inducteurs	induction maximale - réduction de section de la bobine supra	[Russenschuck 94]
cuisson par induction	inducteurs	induction maximale diminution du champ rayonné	[Leschi 94]
mailleur adaptatif	nœuds du maillage	meilleure qualité de maillage	[Saludjian 97]
matériaux hétérogènes anisotropes	propriétés physiques	comportement du matériau imposé	[Dyck 96]

Introduction générale

Outre les applications de conception automatique et d'optimisation proprement dites, les techniques d'optimisation sont largement utilisées pour la résolution de problèmes inverses [Bucci 90][Guarnieri 90][Hoole 91][Chat-uthai 96][Formisano 96].

L'essor de l'informatique et des techniques d'intelligences artificielles a conduit ces dernières années à un développement sans précédent des procédés d'optimisation mathématiques qui peuvent aujourd'hui prendre en compte des dizaines de paramètres de conception. En particulier, les méthodes *stochastiques* prennent peu à peu le pas sur les techniques *déterministes* conventionnelles : d'une part, elles permettent de localiser l'optimum d'une fonction dans l'espace des paramètres sans avoir recours aux dérivées de la fonction par rapport à ces paramètres ; d'autre part elles ne se laissent pas piéger par un optimum local et réussissent le plus souvent à déterminer l'optimum global de la fonction considérée. Parmi les techniques d'optimisation stochastiques, une attention particulière a été portée sur les méthodes dites *évolutionnistes* et qui ont connu depuis le début des années soixante une croissance exponentielle.

Notre travail s'inscrit dans le cadre du développement et de la caractérisation d'une classe de méthodes évolutionnistes qualifiées de méthodes génétiques de nichage. L'association de ces techniques avec les outils de modélisation numérique doit permettre l'amélioration des performances des dispositifs électrotechniques.

Le chapitre I dresse un état de l'art des techniques d'optimisation mathématiques, tant en ce qui concerne les méthodes de résolution que les procédés réalisant la prise en compte les contraintes de conception. Nous insistons aussi sur les méthodes d'optimisation vectorielles qui permettent de traiter des problèmes à objectifs multiples.

Le chapitre II introduit les concepts fondamentaux de l'algorithme génétique standard et analyse les divers opérateurs qui conditionnent la recherche en direction de l'optimum global.

Le chapitre III est consacré à l'étude des méthodes de nichage. Ces techniques associées à l'algorithme génétique permettent de rechercher des solutions optimales multiples aussi bien locales que globales. Pour cette raison, elles seront aussi qualifiées de multimodales.

Le chapitre IV illustre l'intérêt des méthodes de nichage par deux problèmes tests d'électromagnétisme spécifiquement définis de façon à présenter un ensemble de solutions optimales, et ce dans le but de tester l'efficacité des diverses méthodes présentées au chapitre III.

Enfin, le chapitre V propose une méthode originale de conception de forme optimale d'électrodes. Le procédé développé consiste à identifier la forme de l'électrode à une ligne équipotentielle, respectant un gabarit géométrique donné, obtenue à partir d'un système de charges fictives. La conception d'une forme d'électrode se résume alors à la résolution d'un problème d'optimisation sous contraintes, où les paramètres sont les valeurs et les positions des charges.

CHAPITRE I

LES TECHNIQUES D'OPTIMISATION MATHÉMATIQUES

Depuis toujours, l'homme tente d'améliorer sa vie au quotidien ; chaque jour, il résout sans parfois s'en rendre compte des problèmes d'optimisation : il planifie ses occupations, s'efforce de trouver les prix les plus bas en cas d'achat, cherche à minimiser le temps de conduite et la consommation en essence lors d'un long trajet en voiture,... enfin, il attache souvent de l'importance au confort avant tout. Ce dernier point montre là, toute la *problématique* de l'optimisation car cet objectif est une notion très *subjective* et *floue*, donc difficile à exprimer. Pourtant, les mathématiciens ont cherché très tôt à formuler des problèmes de la vie quotidienne en terme mathématique pour les résoudre d'une manière plus rigoureuse. Les fondements mathématiques de l'optimisation sont liés à l'apparition des méthodes de calcul intégral et différentiel. Les bases formelles ont été établies au XVII^{ème} par *Newton* (1642-1726) et *Leibniz* (1646-1716). Elles ont ensuite connu un essor important avec les travaux de *Mc Laurin* (1698-1746), *Euler* (1707-1783), *Lagrange* (1736-1813) et *Cauchy* (1789-1857) qui ont débouché sur les premières méthodes analytiques. Durant ces dernières années, l'augmentation des performances des calculateurs a permis un développement sans précédent des techniques d'optimisation numériques. Celles-ci rencontrent aujourd'hui un succès grandissant auprès de l'ingénieur amené à concevoir des dispositifs de plus en plus performants, de meilleure qualité, et de plus grande rentabilité.

Ce chapitre présente l'état de l'art des techniques d'optimisation mathématiques. Après avoir introduit la démarche systématique à adopter en optimisation, nous aborderons brièvement les méthodes de résolution en spécifiant les caractéristiques de chacune d'elles. Nous nous intéresserons par la suite aux procédés permettant de prendre en compte les contraintes de conception ainsi qu'aux méthodes d'optimisation à objectifs multiples.

I.1. DEMARCHE SYSTEMATIQUE EN OPTIMISATION

L'optimisation est souvent résumée aux techniques de résolution mathématiques auxquelles on impute par la suite les échecs que l'on rencontre. Or, comme pour la plupart des problèmes que l'ingénieur est amené à résoudre, l'optimisation doit faire l'objet d'une démarche systématique qui comporte quatre phases fondamentales résumées à la Figure 1.

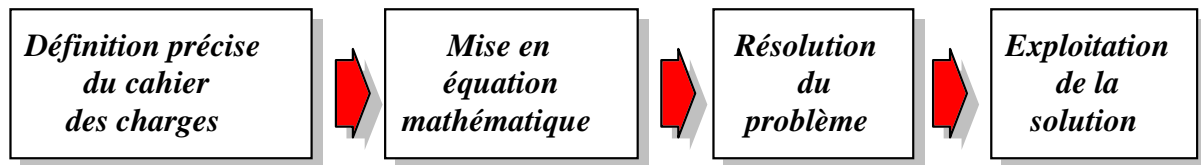


Figure 1. Démarche systématique en optimisation

I.1.1. Définition précise du cahier des charges

Cette phase impose une caractérisation rigoureuse du dispositif à optimiser. Elle consiste à définir de façon précise :

1. le comportement et les performances souhaités (*objectifs*)
2. les paramètres de conception qui permettront d'influencer les performances (*variables de conception* aussi appelées variables *objets* ou variables de *projet*)
3. les éventuelles *contraintes* liées à la fabrication ou à l'utilisation du dispositif.

Cette phase est fondamentale dans le processus d'optimisation parce qu'elle conditionne les étapes suivantes. Elle n'est pas facile à aborder car le *choix* des variables objets n'est jamais *unique*.

Il est généralement recommandé de définir un problème le plus simplement possible en choisissant un *minimum* de *paramètres*. Un problème bien posé et un choix judicieux des paramètres facilitent grandement sa résolution.

I.1.2. Mise en équation mathématique

Cette phase consiste en la traduction du problème physique en un problème mathématique équivalent. C'est sans doute l'étape la plus délicate du processus d'optimisation car là aussi, la formulation mathématique d'un problème n'est jamais *unique*, en particulier les définitions du critère d'optimalité et de la fonction caractérisant les performances du système.

I.1.3. Résolution du problème d'optimisation

La recherche de l'optimum d'un problème est réalisée à l'aide de méthodes d'optimisation mathématiques. Nous aborderons ce point plus en détail dans les paragraphes suivants.

I.1.4. Exploitation de la solution

Une fois le problème résolu, il est impératif d'évaluer la qualité de la solution et en cas d'échec de s'interroger sur les choix adoptés lors des différentes phases. Comme nous l'avons signalé plus haut, on impute trop souvent l'échec de l'optimisation à la méthode de recherche employée pour la localisation de l'optimum. Il faut aussi revenir sur les phases 1 et 2 car la résolution d'un problème mal formulé (tant au niveau du cahier des charges que du problème mathématique lui-même) ne peut conduire qu'à une solution erronée, quelle que soit la technique de résolution employée.

I.2. FORMULATION D'UN PROBLEME D'OPTIMISATION

Un problème d'optimisation de dimension n peut être formulé de façon générale suivant l'équation (1).

$$(P) \left\{ \begin{array}{ll} \min_{\mathbf{x} \in I} f_{obj}(\mathbf{x}) & \\ g_j(\mathbf{x}) \leq 0 & 1 \leq j \leq m \\ h_k(\mathbf{x}) = 0 & 1 \leq k \leq l \\ x_{i \min} \leq x_i \leq x_{i \max} & 1 \leq i \leq n \end{array} \right. \quad (1)$$

où \mathbf{x} est un vecteur à n composantes représentant les variables *objets* du problème.

I est l'espace des paramètres (ou espace de *recherche*) borné par les contraintes de domaine

f_{obj} est le critère à minimiser souvent appelé *fonction objectif*.¹

$x_{i \min}$ et $x_{i \max}$ désignent les *contraintes de domaine*.

$g_j(\mathbf{x})$ et $h_k(\mathbf{x})$ sont respectivement des *contraintes d'inégalité* et d'*égalité*.

Par la suite, nous ne considérerons que des contraintes d'inégalité. En effet, nous remarquons qu'une contrainte d'égalité $h_k(\mathbf{x}) = 0$ est équivalente à deux contraintes d'inégalité $h_k(\mathbf{x}) \leq 0$ et $-h_k(\mathbf{x}) \leq 0$.

L'ensemble des régions de l'espace de recherche où les contraintes de conception sont vérifiées est dénommée espace *réalisable* ou domaine *admissible*. Inversement, l'espace *irréalisable* (non réalisable ou *non admissible*) ou le domaine *interdit*, désigne l'ensemble des zones de l'espace où les contraintes sont *violées*.

¹ On qualifie aussi le critère d'optimisation de fonction de *forme* (*fitness*), de fonction *coût* en cas de minimisation et de fonction *profit* en cas de maximisation.

I.3. RESOLUTION D'UN PROBLEME NON CONTRAINT

I.3.1. Difficulté d'un problème non contraint

La difficulté d'un problème non contraint est liée à la nature de la fonction objectif. Celle-ci peut être *unidimensionnelle* ou *multidimensionnelle*, *convexe* ou *concave*, *linéaire* ou non linéaire, *continue* ou *discontinue*, *différentiable* ou non différentiable,... Elle dépend par ailleurs du nombre de *modes* que la fonction peut présenter.

Un point \mathbf{x}^* de l'espace est un minimum local si $\forall \mathbf{x} \in I$ tel que $d(\mathbf{x}, \mathbf{x}^*) \leq \varepsilon$,

$$f_{obj}(\mathbf{x}^*) \leq f_{obj}(\mathbf{x}) \quad (2)$$

où $d(\mathbf{x}, \mathbf{x}^*)$ désigne la distance entre le point \mathbf{x} et le point \mathbf{x}^* . Cette relation signifie que dans le voisinage de \mathbf{x}^* défini par une taille ε , il n'existe aucun point pour lequel $f_{obj}(\mathbf{x})$ est strictement inférieure à $f_{obj}(\mathbf{x}^*)$. S'il y a égalité entre les deux termes de l'équation (2) pour plusieurs points \mathbf{x} du voisinage de \mathbf{x}^* , le minimum est un minimum local *faible* (il révèle la présence d'un *plateau*). C'est un minimum local *fort* dans le cas contraire (il représente alors un *mode* de la fonction).

Une condition nécessaire pour que \mathbf{x}^* soit minimum local s'écrit :

$$\begin{cases} \nabla f_{obj}(\mathbf{x}^*) = 0 \\ \mathbf{F}_{xx}(\mathbf{x}^*) \text{ non négative} \end{cases} \quad (3)$$

où ∇f_{obj} est le gradient de la fonction objectif et $\mathbf{F}_{xx} = \nabla^2 f_{obj}$ la matrice des ses dérivées partielles (qualifiée aussi de *Hessien*). La condition de nullité du gradient est dite *condition de stationnarité*.

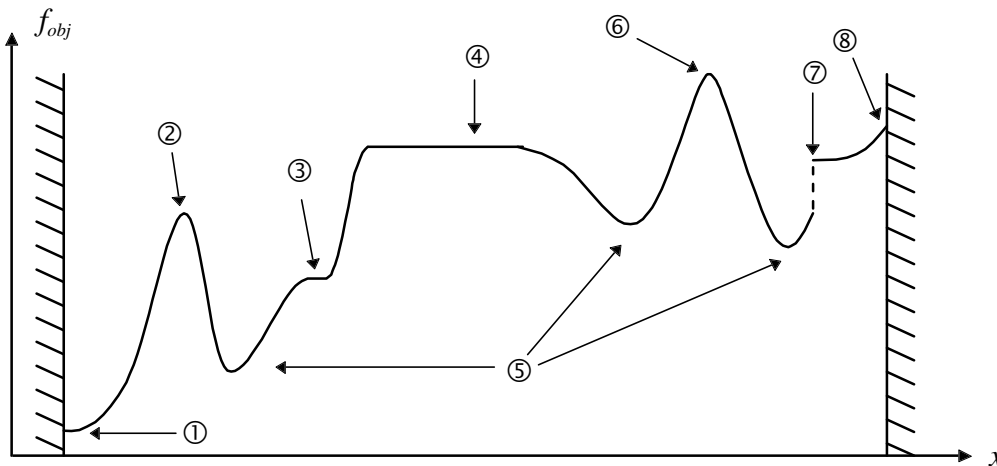
Une condition suffisante pour que \mathbf{x}^* soit minimum local s'écrit :

$$\begin{cases} \nabla f_{obj}(\mathbf{x}^*) = 0 \\ \mathbf{F}_{xx}(\mathbf{x}^*) \text{ positive} \end{cases} \quad (4)$$

Il est à noter que les conditions (3) et (4) ne s'appliquent pas à des points situés sur les frontières de l'espace de recherche et sont valables uniquement pour des fonctions différentiables.

Lorsqu'une fonction admet plusieurs minima locaux, elle est dite *multimodale* (elle est *unimodale* dans le cas contraire). Le plus petit minimum local d'une fonction multimodale, ou *minimum minimorum*, est aussi appelé *minimum global*. Il doit satisfaire en tout point de l'espace de recherche la relation (2).

La Figure 2 présente, à titre d'exemple, une distribution possible des optima d'une fonction objectif unidimensionnelle et multimodale, ainsi que certains points particuliers pouvant poser des difficultés aux méthodes de résolution.



- | | |
|---------------------------------------|-----------------------------------|
| ① minimum global sur une frontière | ⑤ minima locaux |
| ② maximum local | ⑥ maximum global |
| ③ point d'inflexion | ⑦ saut ou discontinuité |
| ④ plateau (ensemble de maxima locaux) | ⑧ maximum local sur une frontière |

Figure 2. Points singuliers d'une fonction unidimensionnelle multimodale

I.3.2. Caractéristiques des méthodes de résolution

I.3.2.1. Modèle de la boîte noire

Une méthode de résolution peut être caractérisée selon le modèle de la *boîte noire*, illustré à la Figure 3. Elle doit permettre d'obtenir l'optimum d'un problème à partir du calcul de la fonction objectif en certains points. Les mécanismes d'exploration de l'espace, spécifiques à la méthode de résolution, sont conditionnés par des *paramètres de contrôle* (nombre de points de calcul, directions de recherche,...) et des *conditions initiales* (variables objets initiales, valeur initiale des paramètres de contrôle,...).

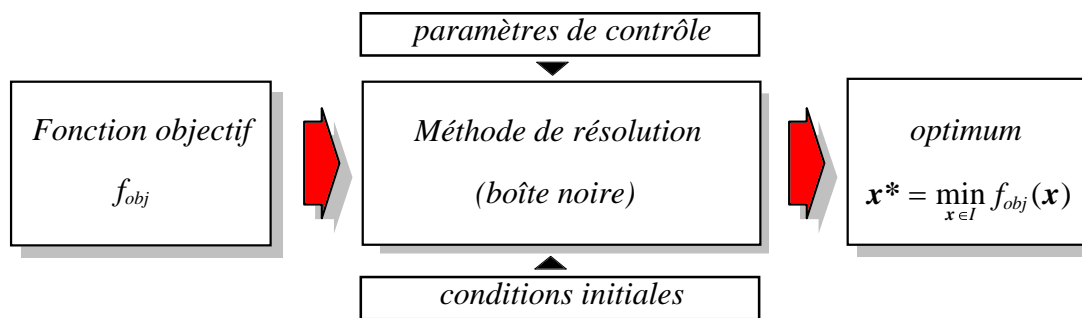


Figure 3. Caractéristiques d'une méthode de résolution - modèle de la boîte noire

I.3.2.2. Efficacité d'une méthode de résolution

L'efficacité d'une méthode de résolution est liée à sa *sensibilité* et à sa *robustesse* par rapport aux paramètres de contrôle et aux conditions initiales.

Lorsque les paramètres de contrôle (ou les conditions initiales) doivent prendre une valeur bien précise pour que la méthode de résolution converge vers l'optimum d'une fonction donnée, la méthode est *sensible* par rapport à ces mêmes paramètres (respectivement par rapport aux conditions initiales).

Une méthode de résolution est *robuste* si pour une même valeur des paramètres de contrôle et des conditions initiales, elle est capable de trouver l'optimum de fonctions très distinctes.

Une méthode idéale devrait être totalement insensible aux conditions initiales et aux paramètres de contrôle, et converger vers l'optimum quelle que soit la fonction testée.

I.3.2.3. Opérateurs de recherche fondamentaux

La recherche de l'optimum d'une fonction est généralement réalisée à l'aide de deux opérateurs fondamentaux : l'*exploration* et l'*exploitation*.

- L'*exploration* a pour objectif de localiser un optimum de la fonction, c'est à dire de rechercher une vallée (en cas de minimisation) ou un sommet (en cas de maximisation)
- L'*exploitation* renforce l'exploration en permettant la descente au creux de la vallée ou l'ascension du sommet.

Le succès et l'efficacité d'une technique de résolution dépendent la plupart du temps d'un *compromis* entre l'exploration et l'exploitation. Certaines méthodes toutefois n'utilisent qu'un seul de ces opérateurs pour parvenir à l'optimum.

I.3.2.4. Spécificités liées au mode de recherche de l'optimum

Lorsque l'évolution de la méthode de résolution est prévisible et ne laisse aucune place au hasard, celle-ci est qualifiée de *déterministe*. En revanche, les méthodes dites *stochastiques* s'appuient sur des mécanismes de transitions probabilistes qui peuvent conduire à des résultats différents pour des conditions initiales et des paramètres de contrôle identiques. Les méthodes déterministes utilisent essentiellement des opérateurs d'exploitation alors que les méthodes stochastiques reposent principalement sur l'exploration.

Les méthodes déterministes sont presque toujours des méthodes *locales*, c'est à dire qu'elles convergent vers un optimum dépendant uniquement du point de départ, qu'il soit local ou global. A l'opposée, les techniques stochastiques ou *stochasto-déterministes* sont plutôt des méthodes *globales* qui permettent de localiser l'optimum global des fonctions.

Enfin, les méthodes de résolution sont qualifiées de *multimodales* si elles sont capables de trouver l'ensemble des optima d'une fonction.

I.3.2.5. Ordre d'une méthode de résolution

Les méthodes de résolution peuvent être classées à partir de leur *ordre* selon qu'elles nécessitent ou non le calcul des dérivées de la fonction objectif par rapport aux paramètres. Une méthode est dite d'ordre 0 (ou *directe*) si elle utilise uniquement la connaissance de la fonction elle même. Elle est d'ordre 1 si elle requiert le calcul des dérivées premières et d'ordre 2 s'il lui faut aussi accéder aux dérivées secondes.

I.3.3. Etat de l'art des méthodes de résolution

Dans cette partie, nous exposons brièvement les techniques de résolution les plus populaires. Le lecteur est invité à se reporter aux diverses références citées pour plus de détails.

I.3.3.1. Méthodes d'optimisation unidimensionnelles

Les méthodes d'optimisation unidimensionnelles peuvent être classées en deux groupes :

- les méthodes de *subdivision d'intervalles*

Elles consistent à se rapprocher de l'optimum par réductions successives de l'intervalle de recherche complet, en tenant compte de la valeur de la fonction aux extrémités de chaque sous-intervalle exploré séquentiellement [Rao 96][Schwefel 95].

La méthode la plus simple de *dichotomie* classique utilise un facteur de réduction constant de 1/2. Dans la méthode de *Fibonacci*, le facteur de réduction est adapté au cours de la recherche : il est défini comme le rapport de deux termes consécutifs de la suite de Fibonacci (ce facteur tend vers le nombre d'or égal à $2 / (1 + \sqrt{5}) \approx 0.618$). L'intérêt de la méthode réside dans le fait qu'à chaque itération, la valeur de la fonction est évaluée en un seul point pour la détermination de l'intervalle suivant. L'autre extrémité du nouvel intervalle de recherche est déduit des points testés lors des itérations précédentes. La *section dorée* ou méthode du *nombre d'or* est très similaire à la méthode de Fibonacci. Elle utilise un facteur de réduction égal au nombre d'or.

Les techniques de subdivision d'intervalles peuvent s'appliquer à des fonctions éventuellement discontinues. Elles nécessitent que ces dernières soient unimodales. Dans le cas contraire, elles ne garantissent pas la convergence vers l'optimum global.

Par ailleurs, l'optimum ne peut être localisé "exactement". Il est seulement possible d'obtenir sa valeur dans un certain intervalle d'incertitude (l'intervalle obtenu à la dernière itération à la suite des subdivisions successives). La précision de la recherche est fonction de l'intervalle de départ et du nombre d'itérations. Elle peut être améliorée en augmentant le nombre de subdivisions.

- les méthodes d'*interpolations*

Les méthodes *Lagrangiennes* s'efforcent de se rapprocher de l'optimum par réductions successives de l'intervalle de recherche, en interpolant la fonction par un polynôme d'ordre n . La technique d'interpolation quadratique de la fonction par un polynôme d'ordre 2 est la plus populaire [Schwefel 95].

D'autres stratégies comme la méthode de *Regula-falsi* et la méthode de *Newton-Raphson* recherchent séquentiellement le passage par zéro de la dérivée de la fonction objectif à partir de la connaissance du gradient (et parfois de la dérivé seconde) [Rao 96][Schwefel 95]. La convergence vers l'optimum est très rapide mais exclusivement locale. La réussite dépend fortement du choix du point de départ (celui-ci doit être proche de l'optimum si la fonction est multimodale).

I.3.3.2. Méthodes d'optimisation multidimensionnelles

Une classification des méthodes d'optimisation multidimensionnelles est donnée à la Figure 4.

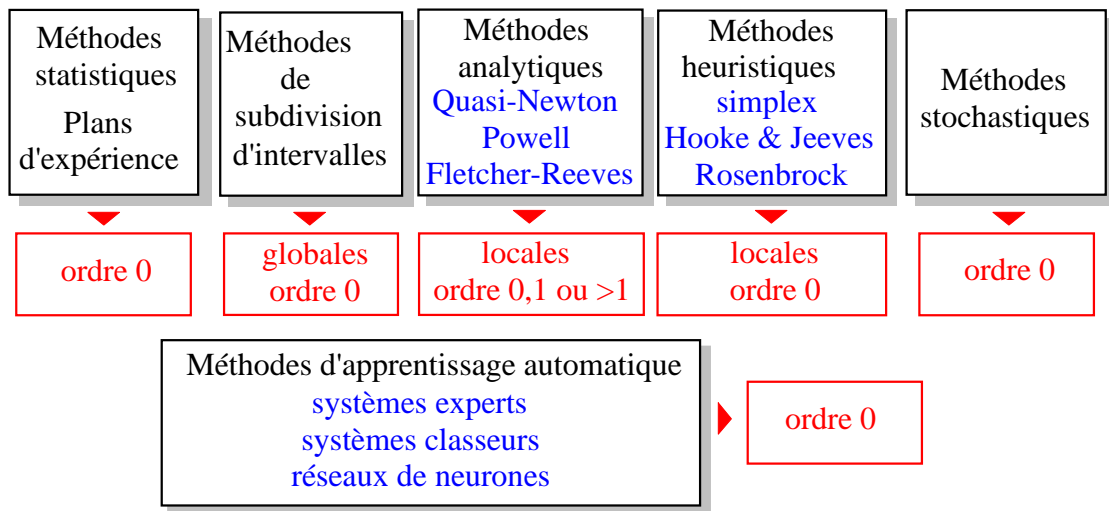


Figure 4. Méthodes d'optimisation multidimensionnelles

- Les *méthodes statistiques* et les *plans d'expérience* sont habituellement employés lorsque le nombre d'évaluations de la fonction objectif est très restreint [Sado 92][Rao 96]. Elles restent toutefois limitées et ne permettent pas de converger précisément vers l'optimum.
- Les *méthodes de subdivision d'intervalles* sont basées sur le même principe que les techniques unidimensionnelles présentées précédemment [Messine 98]. Les plus évoluées sont capables de localiser tous les optima de la fonction avec un certain degré d'incertitude lié à la taille des subdivisions dans l'espace. Elles sont généralement coûteuses en nombre d'évaluations de l'objectif lorsque le nombre de paramètres est élevé.
- Pour déterminer un optimum, les *méthodes analytiques* se basent sur la connaissance d'une direction de recherche donnée par le gradient de la fonction objectif par rapport aux paramètres. Par exemple, la méthode de *Cauchy* ou méthode de la *plus grande pente* (*steepest descent/ascent*) et la méthode du *gradient conjugué* (aussi appelé méthode de *Fletcher Reeves*), sont d'ordre 1. D'autres techniques de gradient nécessitent le calcul du Hessien comme les méthodes de type *Quasi-Newton*. En revanche, les stratégies de *Powell* et *Stewart* reposent sur une estimation du gradient à partir de la valeur de l'objectif en certains points. Si ces méthodes sont intéressantes en raison de leur grande rapidité de convergence, elles ont plusieurs inconvénients :
 1. La fonction objectif doit être continue et dérivable
 2. Lorsque le gradient et le Hessien de la fonction ne sont connus analytiquement, leur détermination par des méthodes de différences finies est toujours délicate à cause de problèmes liés au choix du pas de variation et des erreurs d'arrondis.

- 3. Ces méthodes nécessitent la résolution de systèmes matriciels pouvant être mal conditionnés.
- 4. La convergence est exclusivement locale. Le point de départ doit donc être proche de l'optimum (La sensibilité par rapport aux conditions initiales est importante)
- Les méthodes *heuristiques* (parfois appelées méthodes *géométriques*) explorent l'espace par essais successifs en recherchant les directions les plus favorables. On emploie le plus souvent la stratégie de *Hooke et Jeeves*, la méthode de *Rosenbrock*, ou la méthode du *simplex* de *Nelder et Mead* [Rao 96][Schwefel 95]. Toutes ces techniques sont déterministes et locales mais elles sont beaucoup plus robustes que les méthodes analytiques classiques, en particulier si la fonction objectif est discontinue ou bruitée. Par contre, elles deviennent prohibitives lorsque le nombre de paramètres est élevé.
- Les méthodes *stochastiques* explorent des points de l'espace suivant des processus aléatoires qui permettent de progresser avec une probabilité suffisante vers l'optimum global. Nous aborderons ces méthodes de façon plus approfondie dans le paragraphe suivant.
- Enfin, les *méthodes d'apprentissage automatique* ne sont pas des techniques de résolution proprement dites, mais elles sont parfois couplées aux méthodes précédentes pour augmenter l'efficacité de la convergence [Ebner 98].

I.3.3.3. Méthodes d'optimisation stochastiques

Les méthodes stochastiques les plus employées sont illustrées à la Figure 5.

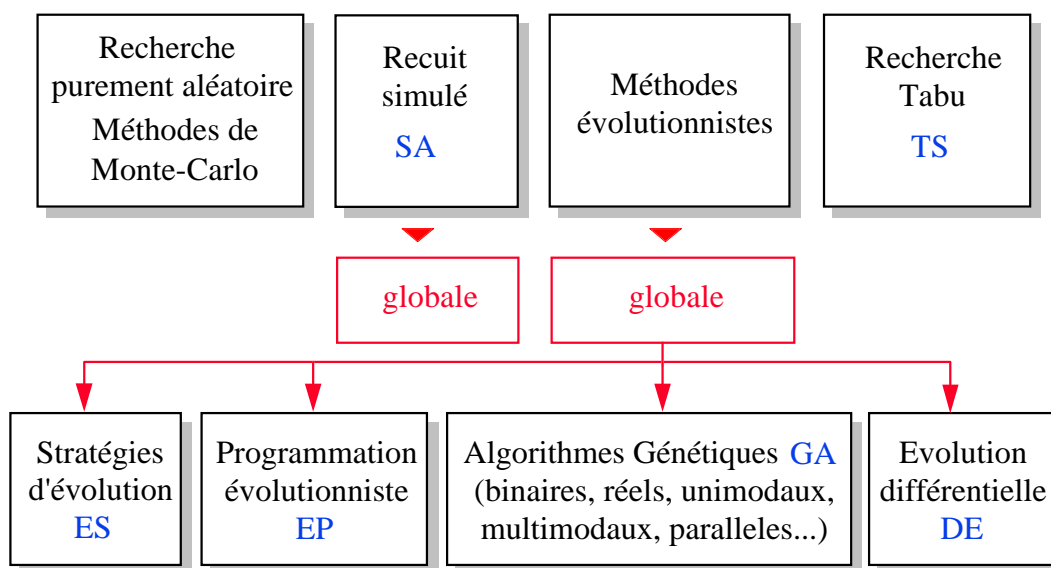


Figure 5. Méthodes d'optimisation stochastiques

Parmi les méthodes stochastiques, on distingue les techniques de recherche purement aléatoires (souvent regroupées dans la classe des méthodes de type *Monte-Carlo*) des stratégies possédant aussi des mécanismes déterministes (méthodes *stochasto-déterministes*).

- Les méthodes d'exploration purement aléatoire (*random walk methods*) peuvent être utilisées, en première approche, pour avoir des renseignements utiles sur la forme de la fonction. Elles permettent par exemple de choisir de façon plus appropriée le point de départ et les conditions initiales d'un algorithme de recherche local. Toutefois, même dans ce cas d'utilisation, elles ne garantissent aucunement la localisation de l'optimum global.
- La recherche Tabu (*Tabu search*) est une méta-heuristique qui a été développée dans le même esprit pour éviter aux méthodes d'optimisation locales d'être piégées par des optima locaux. Elle est néanmoins utilisée la plupart du temps avec des techniques d'optimisation stochastiques globales telles que le recuit simulé [Battiti 94][Vasconcelos 94] (voir ci-après).

Les méthodes d'optimisation stochastiques globales reposent sur des mécanismes stochasto-déterministes qui leur permettent d'éviter les pièges constitués par les optima locaux. Pour des fonctions objectifs multimodales, la probabilité de convergence de ces méthodes vers l'optimum global est plus élevée qu'avec les techniques déterministes conventionnelles.

- Le *recuit simulé* a été introduit par *Kirkpatrick* pour résoudre des problèmes d'analyse combinatoire [Kirkpatrick 83]. Les concepts fondamentaux de cette technique sont tirés de l'analogie entre l'optimisation et la thermodynamique du recuit d'un solide établie par *Metropolis* [Metropolis 53]. Cette analogie réside dans le fait que le paramètre de contrôle de l'algorithme peut être assimilé à une température. Le principe du recuit simulé est le suivant : pour chaque mouvement aléatoire du point de recherche \mathbf{x}_0 , la nouvelle solution \mathbf{x} est acceptée si elle conduit à une amélioration de l'objectif. Dans le cas contraire elle est acceptée avec une probabilité :

$$w = \exp\left(-\frac{f_{obj}(\mathbf{x}) - f_{obj}(\mathbf{x}_0)}{T}\right) \quad (5)$$

où T est le paramètre de contrôle de l'algorithme souvent appelé "température" en raison de l'analogie avec le processus physique du recuit.² En pratique, la température est adaptée au

² Dans le processus physique du recuit, la probabilité de se trouver dans un état d'énergie E à la température T , est donnée par la loi de Boltzmann $w = 1/z(T) \cdot \exp(-E/kT)$ où $z(T)$ est une fonction de partition [Aarts 90] et k est la constante de Boltzmann.

cours de la recherche et réduite séquentiellement à chaque itération $T_{k+1} = \alpha T_k$ en fonction d'un facteur de réduction α compris entre 0 et 1 (généralement $\alpha \in [0.8, 0.99]$). Pour chaque valeur de température, on exécute le processus de recuit L_k fois jusqu'à équilibre. L'algorithme est arrêté lorsqu'il n'y a plus d'amélioration sensible du critère. L'efficacité du recuit simulé dépend du procédé de perturbation des paramètres et du schéma de refroidissement. Celui-ci est caractérisé par l'ensemble des paramètres de contrôle et des conditions initiales qui gouvernent le recuit :

1. La valeur initiale de la température T_0
2. La valeur du facteur de réduction de la température
3. Le nombre d'exécutions du processus de recuit L_k pour chaque valeur de température
4. Le critère d'arrêt de l'algorithme

La rapidité et le succès de la convergence dépendent fortement de la valeur de ces paramètres. Un choix mal approprié peut entraîner la convergence de l'algorithme vers un optimum local. Des variantes du recuit simulé classique ont été proposées pour augmenter la performance de la méthode et diminuer la sensibilité de l'algorithme par rapport aux conditions initiales et aux paramètres de contrôle [Yiao 91][Ingber 92].

- Enfin, les méthodes *évolutionnistes* font partie de la dernière grande classe de méthodes d'optimisation stochastiques. Elles reposent sur une analogie avec la théorie de l'évolution naturelle des espèces de Darwin (1809-1882) selon laquelle, les individus les mieux adaptés à leur environnement survivent et peuvent se reproduire pour donner des enfants encore mieux adaptés de génération en génération [Darwin 59]. Contrairement aux techniques d'optimisation précédentes qui explorent l'espace à partir d'un point unique, les méthodes évolutionnistes partent d'un ensemble de configurations (ou population d'individus) et le font évoluer à partir d'opérateurs à transitions aléatoires, selon le principe de la Figure 6. L'*exploitation* est réalisée par l'opérateur de *sélection* qui retient les individus les plus performants. L'*exploration* est effectuée à partir de l'opérateur *d'évolution* qui oriente la recherche vers de nouveaux points de l'espace.

Les algorithmes évolutionnistes remontent à l'introduction des algorithmes génétiques (GA) par Holland [Holland 75]. Parallèlement, Reshenberg et Schwefel d'un côté et Fogel de l'autre mettaient au point deux méthodes assez similaires, les *stratégies d'évolution* (ES) et la *programmation évolutionniste* (EP) [Schwefel 95][Bäck 96]. Depuis, ces techniques ont connu un essor spectaculaire, en particulier ces dernières années avec la prolifération des travaux et des congrès internationaux (FOGA, PPSN, ICEC, ICGA...)

dans ce domaine. Une nouvelle méthode prometteuse, qualifiée d'évolution différentielle (DE), est apparue récemment [Storn 96].

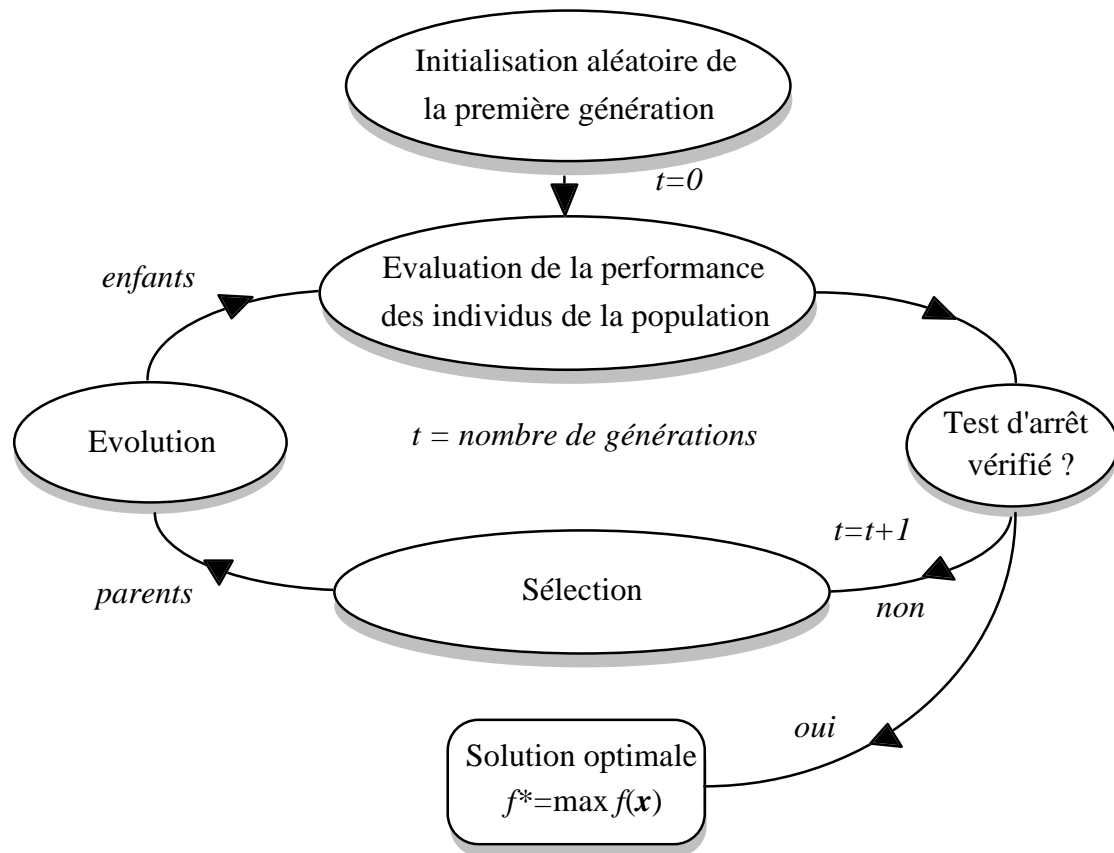


Figure 6. Principe d'une méthode évolutionniste

Les différences entre ces méthodes sont liées à la représentation des individus et aux modes d'évolution de la population. Les algorithmes génétiques utilisent un codage des paramètres de la fonction à optimiser alors que les autres techniques se servent directement de la valeur des variables objets. Chacune des méthodes enfin est caractérisée par un opérateur d'évolution particulier (les opérateurs de *sélection* ou de *compétition* entre individus restant assez similaires quelles que soient les techniques employées). Les algorithmes génétiques et l'évolution différentielle reposent essentiellement sur un mécanisme de croisement (*crossing-over* que l'on réduit au terme *crossover*) qui permet la génération de nouvelles configurations par recombinaison de solutions existantes. Les stratégies d'évolution et la programmation évolutionniste sont pour leur part basées principalement sur un procédé de *mutation* qui régit l'évolution de la population par perturbations successives de chaque solution.

Les méthodes évolutionnistes s'affirment peu à peu comme les techniques de résolution les plus robustes. En plus de leur caractère global, elles s'adaptent très bien au modèle de la boîte noire : elles peuvent être appliquées à des problèmes très divers en gardant une

relative indépendance par rapport au processus à optimiser. Ce n'est pas le cas par exemple des méthodes de type gradient qui nécessitent, la plupart du temps, d'être insérées dans le processus d'optimisation pour accéder aux dérivées de l'objectif par rapport aux paramètres.

I.4. PRISE EN COMPTE DES CONTRAINTES DE CONCEPTION

I.4.1. Méthodes de transformation

I.4.1.1. Méthodes de pénalités intérieure et extérieure

Les méthodes de transformation consistent à introduire les contraintes de conception dans la fonction que l'on cherche à optimiser. Ainsi le problème contraint (6),

$$\begin{cases} \min_{\mathbf{x} \in I} f_{obj}(\mathbf{x}) \\ g_j(\mathbf{x}) \leq 0 \quad j = 1 \dots m \end{cases} \quad (6)$$

peut être transformé en un problème équivalent non contraint, formulé par (7),

$$\min_{\mathbf{x} \in I} \phi_{obj}(\mathbf{x}, r) = f_{obj}(\mathbf{x}) + r \sum_{j=1}^m P_j(g_j(\mathbf{x})) \quad (7)$$

où r est un *paramètre* (ou *coefficient*) de *pénalité*. La *fonction de pénalité* P_j est construite de manière à favoriser la sélection des points admissibles par rapport aux points non admissibles, par la méthode d'optimisation sans contrainte qui sera utilisée. Les méthodes de pénalité sont divisées en deux classes selon la nature de la fonction de pénalité :

Les méthodes de pénalité *intérieures* utilisent une fonction de pénalité construite de telle façon que la possibilité de réalisation soit garantie dans l'ensemble du processus d'optimisation. De telles méthodes sont également appelées méthodes à barrière car la fonction de pénalité forme une barrière infinie le long de la frontière du domaine réalisable. Les fonctions de pénalités intérieures les plus employées dans la littérature sont [Rao 96] :

la fonction inverse :

$$P_j(g_j(\mathbf{x})) = \frac{-1}{g_j(\mathbf{x})} \quad (8)$$

la fonction logarithmique :

$$P_j(g_j(\mathbf{x})) = -\log(-g_j(\mathbf{x})) \quad (9)$$

L'utilisation des méthodes à barrière impose que les paramètres soient situés dans l'espace réalisable. La recherche de points admissibles peut être effectuée à l'aide d'une

méthode stochastique par échantillonnage aléatoire mais celle-ci peut conduire à des temps de calcul prohibitifs lorsque les contraintes sont sévères et que l'espace réalisable est très réduit par rapport à l'espace irréalisable. Ces méthodes à barrière, outre le fait qu'elles requièrent des jeux de paramètres réalisables, sont relativement difficiles à mettre en œuvre.

Les méthodes de pénalités extérieures ne présentent pas cet inconvénient car elles sont valables aussi bien à l'intérieur qu'à l'extérieur de l'espace réalisable. Elles augmentent la pénalisation des solutions à mesure que l'on s'éloigne de l'espace réalisable. Les fonctions de pénalité extérieures les plus populaires sont les suivantes [Rao 96] :

$$P_j(g_j(\mathbf{x})) = \max[0, g_j(\mathbf{x})] \quad (10)$$

$$P_j(g_j(\mathbf{x})) = \max[0, g_j(\mathbf{x})]^2 \quad (11)$$

Avec des méthodes de pénalités extérieures, la continuité de la fonction de transformation est assurée si on est capable de calculer la fonction objectif dans l'espace irréalisable. Pour des problèmes réels, cela n'est malheureusement pas toujours le cas. Nous reviendrons sur ce point au paragraphe I.4.1.4.

I.4.1.2. Techniques de minimisation séquentielles sans contraintes

Lorsque l'on utilise des méthodes de résolution analytiques, les problèmes de conditionnement de la fonction $\phi_{obj}(\mathbf{x}, r)$ imposent d'avoir recours à une méthode de minimisation séquentielle sans contraintes (*Sequential Unconstrained Minimization Techniques*). Dans ce type de méthodes, le problème d'optimisation sous contraintes est transformé en une série de problèmes non contraints pour lesquels le paramètre de pénalité est adapté à chaque itération. La méthode des multiplicateurs de Lagrange ou méthode du Lagrangien augmenté est la plus connue de ces techniques. Nous ne la détaillerons pas ici, mais nous invitons le lecteur à se reporter aux références suivantes pour plus d'informations [Vasconcelos 94][Rao 96]. Nous signalons toutefois que la méthode du Lagrangien augmenté requiert généralement le calcul des dérivées de la fonction objectif et des contraintes par rapport aux paramètres pour fixer la valeur des multiplicateurs de Lagrange. Ceci peut s'avérer problématique, pour les raisons que nous avons évoquées précédemment.

L'utilisation de méthodes d'optimisation stochastiques, moins sensibles aux discontinuités des fonctions et aux valeurs du paramètre de pénalité, évitent d'avoir recours aux techniques de minimisation séquentielles sans contraintes. Nous pouvons alors employer directement une méthode de pénalité extérieure avec un coefficient de pénalité constant au cours de la recherche.

I.4.1.3. Pénalisation radicale des solutions

La méthode de pénalisation radicale (ou *death penalty method*) est un processus de pénalisation très populaire dans le domaine de l'optimisation évolutionniste [Michalewicz 96]. Il s'agit d'écarter les solutions non réalisables en attribuant à la fonction de transformation une valeur très élevée en cas de minimisation ($\phi_{obj} \rightarrow +\infty$) ou une valeur nulle en cas de maximisation ($\phi_{obj} = 0$). La probabilité de survie de ces solutions, déterminée par les mécanismes de sélection, est alors quasi-nulle. Cette méthode est séduisante en raison de sa grande simplicité. Elle peut être appliquée avec succès lorsque l'espace de recherche est convexe et qu'il constitue une part importante de l'espace complet. Dans le cas contraire, cette approche a de sérieuses limitations, les solutions situées dans l'espace irréalisable ne pouvant être améliorées en raison de l'absence de directions données par la méthode de pénalisation.

I.4.1.4. Continuité de la fonction de transformation

Pour des problèmes physiques réels, il n'est pas toujours possible de calculer la fonction objectif lorsque certaines contraintes sont violées. On comprendra aisément que lorsque le dispositif n'est pas réalisable au sens physique du terme, il devient assez difficile de caractériser ses performances! De ce fait, il n'est plus possible, avec une méthode de pénalité extérieure, d'assurer la continuité de la fonction de transformation au niveau de la frontière entre l'espace réalisable et l'espace irréalisable. Celle-ci se réduit au terme de pénalisation dans l'espace irréalisable comme le montre l'équation (12)

$$\phi_{obj}(\mathbf{x}, r) = \begin{cases} f_{obj}(\mathbf{x}) & \text{si } \mathbf{x} \text{ réalisable} \\ r \sum_{j=1}^m P_j(g_j(\mathbf{x})) & \text{sinon} \end{cases} \quad (12)$$

L'utilisation de méthodes de résolution appropriées à l'optimisation de fonctions discontinues est impérative. Par ailleurs, il faut s'assurer que le terme de pénalité reste toujours assez important pour éviter la convergence vers un point de l'espace irréalisable. On pourra par exemple retenir la fonction de transformation suivante :

$$\phi_{obj}(\mathbf{x}, r) = \begin{cases} f_{obj}(\mathbf{x}) & \text{si } \mathbf{x} \text{ réalisable} \\ r_0 + r \sum_{j=1}^m P_j(g_j(\mathbf{x})) & \text{sinon} \end{cases} \quad (13)$$

où $r_0 \geq \max_{\mathbf{x} \in I} f_{obj}(\mathbf{x})$ assure une pénalisation suffisante pour que la fonction $\phi_{obj}(\mathbf{x}, r)$ ait toujours des valeurs inférieures dans le domaine réalisable.

I.4.1.5. Mise à l'échelle des contraintes

Avec les méthodes de pénalité intérieure et extérieure, il est préférable que les contraintes aient des "dimensions" comparables. Dans le cas contraire, une ou plusieurs contraintes pourraient dominer et masquer les autres lors du processus d'optimisation. Pour des problèmes d'optimisation réels, nous nous trouvons souvent dans ce cas de figure. Il est alors impératif d'effectuer une normalisation des contraintes pour les rendre sensiblement comparables. Nous pouvons par exemple chercher à exprimer chaque contrainte sous la forme :

$$g_j(\mathbf{x}) = \delta_j(\mathbf{x}) - \delta_{j\max} \leq 0 \quad (14)$$

et les normaliser par le terme $\delta_{j\max}$. Les nouvelles contraintes du problème s'écrivent :

$$g'_j(\mathbf{x}) = \frac{g_j(\mathbf{x})}{\delta_{j\max}} = \frac{\delta_j(\mathbf{x})}{\delta_{j\max}} - 1 \leq 0 \quad (15)$$

I.4.2. Optimisation séquentielle des contraintes

La procédure consiste à utiliser à tour de rôle chaque contrainte comme fonction objectif d'un problème d'optimisation avec contraintes (les autres contraintes). Cette approche a été mise en œuvre avec succès dans le domaine évolutionniste [Schoenauer 93] [Michalewicz 96]. Elle semble particulièrement bien adaptée aux méthodes d'évolution qui progressent à partir d'une population de points. En pratique, on optimise la première contrainte indépendamment des autres jusqu'à ce qu'un certain pourcentage de la population satisfasse celle-ci. Puis, on choisit la deuxième contrainte comme fonction objectif du problème et on recommence l'opération en tenant compte cette fois de la première contrainte. L'opération est répétée ainsi de suite jusqu'à ce que toutes les contraintes soient vérifiées pour une partie de la population. Le critère d'optimisation est alors pris comme fonction objectif finale et le problème est résolu en considérant l'ensemble des contraintes. Le pouvoir de cette méthode réside dans la mémoire acquise par la population à chaque itération du processus. Cette mémoire d'apprentissage (*behavioral memory*) permet à l'algorithme d'évolution de garder des informations utiles au cours de l'optimisation séquentielle de chaque contrainte. Bien que cette méthode paraisse plus lourde et plus coûteuse en temps de calcul, elle est assez intéressante lorsque l'on est en présence de nombreuses contraintes conflictuelles ou difficiles à respecter. Par ailleurs, l'utilisation d'une méthode de pénalisation radicale (pour des solutions violant des contraintes optimisées auparavant) est rendue possible, puisque à chaque étape du processus, un pourcentage suffisant de la population satisfait les contraintes du problème.

L'optimisation séquentielle des contraintes peut également être réalisée à l'aide d'une méthode de pénalité extérieure. Au lieu de choisir un coefficient de pénalité identique pour toutes les contraintes, nous pouvons attribuer à chacune d'elles un poids particulier. Il suffit d'écrire la fonction de transformation sous la forme suivante :

$$\phi_{obj}(\mathbf{x}, r) = f_{obj}(\mathbf{x}) + \sum_{j=1}^m r_j P_j(g_j(\mathbf{x})) \quad (16)$$

où la valeur des coefficients de pénalité r_j est adaptée suivant la nature des contraintes associées. On affectera par exemple un poids plus important aux contraintes les plus sévères à vérifier en premier lieu.

I.4.3. Transformation d'un problème contraint en un problème équivalent non contraint à objectifs multiples

Il est parfois difficile de discerner une contrainte d'un objectif. En général, une contrainte doit absolument être vérifiée pour que le système à optimiser soit réalisable physiquement alors qu'un objectif est un état de fonctionnement souhaité pour le système. Cette notion n'étant pas toujours claire, il est quelquefois possible de transformer des contraintes en objectifs. Dans ce cas, le problème contraint (6) peut être exprimé sous la forme d'un problème équivalent non contraint à objectifs multiples, où les objectifs sont inclus dans un vecteur à $m+1$ composantes :

$$\mathbf{F} = \begin{bmatrix} f_{obj}(\mathbf{x}) \\ f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{bmatrix} \quad \text{avec } f_j(\mathbf{x}) = f(g_j(\mathbf{x})) \text{ pour } j = 1..m \quad (17)$$

I.5. OPTIMISATION A OBJECTIFS MULTIPLES

L'optimisation à objectifs multiples ou optimisation *multi-critères* a pour objet la minimisation ou la maximisation simultanée de plusieurs fonctions, à partir des mêmes variables objets. Ces problèmes entrent dans la catégorie des problèmes d'optimisation vectoriels (*Vector Optimization Problems* ou VOP)

I.5.1. Formulation d'un problème d'optimisation vectoriel

Un problème d'optimisation vectoriel contraint peut être exprimé de la façon suivante :

$$\left\{ \begin{array}{l} \text{"min"}_{\mathbf{x} \in I} \mathbf{F} = \begin{bmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_l(\mathbf{x}) \end{bmatrix} \\ g_j(\mathbf{x}) \leq 0 \quad j = 1..m \end{array} \right. \quad (18)$$

où les fonctions $f_k(\mathbf{x})$ et $g_j(\mathbf{x})$ désignent respectivement les objectifs et les contraintes du problème. L'opérateur *min* doit permettre la minimisation simultanée des composantes du vecteur \mathbf{F} .

I.5.2. Optimum de Pareto

La difficulté principale d'un problème d'optimisation vectoriel est lié à la présence de conflits entre les divers objectifs. En effet, les solutions optimales pour un objectif donné ne correspondent généralement pas à celles des autres objectifs pris indépendamment. De ce fait, il n'existe, la plupart du temps, aucun point de l'espace des paramètres où toutes les fonctions objectifs sont optimales simultanément. Une solution du problème vectoriel représente un compromis entre les différents critères. Le meilleur compromis est défini de telle manière qu'il ne soit plus possible d'améliorer une composante du vecteur critère sans détériorer les autres. Dans ce cas la solution est dite *Pareto-optimale*.

Formellement, \mathbf{x}^* est un optimum de *Pareto* s'il n'y a aucune solution \mathbf{x} pour laquelle :

$$f_k(\mathbf{x}) < f_k(\mathbf{x}^*) \quad \forall k \in [1, l] \quad (19)$$

La solution d'un problème d'optimisation vectoriel est représentée par un ensemble d'optima de *Pareto* appelé frontière de *Pareto*. La Figure 7 illustre une interprétation géométrique de la frontière de Pareto dans le cas d'un problème d'optimisation vectoriel à deux objectifs.

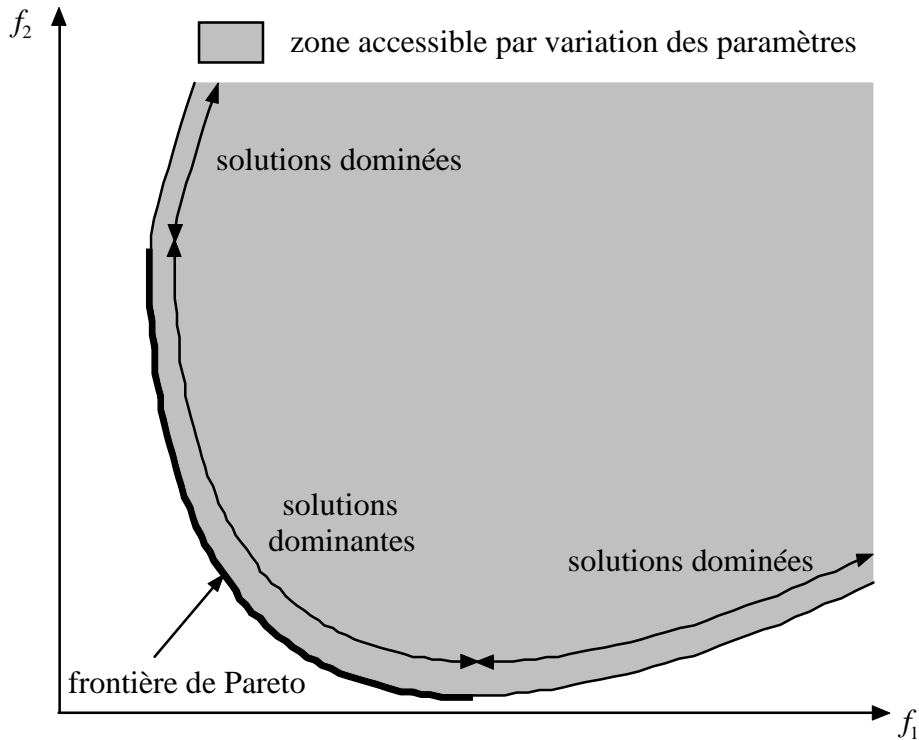


Figure 7. Frontière de Pareto d'un problème d'optimisation à deux objectifs

La résolution complète d'un problème multi-critères nécessite la mise en œuvre de méthodes d'optimisation multimodales capables d'identifier la frontière de *Pareto*. Cependant, dans le cadre d'une approche simplifiée, il est aussi possible de rechercher une solution unique en transformant le problème d'optimisation vectoriel en un problème d'optimisation scalaire. La méthode de pondération ou la méthode de *Marglin* sont habituellement employées pour effectuer ce type de transformation.

I.5.3. Méthode de pondération

Cette méthode consiste à affecter un poids à chaque critère et à définir comme fonction objectif du problème scalaire équivalent la somme pondérée des critères :

$$f_{obj}(\mathbf{x}) = \sum_{k=1}^l w_k f_k(\mathbf{x}) \quad (20)$$

où w_k désignent les *poids* (ou *facteurs*) de pondération. Généralement on impose $\sum_{k=1}^l w_k = 1$.

La méthode de pondération est relativement simple mais il est assez difficile de fixer *a priori* les valeurs des facteurs de pondération associés à chaque objectif. Il est par ailleurs impératif de normaliser les objectifs lorsque ceux-ci n'ont pas les mêmes dimensions. La mise

à l'échelle peut être effectuée à partir de la valeur maximale $f_{k\max}$ prise par chaque objectif lorsque celle-ci est connue. La relation (20) devient alors :

$$f_{obj}(\mathbf{x}) = \sum_{k=1}^l w_k \frac{f_k(\mathbf{x})}{f_{k\max}} \quad (21)$$

La Figure 8 illustre l'interprétation géométrique de la méthode de pondération dans le cas d'un problème à deux objectifs.

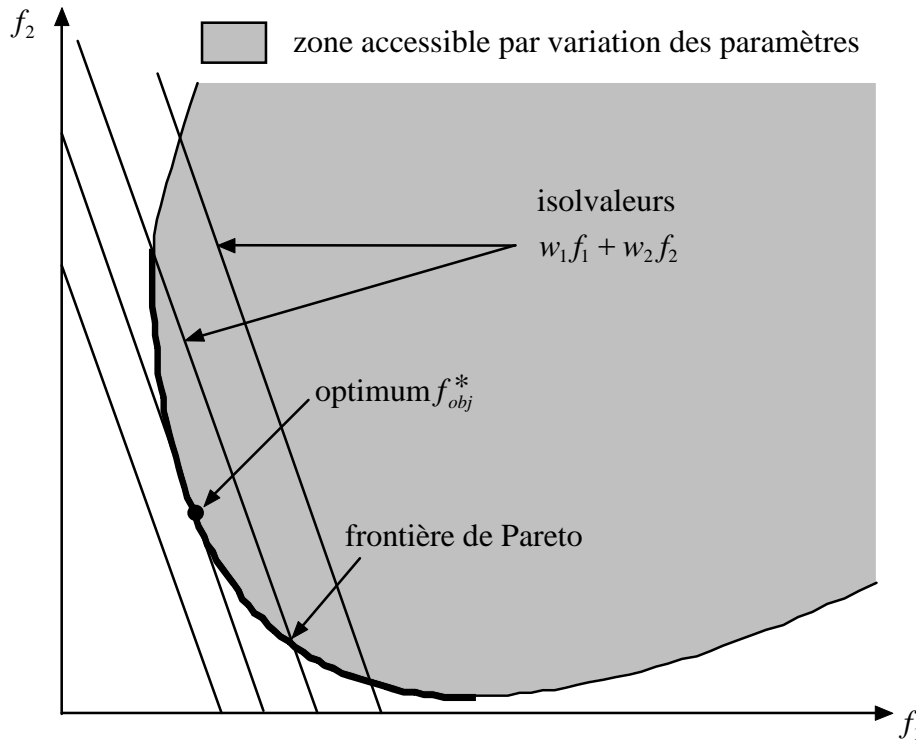


Figure 8. Interprétation géométrique de la méthode de pondération

I.5.4. Méthode de Marglin

Nous avons déjà signalé au paragraphe I.4.3, qu'il était possible de transformer un problème d'optimisation sous contraintes en un problème à objectifs multiples. Inversement, la méthode de *Marglin* permet de formuler un problème multi-critères en un problème scalaire contraint [Messine 98]. Elle consiste à prendre, comme fonction objectif du problème scalaire, une seule composante du vecteur critère, et à transformer les autres en contraintes d'inégalité suivant la relation (22) :

$$\begin{cases} \min_{\mathbf{x} \in I} f_1(\mathbf{x}) \\ g_j(\mathbf{x}) \leq 0 & j = 1..m \\ f_k(\mathbf{x}) - f_{k\max} \leq 0 & k = 2..l \end{cases} \quad (22)$$

où $f_{k\max}$ est la valeur maximale admissible pour le k ième objectif. Le principal défaut de la méthode réside dans le choix des valeurs maximales associées à chaque objectif. Si ces valeurs sont prises trop grandes, la solution du système sera sensiblement la solution du problème monocritère avec la fonction objectif utilisée. Par contre, lorsque ces valeurs sont trop faibles, les contraintes risquent de ne pas être satisfaites simultanément. Néanmoins, il est généralement plus facile de fixer des seuils pour les objectifs dans la méthode de *Marglin* que des poids dans la technique de pondération.

La Figure 9 illustre l'interprétation géométrique de la méthode de *Marglin* dans le cas d'un problème à deux objectifs.

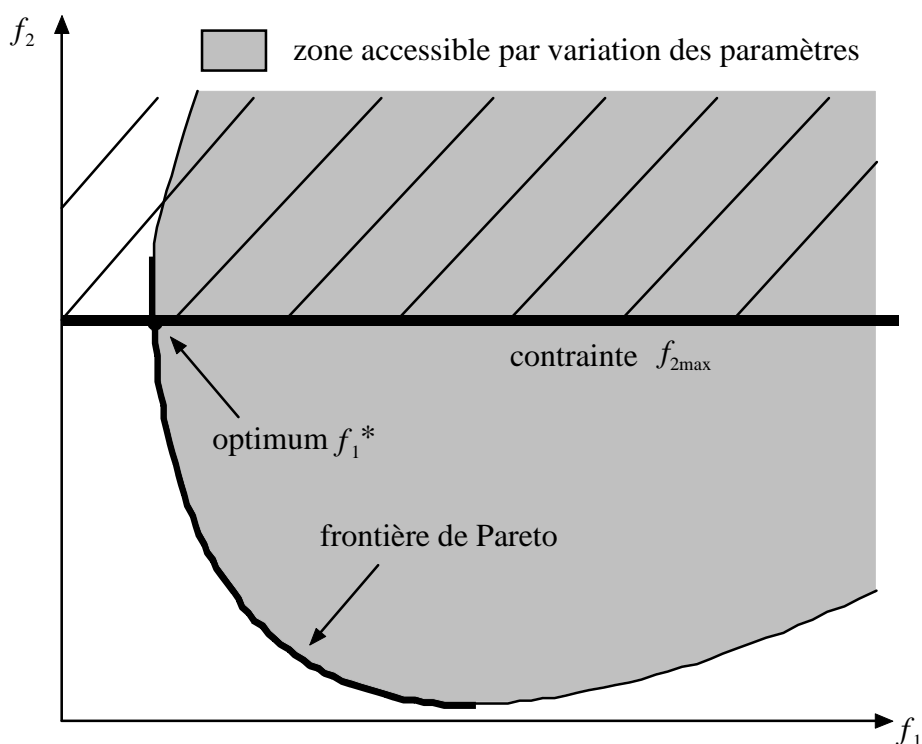


Figure 9. Interprétation géométrique de la méthode de *Marglin*

I.5.5. Optimisation à objectifs multiples par la logique floue

Le concept de logique floue, introduit par *Zadeh* [Zadeh 65] en 1965, constitue une extension de la théorie classique des ensembles.

I.5.5.1. Notion de sous-ensemble flou

Dans la théorie classique des ensembles, un sous-ensemble A de B est défini par une fonction d'appartenance $\mu_A(x)$ qui caractérise tout élément x appartenant à B . Cette fonction prend la valeur 1 si x appartient à A et la valeur 0 dans le cas contraire :

$$\mu_A(x) = \begin{cases} 0 & \text{si } x \notin A \\ 1 & \text{si } x \in A \end{cases} \quad (23)$$

En logique floue, un sous-ensemble flou A de B est défini par une fonction d'appartenance $\mu_A(x)$ pouvant prendre différentes valeurs comprises entre 0 et 1, selon le degré d'appartenance de l'élément x au sous-ensemble A .

$$\mu_A(x) \in [0,1] \quad (24)$$

La fonction d'appartenance est voisine de 1 lorsque l'élément x de B a tendance à appartenir au sous-ensemble A . Elle est voisine de 0, lorsque au contraire, ce dernier a tendance à ne pas appartenir à A . Enfin, comme en logique classique, elle prend la valeur 1 dans le cas où x appartient *incontestablement* à A et la valeur 0 si x n'appartient *incontestablement* pas à A .

La Figure 10 présente une fonction d'appartenance possible caractérisant l'état de chaleur "chaud". Nous pouvons affirmer incontestablement qu'il fait froid (c'est à dire *pas* chaud) de -10°C à 0°C ($\mu(T) = 0$) et chaud à partir de 30°C ($\mu(T) = 1$). De 0°C à 30°C , la sensation de chaleur est progressive, la fonction d'appartenance $\mu(T)$ augmente donc (linéairement dans l'exemple proposé) de 0 vers 1. L'état "*moyennement*" chaud est situé à la température de 15°C .

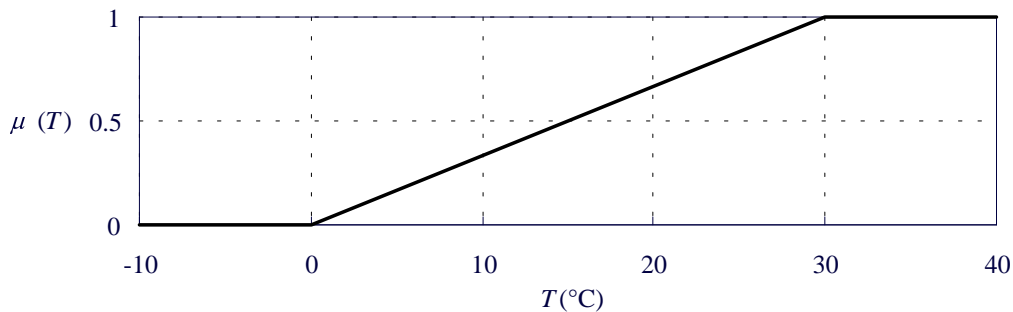


Figure 10. Exemple de fonction d'appartenance caractérisant l'état de chaleur "chaud"

I.5.5.2. Représentation des fonctions d'appartenance

Les fonctions d'appartenance les plus utilisées sont des fonctions à formes triangulaire, trapézoïdale ou gaussienne. La Figure 11 et la Figure 12 illustrent respectivement des formes de fonctions d'appartenance symétriques et non symétriques.

Nous définissons également une fonction d'appartenance particulière, le *singleton*, caractérisée par $\mu(x) = 1$ lorsque $x = x_0$ et $\mu(x) = 0$ ailleurs. Un exemple de singleton est présenté à la Figure 13.

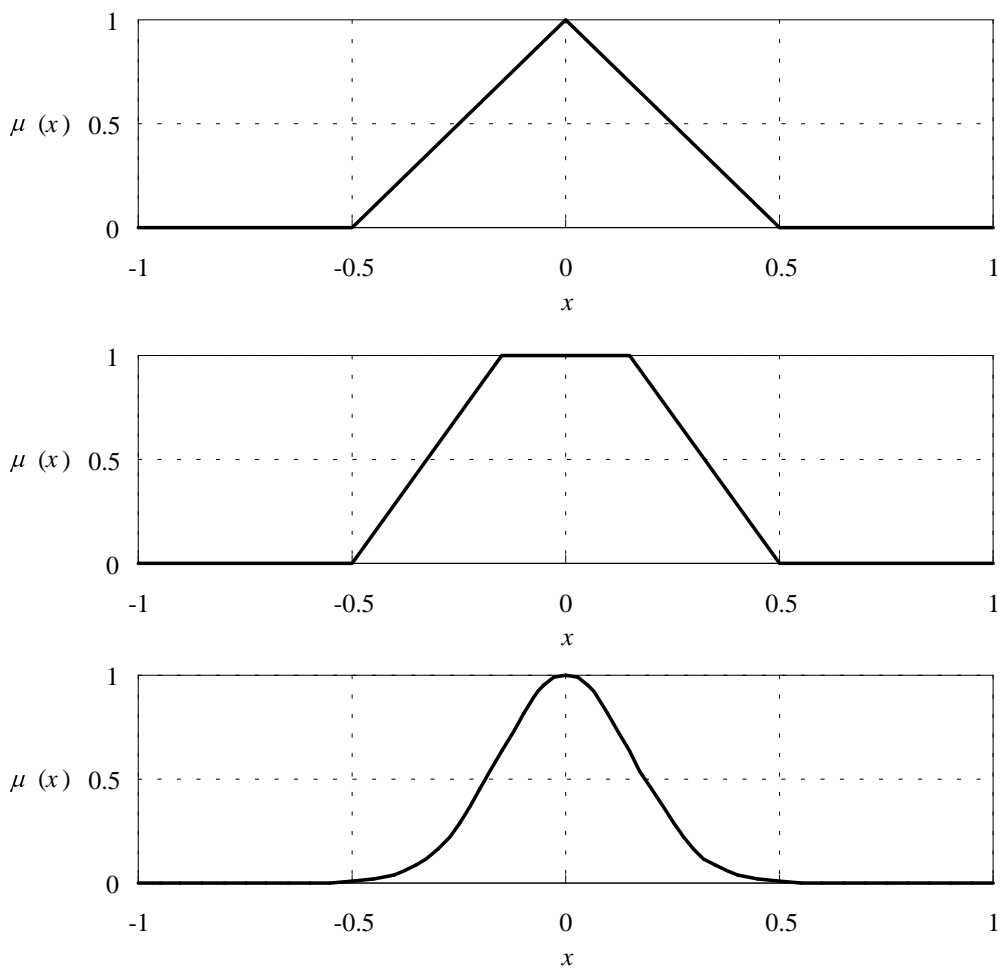


Figure 11. Exemple de fonctions d'appartenance symétriques

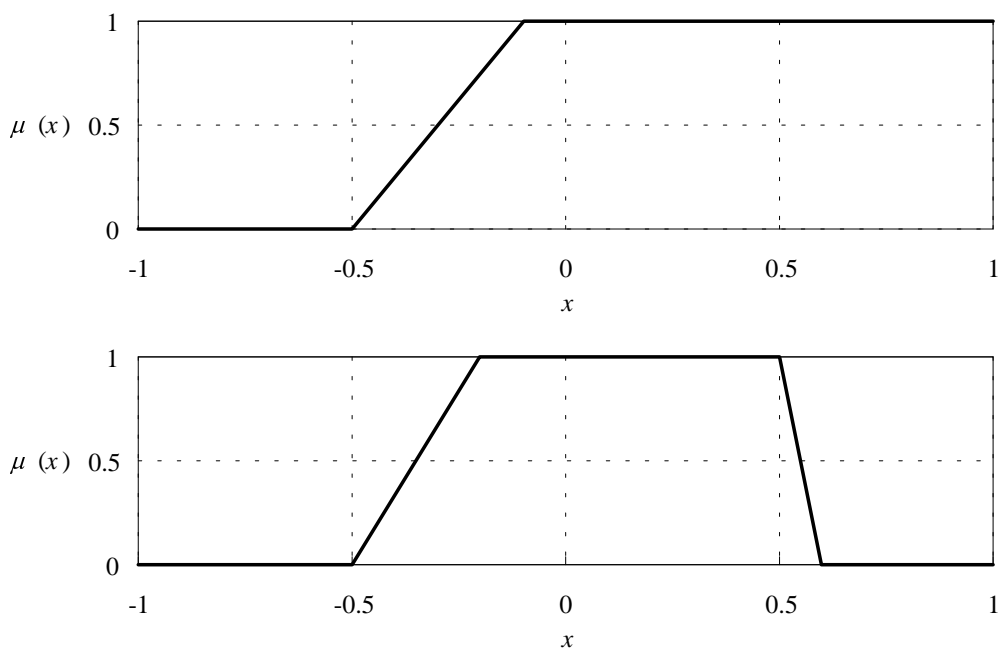


Figure 12. Exemples de fonctions d'appartenance non symétriques

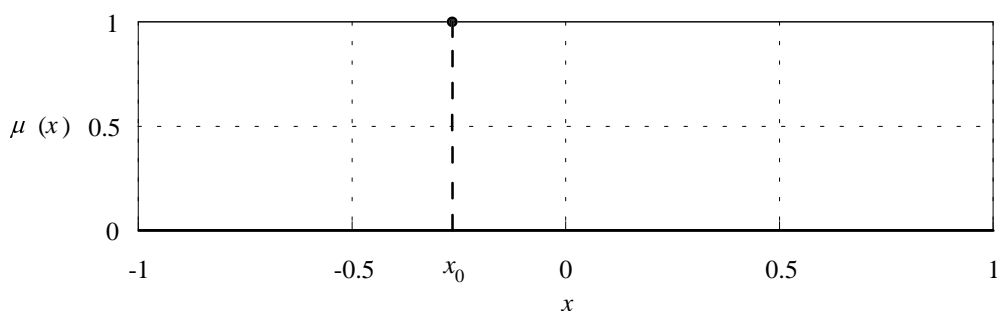


Figure 13. Fonction d'appartenance *singleton*

I.5.5.3. Opérations sur les sous-ensembles flous

Les principales opérations utilisées en logique floue sont résumées dans le Tableau 2.

Tableau 2. Opérations en logique floue

<i>opération</i>	<i>Symbolique</i>	<i>fonction d'appartenance</i>
inclusion de deux sous-ensembles	$A \subset B$	$\mu_A(x) \leq \mu_B(x) \quad \forall x$
égalité de deux sous-ensembles	$A = B$	$\mu_A(x) = \mu_B(x) \quad \forall x$
intersection de deux sous-ensembles	$A \cap B$	$\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x))$
union de deux sous-ensembles	$A \cup B$	$\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x))$
complément d'un sous-ensemble	\bar{A}	$\mu_{\bar{A}}(x) = 1 - \mu_A(x)$

Notons que l'*intersection* et l'*union* de deux sous-ensembles désignent respectivement le ET (ou multiplication) logique et le OU (ou addition) logique. Le choix du *min* est du *max* pour ces opérateurs est justifié par le fait que ceux-ci vérifient une grande partie des propriétés de la logique classique. De façon générale, les opérateurs intersection et d'union se réfèrent respectivement à deux familles de normes appelées normes triangulaires (T-normes) et conormes triangulaires (T-conormes) [Driankov 93]. Les principales T-normes et T-conormes de la littérature sont présentées dans le Tableau 3.

Tableau 3. T-norme et T-conorme classiques

<i>T-norme</i>	<i>T-conorme</i>
$\min(x,y)$	$\max(x,y)$
xy	$x+y-xy$
$\max(x+y-1,0)$	$\min(x+y,1)$

I.5.5.4. Résolution de problèmes à objectifs multiples par la logique floue

Le principe d'application de la logique floue à la résolution d'un problème à objectifs multiples est illustré à la Figure 14 [Chiampi 96].

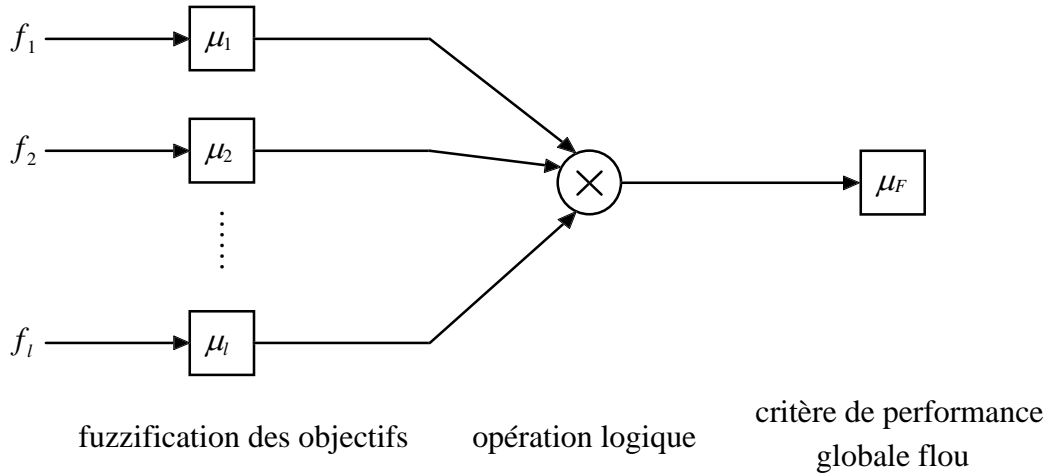


Figure 14. Principe de l'optimisation multi-critères par la logique floue

Chaque objectif f_k est *fuzzifié* à l'aide d'une fonction d'appartenance μ_k et regroupé dans un indicateur flou μ_F caractérisant la performance globale du système. Les fonctions d'appartenance doivent être définies de telle manière qu'elles "tendent" vers 1 lorsque l'objectif correspondant tend à être obtenu.

Par suite, l'ensemble des objectifs est pris en compte dans l'indicateur flou de performance globale, au moyen de l'opérateur d'intersection :

$$\mu_F(\mathbf{x}) = \min_{k=1..l} \{ \mu_k(f_k(\mathbf{x})) \} \quad (25)$$

Le problème d'optimisation à objectifs multiples est ainsi transformé en un problème scalaire équivalent, formulé terme de maximisation, dont la fonction objectif est donnée par $\mu_F(\mathbf{x})$. La solution optimale du problème \mathbf{x}^* s'obtient selon (26) :

$$\mathbf{x}^* = \max_{\mathbf{x} \in I} \{ \mu_F(\mathbf{x}) \} = \max_{\mathbf{x} \in I} \left\{ \min_{k=1..l} \{ \mu_k(f_k(\mathbf{x})) \} \right\} \quad (26)$$

L'approche de l'optimisation multi-critères par la logique floue est bien appropriée lorsque l'on a une idée précise sur les objectifs que l'on souhaite obtenir et sur la tolérance admissible pour chacun d'eux. Par contre, si on cherche à minimiser plusieurs critères simultanément, il n'est pas toujours facile de déterminer *a priori* l'ordre de grandeur optimal de chaque critère. La définition des fonctions d'appartenance est alors problématique.

Comme pour le choix des références dans la technique de pondération et la méthode de *Marglin*, il peut s'avérer intéressant d'optimiser chaque objectif indépendamment des autres pour avoir une idée sur la performance optimale qui les caractérise et fixer de façon plus judicieuse ces mêmes références. Toutefois, la procédure devient coûteuse puisqu'il faut d'abord autant de résolution de problèmes scalaires que d'objectifs, avant de pouvoir déterminer la solution du problème vectoriel complet.

I.6. CONCLUSION

Dans ce chapitre, nous avons dressé un état de l'art des méthodes d'optimisation mathématiques. Parmi l'ensemble des techniques présentées, nous avons choisi de nous intéresser en particulier aux algorithmes génétiques en raison de leur capacité à traiter des fonctions objectifs non convexes, discontinues ou multimodales. Les problèmes rencontrés en génie électrique présentent habituellement ces caractéristiques [Hoole 91][Gottvald 92] [Simkin 92][Alotto 96]. Les deux prochains chapitres sont consacrés aux algorithmes génétiques. Le chapitre II établit les principes fondamentaux de l'algorithme génétique standard tandis que le chapitre III traite de méthodes génétiques multimodales.

CHAPITRE II

L'ALGORITHME GENETIQUE STANDARD

Les Algorithmes Génétiques (GA) sont des techniques d'optimisation stochastiques qui tentent d'imiter les processus d'évolution naturelle des espèces et de la génétique. Ils agissent sur une population d'individus assujettis à une sélection darwinienne : les individus (ou *parents*) les mieux adaptés à leur environnement survivent et peuvent se reproduire. Ils sont alors soumis à des mécanismes de recombinaisons analogues à ceux de la génétique. Des échanges de gènes entre parents résulte la création de nouveaux individus (ou *enfants*), qui permettent de tester d'autres configurations de l'espace de recherche.

Les Algorithmes Génétiques diffèrent fondamentalement des autres méthodes dans la recherche de l'optimum [Goldberg 89]:

- Ils utilisent un *codage* des paramètres de la fonction à optimiser et non les paramètres eux-mêmes.
- Ils agissent sur un *ensemble de configurations* et non sur un point unique.
- Ils n'utilisent que *les valeurs* de la fonction à optimiser, pas sa dérivée ou une autre connaissance auxiliaire.
- Ils utilisent des règles de transitions *probabilistes* (non déterministes).

Ce chapitre présente les bases fondamentales de l'algorithme génétique et propose un état de l'art des opérateurs génétiques qui interviennent dans le processus de recherche de l'optimum global.

II.1. GENERALITES

II.1.1. Formulation du problème d'optimisation

En raison de l'analogie avec la théorie de l'évolution, l'algorithme génétique est naturellement formulé en terme de maximisation :

$$\max_{x \in I} F_{obj}(\mathbf{x}) \quad (27)$$

où I désigne l'espace de recherche et F_{obj} la fonction à maximiser.

Par ailleurs, la fonction F_{obj} doit avoir des valeurs positives sur l'ensemble du domaine I . (cette condition est imposée par l'utilisation de certains opérateurs génétiques que nous décrirons par la suite). Dans le cas contraire, il convient d'ajouter aux valeurs de F_{obj} une constante positive F_{min} suffisamment grande conformément à l'équivalence de (27) et (28).

$$\max_{x \in I} F_{obj}(\mathbf{x}) + F_{min} \quad (28)$$

Dans de nombreux problèmes, l'objectif est exprimé sous forme de minimisation d'une fonction f_{obj} ,

$$\min_{x \in I} f_{obj}(\mathbf{x}) \quad (29)$$

Le passage d'un problème de minimisation à un problème de maximisation est obtenu par transformation de la fonction f_{obj} . Il est aisé de montrer par exemple que (29) est équivalent à (30),

$$\max_{x \in I} F_{obj}(\mathbf{x}) = F_{max} - f_{obj}(\mathbf{x}) \quad (30)$$

où $F_{max} \geq \max_{x \in I} f_{obj}(\mathbf{x})$ pour éviter que la nouvelle fonction prenne des valeurs négatives sur son domaine de définition.

Notons que le choix de la fonction F_{obj} n'est pas unique. On rencontre également dans la littérature la fonction de transformation suivante [Goldberg 89][Saludjian 97] :

$$F_{obj}(\mathbf{x}) = \frac{1}{1 + f_{obj}(\mathbf{x})} \quad (31)$$

II.1.2. Terminologie, définitions et notations employées

II.1.2.1. Terminologie et définitions

L'*environnement* ou *espace de recherche* définit l'ensemble des configurations possibles des paramètres de la fonction à optimiser. Un *individu* représente un point de cet espace, la *population* désignant un groupe d'individus.

On qualifie souvent la fonction à optimiser de fonction d'*adaptation*. Elle mesure la *performance* des individus dans leur environnement.

Le *chromosome* est une chaîne d'allèles qui contient les caractéristiques génétiques d'un individu et code les paramètres de sa fonction d'adaptation. Le *gène* ou *allèle* est la partie élémentaire du chromosome. Il représente un trait de caractère ou une fonction particulière. Si le chromosome est codé en binaire, un gène sera symbolisé par un bit auquel seront associées plusieurs de ces caractéristiques (valeur, position ou poids dans la chaîne, ...).

Le *génotype* d'un individu caractérise la structure du chromosome tandis que le *phénotype* désigne les valeurs des paramètres réels et de la fonction objectif qui lui sont associées.

II.1.2.2. Notations

Dans la suite,

$$P^t = (a_1^t, a_2^t, \dots, a_N^t) \in I^N \quad (32)$$

désignera la population à la génération t , N le nombre total d'individus (soit la taille de la population) et I l'espace de recherche.

La fonction $f: I \rightarrow \mathfrak{R}$ donne la performance des individus. On supposera dans ce qui suit que les individus sont classés par ordre décroissant de leur performance, a_1^t étant le meilleur individu de la population P^t . On définit alors l'application $\text{rang}: I \rightarrow \{1, 2, \dots, N\}$ de la manière suivante :

$$\forall i \in \{1, 2, \dots, N\} : \text{rang}(a_i^t) = i \quad (33)$$

II.1.3. Le codage des paramètres

Selon *Goldberg*, l'utilisateur doit choisir le plus petit alphabet qui permette une expression naturelle des paramètres du problème (principe des alphabets minimaux) [Goldberg 89]. C'est pourquoi, l'alphabet binaire $\{0,1\}$ est particulièrement bien adapté à la représentation des paramètres.

II.1.3.1. Le codage binaire standard

Pour chaque paramètre x_i situé dans l'intervalle $[x_{i\min}, x_{i\max}]$, on associe une chaîne binaire $b_0b_1\dots b_{l_{x_i}-1}$ définie sur l_{x_i} bits. A cette chaîne correspond une valeur entière naturelle,

$$N(x_i) = \sum_{i=0}^{l_{x_i}-1} 2^{l_{x_i}-i-1} b_i \quad (34)$$

Le paramètre réel x_i de l'espace de recherche relatif à $N(x_i)$ est obtenu par mise à l'échelle linéaire :

$$x_i = x_{i\min} + \frac{x_{i\max} - x_{i\min}}{2^{l_{x_i}} - 1} N(x_i) \quad (35)$$

Cette méthode de codage est relativement facile à implanter mais elle présente l'inconvénient de limiter la précision des paramètres à une valeur ε_i correspondant à l'écart entre deux configurations réelles adjacentes obtenues, pour une variation du bit le moins significatif (voir Figure 15).

$$\varepsilon_i = \frac{x_{i\max} - x_{i\min}}{2^{l_{x_i}} - 1} \quad (36)$$

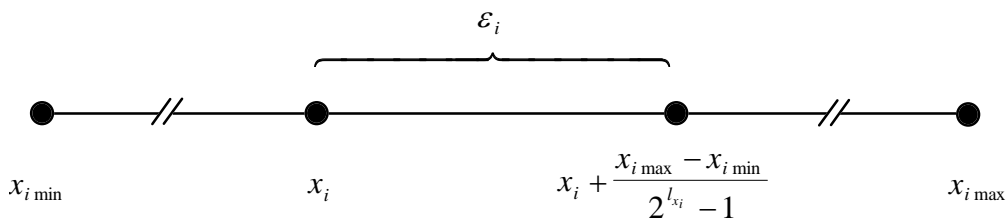


Figure 15. Précision maximale de la méthode de codage binaire standard

La longueur l_{x_i} de la sous-chaîne binaire nécessaire pour obtenir une précision ε_i pour le paramètre réel x_i correspond donc au plus petit entier naturel tel que :

$$l_{x_i} \geq \frac{1}{\ln 2} \ln \left(\frac{x_{i \max} - x_{i \min}}{\varepsilon_i} + 1 \right) \quad (37)$$

La longueur totale du chromosome est donnée par :

$$l = \sum_{i=1}^n l_{x_i} \quad (38)$$

où n est le nombre total de paramètres. En codant chaque paramètre sur une longueur fonction de la précision requise, on pourra adapter la taille du chromosome de manière à optimiser le rapport temps de calcul et précision de la solution finale.

II.1.3.2. Le code Gray

Avec le codage binaire, deux configurations proches dans l'espace des paramètres peuvent avoir des chromosomes très distincts. Par exemple, les chaînes "01111" et "10000" correspondent à deux configurations réelles voisines alors qu'elles diffèrent de 5 bits. Cette caractéristique peut s'avérer pénalisante pour la recherche locale en raison de la difficulté liée au franchissement de ces "falaises de Hamming" (*Hamming cliffs*). L'utilisation du code Gray a été recommandée pour répondre à ce problème [Caruana 88][Bäck 93]. En effet, avec ce code, les entiers adjacents ne diffèrent que d'un bit (distance de Hamming de 1). Le passage entre deux configurations réelles voisines devient beaucoup plus facile puisqu'il suffit de modifier un seul bit dans le chromosome (nous verrons ultérieurement que nous pouvons obtenir cet effet avec l'opérateur de mutation).

Le Tableau 4 présente les codes binaire et Gray pour des chaînes de 4 bits.

Le passage du code binaire au code Gray est effectué de la manière suivante :

$$g_i = \begin{cases} b_i & \text{si } i = 0 \\ b_{i-1} \oplus b_i & \text{si } i \geq 1 \end{cases} \quad (39)$$

où \oplus représente l'addition modulo 2.

La transformation inverse s'obtient selon (40).

$$b_i = \bigoplus_{j=0}^i g_j \quad (40)$$

Si on considère que le chromosome est représenté en code Gray, on effectuera d'abord la transformation (40) avant d'appliquer le décodage binaire standard.

Tableau 4. Code binaire et code Gray sur 4 bits

<i>Entier</i>	<i>Binaire</i>	<i>Gray</i>
0	0000	0000
1	0001	0001
2	0010	0011
3	0011	0010
4	0100	0110
5	0101	0111
6	0110	0101
7	0111	0100
8	1000	1100
9	1001	1101
10	1010	1111
11	1011	1110
12	1100	1010
13	1101	1011
14	1110	1001
15	1111	1000

II.1.3.3. Le codage dynamique des paramètres

Pour résoudre le problème de précision inhérent au décodage binaire standard et améliorer la recherche locale, un codage dynamique des paramètres (*Dynamic Parameter Encoding*) a été proposé [Schraudolph 92]. La procédure de décodage consiste à modifier (35) suivant (41).

$$x_i = x_{i\min} + \frac{x_{i\max} - x_{i\min}}{2^{l_{x_i}} - 1} (N(x_i) + dN(x_i)) \quad (41)$$

où $dN(x_i)$ est une variable réelle aléatoire à densité uniforme prise dans l'intervalle [0,1].

L'introduction de $dN(x_i)$ supprime les discontinuités de précision (voir Figure 15) entre deux configurations réelles adjacentes, obtenues pour une variation du bit le moins significatif. Le décodage dynamique augmente le nombre de points pouvant être visités, améliorant ainsi la qualité de la recherche.

Comme à une configuration donnée du chromosome correspond une gamme de paramètres possible, il devient nécessaire de mémoriser les caractéristiques génétiques des individus après décodage.

II.1.4. Les schémas

II.1.4.1. Principe

La similarité entre individus d'une même population nous amène à définir la notion de schéma. Un schéma est une séquence sur l'alphabet binaire $\{0,1\}$ augmenté du symbole indifférent *, ce dernier pouvant prendre indifféremment la valeur 0 ou 1.

Par exemple, le schéma $H = 0**10$ désigne les chaînes binaires de longueur $l = 5$ possédant un 0 sur la première et la dernière position et un 1 sur l'avant-dernière. Ce schéma représente donc l'ensemble des motifs similaires suivants :

$$H1 = 00010$$

$$H2 = 00110$$

$$H3 = 01010$$

$$H4 = 01110$$

II.1.4.2. Caractéristiques d'un schéma

Un schéma est caractérisé par :

- son *ordre*
- sa *longueur utile* ou longueur de définition

L'ordre d'un schéma H noté $o(H)$ est le nombre de positions *instanciées* (ou nombre de positions *fixes*, soit le nombre total de 0 et de 1) dans le motif considéré.

$$\text{exemple : } o(0**1*1*) = 3$$

La longueur utile du schéma H notée $\delta(H)$ désigne la distance entre la première et la dernière position instanciée dans le motif considéré.

exemple : $\delta(0^{**}1^{*}1^{*}) = 6 - 1 = 5$

II.1.4.3. Interprétation géométrique des schémas

La prise en compte de la géométrie de l'espace binaire a conduit à une analyse des schémas en terme d'hyperplans. Considérons l'ensemble des schémas de longueur $l = 3$. L'espace de recherche est un cube dont les sommets représentent les schémas d'ordre 3, les arêtes les schémas d'ordre 2 et les plans les schémas d'ordre 1. Le domaine entier est symbolisé par le schéma $***$ d'ordre 0 (voir Figure 16).

Les notions de points (ou sommets), lignes (ou arêtes) et plans, valables pour des schémas de longueur $l = 3$, sont généralisables en hyperplans de dimension variable pour un hypercube de dimension quelconque.

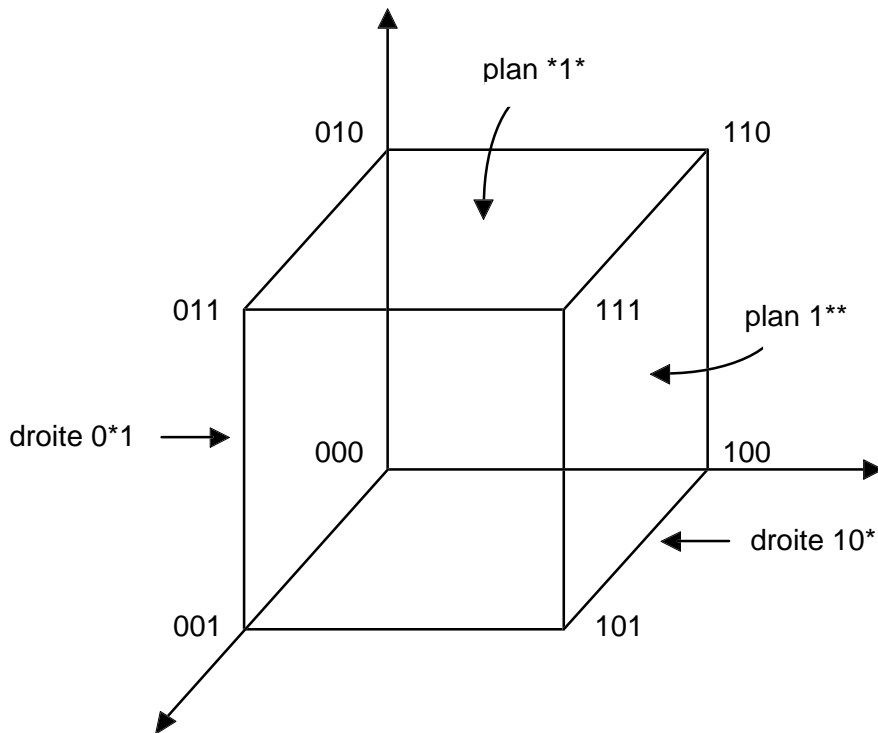


Figure 16. Représentation des schémas dans un espace tridimensionnel

II.1.4.4. Localisation des schémas dans l'espace de recherche réel

Chaque schéma ou hyperplan correspond à une ou plusieurs régions de l'espace de recherche réel. La Figure 17 illustre la localisation des schémas $H = 0^{*}0$ et $H = ^{*}1^{*}$ pour une fonction à un seul paramètre x .

La Figure 18 montre les quatre zones qui correspondent au schéma $H=*1**00$ dans l'espace des paramètres pour une fonction de deux variables x et y (les trois premiers bits du schéma codent le paramètre x et les trois derniers le paramètre y).

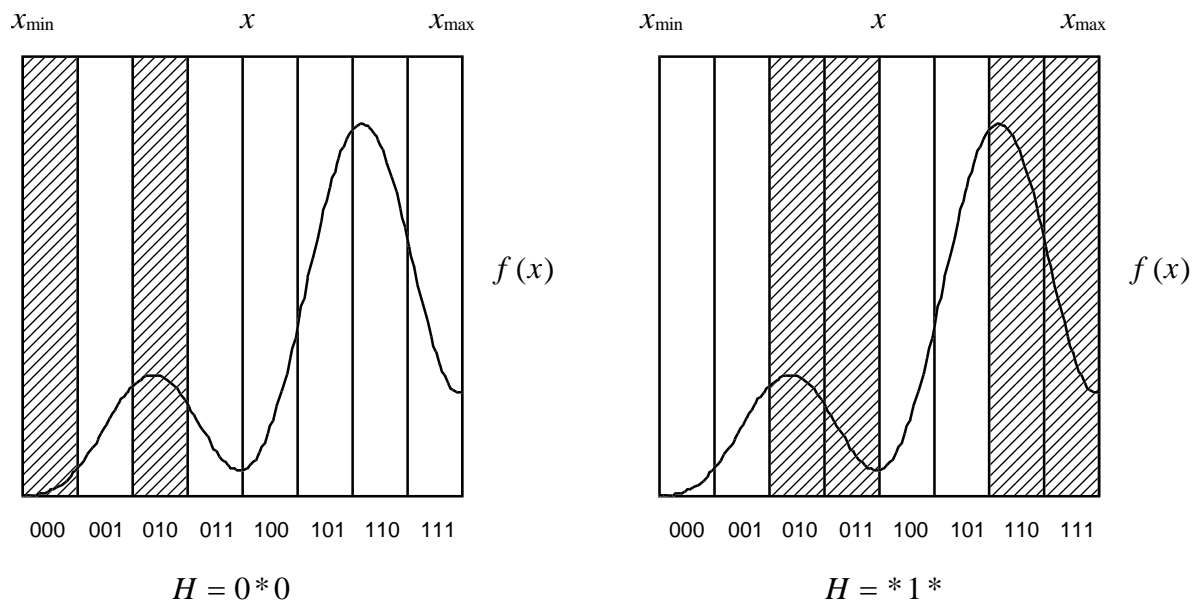


Figure 17. Localisation des schémas dans l'espace de recherche réel

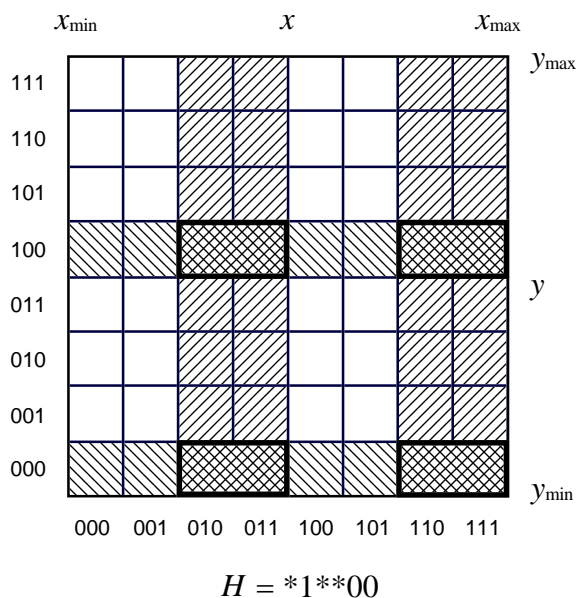


Figure 18. Localisation du schéma $H = *1**00$ dans l'espace des paramètres

II.2. STRUCTURE DE L'ALGORITHME GENETIQUE

L'organigramme fonctionnel de la Figure 19 illustre la structure de l'algorithme génétique standard. Nous détaillerons les diverses phases qui le constituent et présenterons les mécanismes associés à chacune d'entre elles dans les sections suivantes.

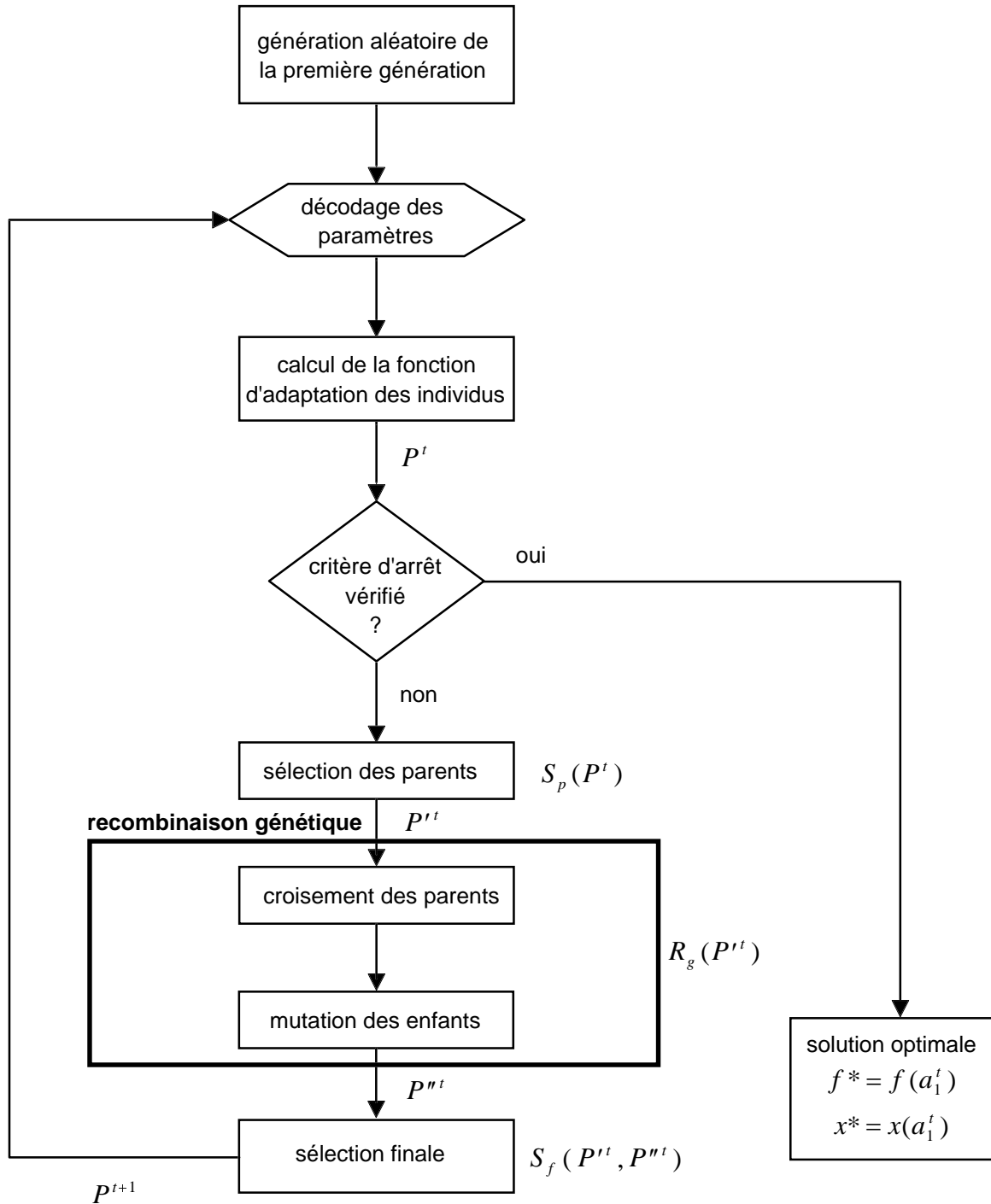


Figure 19. Structure de l'algorithme génétique

II.3. INITIALISATION DE LA POPULATION

Les chromosomes des individus de la première génération sont habituellement initialisés de façon uniformément aléatoire, en respectant l'équiprobabilité d'obtenir une valeur égale à 0 ou 1 pour chaque bit. D'autres techniques d'initialisation ont été proposées mais ces dernières restent très peu employées [Kallel 97].

II.4. LA SÉLECTION DES PARENTS

La sélection des parents est un mécanisme qui fixe à partir d'une génération donnée, quels individus pourront se reproduire pour former la génération suivante. Lors de cette phase, les individus les plus forts sont généralement dupliqués et forment les parents de la génération en cours, alors que les plus faibles disparaissent sans avoir la possibilité de se reproduire. Si P^t est la génération courante, la population à la suite de la sélection S_p sera notée :

$$P^{t'} = S_p(P^t) = (a_1^{t'}, a_2^{t'}, \dots, a_N^{t'}) \quad (42)$$

II.4.1. Les méthodes de sélection

Elles sont notamment caractérisées par la probabilité $p_s(a_i^t)$ qu'un individu a_i^t de la population soit retenu pour participer à la recombinaison génétique.

II.4.1.1. La sélection proportionnelle RWS

La sélection proportionnelle (*Proportional Selection* ou *Roulette Wheel Selection*) consiste à dupliquer chaque individu proportionnellement à son adaptation dans son milieu. Ceci peut être effectué aisément en procédant à des tirages aléatoires consécutifs où chaque individu a une probabilité d'être sélectionnée égale à [Holland 75]:

$$p_s(a_i^t) = \frac{f(a_i^t)}{\sum_{j=1}^N f(a_j^t)} \quad (43)$$

En pratique, on utilise une roue de loterie divisée en N secteurs (représentant les individus à sélectionner) dont la surface est proportionnelle à la probabilité de sélection correspondante. N tirages aléatoires sont ensuite réalisés pour sélectionner N individus. Le nombre d'occurrences (où nombre de copies) attendu pour chaque individu après sélection s'écrit :

$$n(a_i^t) = N p_s(a_i^t) \quad (44)$$

Chapitre II

Cette sélection, uniquement basée sur la performance, présente l'inconvénient de favoriser injustement la prolifération de "super-individus" et de faire disparaître de nombreux génotypes dans la population¹. Pour éviter une convergence prématurée de l'algorithme et préserver la diversité des individus, il est recommandé de réaliser un réajustement de la fonction d'adaptation avant la sélection.

La technique de réajustement la plus répandue est la mise à l'échelle linéaire [Goldberg 89]. Celle-ci consiste à comprimer le domaine de variation de la fonction d'adaptation tout en gardant une valeur moyenne identique ($f'_{moy} = f_{moy}$). Au maximum de la performance correspond une nouvelle valeur $f'_{max} = \beta f_{moy}$ avec $1.2 \leq \beta \leq 2$ (typiquement $\beta = 2$). La fonction d'adaptation après mise à l'échelle prend la forme suivante :

$$f' = af + b \quad (45)$$

avec
$$a = \frac{\beta - 1}{f_{max} - f_{moy}} f_{moy} \quad (46)$$

et
$$b = \frac{f_{max} - \beta f_{moy}}{f_{max} - f_{moy}} f_{moy} \quad (47)$$

Toutefois, il faut effectuer le test de non-négativité (48) avant la mise à l'échelle. En effet, si $f_{max} < \beta f_{moy}$ la compression du domaine de performance est déjà implicite et dans ce cas, la mise à l'échelle étend au contraire les valeurs de f , diminuant l'influence des individus trop "faibles".

$$f_{min} > \frac{\beta f_{moy} - f_{max}}{\beta - 1} \quad (48)$$

Si l'inéquation précédente n'est pas vérifiée, la fonction d'adaptation mise à l'échelle prendra des valeurs négatives. Il convient donc de limiter l'extension de la borne inférieure à $f'_{min} = 0$ tout en gardant $f'_{moy} = f_{moy}$. Les coefficients du réajustement linéaire s'expriment alors par :

$$a = \frac{f_{moy}}{f_{moy} - f_{min}} \quad (49)$$

$$b = -\frac{f_{moy}}{f_{moy} - f_{min}} f_{min} \quad (50)$$

¹ Le lecteur pourra se reporter à [Baker 87] pour une étude complète du biais dans les méthodes de sélection.

Pour résumer, le mécanisme de la mise à l'échelle linéaire peut être représenté par la figure 6.

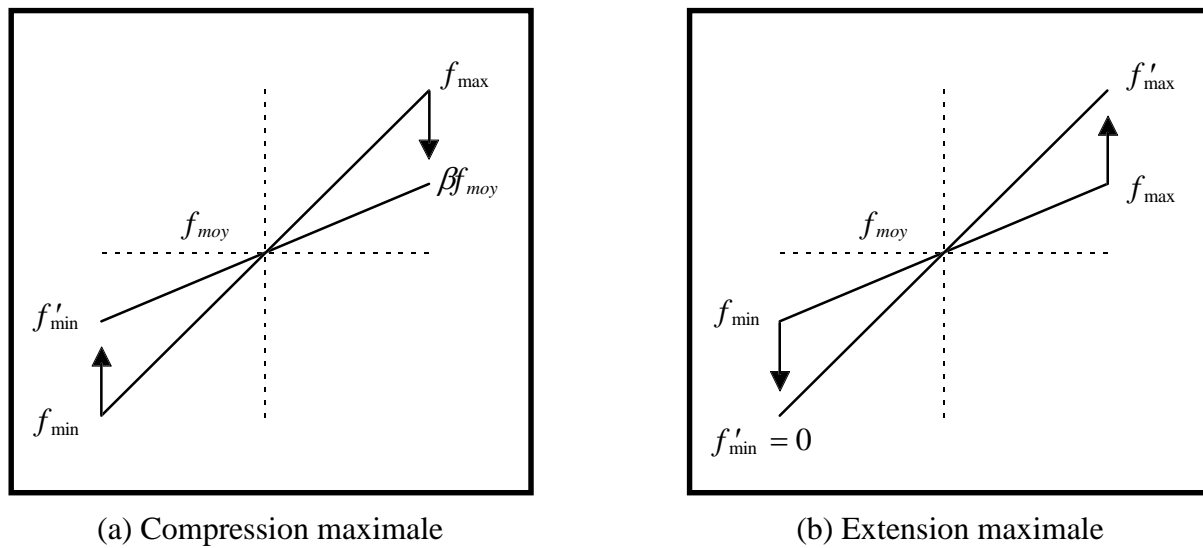


Figure 20. Principe de la mise à l'échelle linéaire

Il existe d'autres types de réajustement de la fonction d'adaptation notamment la troncature sigma qui supprime les individus de faible performance avant une éventuelle mise à l'échelle linéaire :

$$f' = f - (f_{moy} - c\sigma) \quad (51)$$

où $1 \leq c \leq 3$ et σ est l'écart-type de l'adaptation, calculé sur la population. Les résultats négatifs sont forcés à zéro.

Notons enfin le changement d'échelle en puissance qui modifie la fonction d'adaptation de la façon suivante :

$$f' = f^\beta \quad (52)$$

où la constante β peut être adaptée au cours des générations pour accentuer la sélection des individus situés à proximité des sommets [Goldberg 89].

II.4.1.2. La sélection proportionnelle à reste stochastique

Avec la sélection à reste stochastique (*Stochastic Remainder Selection*), le nombre de copies $n(a'_i)$ d'un individu est directement fixé par le rapport de sa performance avec la performance moyenne de la population [Baker 87] :

$$n(a_i^t) = \text{partie entière} \left(\frac{f(a_i^t)}{f_{\text{moy}}} \right) \quad (53)$$

où $f_{\text{moy}} = \frac{1}{N} \sum_{i=1}^N f(a_i^t)$ est l'adaptation moyenne de la population.

Dans un premier temps, on fixe $\alpha = \sum_{i=1}^N n(a_i^t)$ individus à partir de (53) puis on complète la population en associant à chaque individu une probabilité d'être sélectionné égale à :

$$p_s(a_i^t) = \frac{1}{N - \alpha} \left(\frac{f(a_i^t)}{f_{\text{moy}}} - \text{partie entière} \left(\frac{f(a_i^t)}{f_{\text{moy}}} \right) \right) \quad (54)$$

Comme pour la sélection RWS, il est judicieux de faire précéder la phase de sélection d'une procédure de réajustement de la fonction d'adaptation. Le biais¹ qui résulte de cette technique est quasi nul.

II.4.1.3. La sélection stochastique universelle

Contrairement à la méthode de sélection proportionnelle où il faut réaliser N tirages aléatoires pour sélectionner N individus, la sélection stochastique universelle (*Stochastic Universal Selection*) ne nécessite qu'un seul tirage pour choisir tous les parents d'une génération [Baker 87]. A partir d'une variable aléatoire θ , prise uniformément dans l'intervalle $[0, f_{\text{moy}}]$, on définit deux séries de pointeurs p_u et p_v de la manière suivante :

$$p_u = \theta + (u - 1)f_{\text{moy}} \quad u = 1, 2, \dots, N \quad (55)$$

$$p_v = \sum_{j=1}^v f(a_j^t) \quad v = 1, 2, \dots, N \quad (56)$$

Le pseudocode de la sélection stochastique universelle est donné ci-après :

```

u = 1
v = 1
pour i = 1..N
    tant que p_u < p_v
        sélectionner a_i^t
        incrémenter u
    fin tant que
    incrémenter v
fin pour
    
```

Cette sélection reproduit aussi les individus proportionnellement à leur fonction d'adaptation mais contrairement aux techniques précédentes son biais¹ est totalement nul.

II.4.1.4. La sélection linéaire par rapport au rang

Avec la sélection linéaire par rapport au rang (*Linear Ranking Selection*), un individu est choisi aléatoirement avec une probabilité proportionnelle à son rang [Baker 85] :

$$p_s(a_i^t) = \frac{1}{N} \left(\eta_+ - (\eta_+ - \eta_-) \frac{i-1}{N-1} \right) \quad (57)$$

$$\text{où } \eta_- = 2 - \eta_+ \text{ et } 1 \leq \eta_+ \leq 2$$

Un réajustement préalable de la fonction d'adaptation n'est pas nécessaire, la méthode étant beaucoup plus "juste" dans la mesure où elle n'utilise pas directement la performance des individus.

II.4.1.5. La sélection uniforme par rapport au rang

La sélection uniforme par rapport au rang (*Uniform Ranking Selection*) consiste à choisir de façon équiprobable les individus de rang inférieur ou égal à μ ($\mu \leq N$). Les autres individus sont exclus de la population et ne peuvent participer à la reproduction [Bäck 91]. La probabilité de sélection s'exprime par :

$$p_s(a_i^t) = \begin{cases} \frac{1}{\mu} & \text{si } 1 \leq i \leq \mu \\ 0 & \text{si } \mu < i \leq N \end{cases} \quad (58)$$

Cette méthode est équivalente à la sélection par troncation (*truncation selection*) qui duplique uniquement un pourcentage des meilleurs individus de la population [Mühlenbein 94]. Elle est aussi connue sous le nom de sélection (μ, λ) , λ désignant la taille de la population ($\lambda = N$), utilisée notamment dans les stratégies d'évolution [Schwefel 95].

Ce mode de sélection peut être combiné à (43) ou (57) pour donner respectivement naissance à des sélections (μ, λ) proportionnelle ou linéaire par rapport au rang [Bäck 91]. A notre connaissance, ces deux techniques restent toutefois très peu employées.

II.4.1.6. La sélection par tournoi

Lors de la sélection par tournoi (*Tournament Selection*), k individus de la population sont choisis aléatoirement et celui, dont la performance est la plus élevée, est retenu pour participer à la reproduction. L'opération répétée autant de fois qu'il y a d'individus à sélectionner.

La probabilité qu'un individu de rang i soit sélectionné après compétition est donnée ci-après :

$$p_s(a_i^t) = \begin{cases} \frac{C_{N-i}^{k-1}}{C_N^k} & \text{si } 1 \leq i \leq N - k + 1 \\ 0 & \text{si } N - k + 2 \leq i \leq N \end{cases} \quad (59)$$

Dans le cas particulier du tournoi de deux individus ($k = 2$), qualifié de tournoi binaire probabiliste (*Binary Tournament Selection*), (59) se réduit à (60).

$$p_s(a_i^t) = \frac{2}{N} \frac{N-i}{N-1} \quad (60)$$

Cette formulation permet de constater que la sélection par tournoi de deux individus est équivalente à la sélection linéaire par rapport au rang de coefficient $\eta_+ = 2$.

II.4.2. Les caractéristiques des mécanismes de sélection

Les mécanismes de sélection dans les algorithmes génétiques jouent un rôle très important d'une part pour diriger les recherches vers les meilleurs individus, d'autre part pour maintenir la diversité des génotypes dans la population. Ils sont directement liés au compromis entre une vitesse de convergence élevée et une forte probabilité de trouver un optimum global dans le cas d'un problème multimodal.

II.4.2.1. La sélection dynamique et la sélection statique

Les probabilités de sélection d'un individu de rang i peuvent être dépendantes de sa fonction d'adaptation et donc varier au cours des générations. On parlera alors de sélection *dynamique* (cas de la sélection proportionnelle par exemple). Inversement, une sélection est dite *statique* si pour un individu de rang i , les probabilités de sélection restent inchangées dans le temps (cas des méthodes de sélection liées au rang)

Formellement une sélection est statique si et seulement si

$$\forall i \in \{1, 2, \dots, N\} \forall (t_1, t_2) \in \mathbb{N}^2 : p_s(a_i^{t_1}) = p_s(a_i^{t_2}) \quad (61)$$

Elle est dynamique lorsque la relation précédente n'est pas vérifiée.

II.4.2.2. La sélection préservatrice et la sélection extincitrice

Une sélection est *préservatrice* si tout individu de la population a la possibilité de participer à la reproduction pour former la génération suivante (cas de la sélection proportionnelle ou de la sélection linéaire par rapport au rang). Inversement, une sélection est

extinctrice si certains individus de la population ont une probabilité nulle de se reproduire (cas de la sélection par troncation et de la sélection par tournoi)

Une procédure de sélection est extinctrice si et seulement si :

$$\forall t \geq 0 \exists i \in \{1, 2, \dots, N\} : p_s(a_i^t) = 0 \quad (62)$$

Une procédure de sélection est préservatrice si et seulement si :

$$\forall t \geq 0 \forall i \in \{1, 2, \dots, N\} : p_s(a_i^t) \neq 0 \quad (63)$$

II.4.2.3. La sélection élitiste

Une sélection est élitiste si elle garantit la survie du meilleur individu de la population. Elle assure alors nécessairement la relation suivante :

$$\forall t > 1 : f(a_1^{t'}) \geq f(a_1^t) \quad (64)$$

Des différentes méthodes présentées, seules la sélection à reste stochastique et la sélection stochastique universelle sont élitistes. Avec les autres procédures, la probabilité de sélectionner un individu est généralement élevée mais elle reste toujours inférieure à un. Il est donc tout à fait possible que le meilleur individu de la population ne soit pas retenu dans la population intermédiaire de parents.

II.4.2.4. L'intensité de la sélection

L'intensité de la sélection I_s , (*selection intensity* ou *selection pressure*) définie par *Mühlenbein* et *Schlierkamp-Vosen*, mesure l'amplitude de la sélection sur la population [Mühlenbein 91]. Si l'intensité de la sélection d'une méthode est connue et si la performance des individus de la population à la génération t est distribuée selon une loi normale $N(\mu_t, \sigma_t)$, la valeur moyenne de l'adaptation des individus après sélection peut être estimée suivant (65).

$$\mu_{t+1} = \mu_t + \sigma_t I_s \quad (65)$$

Ce modèle suppose que la performance des individus suit une loi normale. En pratique, cela est vrai pour de nombreux problèmes car les opérateurs génétiques ont un effet de "normalisation" sur l'adaptation des individus [Mühlenbein 91][Miller 95a].

L'intensité de la sélection des principales techniques employées est donnée dans le Tableau 5.

Tableau 5. Intensité de sélection des principales techniques utilisées

Méthode de sélection	Paramètres	Intensité de la sélection I_s
sélection par tournoi	k	$\mu_{k:k}^2$
sélection (μ, λ)	μ, λ	$\frac{1}{\mu} \sum_{i=\lambda-\mu+1}^{\lambda} \mu_{i:\lambda}$
sélection linéaire par rang	η_+	$(\eta_+ - 1) \frac{1}{\sqrt{\pi}}$
sélection proportionnelle sélection stochastique universelle	μ_t, σ_t	$\frac{\sigma_t}{\mu_t}$

Signalons enfin que la sélection par tournoi de deux individus a pour intensité $I_s = \mu_{2:2} = \frac{1}{\sqrt{\pi}} \approx 0.5642$ (qui est identique à l'intensité de la sélection linéaire par rapport au rang avec $\eta_+ = 2$)

II.4.3. Effet de la sélection

L'importance du rôle de la sélection peut être mesurée par son effet sur la propagation d'un schéma H dans une population de N individus entre deux générations successives. On suppose que la sélection est de type proportionnelle (43) et que le nombre d'exemplaires du schéma H est égal à $m(H, t)$ à la génération t . A la génération $t+1$, le nombre attendu de représentants de H s'écrit :

$$m(H, t+1) = \frac{f(H)}{f_{moy}} m(H, t) \tag{66}$$

où $f(H)$ et f_{moy} désignent respectivement l'adaptation moyenne des individus possédant le schéma H à la génération t et la moyenne de la fonction d'adaptation calculée sur l'ensemble de la population à la même génération.

L'équation (66) montre que les schémas présents dans les individus de performance supérieure à la moyenne seront d'avantage copiés à la génération suivante. Inversement, les schémas se trouvant dans les individus de performance inférieure à f_{moy} seront moins dupliqués.

² $\mu_{i:j}$ désigne la statistique d'ordre [Miller 95a]

En supposant qu'à chaque génération, les représentants d'un schéma H ont une adaptation supérieure à la moyenne d'une quantité cf_{moy} , on peut réécrire l'équation d'évolution (66) de la façon suivante :

$$m(H, t + 1) = (1 + c)m(H, t) \quad (67)$$

En supposant que la constante c est stationnaire³ de la génération initiale à la génération t (67) devient :

$$m(H, t + 1) = (1 + c)^t m(H, 0) \quad (68)$$

La relation précédente montre que la sélection augmente (ou diminue) *exponentiellement* le nombre de schémas qui ont une performance en dessus (respectivement en dessous) de la moyenne au cours des premières générations.

La sélection oriente l'exploration vers les meilleurs individus déjà identifiés mais ne fait apparaître aucune configuration nouvelle. La reconnaissance d'autres points de l'espace est assurée par les opérateurs de recombinaison génétique.

II.5. LES OPERATEURS DE RECOMBINAISON GENETIQUE

Les opérateurs de recombinaison génétique garantissent la diversification des individus dans la population en assurant le brassage des matériels génétiques au cours des générations.

La recombinaison génétique notée R_g , modifie la population comme suit :

$$P^{n^t} = R_g(P^{t^t}) = (a_1^{n^t}, a_2^{n^t}, \dots, a_N^{n^t}) \quad (69)$$

Les deux opérateurs fondamentaux de la recombinaison génétique sont le croisement et la mutation.

II.5.1. Le croisement

Le *croisement* est un processus aléatoire de probabilité p_c appliqué séquentiellement à des couples de parents pris au hasard dans la population. Il consiste à échanger une partie du matériel génétique des parents pour former deux nouveaux individus (enfants).

³ Cette hypothèse n'est évidemment valable que lors des premières générations, la constante c diminuant pour tendre vers 0 lorsque l'algorithme a convergé

Chapitre II

Les parents après croisement peuvent être retirés de la population de reproducteurs (*croisement sans remplacement*) ou bien être gardés pour avoir une nouvelle chance de se reproduire (*croisement avec remplacement*). C'est la première solution qui est généralement adoptée.

II.5.1.1. Le croisement à un site

Considérons deux chaînes binaires de longueur l qui représentent les chromosomes de deux parents reproducteurs. Lorsque le croisement a lieu, un site de croisement est choisi entre les positions 1 et $l-1$ du chromosome. Le mécanisme de croisement consiste alors à échanger les gènes de chaque parent entre le site sélectionné et la position finale l des deux chaînes, comme le montre la Figure 21.

L'effet d'un opérateur de recombinaison génétique peut être perçu à travers la propagation des schémas comme lors de la sélection des parents. Il convient alors de déterminer le taux de *destruction* (ou de *survie*) d'un schéma quelconque sous l'opérateur considéré.

Pour le croisement à un site, la probabilité qu'un schéma de longueur utile $\delta(H)$ soit détruit est donnée par :

$$P_d(H) = \frac{\delta(H)}{l-1} \quad (70)$$

où l désigne la longueur du chromosome.

Si le croisement est appliqué avec une probabilité p_c , le taux effectif de destruction d'un schéma devient :

$$P_{d,eff}(H) \leq p_c P_d(H) \quad (71)$$

ce qui correspond à une probabilité de survie :

$$P_{s,eff}(H) \geq 1 - p_c P_d(H) \quad (72)$$

La relation précédente est générale et ne dépend pas du type de croisement. Dans le cas du croisement standard à un site, elle peut être réécrite en prenant en compte (70).

$$P_{s,eff}(H) \geq 1 - p_c \frac{\delta(H)}{l-1} \quad (73)$$

L'équation (73) montre que $P_{S,eff}(H)$ diminue lorsque la longueur utile du schéma augmente. Le croisement à un site préserve essentiellement des schémas de longueur utile courte.

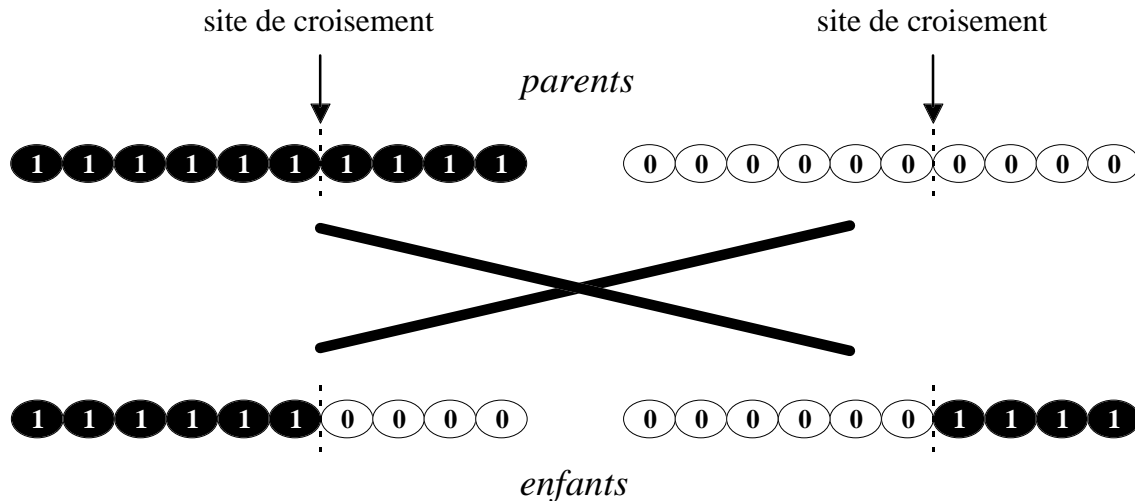


Figure 21. Principe du croisement à un site

II.5.1.2. Le croisement multi-sites

Le croisement peut être généralisé à n sites, la permutation des gènes étant effectuée entre chaque site. Si le nombre de sites est impair, on fixe un emplacement supplémentaire correspondant à la dernière position dans la chaîne comme pour le croisement à un site. Le mécanisme du croisement généralisé est illustré à la Figure 22.

Pour le croisement multi-sites, la probabilité de survie d'un schéma peut être exprimée récursivement par rapport à son ordre $o(H)$ et en fonction de sa longueur utile $\delta(H)$. Nous ne spécifions pas ici l'expression assez complexe de cette probabilité mais nous invitons le lecteur à se reporter à [Spears 90] et [De Jong 92] pour plus d'informations.

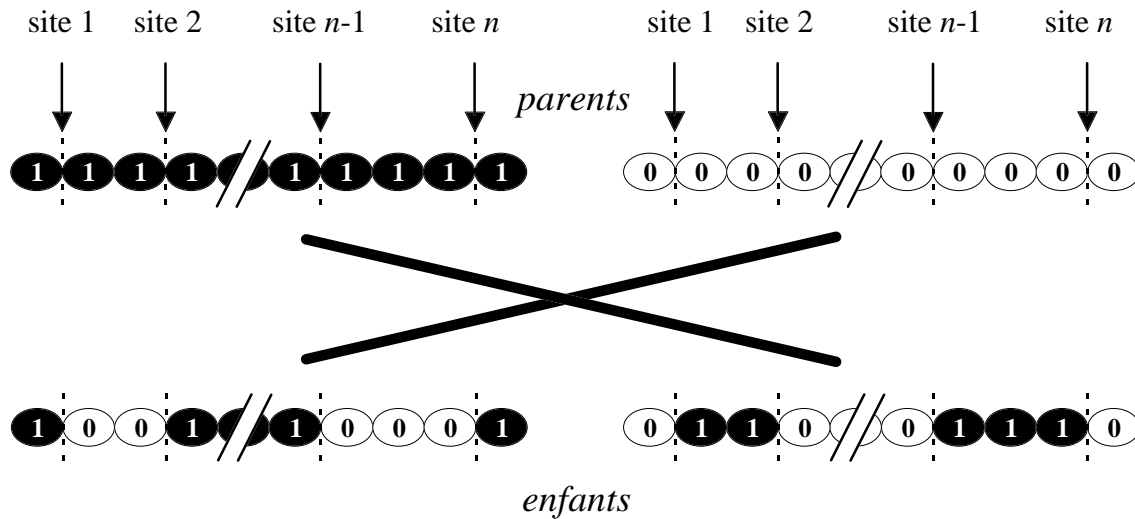


Figure 22. Principe du croisement multi-sites

II.5.1.3. Le croisement uniforme

Le croisement uniforme [Syswerda 89] est obtenu à partir d'un masque binaire initialisé aléatoirement et possédant un nombre de bits égal au nombre de gènes des individus de la population. Le premier enfant est créé en prenant les gènes du premier parent lorsque les bits correspondant dans le masque valent "1" et les gènes du deuxième parent si ces derniers valent "0". Le deuxième enfant s'obtient de la même manière en complémentant le masque.

La Figure 23 illustre le processus du croisement uniforme.

Bien que le nombre effectif de sites ne soit pas fixe à chaque recombinaison, on estime que le croisement uniforme produit une moyenne de $l/2$ sites, l étant toujours la longueur totale du chromosome.

La probabilité de survie d'un schéma soumis au croisement uniforme est donnée par :

$$P_s(H) = \frac{1}{2^{o(H)-1}} \quad (74)$$

Dans le cas du croisement uniforme, les taux de survie et de destruction d'un schéma particulier ne sont pas liés à sa longueur utile mais à son ordre. Il devient alors difficile de quantifier l'effet de cet opérateur.

Signalons pour finir que le masque de croisement est généralement construit de manière uniformément aléatoire en respectant l'équiprobabilité d'obtenir une valeur égale à 0 ou 1 pour chaque bit. Toutefois, un croisement uniforme paramétré fonction d'une probabilité

supplémentaire p_0 , liée à l'apparition d'une des deux valeurs de l'alphabet binaire, a été proposé pour augmenter l'efficacité de l'opérateur [Spears 91]. Lorsque $p_0 = 0.5$, on retrouve évidemment le croisement uniforme standard.

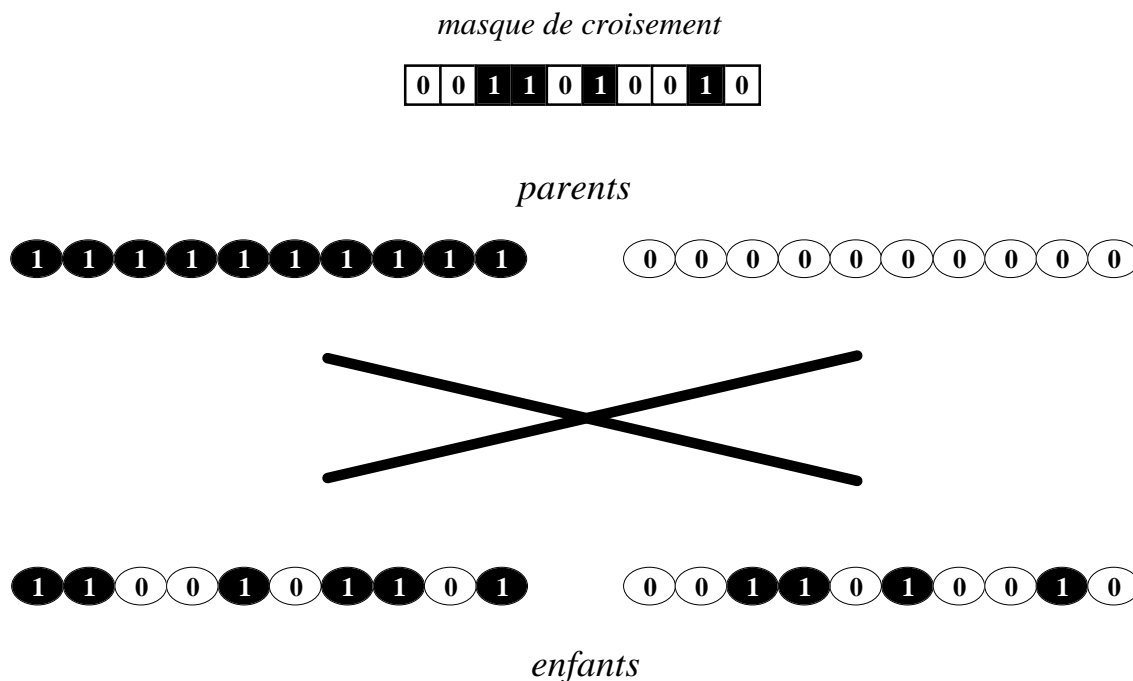


Figure 23. Principe du croisement uniforme

II.5.1.4. Quel type de croisement appliquer ?

Lequel des différents croisements présentés précédemment est le plus performant ? Cette question n'a à ce jour pas trouvé de réponse. Le débat qui doit permettre de trancher sur la meilleure méthode est toujours sujet à controverse.

Syswerda argumente en faveur du croisement uniforme [Syswerda 89]. Sous croisement uniforme, la probabilité de destruction d'un schéma d'ordre $o(H)$ quelconque est indépendante de sa longueur utile $\delta(H)$. Pour le croisement à un ou deux sites, c'est cette même longueur utile qui détermine le taux de destruction et non son ordre. D'après *Syswerda* cela signifie que les schémas de longueur utile importante sont comparativement moins perturbés que ceux de longueur utile faible. La quantité totale de schémas détruit est donc sensiblement plus faible.

Spears et *De Jong* soulignent au contraire les effets néfastes du croisement uniforme et prennent position en faveur du croisement à deux sites [Spears 90]. Toutefois, ils montrent les faiblesses de celui-ci en particulier lorsque la population a largement convergé. Le croisement à deux sites s'avère alors médiocre car incapable de faire apparaître des individus à

caractéristiques nouvelles. Le taux de production de nouveaux individus par le croisement uniforme est supérieur à celui obtenu pour tout autre type de croisement.

Ces constatations ont conduit *Spears* à la création d'un croisement adaptatif dont le mécanisme est le suivant [Spears 95] : un bit de décision est ajouté en dernière position du chromosome des individus de la population. Le croisement à appliquer est fixé en fonction de la valeur des bits de décision des deux parents reproducteurs :

- si les bits de décision des deux parents valent 0, ceux-ci sont recombinaés à l'aide d'un croisement uniforme.
- si les bits de décision des deux parents valent 1, ceux-ci sont recombinaés à l'aide d'un croisement à deux sites.
- si les bits de décision des deux parents diffèrent, le type de croisement à appliquer est fixé aléatoirement de façon équiprobable parmi le croisement uniforme et le croisement à deux sites.

Cette procédure est totalement adaptative dans la mesure où c'est l'algorithme lui-même qui décide du croisement à effectuer. Ce mécanisme, couplé à la sélection, favorise l'opérateur qui produit les enfants les mieux adaptés.

II.5.2. La mutation

II.5.2.1. Principe

La mutation est une altération aléatoire d'un gène d'un individu. L'opérateur de mutation consiste à compléter la valeur d'un bit du chromosome avec une probabilité p_m (voir Figure 24). Le processus est exécuté bit à bit.

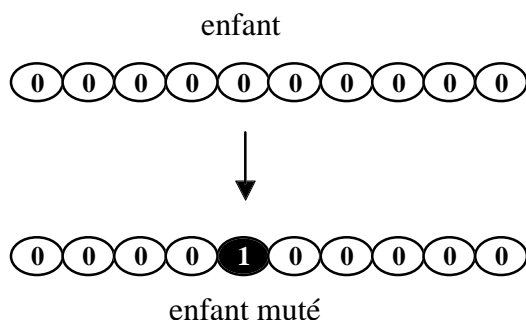


Figure 24. Principe de la mutation

La combinaison de la sélection et du croisement est normalement nécessaire pour assurer l'évolution de la population et la convergence de l'algorithme vers l'optimum. Mais, il arrive parfois que des informations essentielles contenues dans les gènes des individus disparaissent au cours des générations. Le rôle principal de la mutation est de protéger les individus contre cette perte irrémédiable de matériel génétique. Prise indépendamment, la mutation constitue un opérateur de recherche secondaire qui favorise l'apparition de génotypes nouveaux.

Pour qu'un schéma survive à l'opération de mutation, il faut que toutes ses positions instanciées soient conservées. Comme la probabilité de survie d'un allèle quelconque est $1 - p_m$ et que le nombre de positions fixes dans le schéma est $o(H)$, la probabilité de survie d'un schéma à la mutation est :

$$P_s(H) = (1 - p_m)^{o(H)} \quad (75)$$

Pour de très faibles taux de mutation, l'expression précédente peut être évaluée approximativement comme suit :

$$P_s(H) = 1 - o(H)p_m \quad (76)$$

II.5.2.2. Taux optimal de mutation et mutation adaptative

Pour ne pas détériorer les performances de l'algorithme génétique, *Goldberg* conseille une fréquence de mutation tous les 1000 bits [Goldberg 89]. D'autres préconisent plutôt la valeur suivante pour p_m [Mühlenbein 92][Bäck 93] :

$$p_m^* = \frac{1}{l} \quad (77)$$

où l est la longueur du chromosome.

Schaffer et al. proposent pour leur part une formule empirique qui exprime le taux optimal de mutation en fonction de la longueur du chromosome et de la taille de la population [Schaffer 89] :

$$\ln N + 0.93 \ln p_m^* + 0.45l = 0.56 \quad (78)$$

Le taux de mutation est déterminé approximativement à partir de (78) suivant

$$p_m^* = \frac{1}{N\sqrt{l}} \quad (79)$$

Signalons enfin que plusieurs méthodes de mutation adaptatives ont été développées. Curieusement, certaines augmentent le taux de mutation au cours des générations [Davis

85][Vasconcelos 97] alors que d'autres, au contraire, le diminuent. Ainsi il est courant de diminuer la probabilité de mutation exponentiellement en fonction du nombre de génération t [Fogarty 89][Hesser 91]. Par exemple,

$$p_m^*(t) = \sqrt{\frac{\alpha}{\beta}} \frac{e^{-\gamma t/2}}{N\sqrt{t}} \quad (80)$$

où α, β, γ sont des constantes [Hesser 91].

Une dernière technique adaptative proposée par *Th. Bäck* consiste à muter chaque paramètre de la fonction à optimiser avec une probabilité dépendant d'informations contenues dans les gènes des individus [Bäck 92a][Bäck 92b]. Pour cela, on ajoute au chromosome autant de sous-chaînes binaires qu'il y a de paramètres. Ces chaînes codent les taux de mutations associés aux paramètres (le mécanisme complexe de codage est décrit en détail dans [Bäck 92b]). Ce mode de mutation conduit à un brassage chromosomique important qui doit être contrôlé par des procédures de sélection extinctrices pour supprimer les "mauvais" mutants engendrés.

II.5.3. Le réordonnement des gènes

Les mécanismes de codage, de croisement et de mutation présentés précédemment ne sont pas forcément adaptés lorsque l'on cherche à résoudre des problèmes combinatoires (*order based problems*) comme la recherche d'itinéraires par exemple. L'exemple le plus connu est le cas du problème du voyageur de commerce (*Traveling Salesman Problem*) où un voyageur connaissant un ensemble de villes ainsi que les distances entre ces villes, doit découvrir le chemin le plus court de façon à visiter toutes les villes (sans passer deux fois dans la même ville!). Le codage binaire n'est pas très approprié à la résolution de ce type de problème. Il est beaucoup plus judicieux d'utiliser un codage par représentation ordinale ou par représentation de chemins [Goldberg 89]. Néanmoins cela impose par la suite la définition de procédures de recombinaison spécifiques, qualifiées de procédures de réordonnement (*ordering operators*). Nous invitons le lecteur à se reporter à [Kargupta 92] pour plus d'informations.

II.6. LA SELECTION DES INDIVIDUS D'UNE NOUVELLE GENERATION

A la suite de la recombinaison génétique, la population compte $2N$ individus (N parents et N enfants). Il faut donc éliminer N individus pour constituer la génération suivante. C'est le rôle de la sélection finale S_f , qui agit sur les populations de parents et d'enfants d'une génération pour créer la nouvelle génération.

$$P^{t+1} = S_f(P^t, P^{n^t}) = (a_1^{t+1}, a_2^{t+1}, \dots, a_N^{t+1}) \quad (81)$$

Pour effectuer cette sélection entre parents et enfants, plusieurs stratégies sont possibles.

II.6.1. La sélection par descendance

Avec la sélection par descendance ou sélection générationnelle (*Generational Selection*), il n'y a aucune compétition entre parents et enfants. La population de la nouvelle génération est obtenue par descendance, les enfants remplaçant automatiquement leurs parents quelle que soit leur adaptation.

$$P^{t+1} = S_f(P^t, P^{n^t}) = P^{n^t} \quad (82)$$

L'inconvénient de ce mode de sélection est que l'on risque de voir disparaître les caractéristiques génétiques des parents les mieux adaptés si elles n'ont pas été totalement transmises lors de la recombinaison génétique.

II.6.2. Les sélections élitistes

Les sélections élitistes au contraire préservent certains parents des effets destructeurs du croisement et de la mutation. La stratégie élitiste classique, couramment employée, garantit la survie du meilleur parent à chaque génération [De Jong 75].

D'autres variantes consistent à réaliser des compétitions entre parents et enfants pour déterminer les "survivants" de la génération (comme dans le cas de la sélection par tournoi). Les enfants peuvent par exemple être insérés dans la population si et seulement si leur performance est supérieure à celle de leurs parents à rang équivalent :

$$a_i^{t+1} = \begin{cases} a_i^{n^t} & \text{si } a_i^{n^t} > a_i^t \\ a_i^t & \text{sinon} \end{cases} \quad (83)$$

Il est aussi possible de retenir après croisement les deux meilleurs individus entre les deux parents et leurs deux descendants [Lobo 96]. On peut enfin choisir un plan d'élitisme sévère en gardant les N meilleurs individus parmi la population intermédiaire de parents et d'enfants.

Ces deux procédures de sélection sont qualifiées de stationnaires (*Steady State Selection*) car une partie seulement des individus de la population est remplacée à chaque génération.

Les stratégies élitistes améliorent considérablement les performances de l'algorithme pour certaines fonction mais s'avèrent assez inefficaces pour d'autres. Elles augmentent la vitesse

de convergence en favorisant l'exploitation et la recherche locale. Par contre, la disparition continue d'individus de faible performance conduit à un appauvrissement des génotypes qui peut entraîner une convergence prématurée vers un optimum local.

II.7. LA CONVERGENCE DES ALGORITHMES GENETIQUES

II.7.1. Analyse de la convergence

II.7.1.1. Le théorème des schémas

Nous pouvons obtenir approximativement l'effet cumulé de la sélection, du croisement (avec ou sans réordonnement) et de la mutation à partir des équations (66), (72) et (76) en supposant l'indépendance des différents opérateurs :

$$m(H, t + 1) \geq m(H, t) \frac{f(H)}{f_{moy}} (1 - p_c P_d(H) - o(H) p_m) \quad (84)$$

Cette équation représente l'évolution des schémas entre deux générations successives et caractérise le fonctionnement de l'algorithme génétique. Elle est plus connue sous le nom de théorème des schémas ou théorème fondamental des algorithmes génétiques [Holland 75][Goldberg 89]. Elle montre que l'efficacité de l'algorithme génétique réside dans son aptitude à recombinaison des schémas d'adaptation élevée pour former des individus de plus en plus performants. Le théorème des schémas peut être réécrit sous une forme plus générale [Kargupta 92] :

$$G = \frac{m(H, t + 1)}{m(H, t)} \geq \phi (1 - \xi) \quad (85)$$

où G est le taux de croissance d'un schéma, ϕ est représentatif de l'intensité de la sélection et ξ désigne le taux de destruction global par les différents opérateurs de recombinaison génétique.

La capacité de convergence de l'algorithme génétique vers l'optimum recherché est fortement liée à l'équilibre entre l'intensité de la sélection et le taux de destruction des schémas [Goldberg 93]. Avec une sélection exigeante (ϕ élevé) et une faible recombinaison (ξ faible) l'algorithme a de forte chance de converger vers une solution locale. De même, avec une sélection peu sévère (ϕ faible) et une forte recombinaison (ξ élevée), il risque de ne jamais pouvoir approcher l'optimum. Cet équilibre correspond au compromis entre l'exploration et l'exploitation que l'on retrouve dans toute méthode de recherche globale. Nous voyons ici, la

nécessité de contrôler les paramètres stochastiques des différents opérateurs génétiques afin de réaliser le meilleur compromis possible et d'augmenter l'efficacité de l'algorithme.

II.7.1.2. Parallélisme implicite

Pour atteindre l'optimum recherché, l'algorithme génétique doit être en mesure d'explorer de nombreuses régions de l'espace de recherche. Son efficacité est liée à son aptitude à créer, par recombinaison, de nouveaux schémas représentant les régions inexplorées. *Holland* a quantifié et estimé que le nombre de schémas effectivement traités à chaque génération était d'environ N^3 , N désignant la taille de la population. Ce résultat qu'il nomme *parallélisme implicite* montre la supériorité de l'exploration génétique par rapport aux techniques d'optimisation purement aléatoires [Holland 75][Bertoni 92]. Les algorithmes génétiques sont capables d'analyser en parallèle plusieurs solutions dans l'espace de recherche (voir Figure 17 et Figure 18) et d'orienter, grâce à la sélection, l'exploration vers les zones de forte adaptation. Ils évoluent vers l'optimum en recombinaison des schémas fortement adaptés mais de longueur utile faible (que l'on qualifie de "blocs de construction" ou *building blocs*).

II.7.2. Critères de convergence

Un aspect important des algorithmes génétiques, pourtant peu abordé à ce jour, concerne le choix du critère de convergence. Quand peut-on considérer que l'algorithme a convergé vers l'optimum global ?

Le problème est généralement éludé en imposant un nombre de génération maximal t_{\max} et en arrêtant la recherche lorsque $t = t_{\max}$. On estime alors que l'algorithme a convergé et que l'individu de plus forte performance dans la population $P^{t_{\max}}$ correspond à la solution recherchée.

Une méthode plus "rigoureuse" consiste à supposer que l'algorithme converge vers l'optimum lorsque l'adaptation d'une partie (ou de l'ensemble) de la population se rapproche de celle du meilleur individu (voir Figure 25). On peut considérer que cet événement se produit à la génération t pour laquelle,

$$1 - \frac{f_{\text{moy}}}{f_{\text{max}}} \leq \varepsilon \quad (86)$$

où ε est la précision requise sur la convergence, $f_{\text{max}} = f(a_i^t)$ la performance du meilleur individu de la population à la génération t , et f_{moy} la moyenne de l'adaptation calculée sur l'ensemble de la population ou sur une partie correspondant à un pourcentage des représentants les plus performants.

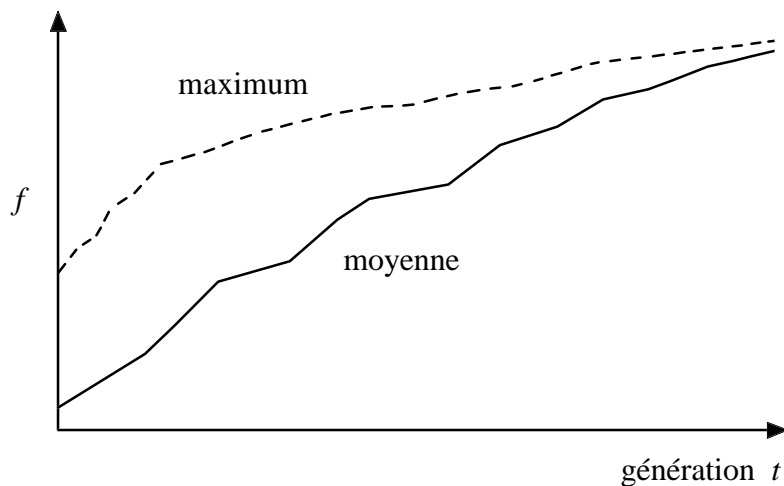


Figure 25. Convergence de l'algorithme génétique

Nous pouvons aussi supposer que l'algorithme a convergé lorsque le meilleur individu de la population n'évolue plus. Il suffit pour cela de tester la variation de sa performance sur une série de générations [Vasconcelos 97] :

$$\left| \Delta t f(a_1^t) - \sum_{i=1}^{\Delta t} f(a_1^{t-i}) \right| \frac{1}{\varepsilon + f(a_1^t)} \leq \varepsilon \quad (87)$$

Ces critères ne sont pourtant absolument pas fiables dans la mesure où l'algorithme peut converger et se stabiliser autour d'une solution dans l'attente d'une mutation qui le dirigera vers une autre région plus intéressante. Le choix d'un critère d'arrêt idéal reste donc sans réponse.

II.7.3. Taux de réussite

En raison du caractère stochastique de l'exploration génétique, il est habituel d'exécuter plusieurs fois le même algorithme sur un même problème test. Un taux de succès (ou taux de réussite) rend compte de son efficacité. On peut par exemple définir le taux de réussite comme le rapport du nombre de fois où l'algorithme a convergé vers la solution optimale sur le nombre total d'exécutions. La dispersion de résultats, due à la nature stochastique des algorithmes, peut être prise en compte au moyen de grandeurs statistiques classiques (écart-type, facteur de dispersion, intervalles de confiance, ...)

II.8. HYBRIDATION DES ALGORITHMES

L'efficacité de la recherche génétique peut être augmentée par hybridation des algorithmes avec des méthodes de recherche locale déterministe de type gradient ou

heuristique [Vasconcelos 94]. Toutefois, l'hybridation pose un problème identique à celui de la convergence, exposé au paragraphe II.7.2. Il est très délicat de déterminer le moment opportun où l'on va commuter de l'algorithme génétique global vers la méthode d'optimisation locale. Des critères similaires à ceux présentés au paragraphe II.7.2 peuvent être employés pour fixer l'instant de commutation, mais ceux-ci n'offrent aucune garantie d'éviter la convergence prématurée vers une solution locale. Si l'algorithme hybride est globalement plus robuste et précis, rien n'indique qu'il surpasse l'algorithme génétique élitiste classique (non hybridé) dans tous les cas [Lobo 96].

II.9. ALGORITHMES GENETIQUES A CODAGE REEL

II.9.1. Codage des paramètres

Le théorème fondamental des schémas [Holland 75], l'hypothèse des blocs de constructions (building blocs) [Goldberg 89] et la notion de problèmes trompeurs [Goldberg 91][Goldberg 92] constituent les bases théoriques de l'algorithme génétique à codage binaire. Pourtant, l'apparition du codage réel remet aujourd'hui en cause ces principes fondamentaux.

Les algorithmes génétiques à codage réel sont similaires à ceux présentés précédemment, tant en ce qui concerne leur structure que leurs mécanismes de sélection. Par contre, ils sont caractérisés par des opérateurs de recombinaison génétiques qui utilisent directement les valeurs des paramètres réels pour évoluer vers l'optimum. Le chromosome des individus est donc représenté par les n variables objets dont l'intervalle de variation est défini par les contraintes de domaine du problème.

II.9.2. Le croisement

Deux types de croisement différents sont employés par les algorithmes génétiques à codage réel :

- Le croisement *discret* est assez similaire aux méthodes de croisement utilisées dans le cas du codage binaire. Il consiste à permuter les variables de chaque parent entre des sites du chromosome choisis aléatoirement. Le croisement discret peut être à un site, multi-sites ou uniforme comme le montre la Figure 26. Cette technique a un taux d'exploration relativement limité car elle ne fait apparaître aucune valeur nouvelle pour les paramètres.

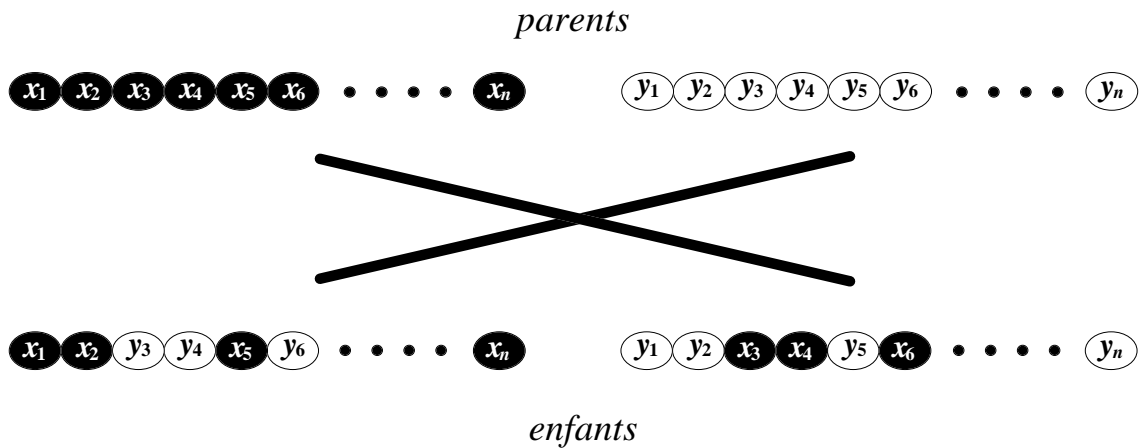


Figure 26. Principe du croisement discret

- Le croisement *continu* assure l'exploration de l'espace en testant des nouvelles valeurs de variables objets. Il est appliqué paramètre par paramètre en utilisant le mécanisme suivant :

$$\begin{cases} x'_i = x_i + a_1(y_i - x_i) \\ y'_i = y_i + a_2(x_i - y_i) \end{cases} \quad (88)$$

où x_i et y_i désignent les paramètres (de même rang) des deux parents, x'_i et y'_i étant les paramètres (de rang correspondant) des deux enfants obtenus. a_1 et a_2 sont deux variables prises dans l'intervalle $[0,1]$. Lorsque a_1 et a_2 sont deux variables aléatoires distinctes, le croisement est continu *asymétrique*. Lorsque a_1 et a_2 désignent une même valeur aléatoire, le croisement est continu *symétrique*. Enfin, si $a_1 = a_2 = 1/2$, le croisement est continu *arithmétique*. Il conduit alors à un unique enfant représentant une moyenne arithmétique des parents. Dans, ce cas il est nécessaire d'effectuer un nombre de croisement deux fois plus important si nous voulons obtenir autant d'enfants que de parents à chaque génération.

II.9.3. La mutation

La mutation consiste à perturber une variable du problème. Elle est généralement appliquée paramètre par paramètre avec des probabilités plus élevées que pour le codage binaire standard. Comme pour l'opérateur de croisement, plusieurs modes de mutation sont possibles :

- La mutation *uniforme* ou mutation *globale* modifie un paramètre en choisissant une nouvelle valeur de façon uniformément aléatoire dans l'intervalle défini par les contraintes de domaine.

- La mutation *non uniforme* ou mutation *locale* consiste à perturber de façon locale la valeur d'un paramètre. On peut par exemple ajouter ou retrancher à la variable aléatoire Δx dont l'amplitude diminue avec le nombre de générations [Saludjian 97] :

$$x'_i = x_i \pm \Delta x_i \quad (89)$$

Δx peut aussi être une variable à densité gaussienne dont l'amplitude est adaptée à chaque génération, comme pour les stratégies d'évolution [Schwefel 95] ou la programmation évolutionniste [Bäck 96].

Les opérateurs de mutation locale confèrent à l'algorithme génétique à codage réel un mécanisme de recherche locale (*Hillclimbing*) implicite qui évite d'avoir recours à l'hybridation avec une méthode d'optimisation déterministe. Contrairement à l'algorithme génétique à codage binaire, la mutation joue cette fois un rôle prépondérant dans l'exploration des nouvelles solutions.

II.9.4. L'algorithme génétique recombinateur (BGA)

Un des inconvénients des opérateurs de mutation présentés précédemment réside dans le fait qu'il est nécessaire d'adapter l'amplitude des perturbations appliquées aux variables au cours des générations. Cette adaptation n'est pas toujours facile et le choix de l'amplitude initiale est souvent problématique. *Mühlenbein* et *Schlierkamp-Voosen* ont mis au point un algorithme robuste qui s'affranchit de ce problème [Mühlenbein 93]. L'algorithme génétique "recombinateur" (*Breeder Genetic Algorithm, BGA*) possède un opérateur δ qui régit les perturbations des paramètres aussi bien dans le cas du croisement que dans la mutation. Une variable mutée s'écrit :

$$x'_i = x_i \pm \Delta x_i \cdot \delta \quad (90)$$

où Δx_i est généralement fixée à la moitié du domaine de définition de la variable correspondante. Nous avons pour notre part étendu le pouvoir d'exploration de l'algorithme en imposant $\Delta x_i = x_{i \max} - x_{i \min}$. Le signe \pm dans l'équation (90) est choisi avec une probabilité de 0.5. δ est calculé à partir d'une distribution qui favorise les valeurs faibles. La mutation *discrète* utilise un facteur δ défini de la façon suivante :

$$\delta = \sum_{i=0}^{k-1} \alpha_i \cdot 2^{-i} \quad \alpha_i = \{0,1\} \quad (91)$$

où k est le coefficient de précision (généralement k est pris égal à 16). α_i est fixée à 1 avec une probabilité $p_\delta = 1/k$ dans le terme somme. En moyenne il y a juste une valeur j

($0 \leq j \leq k - 1$) pour laquelle $\alpha_i = \alpha_j \neq 0$. On a alors $\delta = 2^{-j}$. Pour de nombreuses raisons, il est plus intéressant d'utiliser une mutation *continue* basée sur le même principe. Nous pouvons par exemple, exprimer le facteur δ de la façon suivante [Schlierkamp 94] :

$$\delta = 2^{-k \cdot \alpha} \quad (92)$$

où α est une variable aléatoire à densité uniforme dans l'intervalle $[0,1]$.

La recombinaison en ligne (*BGA line recombination*) permet de tester de nouveaux points dans une direction définie par deux parents. Soient $\mathbf{x} = (x_1, x_2, \dots, x_n)$ et $\mathbf{y} = (y_1, y_2, \dots, y_n)$ les chromosomes de deux parents (\mathbf{x} étant le parent le *mieux adapté*), l'enfant résultant \mathbf{x}' aura des composantes caractérisées par :

$$x'_i = x_i \pm \Delta x_i \cdot \delta \cdot \frac{y_i - x_i}{\|\mathbf{x} - \mathbf{y}\|} \quad (93)$$

où $\|\mathbf{x} - \mathbf{y}\|$ désigne la norme euclidienne du vecteur $\mathbf{x} - \mathbf{y}$, δ et Δx_i étant définis comme précédemment. *Mühlenbein* et *Schlierkamp Voosen* préconisent de fixer le choix du signe "-" dans l'équation (93) avec une probabilité de 0.9 pour favoriser l'exploration dans la direction descendante. Néanmoins, nous avons décidé de respecter l'équiprobabilité de recherche dans les deux directions comme dans le cas de la mutation, pour limiter les tendances et les biais éventuels et se démarquer des croisements de type continu utilisés avec le codage réel. En effet, les opérateurs de croisement que nous avons définis au paragraphe précédent, ne sont pas "étendus", c'est à dire qu'ils recherchent uniquement des solutions à l'intérieur d'un hypercube ayant les deux parents comme sommets extrêmes. La recombinaison en ligne est similaire à un croisement étendu (extended recombination [Mühlenbein 93]) dans la mesure où des points peuvent être testés en dehors de cet hypercube.

Si la nouvelle valeur du paramètre se trouve à l'extérieur du domaine de recherche, aussi bien pour la mutation que pour le croisement, celle-ci est modifiée et positionnée sur la frontière la plus proche ($x'_i = x_{i \max}$ ou $x'_i = x_{i \min}$).

II.9.5. Performances comparées

Nous présentons dans cette partie une comparaison des performances de l'algorithme génétique binaire avec celles d'un algorithme génétique à codage réel classique et du BGA. La démarche est identique à celle adoptée pour déterminer les opérateurs génétiques les plus efficaces dans le cas de l'algorithme génétique binaire [Annexe III].

- Le BGA est légèrement différent de celui employé par *Mühlenbein* et *Schlierkamp-Voosen*. Il utilise une procédure de sélection par tournoi de deux individus, un élitisme classique (préservation du meilleur individu) et des opérateurs de recombinaison analogues à ceux définis au paragraphe précédent (croisement discret, recombinaison en ligne et mutation continue caractérisées par $k=16$ et $\Delta x_i = x_{i_{\max}} - x_{i_{\min}}$). Les probabilités de croisement et de mutation sont respectivement $p_c = 1$ et $p_m = 0.03$.
- L'algorithme génétique à codage réel est utilisé avec une sélection par tournoi de deux individus, suivie d'une recombinaison discrète, d'un croisement continu symétrique et d'une mutation analogue à celle définie pour le BGA. Les probabilités de croisement et de mutation sont respectivement $p_c = 1$ et $p_m = 0.03$.
- L'algorithme génétique à codage binaire est appliqué avec les opérateurs de sélection et de recombinaison préconisés à l'Annexe III. Il utilise une sélection stochastique universelle, un croisement uniforme et un mécanisme d'élitisme classique (préservation du meilleur individu). Les chromosomes des individus sont codés en binaire avec décodage dynamique. La longueur de codage adoptée pour chaque paramètre est celle recommandée par *De Jong* [De Jong 75] (voir Annexes II). Les probabilités de croisement et de mutation sont respectivement $p_c = 1$ et $p_m = 0.001$.

Le Tableau 6 présente une comparaison des résultats obtenus avec les différents algorithmes génétiques sur les problèmes de la suite de *De Jong* augmentée de la fonction de *Rastrigin*. La taille de la population est fixée à 30 et le nombre de générations à 1000. L'algorithme est arrêté en cas de convergence à 10^{-3} près par rapport à l'optimum réel. 100 tests sont effectués pour tenir compte de l'aspect stochastique de la recherche. Les grandeurs moyennes des différents critères de performance définis à l'Annexe III sont présentés dans le Tableau 6. L'écart-type correspondant est spécifié entre parenthèses.

Les résultats montrent une supériorité des algorithmes génétiques à codage réel, en particulier pour les fonctions F_5 et F_6 . Le BGA s'affirme d'ailleurs comme le meilleur des trois algorithmes en présentant une vitesse de convergence et un taux de réussite extrêmement élevés pour la plupart des fonctions testées. Seule la fonction F_4 a vraiment posé des problèmes aux divers algorithmes qui ont été incapables de converger précisément vers l'optimum avec la taille de population et le nombre de générations considérés. Il s'avère que le nombre maximal d'évaluations permis est largement insuffisant pour parvenir à l'optimum. En effet, l'algorithme doit d'abord minimiser en présence de bruit les 30 paramètres de la fonction avant de pouvoir minimiser le bruit lui-même. Le faible gradient autour de l'optimum et le bruit ajouté confèrent à ce problème une difficulté particulière.

Chapitre II

Tableau 6. Comparaison d'algorithmes génétiques standards - le nombre d'évaluations de fonction et la solution optimale sont moyennés sur 100 tests ; l'écart-type correspondant est précisé entre parenthèses

Type de GA	Taux de réussite	Nombre d'évaluations de fonction	solution optimale	Meilleure solution trouvée
<i>Fonction F_1 (sphère $n=3$)</i>				
GA Binaire	100%	1460 (856)	0.00057 (0.00032)	0.00002
GA réel	100%	742 (522)	0.00056 (0.00031)	0.00000
BGA	100%	499 (129)	0.00060 (0.00027)	0.00000
<i>Fonction F_2 (Rosenbrock $n=2$)</i>				
GA Binaire	44%	2320 (4873)	0.01384 (0.02836)	0.00002
GA réel	12%	243 (91)	0.09382 (0.08976)	0.00000
BGA	100%	3497 (2097)	0.00065 (0.00030)	0.00000
<i>Fonction F_3 (paliers $n=5$)</i>				
GA Binaire	100%	1022 (953)	0 (0)	0
GA réel	100%	1400 (351)	0 (0)	0
BGA	100%	256 (45)	0 (0)	0
<i>Fonction F_4 (bicarrée bruitée $n=30$)</i>				
GA Binaire	0%	-	0.01892 (0.00715)	0.00449
GA réel	0%	-	0.01320 (0.00458)	0.00469
BGA	0%	-	0.02732 (0.00905)	0.01033
<i>Fonction F_5 (trous de renard $n=2$)</i>				
GA Binaire	26%	843 (1506)	2.37706 (2.27600)	0.99800
GA réel	97%	9011 (6 277)	1.02799 (0.17039)	0.99800
BGA	100%	671 (410)	0.99822 (0.00027)	0.99800
<i>Fonction F_6 (Rastrigin $n=10$)</i>				
GA Binaire	0%	-	15.0549 (4.96895)	4.49136
GA réel	99%	14 384 (5060)	0.01038 (0.09945)	0.00000
BGA	90%	21 674 (4402)	0.00089 (0.00020)	0.00037

II.10. CONCLUSION

Dans ce chapitre nous avons établi les bases fondamentales de l'algorithme génétique standard. Nous avons aussi présenté certaines de ses limitations et souligné l'intérêt du codage réel. Toutefois, nous continuerons à exploiter par la suite les algorithmes génétiques "binaires" car ceux-ci sont à l'origine des méthodes multimodales que nous allons exposer dans le chapitre suivant. Ces méthodes constituent une alternative intéressante à l'hybridation et aux techniques adaptatives complexes. Nous verrons qu'elles permettent à la fois une amélioration de la recherche locale et une augmentation de l'exploration globale, qui évite aux algorithmes d'être piégés par des optima locaux et de converger prématurément.

ALGORITHMES GENETIQUES MULTIMODAUX ET METHODES DE NICHAGE

Maintenir une grande variété de génotypes dans la population est nécessaire pour augmenter les chances de succès de l'exploration génétique. La diversité génétique aide les individus à s'adapter rapidement aux changements de l'environnement et permet à la population de continuer à chercher des solutions constamment meilleures, évitant ainsi le piège des optima locaux. Pour les algorithmes génétiques standards unimodaux, il est difficile de préserver cette diversité car les individus de performance élevée sont dupliqués de façon exponentielle par l'opérateur de sélection. La reproduction entraîne inévitablement une convergence rapide vers les régions optimales mais conduit parallèlement à un appauvrissement de la population. En effet, lorsque l'algorithme a convergé tous les individus de la population (excepté quelques mutants) sont identiques - ils représentent normalement la solution globale du problème en cas de succès. Cette convergence vers une unique solution est restrictive et frustrante car la présence d'un seul individu au voisinage de l'optimum global serait suffisante. Les autres membres de la population pourraient être répartis dans d'autres régions de l'espace et donner ainsi de multiples renseignements utiles sur la forme de la fonction objectif. Par ailleurs, la traduction d'un problème d'optimisation réel avec tous ses aspects (performance "pure" mais aussi sensibilité, facilité de fabrication, prix de revient, ...) sous forme d'une fonction à optimiser n'est pas toujours une chose simple. Dès lors, le concepteur apprécie lorsqu'il étudie l'un des aspects d'être conduit à plusieurs possibilités (plus ou moins parfaites suivant cet aspect là) plutôt qu'à une solution unique. Au lieu de chercher un unique jeu de paramètres correspondant à l'optimum global de la fonction objectif, on en vient ainsi à l'idée de localiser toutes les régions de l'espace des paramètres dans lesquelles cette fonction a de "bonnes" valeurs. Les méthodes de nichage ont été introduites dans les algorithmes génétiques à cet effet, afin de limiter la dérive génétique due à la sélection et d'explorer en parallèle des solutions locales ou globales situées dans des régions éloignées de l'espace.

III.1. METHODES DE NICHAGE

Les méthodes de nichage reposent sur une analogie entre les domaines de recherche en optimisation et les écosystèmes naturels. Dans la nature, les animaux luttent pour leur survie en chassant ou en exploitant les ressources de leur milieu. Chaque espèce évolue de façon à

Chapitre III

remplir une niche écologique. Une espèce représente un groupe d'organismes identiques de caractéristiques biologiques semblables. Dans chaque niche, les ressources naturelles sont limitées et doivent être partagées entre les représentants des espèces qui la peuplent.

Par analogie, les méthodes de nichage des algorithmes génétiques multimodaux tentent d'induire l'émergence naturelle de niches et d'espèces dans l'environnement (l'espace de recherche). Une niche se réfère typiquement à un optimum de la fonction objectif, la qualité de l'optimum représentant la ressource de la niche. Les espèces sont constituées par des groupes d'individus similaires. La mesure de la similarité entre individus est effectuée à partir d'un critère de distance et d'un seuil de dissimilarité (ou seuil de voisinage).

La Tableau 7 résume les correspondances entre le vocabulaire utilisé en optimisation mathématique et les termes homologues employés pour les écosystèmes naturels.

Tableau 7. Analogie entre les espaces mathématiques et la nature

<i>Optimisation</i>	<i>Nature</i>
espace de recherche	écosystème
fonction objectif	ressources
optima	niches
configurations similaires	espèces
distance mathématique	différence biologique

Nous pouvons séparer les méthodes de nichage en deux groupes fondamentaux. Le premier groupe se compose des techniques caractérisées par un voisinage explicite, celles-ci nécessitant un seuil de dissimilarité fixé pour chaque niche. Le second au contraire est constitué par des procédés de nichage à voisinage implicite. Le seuil de dissimilarité n'est pas explicitement requis. La similarité est définie de façon floue en comparant plusieurs individus.

Le Tableau 8 présente les principales méthodes de nichage selon la classification établie précédemment. Nous développerons en détail la plupart de ces méthodes dans les paragraphes suivants. Le lecteur pourra se reporter à [Smith 93a][Smith 93b][Chun 98] pour de plus amples informations sur les Algorithmes Immunitaires (*Immune Algorithms*) et à [Davidor 91][Mahfoud 95] pour une introduction aux Algorithmes Génétiques Ecologiques (*Eco GAs*).

Tableau 8. Les principales méthodes de nichage

<i>Techniques de nichage à voisinage explicite (seuil de dissimilarité entre individus fixé)</i>	<i>Techniques de nichage à voisinage implicite (seuil de dissimilarité entre individus flou)</i>
Les méthodes de partage	Les méthodes de surpeuplement
La méthode d'éclaircissement	Les Algorithmes Génétiques Ecologiques
Les Algorithmes Génétiques Immunitaires	La technique des sous-populations

III.2. CRITERES DE DISTANCE

La distinction des niches dans l'espace de recherche se fait à partir d'un critère de distance qui permet différencier les individus similaires des individus non similaires. La similarité entre individus peut être basée sur une comparaison de leur génotype ou de leur phénotype. La similarité définie au niveau du génotype est liée à la représentation binaire du chromosome des individus. La similarité définie au niveau du phénotype repose directement sur les paramètres réels des individus dans l'espace de recherche.

III.2.1. Distance génotypique

La distance génotypique standard est la distance de Hamming. Cette dernière est définie de la manière suivante, pour deux individus a_1 et a_2 de chromosomes respectifs b_1 et b_2 de longueur l :

$$d(a_1, a_2) = \sum_{i=1}^l b_1(i) \oplus b_2(i) \tag{94}$$

où $b_{1,2}(i)$ désigne le i ème bit du chromosome $b_{1,2}$ et \oplus l'addition modulo 2 (ou exclusif).

La distance de Hamming donne le nombre total de bits qui diffèrent position par position entre les deux chromosomes des individus. Cette distance n'est pas toujours très appropriée au codage des paramètres. En effet, une variation d'un bit entre deux individus conduit à la même distance quelle que soit la position du bit dans la chaîne binaire. Or, le décodage des chromosomes est généralement fonction de la position des bits et fait qu'une variation d'un bit en début de chaîne n'a pas le même effet qu'une variation d'un bit en fin de chaîne. Dans ce cas, la distance de Hamming n'est pas forcément représentative de la similarité dans l'espace réel. C'est pourquoi pour certains types de codage, il est plutôt recommandé d'utiliser une distance liée au phénotype des individus [Deb 89].

III.2.2. Distances phénotypiques

Les distances phénotypiques sont basées sur des comparaisons directes entre les paramètres réels des individus dans l'espace de recherche. Nous pouvons par exemple définir un ensemble de distances d^β normalisées, variant entre 0 et 1,

- $d^\beta(a_1, a_2) = 1$ correspondant à la distance maximale possible entre deux individus a_1 et a_2 de l'espace (dissimilarité maximale).
- $d^\beta(a_1, a_2) = 0$ étant obtenue lorsque les individus a_1 et a_2 sont identiques (similarité maximale).

L'expression générale du critère de distance paramétrique est donnée par l'équation (95).

$$d^\beta(a_1, a_2) = \frac{1}{n^{1/\beta}} \left(\sum_{i=1}^n |\Delta x_i|^\beta \right)^{1/\beta} \quad (95)$$

$$\text{où } \Delta x_i = \frac{x_{1i} - x_{2i}}{x_{i \max} - x_{i \min}} \quad (96)$$

avec

- n nombre de paramètres (ou dimension de l'espace).
- x_{1i} valeur du i ème paramètre de l'individu a_1
- x_{2i} valeur du i ème paramètre de l'individu a_2
- $x_{i \min}$ valeur minimale du i ème paramètre dans l'espace.
- $x_{i \max}$ valeur maximale du i ème paramètre dans l'espace.
- Δx_i écarts normalisés entre paramètres.
- $\beta > 0$ facteur caractérisant la forme de la niche.

Nous constatons alors que $\beta = 1$ donne pour d^β la valeur moyenne des écarts normalisés entre paramètres et $\beta = 2$ correspond à la distance euclidienne moyenne.

Nous définissons par ailleurs une distance d^∞ correspondant au cas $\beta \rightarrow \infty$ de la façon suivante:

$$d^\infty(a_1, a_2) = \max_{i=1..n} \left| \frac{x_{1i} - x_{2i}}{x_{i \max} - x_{i \min}} \right| \quad (97)$$

Cette distance représente le maximum des écarts normalisés suivant toutes les directions.

La Figure 27 montre les formes de niches obtenues selon la valeur du paramètre β pour un domaine 2D, une niche étant caractérisée par une isodistance $d^\beta = \text{Cte}$. Plus la valeur de β est importante, plus le critère de similarité entre individus est sévère et plus la niche est restreinte.

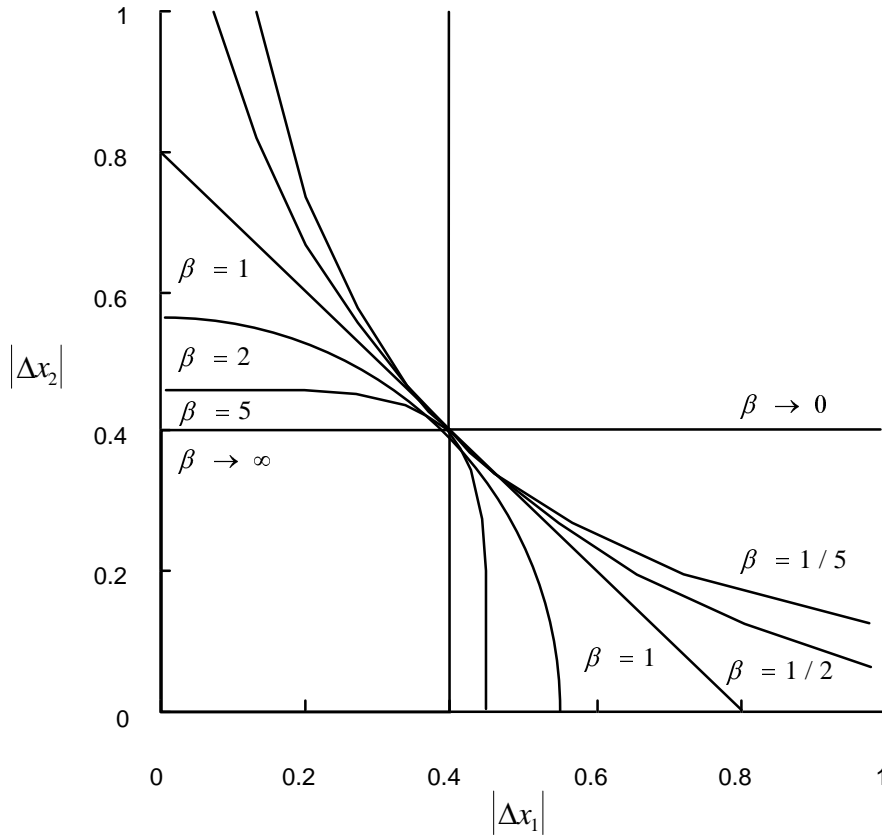


Figure 27. Formes des niches pour un domaine 2D

III.2.3. Distance non paramétrique

Pour des problèmes d'optimisation réels, il est possible d'utiliser un critère de distance qui ne dépend pas explicitement de la valeur des paramètres. La distance peut par exemple être exprimée en fonction d'une grandeur physique qui caractérise les configurations à optimiser. Dans le cas particulier d'un problème d'optimisation vectoriel, il est même intéressant de choisir un critère relatif aux objectifs pour pouvoir identifier la frontière de *Pareto* [Horn 93].

III.3. LES METHODES DE SURPEUPLEMENT

Les méthodes de surpeuplement (*Crowding Methods*) aboutissent à la formation implicite de niches en restreignant les compétitions entre individus au sein de la population globale. Seuls des individus ayant des caractéristiques génétiques les plus proches possibles sont autorisés à lutter pour leur survie.

III.3.1. Le surpeuplement standard

Le surpeuplement initial (*Crowding*) introduit par *De Jong* en 1975 [De Jong 75] est globalement identique à un algorithme génétique standard excepté qu'une partie de la population seulement, est reproduite et meure à chaque génération. Un pourcentage G (*Generation Gap*) fixe la fraction de la population qui sera soumise aux opérateurs génétiques. $G \times N$ individus sont choisis par une méthode de sélection proportionnelle puis subissent les opérations de croisement et de mutation. Les $G \times N$ enfants résultants sont introduits dans la population globale de la façon suivante : pour chaque enfant à insérer, CF individus sont pris au hasard dans la population globale de parents - CF est le facteur de surpeuplement (*Crowding Factor*). L'enfant remplace alors l'élément le plus proche présent dans l'échantillon des CF parents.

Initialement, le critère de similarité utilisé par *De Jong* était la distance de Hamming. Or, comme nous l'avons précédemment signalé, cette distance n'est pas toujours très appropriée. Le codage binaire standard utilisé par *De Jong*, conjointement à cette distance, conduisit à des erreurs stochastiques de remplacement, certains parents étant remplacés par des enfants appartenant à une autre niche. Ceci expliqua en partie l'échec de la méthode pour les fonctions testées. Néanmoins, le surpeuplement standard se révèle incapable de maintenir un grand nombre de niches, même lorsque la proximité est exprimée par une distance liée au phénotype [Deb 89][Mahfoud 95]

III.3.2. La présélection

Cinq ans avant le travail de *De Jong*, *Cavicchio* décrit un mécanisme qu'il nomme présélection et qui a pour but de préserver la diversité des individus [Cavicchio 70]. *Cavicchio* pense qu'il est beaucoup trop coûteux de comparer les caractéristiques génétiques des nouveaux individus créés avec tous les autres éléments de la population. En remarquant que les individus les plus proches des enfants à insérer sont la plupart du temps leurs parents ascendants directs, *Cavicchio* applique la méthode suivante pour conserver des solutions multiples : si un enfant a une adaptation supérieure à celle du plus "mauvais" parent qui l'a engendré, il remplace ce dernier. Dans le cas contraire, il n'est pas introduit dans la population. Tout comme la technique de *De Jong*, la présélection s'est montrée inefficace pour l'optimisation de fonctions multimodales [Mahfoud 95].

III.3.3. Le surpeuplement déterministe

Constatant l'échec des deux méthodes précédentes, *Mahfoud* propose une nouvelle technique qualifiée de *surpeuplement déterministe* (*Deterministic Crowding*) [Mahfoud 95]. Celle-ci s'inspire directement de la présélection de *Cavicchio* en donnant lieu à des compétitions entre parents et enfants descendants. Le mécanisme du surpeuplement déterministe est le suivant :

Dans une population comprenant N individus, $N/2$ paires de parents sont constituées au hasard. Chaque paire est soumise aux opérations de croisement et de mutation. Les enfants entrent ensuite en compétition avec les ascendants directs qui leur ont donné naissance selon deux séries de tournoi possibles. La méthode retient la série qui conduit à des compétitions entre parents et enfants descendants les plus proches.

Un pseudocode du surpeuplement déterministe est donné ci-après,

*pour t générations
faire $N/2$ fois*

1. Sélectionner 2 parents p_1 et p_2 de façon uniformément aléatoire sans remplacement.
 2. Appliquer l'opérateur de croisement (et éventuellement la mutation) pour créer deux enfants e_1 et e_2 .
 3. si $[d(p_1, e_1) + d(p_2, e_2)] \leq [d(p_1, e_2) + d(p_2, e_1)]$
 - si $f(e_1) > f(p_1)$, remplacer p_1 par e_1 .
 - si $f(e_2) > f(p_2)$, remplacer p_2 par e_2 .
- sinon*
- si $f(e_1) > f(p_2)$, remplacer p_2 par e_1 .
 - si $f(e_2) > f(p_1)$, remplacer p_1 par e_2 .

où $f(a_i)$ désigne la performance de l'individu a_i et $d(a_i, a_j)$ la distance entre les individus a_i et a_j .

III.3.4. La sélection par tournoi restreint

La sélection par tournoi (*Tournament Selection*) est comme nous l'avons signalé au chapitre II, une méthode de sélection qui permet de choisir les individus qui pourront se reproduire pour créer la génération suivante. A chaque génération, une compétition a lieu entre des individus pris aléatoirement dans la population et les vainqueurs de ces tournois forment les parents de la nouvelle génération. Ce mécanisme est particulièrement bien approprié à l'optimisation unimodale mais se révèle inefficace lorsque l'on désire maintenir plusieurs solutions, en raison de l'importante dérive génétique qu'il introduit. De ce fait, *Harik* a adapté la sélection par tournoi à l'optimisation multimodale et proposé une technique de nichage appelée sélection par tournoi restreint (*Restricted Tournament Selection*) dont le principe est le suivant [Harik 95] :

- 2 individus p_1 et p_2 sont sélectionnés aléatoirement dans la population.
- Ces individus, croisés et mutés, produisent deux enfants e_1 et e_2 .

Chapitre III

- w éléments sont pris au hasard dans la population globale. A partir de cet échantillon, deux sont sélectionnés, p_1' et p_2' , de telle sorte que leurs caractéristiques génétiques soient les plus proches des enfants respectifs e_1 et e_2 précédemment créés.
- Une compétition a ensuite lieu entre (p_1' et e_1) et (p_2' et e_2). Les vainqueurs des tournois sont réinsérés dans la population.

La procédure est répétée $N/2$ fois par génération.

La sélection par tournoi restreint s'apparente à une méthode de surpeuplement. La taille de la fenêtre de nichage w joue le même rôle que le facteur de surpeuplement CF dans la technique de *De Jong*. Toutefois, par rapport au surpeuplement standard, les nouveaux éléments créés ne sont pas insérés "aveuglément" en remplaçant des individus à caractéristiques génétiques similaires. Ils doivent d'abord entrer en compétition avec des proches voisins avant d'être introduit dans la population globale. Ce degré d'élitisme supplémentaire explique en partie la supériorité de la sélection par tournoi restreint par rapport au surpeuplement standard [Harik 95].

III.4. LES METHODES DE PARTAGE

III.4.1. Le partage standard

III.4.1.1. Principe du partage

La méthode de partage (*Fitness Sharing*) a été initialement introduite par *Goldberg* et *Richardson* en 1987 [Goldberg 87]. Elle consiste à réajuster l'adaptation de chaque individu en fonction des ressources disponibles dans son environnement local, et du nombre de congénères voisins susceptibles de lutter pour ces ressources. Le partage a pour effet de réduire l'adaptation des régions à forte densité de population et d'encourager la recherche vers les régions inexplorées. En pratique, la mise en œuvre de la méthode de partage se fait de la façon suivante :

Chaque individu a_i voit son adaptation f_i réduite d'un facteur m_i correspondant approximativement au nombre d'éléments situés dans son voisinage. La fonction d'adaptation réajustée s'écrit :

$$f_i^s = \frac{f_i}{m_i} \quad (98)$$

où le facteur de nichage m_i (*niche count*) est calculé en sommant les valeurs des contributions des fonctions de partage des N individus de la population :

$$m_i = \sum_{j=1}^N sh(d_{ij}) \quad (99)$$

La fonction de partage $sh(d_{ij})$ mesure la similarité entre individus. Généralement, elle prend pour valeur 1 si les individus a_i et a_j sont identiques, 0 si la distance d_{ij} entre les deux individus dépasse un seuil de dissimilarité fixé, et une valeur intermédiaire à des niveaux intermédiaires de similarité. Le seuil de dissimilarité est aussi appelé seuil de voisinage (ou de proximité), rayon de niche, voire distance de coupure. Il est spécifié par une constante σ_s .

Plusieurs formes de fonctions de partage sont rapportées dans la littérature [Beasley 93][Goldberg 87]. Les plus populaires sont les fonctions de partage en puissance données ci-après.

$$sh(d_{ij}) = \begin{cases} 1 - (d_{ij} / \sigma_s)^\alpha & \text{si } d_{ij} < \sigma_s \\ 0 & \text{ailleurs} \end{cases} \quad (100)$$

où α est un paramètre qui modifie la forme de la fonction de partage (voir Figure 28). α est habituellement fixé à 1, donnant comme fonction résultante la fonction de partage triangulaire.

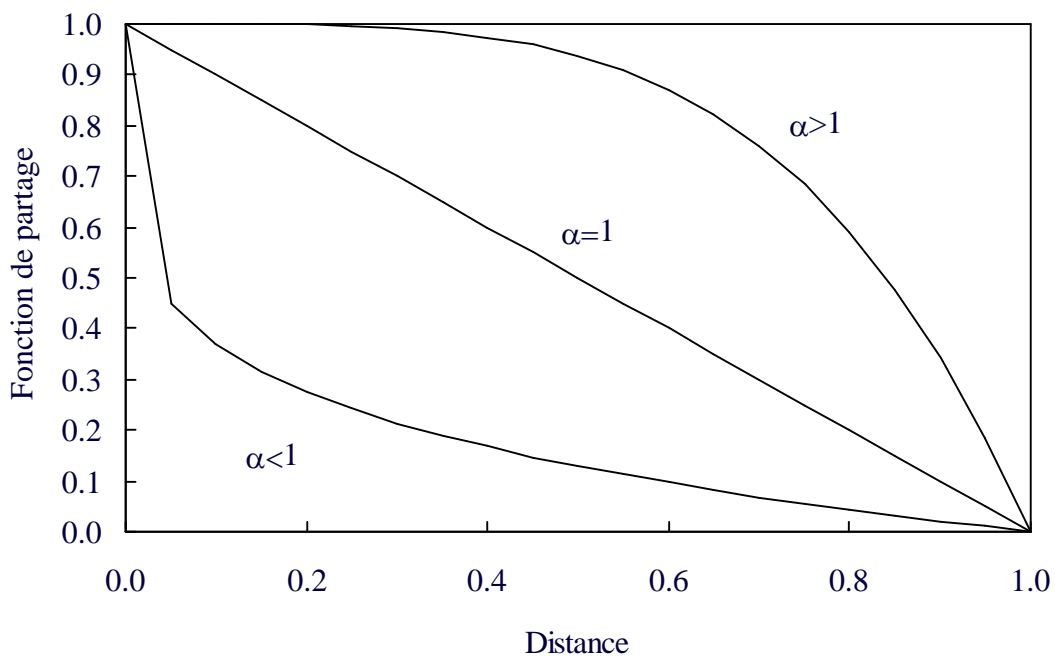


Figure 28. Forme des fonctions de partage en puissance ($\sigma_s = 1$)

Nous mentionnons aussi les fonctions de partage exponentielles, dont l'expression est donnée ci-dessous.

$$sh(d_{ij}) = \begin{cases} 1 - \exp(\ln \xi \cdot (\sigma_s - d_{ij}) / \sigma_s) & \text{si } d_{ij} < \sigma_s \\ 0 & \text{ailleurs} \end{cases} \quad (101)$$

où $\xi > 0$ joue le même rôle que le paramètre α de l'équation (100). La Figure 29 présente la forme des fonctions de partage exponentielles pour différentes valeurs de ξ .

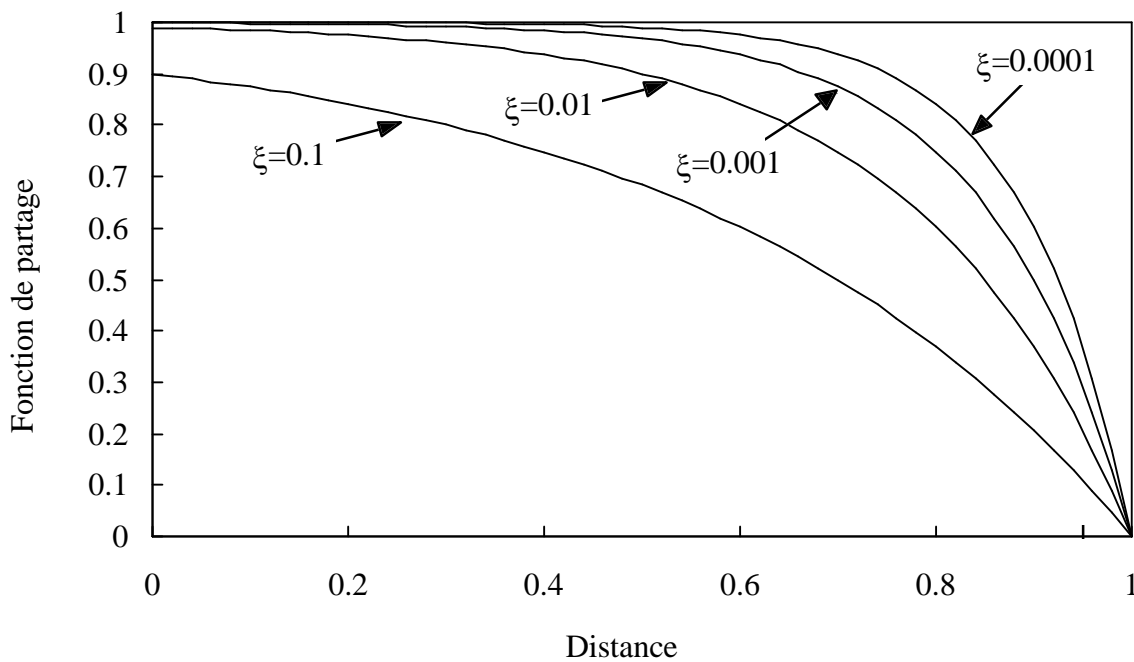


Figure 29. Forme des fonctions de partage exponentielles ($\sigma_s = 1$)

Le facteur de nichage m_i donne donc approximativement le nombre d'éléments de la population situés dans le voisinage de l'élément a_i . En théorie, nous pouvons nous attendre à ce que le mécanisme du partage, représenté par les équations (98) et (99), distribue les individus proportionnellement aux ressources, c'est à dire proportionnellement à la qualité des optima. Dans ce cas idéal, le partage se stabilise lorsque $\forall k \in 1..N - 1$,

$$\frac{f_k}{m_k} = \frac{f_{k+1}}{m_{k+1}} \quad (102)$$

III.4.1.2. Choix du rayon de niche

Un des gros problèmes de la méthode de partage est le choix du rayon de niche. Pour fixer ce dernier, il est nécessaire de connaître (ou de pouvoir estimer) *a priori* la distance entre

les optima. Cette information est généralement inconnue avant l'exploration de l'espace. Pourtant, pour maintenir deux individus a_i et a_j situés sur deux optima d'adaptation différente, il faut obligatoirement [Mahfoud 95] :

$$\sigma_s \leq \frac{d_{ij}}{1 - \min\left(\frac{f_i}{f_j}, \frac{f_j}{f_i}\right)} \quad (103)$$

Une formule empirique a été donnée par *Deb* et *Goldberg* pour estimer le rayon de niche dans le cas général d'un problème de dimension n [Deb 89]. On suppose que les q maxima de la fonction sont entourés par des hypersphères de rayon σ_s , de dimension n , qui ne doivent pas s'intercepter et remplir tout l'espace. Le rayon de niche s'exprime alors par :

$$\sigma_s = \frac{r}{\sqrt[n]{q}} \quad (104)$$

où $r = \frac{1}{2} \sqrt{\sum_{i=1}^n (x_{i\max} - x_{i\min})^2}$ est le rayon de l'hypersphère contenant tout l'espace.

Lorsque l'on utilise un critère de distance normalisé entre 0 et 1 avec des écarts de paramètres rapportés eux aussi dans le même intervalle (voir équation (95)), (104) devient :

$$\sigma_s = \frac{1}{2\sqrt[n]{q}} \quad (105)$$

Dans la méthode de partage standard, le rayon de niche est identique pour tous les individus de la population. Il est donc nécessaire que les optima de la fonction objectif soient à peu près équidistants. Le partage peut échouer dans le maintien de toutes les solutions optimales si la distance estimée entre optima est incorrecte ou si ces derniers ne sont pas repartis uniformément dans l'espace.

III.4.1.3. Taille de la population

Outre les limitations résultant du choix du rayon de niche, la taille de la population minimale nécessaire pour maintenir q niches d'adaptation équivalente pendant G générations avec une probabilité \mathcal{V} a été estimée par *Mahfoud* [Mahfoud 95] à :

$$N = \frac{\ln \frac{1 - \mathcal{V}^{1/G}}{q}}{\ln \frac{q-1}{q}} \quad (106)$$

La Figure 30 illustre la taille de population minimale nécessaire pour maintenir $q=2, 3, 5, 10$ et 20 sommets d'adaptation équivalente avec une probabilité $\gamma = 0.99$ au cours des générations.

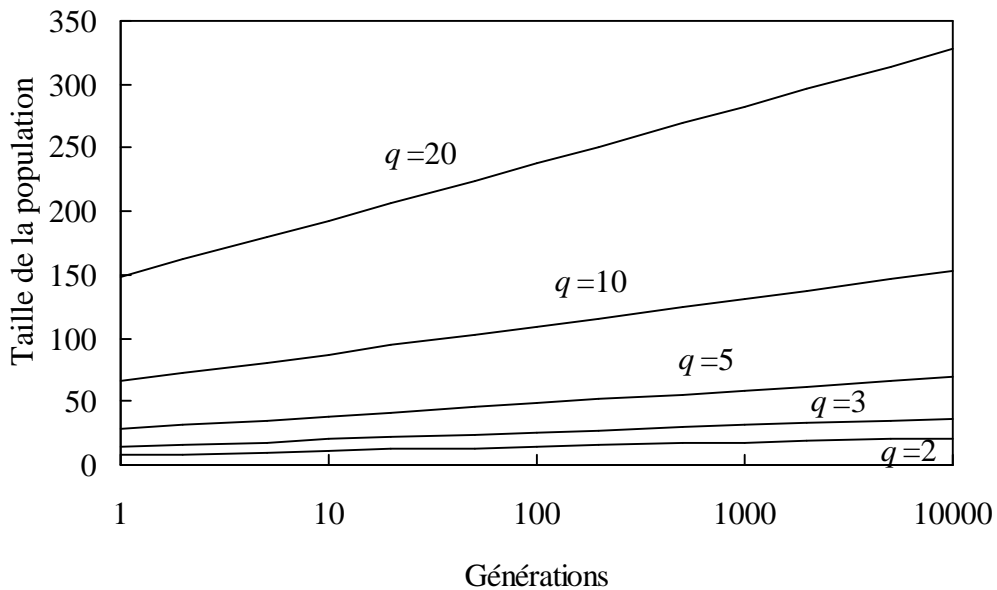


Figure 30. Taille de population minimale nécessaire pour maintenir $q=2, 3, 5, 10, 20$ niches d'adaptation équivalente au cours des générations avec une probabilité $\gamma=0.99$.

III.4.1.4. Sélection et croisement avec la méthode de partage

Le partage doit être appliqué avec la méthode de sélection la moins biaisée possible pour préserver au mieux la diversité des individus. La sélection stochastique universelle et la sélection à reste stochastique sont les plus appropriées. La sélection par tournoi est aussi possible à condition de prendre certaines précautions (mise à jour constante du partage ou *continuously updated tournament selection*) [Oei 91]. Par ailleurs, le partage doit être couplé avec des opérateurs de croisement à recombinaison restreinte. En effet, les croisements entre individus de niches différentes conduisent la plupart du temps à des enfants de faible adaptation (*lethals*) comme le montre la Figure 31. Afin de limiter la formation de ces individus "dégénérés", il est recommandé d'utiliser des opérateurs de croisement à faible taux de destruction des schémas.

Le croisement uniforme paramétré est bien adapté car il allie un fort potentiel de recombinaison et un bon pouvoir d'exploration avec un faible niveau de destruction des schémas. Sous cet opérateur, un paramètre p_0 définit la probabilité d'échange entre deux gènes des parents dans le masque de croisement (pour le croisement uniforme standard $p_0 = 0.5$). Il est toutefois nécessaire de limiter la valeur de cette probabilité car plus celle-ci est élevée plus l'opérateur est destructif.

Le croisement à un site standard est une autre alternative possible à condition d'augmenter suffisamment la longueur des chromosomes (le taux de destruction des schémas de cet opérateur étant inversement proportionnel à cette même longueur).

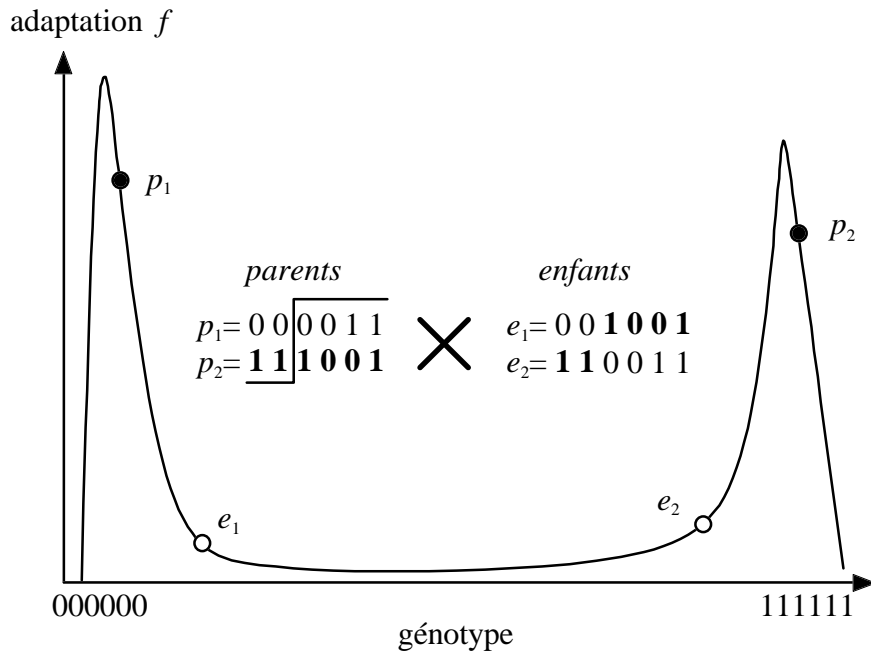


Figure 31. Exemple de croisement conduisant à des individus faiblement adaptés

III.4.1.5. L'union restrictive

Une autre manière de limiter la formation d'individus dégénérés est de pratiquer l'union restrictive (*restrictive mating*), c'est à dire de restreindre les croisements aux croisements entre individus de même niche.

La recombinaison intra-niche (*in breeding*) consiste à croiser uniquement les parents situés dans un même voisinage. Le seuil de voisinage est alors spécifié par une constante σ_m généralement identique au rayon de niche σ_s . En pratique, pour chaque individu à croiser, on cherche un partenaire dans un échantillon de la population (*mating pool*) compris dans un rayon inférieur au seuil de voisinage fixé. Si aucun élément n'est trouvé, on peut sélectionner le deuxième parent arbitrairement dans l'échantillon ou de telle sorte que sa distance avec le premier parent soit minimale [Deb 89][Miller 95b][Yin 93].

La recombinaison en ligne (*line-breeding*) est une autre technique d'union restrictive où chaque parent est recombinaison avec le meilleur individu de la population. Cette méthode intéressante pour les environnements unimodaux se révèle peu efficace lorsque l'on désire maintenir plusieurs solutions optimales. Une variante, proposée par Miller, appelée

Chapitre III

recombinaison en ligne dynamique (*dynamic line-breeding*), consiste à croiser chaque parent avec le meilleur élément de la niche la plus proche [Miller 95b].

Enfin, nous proposons pour notre part une dernière technique d'union restrictive appelée tri d'assortiment [Sareni 98a]. Il s'agit de classer les individus de la population par ordre de proximité avant d'appliquer les opérateurs de croisement et de mutation. Un pseudocode du tri d'assortiment est donné ci-après,

1. Classer les individus de la population par ordre de performance décroissante.
2. *pour* $i=1..N-2$

Chercher l'élément a_k (parmi $i+1..N$ individus) le plus proche de l'individu a_i

$$d_{ik} = \min_{j=i+1..N} d_{ij}$$

Echanger l'élément a_{i+1} avec l'élément a_k .

Fin *pour*

Une fois l'algorithme appliqué, les individus sont croisés et mutés par paire en suivant l'ordre du tri. Notons toutefois que cet algorithme réalise $N(N-1)/2-1$ comparaisons de distance et est donc relativement coûteux.

5.1.6. Mise à l'échelle de la fonction d'adaptation

Le partage peut être grandement amélioré en utilisant une mise à l'échelle de la fonction d'adaptation en puissance

$$f_i^s = \frac{f_i^\beta}{m_i} \quad (107)$$

Ce réajustement permet notamment de mieux différencier les sommets de la fonction objectif. Néanmoins, il reste le problème du choix du paramètre de mise à l'échelle β lié à l'éternel compromis entre l'exploration et l'exploitation. Une valeur de β trop élevée entraîne une convergence rapide vers de supers-individus au détriment des régions inexplorées. Une valeur de β trop faible réduit au contraire l'exploitation et ralentit la recherche en direction des sommets. Pour remédier à ce problème, *Darwen et Yao* conseillent de réajuster la valeur de β au cours des générations [Darwen 95]. Par exemple,

$$\beta = \begin{cases} 1 & \text{si } t < t_{\max} / 4 \\ 2 \left(6 \frac{t}{t_{\max}} - 1 \right) & \text{si } t_{\max} / 4 < t < t_{\max} \end{cases} \quad (108)$$

impose une valeur de $\beta=1$ jusqu'à $t=t_{\max}/4$ (forte exploration, faible exploitation) puis une progression linéaire jusqu'à $\beta=10$ à la dernière génération (faible exploration, forte exploitation).

III.4.2. Le partage dynamique et le partage rapide

La méthode de partage standard est assez coûteuse en temps CPU puisqu'elle requiert N^2 comparaisons de distances par génération. Cette caractéristique est assez pénalisante lorsque l'évaluation du critère de distance prend du temps. Le partage dynamique (*Dynamic Niching*) et le partage rapide (*Fast Sharing*) ont été proposés pour remédier à ce problème [Yin 93][Miller 95b]. Ces deux techniques de nichage, dérivées de la méthode de partage traditionnelle, permettent de réduire la complexité de l'algorithme de $O(N^2)$ à $O(Nq)$ ou $O(2Nq)$, q désignant le nombre de sommets effectivement maintenus durant la recherche. Elles consistent à faire précéder le partage d'une phase de regroupement (*clustering*) qui a pour rôle de rassembler et de classer les individus similaires à l'intérieur de groupes (ou niches) représentant une même sous-population. Une fois la séparation explicite des niches effectuée, chaque individu se trouve affecté à une sous-population donnée. Le partage est alors réalisé en prenant un facteur nichage m_i défini à partir des caractéristiques des sous-populations (le facteur de nichage d'un individu a_i est égal au nombre d'individus appartenant à la même sous-population dans la méthode de partage dynamique ; par contre, pour le partage rapide, il est déterminé en fonction des distances calculées entre chaque individu de la sous-population et le centre de la niche correspondante).

III.4.3. La méthode des sous-populations

La méthode des sous-populations introduite par Spears [Spears 94] est une variante très simplifiée du partage standard. Elle consiste à associer à chaque individu un identificateur (ou label) représentatif de la sous-population à laquelle il appartient. Les labels sont initialisés de façon uniformément aléatoire à la première génération selon le nombre désiré de sous-populations. La sélection favorise la reproduction des individus localisés dans des niches à forte adaptation mais un partage est appliqué pour préserver la diversité des sous-populations. Le facteur de nichage d'un individu est simplement le nombre d'éléments de sa sous-population d'appartenance. Une technique d'union restrictive est facilement mise en œuvre en autorisant uniquement les croisements entre individus de même label, c'est à dire appartenant à une même sous-population.

Cette technique n'offre pourtant aucune garantie de détecter tous les sommets de la fonction objectif, plusieurs sous-populations distinctes pouvant converger vers le même sommet. Cela impose de choisir un nombre de sous-populations très supérieur au nombre de sommets que l'on désire détecter. En l'absence de procédures élitistes, une sous-population peut disparaître lors de la sélection. Il est donc recommandé de réinsérer systématiquement le

meilleur individu de chaque sous-population et d'appliquer la mutation y compris aux labels identificateurs pour préserver la diversité.

III.4.4. Le partage séquentiel

Le partage séquentiel exécute de façon séquentielle un algorithme d'optimisation unimodal en utilisant les connaissances acquises à chaque itération pour éviter de réexplorer des régions où des solutions ont déjà été trouvées [Beasley 93]. Cette méthode de partage est constituée de 3 phases principales comme le montre l'organigramme ci-après.

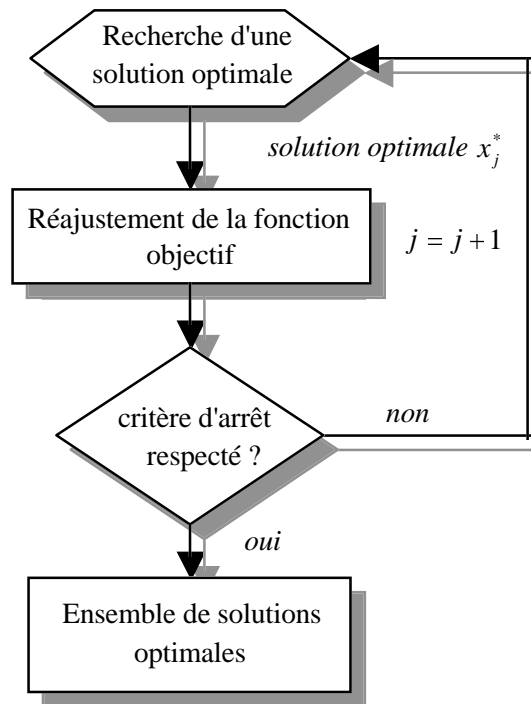


Figure 32. Principe de la méthode de partage séquentielle

III.4.4.1. La recherche d'une solution optimale

Une solution optimale du problème x_t^* est déterminée par une méthode d'optimisation quelconque. La méthode introduite initialement par *Beasley* et al. emploie un algorithme génétique unimodal. Plus généralement, il est possible d'utiliser n'importe quelle technique d'optimisation globale ou locale.

III.4.4.2. Réajustement de la fonction objectif

Une fois la solution x_t^* déterminée, la fonction objectif est modifiée de façon à ne pas réexplorer le voisinage de celle-ci. La nouvelle fonction à optimiser s'écrit,

$$F_{t+1}(x) = F_t(x) \cdot P_{x, x_t^*} \quad (109)$$

$$P_{x, x_t^*} = 1 - sh(d_{x, x_t^*}) \quad (110)$$

où $sh(d_{x_i, x_i^*})$ est la fonction de partage introduite au paragraphe III.4.1.1. La fonction $F_i(x)$ est initialisée avec la fonction objectif du problème $F_0 = f(x)$.

III.4.4.3. Critère d'arrêt

Plusieurs procédures peuvent être envisagées pour stopper l'algorithme :

- On peut fixer *a priori* le nombre de solutions optimales q à rechercher. Il suffit ensuite d'arrêter l'algorithme au bout de q itérations.
- On peut imposer un niveau minimum de qualité pour les solutions optimales de la fonction objectif. L'algorithme est alors exécuté tant que la qualité des solutions optimales détectées reste supérieure au seuil spécifié.

III.4.4.4. Conséquences du partage séquentiel

Une conséquence fâcheuse du partage séquentiel est l'apparition de solutions locales inexistantes à la suite du réajustement de la fonction d'adaptation. La Figure 33 illustre cette caractéristique pour un partage séquentiel en puissance appliqué à la fonction $|\sin(2\pi x)|$.

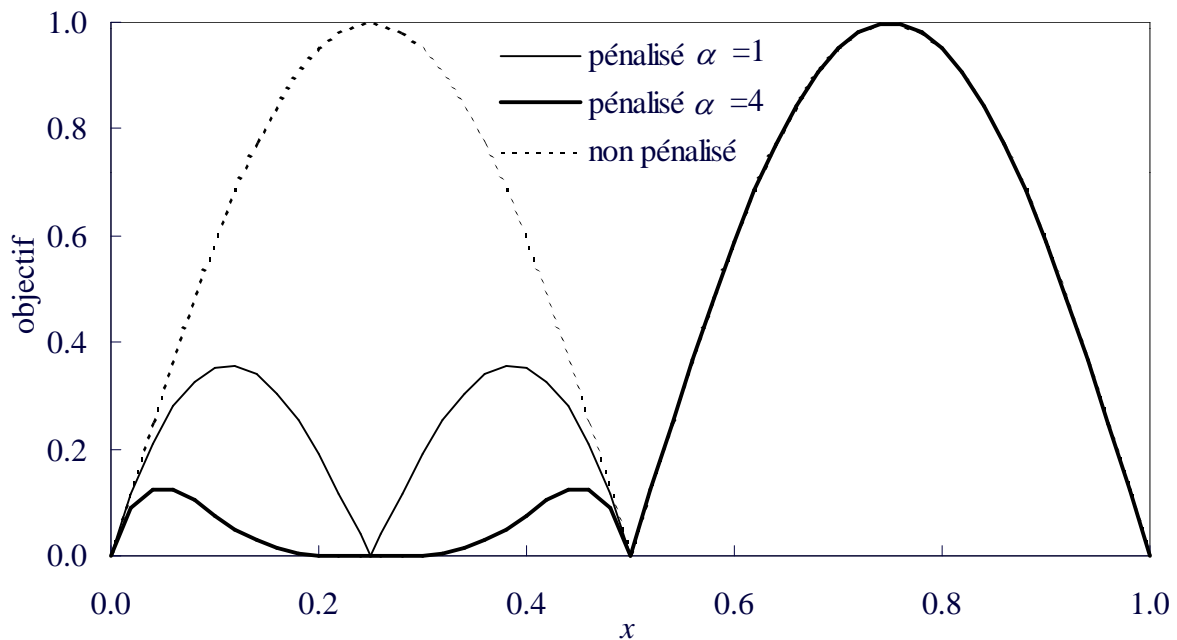


Figure 33. Evolution de la fonction objectif $|\sin(2\pi x)|$ après détection du premier sommet et réajustement par la méthode de partage séquentielle pour diverses valeurs du coefficient de forme α .

Il est possible de limiter l'influence des modes résultant du réajustement en augmentant la pénalisation (c'est à dire la valeur du paramètre α qui modifie la forme de la niche - voir équation (100)) autour des optima détectés les itérations précédentes. Nous pouvons aussi envisager d'éliminer ces solutions locales gênantes en réinitialisant la fonction objectif dans le voisinage considéré. Mais, la discontinuité de la fonction objectif due à cette pénalisation brutale (que l'on peut comparer à la pénalisation radicale utilisée pour le traitement des contraintes) pourrait poser des problèmes en particulier si on utilise, pour la recherche des optima, un algorithme de type gradient.

III.5. LA METHODE D'ECLAIRCISSEMENT

III.5.1. Principe

La méthode d'éclaircissement est voisine du partage standard excepté qu'elle est basée sur le principe des ressources limitées dans l'environnement [Petrowski 96][Petrowski 97a]. Au lieu de partager les ressources entre tous les individus d'une même sous-population, l'éclaircissement les attribue uniquement aux meilleurs représentants. En pratique, la capacité k d'une niche spécifie le nombre maximal d'individus que la niche peut accepter. Après détermination de la performance des individus dans chaque niche, l'éclaircissement préserve les k meilleurs représentants des sous-populations respectives (individus *dominants*) et exclut les autres (individus *dominés*) de la population en réinitialisant leur adaptation. Comme pour la méthode de partage, le voisinage est défini de façon explicite. Les individus appartenant à une même niche sont séparés d'une distance inférieure à un seuil de similarité fixé. Le rayon de niche correspondant est aussi appelé rayon d'éclaircissement.

III.5.2. Implantation de la méthode d'éclaircissement

La méthode d'éclaircissement, comme les méthodes de partage parallèles, n'introduit pas de réelle modification par rapport à l'algorithme génétique standard. Une phase supplémentaire, l'éclaircissement est ajoutée avant l'opération de sélection. Un pseudocode de la procédure d'éclaircissement est donné ci-après :

- a_i - individu de rang i de la population.
- N - taille de la population.
- f_i - adaptation d'un individu a_i .
- σ_s - rayon de niche.
- k - capacité des niches.
- nbD - nombre de dominants dans une niche.
- d_{ij} - critère de distance entre deux individus.

Classer les individus de la population par ordre de performance décroissante.
pour $i=1..N$
 si $f_i \neq 0$
 $nbD = 1$
 pour $j=i+1..N$
 si $d_{ij}(a_i, a_j) \leq \sigma_s$
 si $nbD < k$
 Incrémenter nbD ($nbD = nbD+1$)
 sinon
 Réinitialiser l'adaptation de l'individu a_i ($f_i = 0$)
 Fin pour
 Fin pour

III.5.3. Elitisme

Le principal objectif des techniques élitistes est de préserver les "bonnes" solutions des effets destructifs des différents opérateurs génétiques. Si la procédure d'élitisme standard (qui consiste à conserver le meilleur individu de la population globale au cours des générations) est bien adaptée à l'optimisation des fonctions unimodales, elle se révèle assez inefficace lorsqu'elle est appliquée à des problèmes multimodaux [De Jong 75]. En effet, la dérive génétique supplémentaire qu'elle introduit peut conduire l'algorithme à converger prématurément vers des solutions locales.

De façon à limiter les effets destructeurs des divers opérateurs génétiques, la méthode d'éclaircissement peut être utilisée conjointement avec une stratégie élitiste qui garantit la survie du meilleur individu de chaque sous-population de génération en génération. En pratique nous proposons d'appliquer la méthode suivante pour préserver la diversité et la qualité des sous-populations. A chaque génération, les individus dominants de chaque niche entrent en compétition avec leurs homologues de la génération précédente. Les vainqueurs des tournois résultants sont réinsérés dans la population courante.

III.5.4. Complexité

La complexité de la méthode d'éclaircissement est majorée par $O(qN)$ où q représente le nombre de niches dans l'espace et N la taille de la population [Petrowski 96]. Cette complexité est sensiblement identique à celle des méthodes de partage dynamique et rapide. Elle varie entre $O(N^2)$ et $O(N)$ suivant le nombre de sommets q à identifier et la valeur du rayon d'éclaircissement σ_s .

III.5.5. Capacité des niches

Pour préserver un maximum de niches et garantir l'exploration d'un grand nombre de solutions, il est recommandé de fixer la capacité des niches à 1. Lorsque la taille de la population est largement supérieure au nombre de sommets que l'on désire identifier, nous pouvons augmenter la valeur de la capacité des niches pour répartir un plus grand nombre d'individus sur les sommets. Néanmoins, si la capacité des niches est trop élevée, les individus vont se concentrer en grand nombre dans les niches les mieux adaptées et délaisser les régions de l'espace où peuvent se trouver d'autres solutions locales de plus faible performance.

III.6. CRITERES DE PERFORMANCE

Un algorithme génétique multimodal est censé détecter et maintenir dans sa population le plus grand nombre de solutions, globales et locales. Son efficacité peut être évaluée au moyen des critères de performance suivants :

III.6.1. Nombre total de sommets maintenus

Le nombre total de sommets q effectivement maintenus dans la population finale, traduit de façon assez simple la capacité de l'algorithme multimodal à identifier des solutions optimales et à les conserver au cours des générations. Cependant, ce critère ne donne aucune indication sur la qualité des optima détectés.

III.6.2. Taux de détection des sommets

Le taux de détection des sommets *TDS* (*maximum peak ratio*) est défini par l'équation (29) comme le rapport entre la somme des fonctions d'adaptation des optima maintenus dans la population finale et la somme des fonctions d'adaptation de l'ensemble des optima de l'espace réel [Miller 95b].

$$TDS = \frac{\sum_{k=1}^q f(x_k^*) \Big|_{d\acute{e}tect\acute{e}}}{\sum_{i=1}^{n_{opt}} f(x_i^*) \Big|_{r\acute{e}el}} \quad (111)$$

où q désigne comme précédemment le nombre total de sommets effectivement maintenus dans la population finale et n_{opt} le nombre total de sommets de l'espace réel ($n_{opt} \geq q$).

Un élément de la population est représentatif d'un sommet de l'espace [Miller 95b]

1. S'il est situé dans le voisinage de l'optimum réel correspondant, c'est à dire à une distance inférieure à un rayon de niche donné.
2. Si sa performance est au moins égale à 80% de celle de l'optimum réel.
3. Si sa performance est la plus élevée parmi les membres de la population qui satisfont les deux premiers critères.

Par conséquent, la valeur maximale du taux de détection des sommets *TDS* est égale 1. Elle correspond au cas d'une parfaite identification de tous les optima de la fonction objectif.

Par rapport au critère précédent, le taux de détection des sommets donne une information à la fois sur la quantité et la qualité des solutions optimales identifiées.

III.6.3. Critère de distribution proportionnelle

Le critère de distribution proportionnelle *TDP* (*Chi-Square-Like Deviation*) mesure l'écart global entre la distribution de population courante et une répartition proportionnelle quasi-idéale de la population sur les optima de l'espace [Deb 89][Miller 95b][Yin 93]. Il caractérise la capacité de la méthode multimodale à peupler proportionnellement les niches en fonction de leur adaptation. Son expression est donnée ci-après :

$$TDP = \sqrt{\sum_{i=1}^{q+1} \left(\frac{X_i - \mu_i}{\sigma_i} \right)^2} \quad (112)$$

où la variable X_i représente le nombre d'individus de la population courante situés dans la niche i , μ_i le nombre moyen d'individus localisés dans cette même niche dans le cas d'une distribution proportionnelle idéale et σ_i l'écart-type correspondant. Plus la valeur de *TDP* est faible, plus la répartition des individus s'apparente à une distribution proportionnelle, meilleur est le critère.

L'écart global est calculé en considérant q niches relatives aux q sommets réels de l'espace et une $q+1$ ème niche comprenant l'ensemble des individus n'appartenant à aucun de ces sommets. Un individu est distribué dans une niche représentative d'un sommet de l'espace si et seulement si il satisfait les deux premiers critères énoncés au paragraphe précédent.

Le nombre moyen d'individus μ_i situés dans une niche donnée i et l'écart-type correspondant σ_i , caractérisant la distribution proportionnelle idéale d'une population de taille N , sont définis par les équations (113) et (114).

$$\mu_i = N \frac{f(x_i^*)|_{réel}}{\sum_{k=1}^q f(x_k^*)|_{réel}} \text{ pour } i = 1..q \text{ et } \mu_{q+1} = 0 \quad (113)$$

$$\sigma_i^2 = \mu_i(1 - \mu_i / N) \text{ pour } i = 1..q \text{ et } \sigma_{q+1}^2 = \sum_{k=1}^q \sigma_k^2 \quad (114)$$

Le Tableau 9 présente les valeurs de ces paramètres dans le cas de la fonction M_4 (voir Annexe II) et d'une population de 100 individus.

Tableau 9. Distribution proportionnelle idéale d'une population de taille 100 (fonction M_4)

<i>niche i</i>	x_i^*	$M_4(x_i^*)$	μ_i	σ_i^2
1er sommet	0.080	1.000	28.797	20.504
2ème sommet	0.247	0.949	27.318	19.855
3ème sommet	0.451	0.771	22.194	17.268
4ème sommet	0.681	0.503	14.492	12.392
5ème sommet	0.934	0.250	07.200	06.682
hors sommet	-	-	00.000	76.701

III.7. ALGORITHMES GENETIQUES MULTIMODAUX COMPARES

III.7.1. Comparaison des méthodes de nichage avec l'algorithme génétique standard

III.7.1.1. Efficacité dans la recherche de l'optimum global

Nous comparons les performances des méthodes de nichage avec celles d'un algorithme génétique élitiste standard à partir de tests effectués sur les fonctions de *De Jong* augmentées de la fonction de *Rastrigin*. Dans chaque cas, les probabilités de croisement et de mutation sont respectivement $p_c = 1$ et $p_m = 0.001$. Les chromosomes des individus sont codés en binaire avec décodage dynamique. La longueur de codage adoptée pour chaque paramètre est celle recommandée par *De Jong* [De Jong 75] (voir Annexes II). Les méthodes à voisinage explicites sont utilisées avec un rayon de niche constant $\sigma_s = 0.2$ (fixé indépendamment de la forme des fonctions) et une distance euclidienne normalisée. La taille de la population est fixée à 30 et le nombre de générations à 1000. L'algorithme est arrêté en cas de convergence à 10^{-3} près par rapport à l'optimum réel. 100 tests sont effectués pour tenir compte de l'aspect

stochastique de la recherche. Les grandeurs moyennes des différents critères de performance définis à l'Annexe III sont présentés dans le Tableau 10. L'écart-type correspondant est spécifié entre parenthèses.

Tableau 10. Comparaison des méthodes de nichage avec l'algorithme génétique standard - le nombre d'évaluations de fonction et la solution optimale sont moyennés sur 100 tests ; l'écart-type correspondant est précisé entre parenthèses.

<i>Type de GA</i>	<i>Taux de réussite</i>	<i>Nombre d'évaluations de fonction</i>	<i>solution optimale</i>	<i>Meilleure solution trouvée</i>
<i>Fonction F_1 (sphère $n=3$)</i>				
GA élitiste	100%	1460 (856)	0.00057 (0.00032)	0.00002
DC	100%	2257 (1181)	0.00064 (0.00025)	0.00007
RTS ($w=9$)	100%	4480 (3571)	0.00056 (0.00029)	0.00002
CL ($k=1$)	100%	5063 (3083)	0.00061 (0.00029)	0.00001
SH	0%	-	2.06958 (1.17663)	0.27651
<i>Fonction F_2 (Rosenbrock $n=2$)</i>				
GA élitiste	44%	2320 (4873)	0.01384 (0.02836)	0.00002
DC	99%	4183 (3840)	0.00058 (0.00029)	0.00002
RTS ($w=9$)	56%	2959 (2738)	0.00291 (0.00679)	0.00007
CL ($k=1$)	42%	6034 (7397)	0.01278 (0.02296)	0.00005
SH	23%	12 994 (8945)	2.73224 (3.46554)	0.00009
<i>Fonction F_3 (paliers $n=5$)</i>				
GA élitiste	100%	1022 (953)	0 (0)	0
DC	100%	1786 (2782)	0 (0)	0
RTS ($w=9$)	82%	4020 (5094)	0.18 (0.38612)	0
CL ($k=1$)	93%	9859 (7579)	0.07 (0.25643)	0
SH	27%	15 951 (8108)	2.54 (1.99707)	0
<i>Fonction F_4 (bicarrée bruitée $n=30$)</i>				
GA élitiste	0%	-	0.01892 (0.00715)	0.00449
DC	0%	-	0.34021 (0.28249)	0.04143
RTS ($w=9$)	0%	-	0.18876 (0.10977)	0.02124
CL ($k=1$)	0%	-	0.27076 (0.10135)	0.09361
SH ($\sigma_s = 0.2$)	0%	-	15.5957 (4.96399)	6.67773

Fonction F_5 (trous de renard $n=2$)				
GA élitiste	26%	843 (1506)	2.37706 (2.27600)	0.99800
DC	84%	2320 (3572)	1.03704 (0.09464)	0.99800
RTS ($w=9$)	39%	2984 (5179)	1.14699 (0.13527)	0.99800
CL ($k=1$)	49%	3214 (3918)	1.27444 (0.74992)	0.99800
SH	33%	1680 (2131)	2.53565 (2.01308)	0.99800
Fonction F_6 (Rastrigin $n=10$)				
GA élitiste	0%	-	15.0549 (4.96895)	4.49136
DC	0%	-	7.99642 (3.57429)	1.11891
RTS ($w=9$)	0%	-	11.4458 (4.32843)	4.18006
CL ($k=1$)	0%	-	7.52062 (2.65020)	2.23829
SH	0%	-	42.9159 (10.0281)	9.69183

L'algorithme génétique élitiste standard (GA élitiste) surpasse les méthodes de nichage sur des problèmes unimodaux (F_1 et F_4 notamment) grâce à son opérateur de sélection qui lui permet d'augmenter rapidement l'exploitation en direction de l'optimum. Par contre, pour des problèmes multimodaux (F_5 et F_6 en particulier) ou pour la fonction de *Rosenbrock* (F_2), son efficacité est largement inférieure à celle des méthodes de nichage si l'on excepte les mauvaises performances de la méthode de partage (SH). Celle-ci est incapable de converger précisément avec la taille de population considérée.

Le surpeuplement déterministe (DC) et la sélection par tournoi restreint (RTS) donnent les meilleurs résultats.

La méthode d'éclaircissement (CL) est assez limitée par le faible nombre d'individus utilisés. Lorsque la taille de la population est trop réduite, le mécanisme de pénalisation radicale qui conduit à l'élimination des individus faiblement adaptés, diminue fortement le nombre de génotypes dans la population. L'éclaircissement est alors incapable de rétablir la diversité à partir des individus non pénalisés. Ceci permet d'expliquer la vitesse de convergence peu élevée de l'algorithme et la supériorité des méthodes de surpeuplement dans ce cas particulier.

III.7.1.2. Comportement qualitatif sur des problèmes multimodaux

La Figure 34 illustre, pour quatre fonctions multimodales, la répartition de 30 individus après 30 générations dans le cas d'un algorithme génétique élitiste et de la sélection par tournoi restreint¹

¹ Il est possible d'obtenir des distributions qualitativement similaires à celles de la sélection par tournoi restreint avec les autres méthodes de nichage.

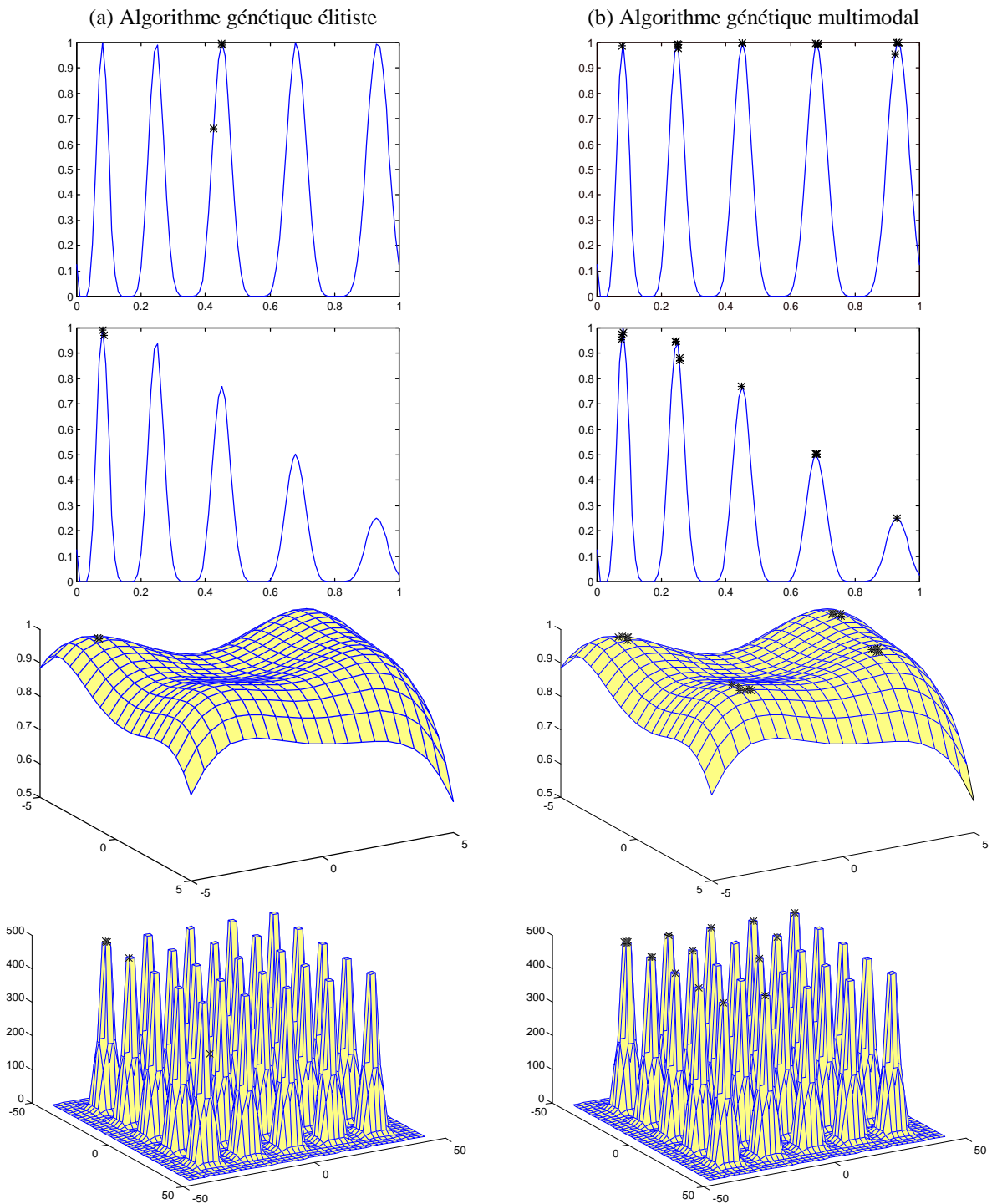


Figure 34. Exemples de répartitions des individus de la population après 30 générations ($N=30$)

Nous vérifions bien qu'une méthode génétique multimodale est capable de répartir sa population sur les optima des fonctions alors qu'un algorithme génétique standard concentre tous ses individus sur l'optimum global.

III.7.2. Résolution de problèmes multimodaux

L'efficacité des principales méthodes génétiques multimodales présentées est évaluée au moyen des critères performance définis dans le paragraphe III.6. Les divers paramètres relatifs aux algorithmes utilisés pour l'optimisation des fonctions tests sont présentés dans le Tableau 11.

Tableau 11. Paramètres utilisés pour l'optimisation des fonctions tests

<i>Fonction</i>	<i>Distance</i>	<i>Rayon de niche σ_s</i>	α	p_m	p_c
M_3	phénotypique d^∞	0.1	1	0	1
M_4	phénotypique d^∞	0.1	1	0	1
M_5	phénotypique d^∞	0.1	1	0.001	1
M_6	phénotypique d^∞	0.05	1	0.001	1
M_7	Hamming	0.2	1	0.001	1

Pour tenir compte de l'aspect stochastique des algorithmes, au moins dix exécutions successives sont lancées avec différentes populations initiales générées arbitrairement. Une moyenne est calculée pour chaque critère sur l'ensemble des exécutions. Deux tests sont effectués pour les fonctions M_3 et M_4 . Le nombre d'évaluations de fonction assez important dans le premier cas (test 1, taille de population $N = 100$ et nombre de générations $G = 200$) alors qu'il est relativement limité dans le second (test 2, taille de population $N = 30$ et nombre de générations $G = 30$). Le codage utilisé pour les tests des fonctions M_3 , M_4 , M_5 et M_6 est le code Gray.

Le taux de mutation est supprimé pour les fonctions M_3 et M_4 afin d'éviter une restauration de matériel génétique perdu. En effet, nous rappelons que le principal rôle de la mutation est de protéger les individus contre la perte éventuelle d'informations génétiques à la suite du croisement. Si nous voulons déterminer l'efficacité des méthodes de nichage et leur capacité à maintenir la diversité au sein de la population, il est nécessaire d'isoler tous les mécanismes qui contribuent à la restauration de cette diversité.

Comme les algorithmes de surpeuplement utilisent implicitement une probabilité de croisement égale à 1, nous imposons aussi cette valeur pour les autres méthodes de nichage. Cela permet en outre d'évaluer les performances dans le cas le plus défavorable au maintien des niches, la probabilité de perturbation des solutions étant maximale.

III.7.3. Etude de la méthode de partage standard

III.7.3.1. Influence du mode de sélection

Nous comparons les performances de la méthode de partage en fonction des trois modes de sélection suivants :

- la sélection stochastique universelle (SUS) [Baker 87]
- la sélection à reste stochastique (SRS) [Baker 87]
- la sélection par tournoi avec mise à jour constante du partage (TS) [Oei 91]

La méthode de partage est combinée à un algorithme génétique doté d'une procédure de croisement à un site. Le Tableau 12 et le Tableau 13 présentent les valeurs des critères de performance obtenues respectivement pour les fonctions M_3 et M_4 selon les différents modes de sélection étudiés.

Tableau 12. Influence du mode de sélection - fonction test M_3
Critères de performance moyennés sur 10 exécutions

Mode de Sélection	Nombre de sommets maintenus q		TDS		TDP moyen		TDP final	
	test 1	test 2	test 1	test 2	test 1	test 2	test 1	test 2
SRS	5	3.6	0.998	0.703	2.615	4.352	2.283	4.255
SUS	5	4.8	0.997	0.935	2.214	2.957	1.672	1.789
TS	5	3.2	0.998	0.678	4.575	5.043	5.076	5.042

Tableau 13. Influence du mode de sélection - fonction test M_4
Critères de performance moyennés sur 10 exécutions

Mode de Sélection	Nombre de sommets maintenus q		TDS		TDP moyen		TDP final	
	test 1	test 2	test 1	test 2	test 1	test 2	test 1	test 2
SRS	5	3.2	0.998	0.628	2.552	4.957	2.514	4.617
SUS	5	4.8	0.997	0.941	2.395	3.461	2.109	2.281
TS	5	3.2	0.987	0.678	5.145	5.114	6.220	4.627

Les résultats obtenus soulignent la supériorité de la SUS quels que soient le nombre d'évaluations de fonction et la taille de la population. TS et SRS se révèlent inefficaces lorsque la population est restreinte (test 2). TS est incapable de former des sous-populations

stables, ce qui conduit à un critère de distribution proportionnel très fluctuant au cours des générations [Sareni 98a].

III.7.3.2. Influence du mode de recombinaison

Nous comparons ici l'influence du mode de croisement sur l'efficacité du partage. Les diverses techniques de recombinaison abordées précédemment, à savoir le croisement à un site, le croisement uniforme standard, le croisement paramétré et le tri d'assortiment, sont étudiées. Le Tableau 14 et le Tableau 15 présentent les valeurs des critères de performance obtenues respectivement pour les fonctions M_3 et M_4 selon les différents modes de recombinaison.

Tableau 14. Influence du mode de recombinaison - fonction test M_3
Critères de performance moyennés sur 10 exécutions

Mode de croisement	Nombre de sommets maintenus q		TDS		TDP moyen		TDP final	
	test 1	test 2	test 1	test 2	test 1	test 2	test 1	test 2
1 site standard	5	4.8	0.997	0.935	2.214	2.957	1.672	1.789
uniforme 0.5	3.8	3.6	0.757	0.696	8.926	5.865	8.145	5.547
uniforme 0.1	5	4.2	0.995	0.830	4.307	4.020	3.717	3.660
uniforme 0.01	5	4.6	0.999	0.889	1.479	3.422	1.041	2.120
tri+uniforme 0.5	5	4.8	0.999	0.947	1.287	2.839	0.740	1.972

Tableau 15. Influence du mode de recombinaison - fonction test M_4
Critères de performance moyennés sur 10 exécutions

Mode de croisement	Nombre de sommets maintenus q		TDS		TDP moyen		TDP final	
	test 1	test 2	test 1	test 2	test 1	test 2	test 1	test 2
1 site standard	5	4.8	0.999	0.941	2.395	3.464	2.109	2.281
uniforme 0.5	4.2	3.2	0.767	0.670	10.58	6.377	9.613	5.415
uniforme 0.1	5	3.8	0.994	0.798	5.144	5.080	4.474	3.883
uniforme 0.01	5	4.4	0.999	0.838	1.355	2.998	1.130	2.335
tri+uniforme 0.5	4.8	4.4	0.984	0.920	2.041	2.712	1.380	1.826

Les résultats montrent les limitations du croisement uniforme standard ($p_0=0.5$, 4^{ème} ligne dans les deux tableaux). Celui-ci est incapable de former des sous-populations stables en raison d'un important taux de destruction des schémas dans la population. Toutefois, combiné avec le tri d'assortiment (dernière ligne), il se révèle supérieur à toutes les autres méthodes de

recombinaison. Le croisement uniforme paramétré de faible probabilité (6^{ème} ligne) est efficace lorsque le nombre d'évaluation de fonctions est élevé mais moins intéressant quand celui-ci est limité. Le croisement standard à un site (3^{ème} ligne) est plus robuste à condition que la longueur du chromosome soit suffisamment grande, le taux de destruction des schémas sous cet opérateur étant inversement proportionnel à cette même longueur.

III.7.4. Comparaisons des principales méthodes de nichage

Nous comparons enfin la performance de la méthode de partage standard associée à un tri d'assortiment avec trois autres techniques de nichage : l'éclaircissement (CL), la sélection par tournoi restreint (RTS) et le surpeuplement déterministe (DC).

Le croisement uniforme standard est utilisé dans les deux méthodes de surpeuplement. L'éclaircissement est combiné à la sélection stochastique universelle et à une stratégie élitiste conforme à celle décrite au paragraphe III.5.3. La capacité des niches est fixée à 10 pour le test 1 et 2 pour le test 2 lors de l'optimisation des fonctions M_3 et M_4 . Les valeurs des critères de performance obtenues pour ces deux fonctions sont présentées dans le Tableau 16 et le Tableau 17.

Tableau 16. Méthodes de nichage comparées - fonction test M_3
Critères de performance moyennés sur 10 exécutions

<i>Méthode de nichage</i>	<i>Nombre de sommets maintenus q</i>		<i>TDS</i>		<i>TDP moyen</i>		<i>TDP final</i>	
	test 1	test 2	test 1	test 2	test 1	test 2	test 1	test 2
tri + partage	5	4.8	0.999	0.947	1.287	2.839	0.740	1.972
CL	5	5	1.000	0.990	0.293	1.045	0.212	0.129
DC	5	5	1.000	0.999	5.683	3.537	5.820	3.155
RTS ($w=30\%N$)	5	4.8	1.000	0.958	2.839	2.743	2.520	3.763

Tableau 17. Méthodes de nichage comparées - fonction test M_4
Critères de performance moyennés sur 10 exécutions

<i>Méthode de nichage</i>	<i>Nombre de sommets maintenus q</i>		<i>TDS</i>		<i>TDP moyen</i>		<i>TDP final</i>	
	test 1	test 2	test 1	test 2	test 1	test 2	test 1	test 2
tri + partage	4.8	4.4	0.984	0.920	2.041	2.712	1.380	1.826
CL	5	4.8	1.000	0.933	1.448	2.397	1.354	1.477
DC	4	4	0.778	0.768	9.582	5.263	10.83	5.264
RTS ($w=30\%N$)	5	5	1.000	0.998	6.516	3.698	6.279	2.924

L'efficacité de la méthode de nichage est liée à sa capacité de trouver des niches nouvelles en produisant de nouveaux individus tout en préservant les niches déjà identifiées. L'éclaircissement semble être la méthode qui réalise le mieux ce compromis. Il produit une grande quantité d'individus nouveaux en recombinaison des éléments de niches différentes et contrôle cette production, parallèlement à la dérive génétique causée par la sélection, en réinitialisant la performance des individus peu adaptés. En outre, la stratégie élitiste préserve les individus dominants des effets destructifs des divers opérateurs génétiques.

L'éclaircissement surpasse toutes autres techniques de nichage en présentant à la fois un fort taux de détection des sommets et un excellent critère de distribution proportionnelle. Les méthodes de surpeuplement ont plus de difficultés à répartir les individus proportionnellement à l'adaptation des niches en raison de leur caractère stationnaire. En effet, elles utilisent une stratégie de remplacement qui minimise les changements dans la population [Miller 95b]. Par conséquent, la distribution de la population dans les diverses niches à la dernière génération dépend fortement de la répartition initiale des individus. C'est pourquoi le critère de distribution proportionnel évolue aussi lentement pour les méthodes de surpeuplement et reste finalement assez médiocre. La Figure 35 illustre cette caractéristique pour la fonction M_4 et le test 1.

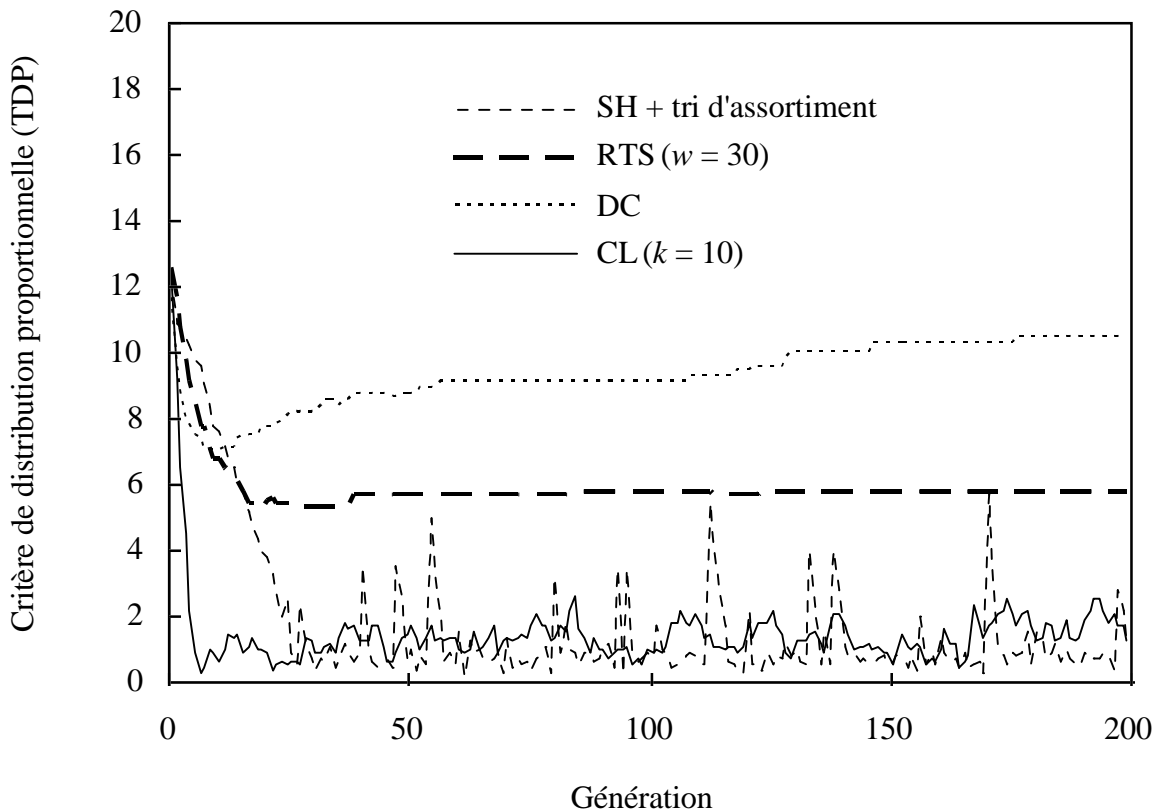


Figure 35. Critère de distribution proportionnelle - fonction M_4 - (test 1 : $N=200$, $G=200$)

Par ailleurs, des erreurs de remplacement peuvent se produire et conduire à l'abandon d'une niche par les individus de la population. Dans le cas du surpeuplement déterministe notamment, il n'est pas impossible que les deux enfants créés soient localisés dans une niche différente de celle de leurs parents respectifs. La dérive résultante peut aboutir à la disparition d'une sous-population et à l'augmentation de la densité de population dans les autres niches.

Ce phénomène est particulièrement visible lors de la recherche des optima de la fonction M_4 . L'algorithme du surpeuplement déterministe détecte dès les premières générations les cinq sommets de la fonction, mais est incapable de les maintenir pendant les générations suivantes. Les individus situés sur le troisième sommet d'abscisse $x=0.451$ migrent progressivement sur le sommet suivant. Ceci se traduit par une augmentation du critère de distribution proportionnelle (voir Figure 35). A la dernière génération, les individus de la population ont totalement déserté le troisième sommet. Le Tableau 18 présente la répartition finale des individus dans les différentes niches, obtenue avec le surpeuplement déterministe pour le test 1. Ces résultats peuvent être comparés à ceux calculés dans le cas d'une distribution proportionnelle idéale.

Tableau 18. Répartition finale des individus de la population (test 1) - fonction test M_4
Résultats moyennés sur 10 exécutions

<i>Niche</i>	<i>Nombre d'individus dans la niche à la dernière génération</i>	<i>Distribution proportionnelle idéale (voir Tableau 9)</i>
1er sommet	30.2	28.797
2ème sommet	21	27.318
3ème sommet	0	22.194
4ème sommet	40	14.492
5ème sommet	8.8	7.200
hors sommet	0	0

La sélection par tournoi restreint semble plus robuste mais son efficacité dépend fortement du choix de la fenêtre de nichage comme le montre l'optimisation sur la fonction M_6 (Tableau 20). La méthode de partage est relativement performante lorsqu'elle est appliquée à ces fonctions "faciles".

Tableau 19. Méthodes de nichage comparées - fonction test M_5 - test 2
Critères de performance moyennés sur 100 exécutions

Méthode de nichage	Nombre de sommets maintenus		TDS	
	moyenne	confiance (95%)	moyenne	confiance (95%)
DC	3.97	± 0.033	0.992	± 0.008
RTS ($w=30\%N$)	4	± 0.000	1.000	± 0.000
CL ($k=1$)	4	± 0.000	0.998	± 0.000
partage standard	4	± 0.000	0.986	± 0.002

Tableau 20. Méthodes de nichage comparées - fonction test M_6 - test 2
Critères de performance moyennés sur 100 exécutions

Méthode de nichage	Nombre de sommets maintenus		TDS	
	moyenne	confiance (95%)	moyenne	confiance (95%)
DC	7.960	± 0.403	0.322	± 0.016
RTS ($w=30\%N$)	7.890	± 0.370	0.318	± 0.015
RTS ($w=50\%N$)	12.190	± 0.373	0.490	± 0.015
RTS ($w=80\%N$)	17.500	± 0.289	0.703	± 0.012
RTS ($w=100\%N$)	21.830	± 0.305	0.874	± 0.012
CL ($k=1$)	16.380	± 0.599	0.654	± 0.024
partage standard	15.530	± 0.585	0.607	± 0.023

Les tests suivants concernent la recherche des optima de la fonction M_7 massivement multimodale et trompeuse. Ce problème a déjà été résolu par *Goldberg* [Goldberg 92] au moyen de la méthode de partage en utilisant une population de très grande taille ($N=5000$) et un réajustement en puissance de la fonction objectif fortement élitiste ($\beta=15$). *Mahfoud* a montré que le surpeuplement déterministe nécessitait environ 100 000 évaluations de fonction et 20 individus par niche pour converger vers les 32 optima globaux. Ces paramètres sont bien évidemment prohibitifs lorsque l'estimation de la qualité d'un individu est coûteuse en temps CPU. C'est pourquoi nous choisissons d'étudier la convergence des diverses techniques multimodales en prenant une taille de population limitée à $N=100$ individus et un nombre de générations $G=200$. Le critère de similarité utilisé pour différencier les sous-populations est la

distance de Hamming normalisée par la valeur maximale correspondante dans l'espace de recherche. Le rayon de niche est fixé à 0.2 conformément à [Goldberg 92].

Nous comparons aussi l'efficacité des méthodes à voisinage explicite (à savoir le partage et l'éclaircissement) avec et sans mise à l'échelle en puissance de la fonction objectif. Deux formes de réajustement sont étudiées. La première, statique, emploie un paramètre de mise à l'échelle constant ($\beta=15$). La seconde, dynamique, modifie la valeur de ce même paramètre au cours des générations. Une valeur de β égale à 1 est prise jusqu'à la 50ème génération, puis celle-ci est linéairement augmentée jusqu'à 15 pendant les générations suivantes.

Le Tableau 21 présente le nombre de sommets effectivement maintenus après 200 générations. Les résultats sont moyennés sur 100 exécutions.

Tableau 21. Méthodes de nichage comparées - fonction test M_7
Critères de performance moyennés sur 100 exécutions

<i>Méthode de nichage</i>	<i>Nombre de sommets maintenus q</i>	<i>Marge de confiance (95%)</i>
RTS ($w=5$)	1.63	± 0.21
RTS ($w=10$)	2.28	± 0.23
RTS ($w=20$)	1.64	± 0.25
RTS ($w=50$)	0.56	± 0.14
DC	0.43	± 0.13
Partage standard	0.00	± 0.00
Partage standard + mise à l'échelle statique	1.48	± 0.17
Partage standard + mise à l'échelle dynamique	3.13	± 0.32
CL ($k=2$) standard	14.10	± 1.59
CL ($k=2$) standard + mise à l'échelle statique	15.58	± 1.66
CL ($k=2$) standard + mise à l'échelle dynamique	14.06	± 1.72

Aucune des méthodes étudiées n'a réussi à maintenir les 32 maxima globaux de la fonction. Néanmoins, l'éclaircissement surpasse toutes les autres techniques de nichage en trouvant et préservant entre 14 et 15 solutions. Comme le signale Goldberg [Goldberg 92], nous remarquons que le partage standard sans mise à l'échelle de la fonction est incapable de détecter un seul maximum global. Un réajustement de la fonction en puissance permet d'augmenter la différenciation des sommets locaux et globaux, et facilite de ce fait l'identification des solutions globales. Mais, l'utilisation d'une procédure de réajustement statique sévère ($\beta=15$) élimine un grand nombre d'individus et réduit donc considérablement

la diversité dans la population. De meilleurs résultats sont obtenus avec une mise à l'échelle dynamique.

Contrairement au partage standard combiné avec une procédure de réajustement de la fonction objectif, l'éclaircissement ne rejette pas un nombre important de solutions prometteuses mais écarte seulement les individus localisés dans le voisinage d'un dominant d'une sous-population. De cette façon, il préserve toujours les bonnes solutions. Par ailleurs, cette méthode ne requiert pas de mise à l'échelle hasardeuse de la fonction d'adaptation. Au contraire, un réajustement dynamique semble plutôt trompeur pour l'éclaircissement notamment pour les individus dominants de deux générations successives qui entrent en compétition à travers la procédure d'élitisme retenue.

III.7.5. Discussion

Le Tableau 22 résume les principales caractéristiques des méthodes multimodales présentées dans ce chapitre :

Les méthodes de nichage peuvent être classées en deux groupes selon qu'elles nécessitent ou non la définition explicite d'un seuil de voisinage pour différencier les individus similaires des individus non similaires dans la population. Les techniques à voisinage implicite telles les méthodes de surpeuplement ne requièrent aucune information sur l'espace à explorer. Par contre, les procédures à voisinage explicite comme le partage ou l'éclaircissement nécessitent de fixer un seuil de similarité, donc de pouvoir estimer *a priori* la distance entre deux optima de l'espace. Ceci est un inconvénient majeur car pour la plupart des problèmes d'optimisation réels, nous n'avons aucune connaissance de l'espace avant son exploration. Les méthodes de séparation de niches ou les techniques à seuil de similarité adaptatif permettent de résoudre partiellement ce problème mais leur implantation est plus complexe [Yin 93][Petrowski 97b].

Lorsque le rayon de niche est judicieusement choisi, l'éclaircissement se révèle supérieur à toutes les autres méthodes de nichage. Le partage, appliqué à des fonctions faciles, est assez efficace à condition d'utiliser la sélection stochastique universelle et des opérateurs de recombinaison peu destructifs pour stabiliser les sous-populations. Par contre, il est inefficace dans le cas général car l'équilibre entre les sous-populations peuplant les différentes niches est difficile à obtenir et la convergence vers les optima nécessite une taille de population très élevée. Il est possible d'augmenter la performance du partage en associant à la fonction d'adaptation une mise à l'échelle en puissance qui permet une meilleure différenciation des optima globaux et locaux. Néanmoins, le choix délicat et hasardeux du paramètre de mise à l'échelle peut entraîner une convergence prématurée de l'algorithme.

Les méthodes de surpeuplement basées sur des règles de tournoi et des stratégies de remplacement ne souffrent pas de cet inconvénient. Bien que leur critère de distribution

proportionnelle ne soit pas très bon, elles restent assez efficaces. La sélection par tournoi restreint est généralement supérieure au surpeuplement déterministe qui peut commettre des erreurs de remplacement entre individus de niches différentes.

Tableau 22. Méthodes de nichage comparées - Avantages et inconvénients

<i>Méthode Multimodale</i>	<i>Avantages</i>	<i>Inconvénients</i>
Recherche locale parallèle (<i>Parallel Hillclimbing</i>)	<ul style="list-style-type: none"> • parallélisme possible • simplicité • rapidité • pas de comparaisons de distances 	<ul style="list-style-type: none"> • aucune garantie de converger vers les solutions globales
Surpeuplement Standard (<i>Crowding</i>)	<ul style="list-style-type: none"> • simplicité • voisinage implicite • parallélisme possible • $O(CF \cdot N)$ comparaisons de distances 	<ul style="list-style-type: none"> • facteur de surpeuplement peu évident à fixer • nombre d'optima détectés faible • mauvaise détection des solutions locales • mauvais critères de performance
Surpeuplement Déterministe (<i>Deterministic Crowding</i>)	<ul style="list-style-type: none"> • simplicité • voisinage implicite • parallélisme possible • $O(N)$ comparaisons de distances • bonne détection des solutions locales 	<ul style="list-style-type: none"> • mauvais critère de distribution proportionnelle • dérive génétique possible due aux éventuelles erreurs de remplacement
Sélection par tournoi restreint (<i>Restricted Tournament Selection</i>)	<ul style="list-style-type: none"> • voisinage implicite • parallélisme possible • $O(CF \cdot N)$ comparaisons de distances • bonne détection des solutions locales 	<ul style="list-style-type: none"> • facteur de surpeuplement peu évident à fixer • mauvais critère de distribution proportionnelle
Partage standard (<i>Fitness Sharing</i>)	<ul style="list-style-type: none"> • parallélisme possible • efficace pour des problèmes "faciles" 	<ul style="list-style-type: none"> • voisinage explicite • $O(N^2)$ comparaisons de distances • équilibre des sous-populations difficile à obtenir • nécessite une taille de population importante et une mise à l'échelle de la fonction objectif
Partage dynamique (<i>Dynamic Niching</i>) Partage rapide (<i>Fast Sharing</i>)	<ul style="list-style-type: none"> • $O(Nq)$ ou $O(2Nq)$ comparaisons de distances • efficace pour des problèmes "faciles" 	<ul style="list-style-type: none"> • voisinage explicite • parallélisme partiellement possible • équilibre des sous-populations difficile à obtenir • nécessite une taille de population importante et une mise à l'échelle de la fonction objectif
Méthode des sous-populations (<i>Subpopulation Schemes</i>)	<ul style="list-style-type: none"> • simplicité • voisinage implicite • parallélisme possible 	<ul style="list-style-type: none"> • mauvaise détection des solutions locales • deux sous-populations peuvent converger une même solution
Partage séquentiel (<i>Sequential Niching</i>)	<ul style="list-style-type: none"> • adaptable à toute méthode d'optimisation locale ou globale 	<ul style="list-style-type: none"> • voisinage explicite • parallélisme impossible • apparition de solutions inexistantes due au réajustement utilisé
Méthode d'éclaircissement (<i>Clearing</i>)	<ul style="list-style-type: none"> • parallélisme possible • élitisme • $O(Nq)$ comparaisons de distances • bons critères de performance 	<ul style="list-style-type: none"> • voisinage explicite • peu efficace pour des tailles de population faibles

III.8. CONCLUSION

Les méthodes génétiques multimodales offrent l'avantage d'une détection multiple de solutions optimales sans toutefois vraiment augmenter la complexité ou le temps de calcul des algorithmes. Le surcoût en temps CPU introduit par l'évaluation des distances entre individus est généralement négligeable par rapport au temps nécessaire pour estimer l'objectif. Ceci est le cas dans la majorité des problèmes réels et notamment en électromagnétisme où le calcul des fonctions objectif est souvent réalisé au moyen d'éléments finis.

Par ailleurs, l'introduction des mécanismes de nichage améliore considérablement l'exploration locale et rend inutile l'hybridation des algorithmes par une méthode d'optimisation locale. Comme nous l'avons signalé au chapitre précédent, il est toujours délicat pour des méthodes hybrides de déterminer le moment propice où l'on va effectuer la commutation entre l'algorithme génétique et la procédure d'optimisation locale.

Signalons enfin que dans les algorithmes génétiques multimodaux, l'équilibre entre la sélection et la recombinaison est naturellement assurée par les mécanismes de nichage. Cela évite de se poser la question du choix des valeurs des probabilités de croisement et de mutation. Une probabilité de croisement de 1 et des taux de mutation généralement faibles garantissent toujours une bonne robustesse à la recherche.

**APPLICATIONS DES ALGORITHMES GENETIQUES
MULTIMODAUX A DES PROBLEMES D'ELECTROMAGNETISME**

Dans ce chapitre, nous illustrons par deux problèmes simples d'électromagnétisme l'intérêt des algorithmes génétiques multimodaux et des méthodes de nichage. Ces problèmes ont été spécifiquement définis de façon à présenter un ensemble de solutions optimales, et ce dans le but de tester l'efficacité des divers algorithmes.

Le premier problème concerne une structure diélectrique stratifiée dont nous désirons fixer le comportement macroscopique. Le deuxième est relatif à l'optimisation d'un pôle d'inducteur.

IV.1. OPTIMISATION D'UN COMPOSITE DIELECTRIQUE

IV.1.1. Définition du problème

On désire déterminer la structure d'un composite diélectrique de telle sorte que sa permittivité effective macroscopique ϵ_{eff} dans la direction du champ appliqué soit égale à ϵ_{obj} . (voir Figure 36). Ce problème est trivial puisque l'on connaît une expression analytique de la permittivité effective [Sareni 96]. Celle-ci est obtenue en considérant l'association en série de deux diélectriques,

$$\epsilon_{eff} = \frac{\epsilon_1 \epsilon_2}{\epsilon_2 f + \epsilon_1 (1 - f)} \quad (115)$$

où ϵ_1 désigne la permittivité du diélectrique 1, ϵ_2 la permittivité du diélectrique 2 et f la concentration en matériau 1.

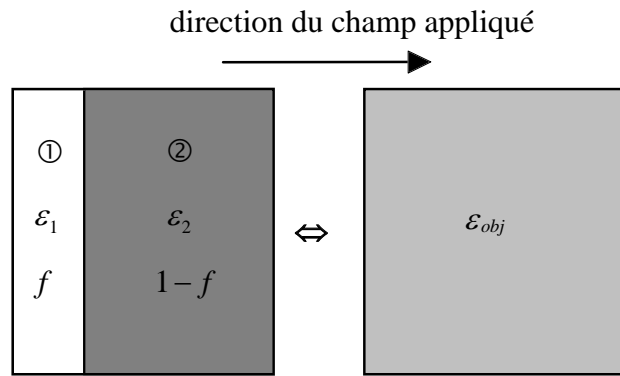


Figure 36. Homogénéisation d'une structure diélectrique stratifiée

Supposons que l'on veuille fabriquer un composite de permittivité effective macroscopique $\varepsilon_{obj} = 1.5$ à partir d'une gamme de matériau 1 de permittivité variant entre 10 et 30, la permittivité du matériau 2 étant imposée ($\varepsilon_2 = 1$) et la concentration en matériau 1 pouvant varier entre 0.1 et 0.9. Le problème d'optimisation à résoudre peut être exprimé de la manière suivante en fonction de deux paramètres ε_1 et f ,

$$\begin{cases} \min_{f, \varepsilon_1} f_{obj} = |\varepsilon_{eff} - 1.5| \\ 0.1 \leq f \leq 0.9 \\ 10 \leq \varepsilon_1 \leq 30 \end{cases} \quad (116)$$

La solution de (116) est évidente puisque d'après (115), nous pouvons déduire un ensemble de couples de solutions telles que $\varepsilon_{eff} = \varepsilon_{obj} = 1.5$ soit $f_{obj} = 0$. La Figure 37 illustre une représentation de la fonction objectif

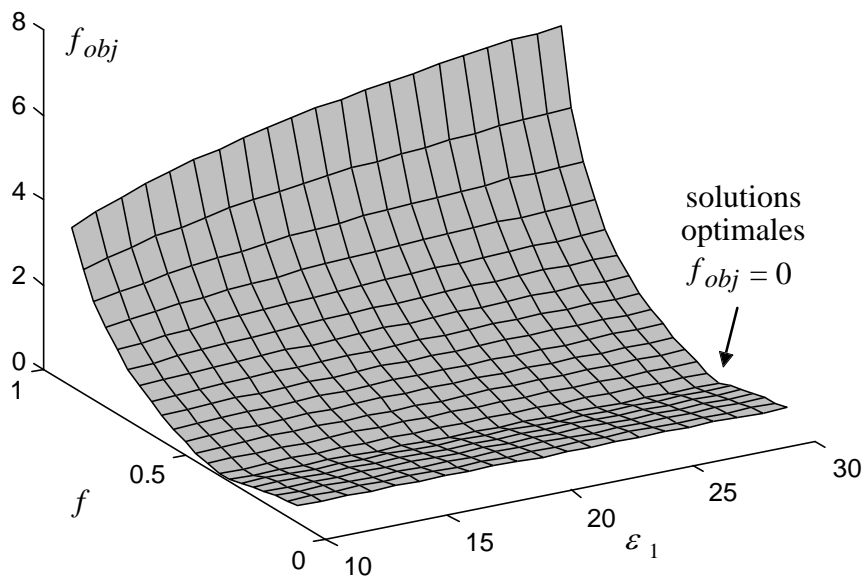


Figure 37. Homogénéisation d'une structure diélectrique stratifiée - représentation de la fonction objectif

Les solutions optimales sont réparties le long d'une ligne $f_{obj}=0$ située à l'intersection de deux surfaces de gradient opposé.

IV.1.2. Résolution par des méthodes génétiques

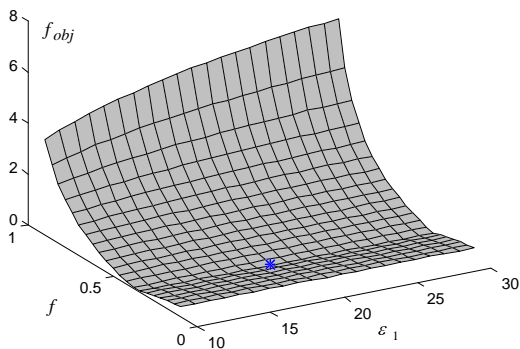
S'il peut être aisé de déterminer *une* solution optimale du problème par un algorithme d'optimisation quelconque, il est sans doute beaucoup plus délicat d'identifier l'*ensemble* des solutions. En effet, comme nous l'avons signalé au chapitre précédent, l'utilisation de méthodes d'optimisation unimodales de façon séquentielle ne garantit pas la détection de tous les optima. L'autre inconvénient de cette approche est la réexploration possible de solutions déjà trouvées et le coût supplémentaire introduit par les multiples exécutions séquentielles de l'algorithme unimodal. L'application des méthodes génétiques multimodales offre une alternative intéressante car nous pouvons nous attendre à ce que les individus de la population soient répartis de façon quasi-uniforme le long de la ligne optimale au bout d'un certain nombre de générations.

Quatre algorithmes génétiques multimodaux sont testés sur ce problème d'homogénéisation à savoir le surpeuplement déterministe, la sélection par tournoi restreint, l'éclaircissement et la méthode de partage. Nous cherchons notamment à déterminer le comportement des diverses techniques multimodales et la répartition qualitative des individus après exécutions des divers algorithmes.

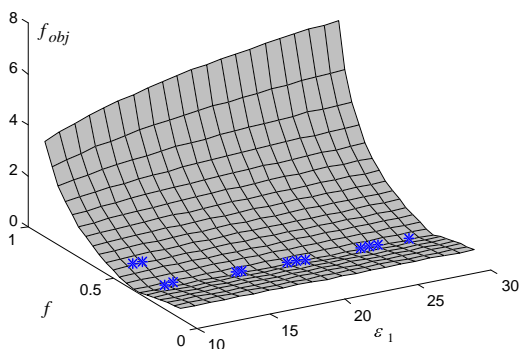
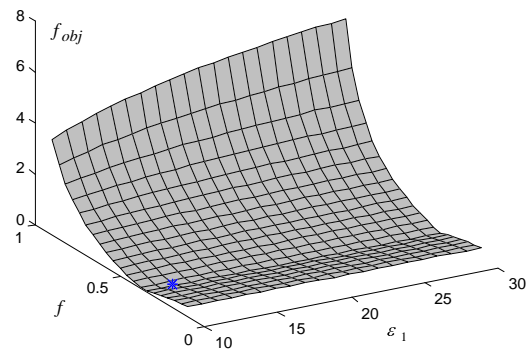
La taille de la population et le nombre de générations sont fixés à 30. Les paramètres sont codés sur 30 bits avec décodage dynamique. Les probabilités de croisement et de mutation sont respectivement $p_c = 1$ et $p_m = 0.001$. Le critère de distance utilisé est de type d^∞ normalisé (voir chapitre III). Le calcul de la fonction objectif est réalisé de façon analytique à l'aide de l'équation (115).

Nous représentons la répartition des individus de la population à la suite de deux exécutions des divers algorithmes (Figure 38, Figure 39 et Figure 40). Nous constatons bien qu'un algorithme génétique standard converge indifféremment vers un point de la ligne des solutions optimales et qu'il concentre toute sa population en cet unique point. Avec un même nombre d'évaluations de la fonction, les méthodes de surpeuplement distribuent la population le long de la ligne des solutions et localisent un grand nombre de configurations optimales distinctes. Cependant, les individus se regroupent par paquets (*clusters*) dans des niches dispersées aléatoirement sur la courbe définie par $f_{obj}=0$. Cette particularité est visible aussi bien pour le surpeuplement déterministe que pour la sélection par tournoi restreint. La méthode d'éclaircissement donne des résultats assez similaires mais les individus de la population sont cette fois distribués de façon quasi-uniforme le long de la ligne des solutions

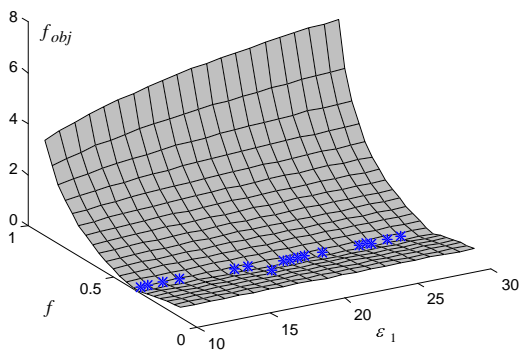
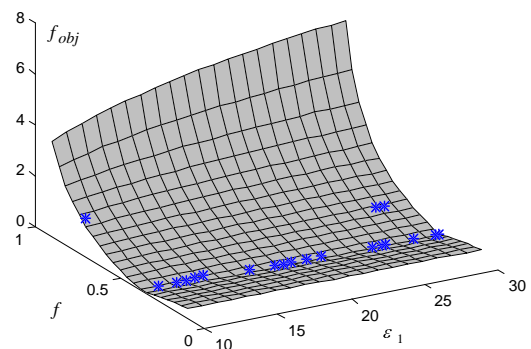
optimales. La distance entre les configurations optimales distinctes est bien fonction du seuil de similarité imposé. Plus le rayon de niche est élevé, plus les solutions détectées sont éloignées. Lorsque ce rayon est important, l'éclaircissement localise un faible nombre de solutions fortement distinctes et répartit le reste de la population de façon quasi-uniforme dans d'autres directions de l'espace. Le comportement de la méthode de partage est assez "chaotique" et désordonné. L'équilibre entre les sous-populations n'a pu être obtenu avec les paramètres utilisés. Seule quelques solutions optimales sont trouvées avec des faibles rayons de niche. L'augmentation du rayon de niche accroît la pénalisation des individus et restreint fortement l'exploitation.



(a) Algorithme génétique standard



(b) Sélection par tournoi restreint



(c) Surpeuplement déterministe

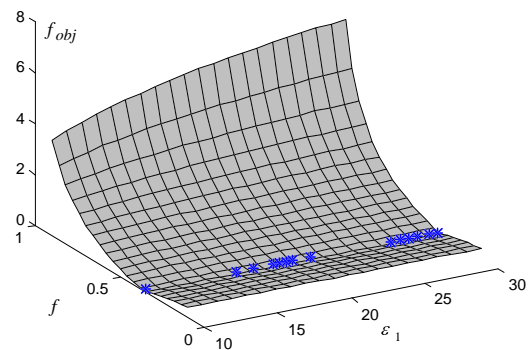
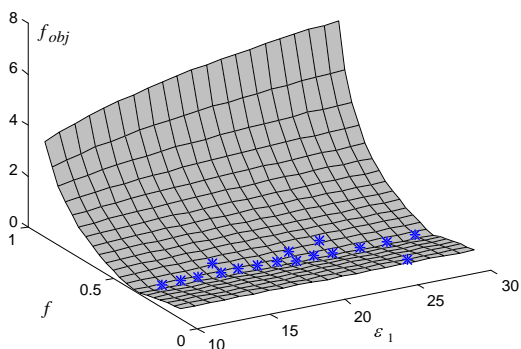
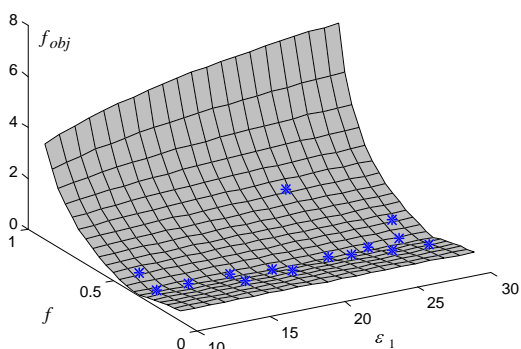
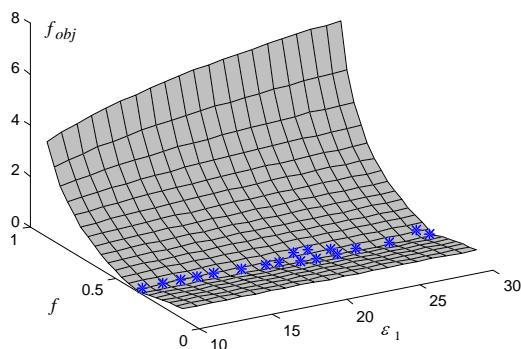


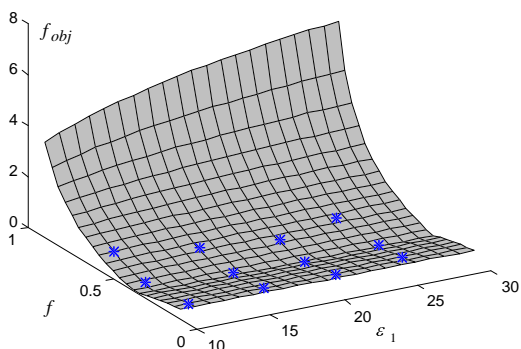
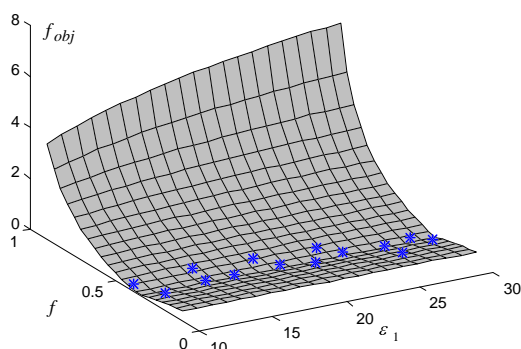
Figure 38. Distribution de la population à la suite de deux exécutions distinctes de méthodes à voisinage implicite - 30 individus 30 générations - Deux exécutions



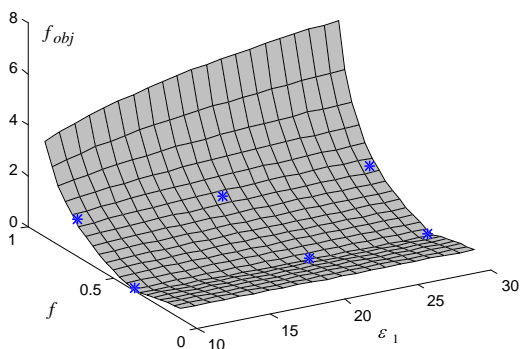
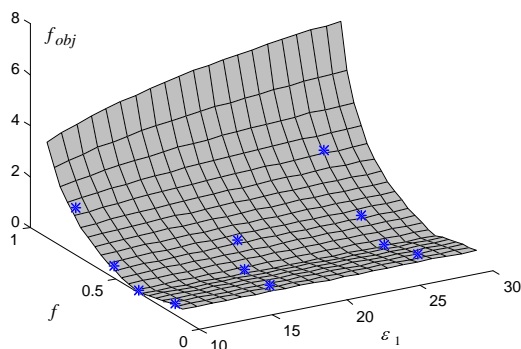
(a) $\sigma_s = 0.05$



(b) $\sigma_s = 0.1$



(c) $\sigma_s = 0.2$



(d) $\sigma_s = 0.4$

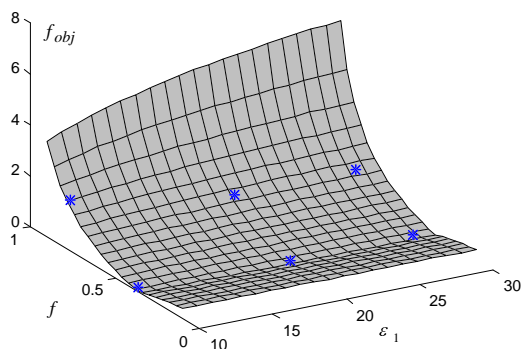
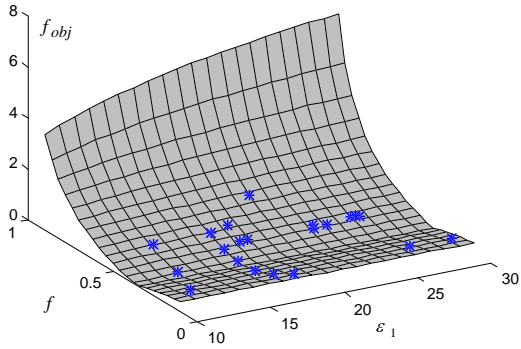
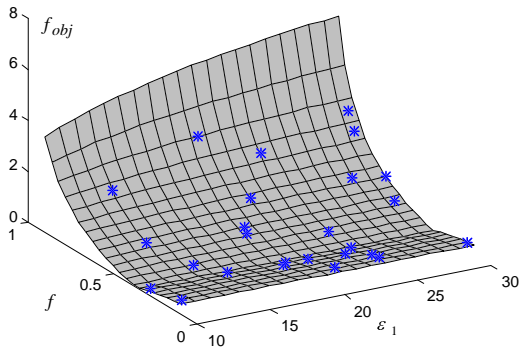
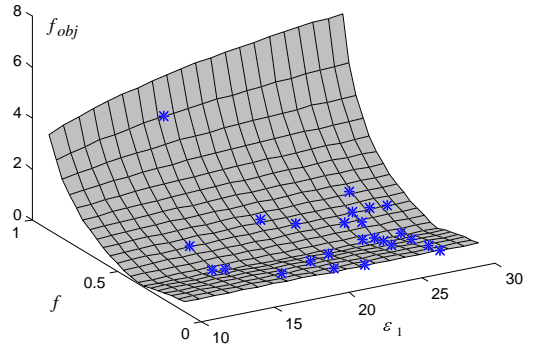


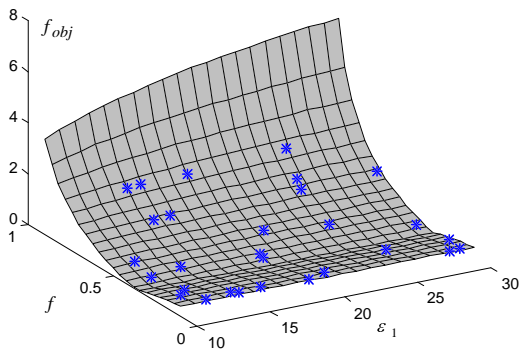
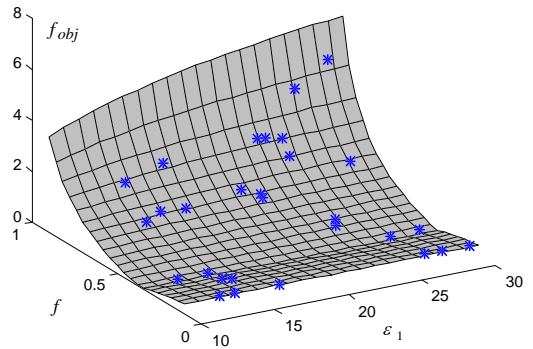
Figure 39. Distribution de la population à la suite de deux exécutions distinctes de la méthode d'éclaircissement en fonction du rayon de niche - 30 individus 30 générations



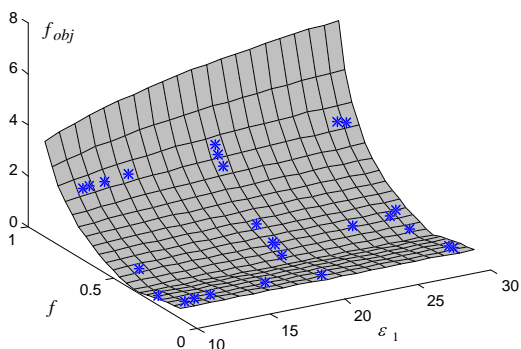
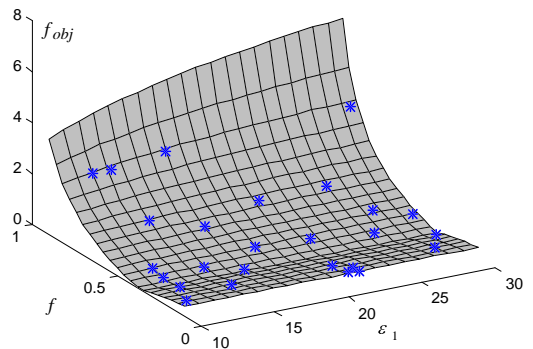
(a) $\sigma_s = 0.05$



(b) $\sigma_s = 0.1$



(c) $\sigma_s = 0.2$



(d) $\sigma_s = 0.4$

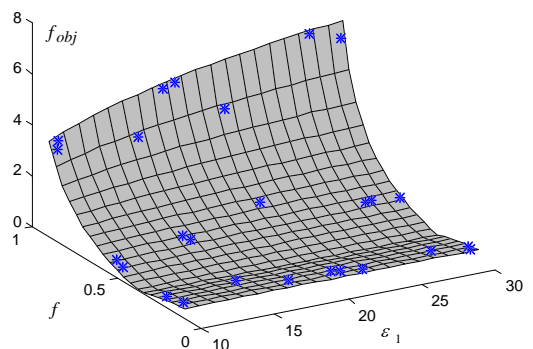


Figure 40. Distribution de la population à la suite de deux exécutions distinctes de la méthode de partage en fonction du rayon de niche - 30 individus 30 générations

Afin d'éviter l'interception du pôle avec le matériau à aimanter et assurer un entrefer au moins égal à 2 mm, les paramètres sont soumis à une contrainte géométrique supplémentaire. Celle-ci garantit que l'abscisse du point P' reste inférieure ou égale à 53 mm

$$x_1 + \sqrt{(x_2^2 + (x_3 - x_1)^2)} \leq 53 \quad (117)$$

L'objectif du problème est de trouver le ou les profils de pôles qui conduisent à une forme d'induction donnée sur l'arc AB. Nous recherchons en particulier les profils de pôle tels que :

1. l'induction soit uniforme sur l'arc AB
2. l'induction soit sinusoïdale sur l'arc AB

IV.2.2. Couplage avec la méthode des éléments finis

Ces deux problèmes sont traités par des algorithmes génétiques couplés à la méthode des éléments finis pour le calcul de l'induction sur l'arc AB. La procédure d'optimisation de forme associe un logiciel de calcul de champ 2D [Flux2D], avec un algorithme génétique selon l'organigramme de la Figure 42.

L'algorithme génétique choisit aléatoirement les paramètres des individus de la population initiale et détermine de génération en génération la valeur des nouveaux paramètres à partir des opérateurs de *sélection* et de *recombinaison*. A chaque génération, la performance des individus de la population est estimée par le logiciel de calcul de champ. Lorsqu'un individu correspond à une configuration non réalisable, un terme de pénalité est directement affecté à sa performance.

Cette approche est relativement intéressante car l'algorithme d'optimisation et le programme de calcul par éléments finis fonctionnent indépendamment. Seules les difficultés de communication entre les deux modules peuvent poser problème.

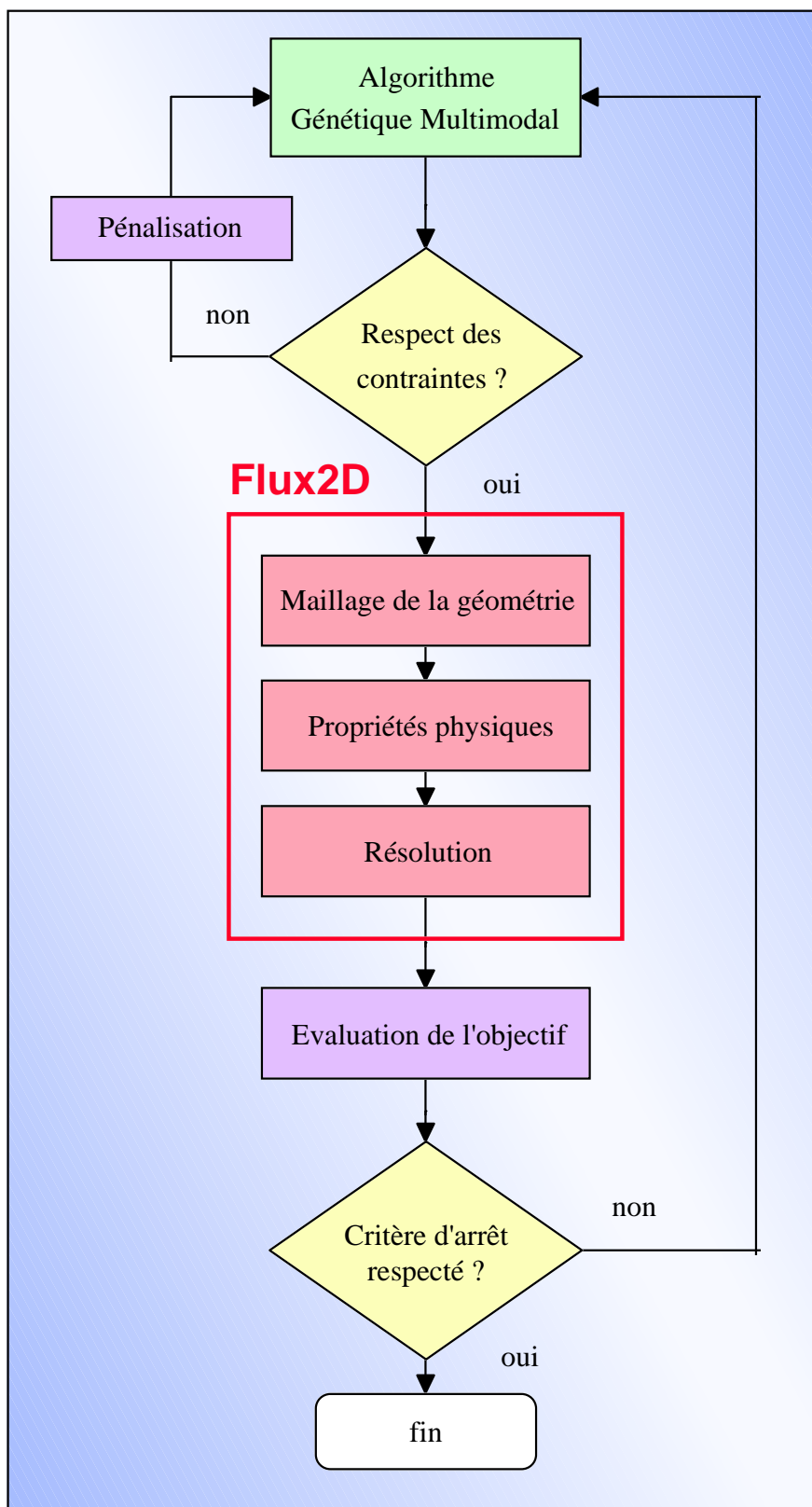


Figure 42. Principe de l'optimisation de forme par éléments finis

IV.2.3. Induction constante sur l'arc AB

Dans un premier temps, nous cherchons à obtenir une induction uniforme sur l'arc AB mais nous ne spécifions volontairement aucune valeur normative pour permettre d'avoir plusieurs solutions optimales dans le domaine de variation des paramètres défini au Tableau 23. Le problème d'optimisation peut être formulé sous la forme suivante :

$$\left\{ \begin{array}{l} \min f_{obj} = \frac{\sigma_B}{\mu_B} \\ x_1 + \sqrt{(x_2^2 + (x_3 - x_1)^2)} \leq 53 \\ -5 \leq x_1 \leq 25 \\ 25 \leq x_2 \leq 45.9 \\ 25 \leq x_3 \leq 26.5 \end{array} \right. \quad (118)$$

où μ_B désigne la moyenne de l'induction sur l'arc AB et σ_B l'écart-type correspondant. L'induction est calculée en 50 points de l'arc AB. La fonction objectif peut être réécrite suivant (119),

$$f_{obj} = \frac{\sqrt{n \sum_{i=1}^n B_i^2 - \left(\sum_{i=1}^n B_i \right)^2}}{\sum_{i=1}^n B_i} \quad (119)$$

où B_i est la valeur de l'induction en un point i de l'arc AB, et $n=50$ le nombre total de points sur l'arc. Lorsque la contrainte géométrique est violée, l'objectif est exprimé par un terme de pénalité :

$$f_{obj} = \lambda_1 + \lambda_2 (x_1 + \sqrt{(x_2^2 + (x_3 - x_1)^2)} - 53) \quad (120)$$

calculé en prenant $\lambda_1 = \lambda_2 = 10$.

Le passage en un problème de maximisation est effectué en modifiant la fonction objectif de la façon suivante :

$$F_{obj} = \frac{1}{1 + f_{obj}} \quad (121)$$

Les méthodes multimodales employées précédemment sont utilisées pour la détection des profils de pôles à induction uniforme sur l'arc AB. A cause du temps important nécessaire pour évaluer l'objectif d'une configuration réalisable (1 à 2 minutes sur une station HP 9000/712), la taille de la population dans les algorithmes génétiques est limitée à 30 de même

que le nombre de générations. Les paramètres sont codés sur 30 bits avec décodage dynamique. Les probabilités de croisement et de mutation sont respectivement $p_c = 1$ et $p_m = 0.001$. Le critère de distance utilisé est de type d^∞ normalisé (voir chapitre III).

Nous présentons par la suite les caractéristiques des solutions optimales obtenues avec les divers algorithmes génétiques dans le Tableau 24. Nous insistons ici sur le caractère qualitatif des résultats en raison du nombre très limité de tests effectués dans chaque cas. Le temps de calcul élevé par test (au minimum 24 heures) a imposé de réduire le nombre d'exécutions de chaque algorithme à 3. Par ailleurs, le logiciel de champ a pu introduire un biais non voulu lors de la recherche des solutions optimales. En effet, la version de Flux2D dont nous disposons ne possède pas de maillage adaptatif. Bien que la géométrie soit modifiée en fonction de la variation des paramètres, le nombre de points sur lesquels le maillage s'appuie reste identique. D'une configuration à l'autre le maillage est simplement déformé et étiré. Lorsque les variables objets ont un domaine de variation large comme c'est le cas pour x_1 et x_2 , des valeurs de paramètres peuvent conduire à un maillage donnant des erreurs numériques importantes. Dans certains cas même, des configurations ne peuvent être calculées avec le maillage défini pour une forme de pôle quelconque. Bien que nous ayons testé un maillage initial sur des configurations correspondant à des valeurs extrêmes des paramètres nous n'avons jamais pu éliminer totalement ces problèmes de conditionnement numérique.

Tableau 24. Caractéristiques des solutions optimales - Meilleurs résultats obtenus après 3 exécutions des divers algorithmes - 30 générations - 30 individus

<i>Algorithme d'optimisation</i>	<i>Nombre de solutions optimales distinctes ($f_{obj} < 0.03$)</i>	<i>niveaux d'induction uniformes détectés</i>	<i>Moyenne des solutions optimales f_{obj}</i>
AG standard	1	1 seul niveau d'amplitude variable	0.002
CL ($\sigma_s = 0.1$)	14	[0.098 T - 0.427 T]	0.014
CL ($\sigma_s = 0.4$)	4	[0.125 T - 0.442 T]	0.009
SH ($\sigma_s = 0.1$)	4	[0.103 T - 0.150 T]	0.024
DC	20	[0.117 T - 0.563 T]	0.011
RTS ($w=10$)	20	[0.122 T - 0.596 T]	0.008

Nous déduisons toutefois des résultats qualitatifs du Tableau 24. Ici encore, un algorithme génétique converge vers une unique solution mais les trois exécutions conduisent à trois solutions différentes. Les méthodes de surpeuplement réussissent bien sur ce problème en

localisant une vingtaine de solutions dont les niveaux d'induction uniformes varient entre 0.1 T et 0.6 T. La méthode d'éclaircissement semble moins performante mais ceci peut s'expliquer par un mauvais choix du rayon de niche et une relative faiblesse lorsque la taille de la population est restreinte. Le partage est très peu efficace en trouvant uniquement quatre solutions de qualité moyenne.

La répartition dans l'espace des paramètres de 20 individus représentant des solutions optimales est illustré à la Figure 43. Nous montrons les niveaux d'induction uniformes sur l'arc AB correspondant à ces solutions à la Figure 44. La Figure 45 présente finalement cinq formes de pôle optimales.

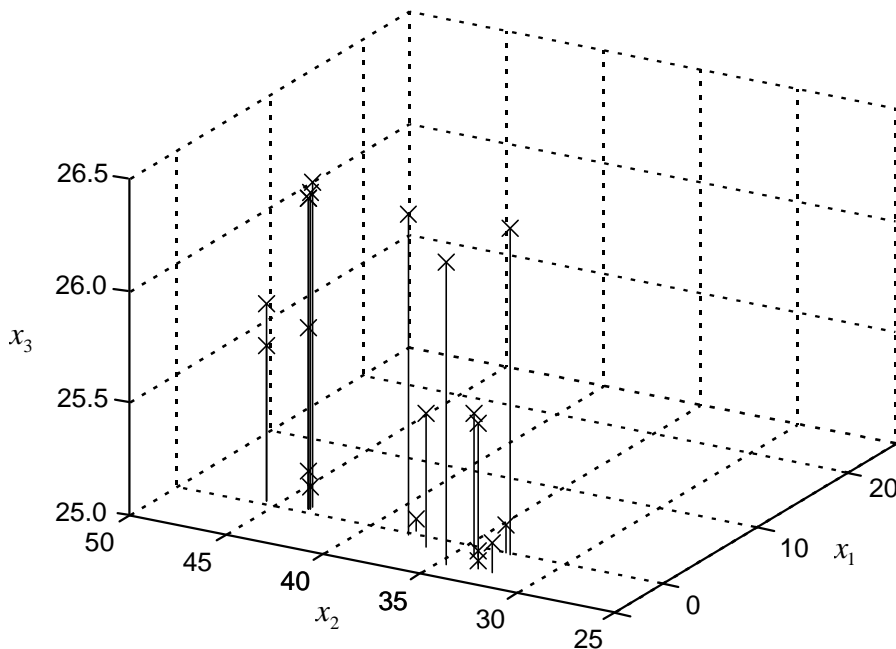


Figure 43. Distribution dans l'espace de recherche de 20 solutions optimales détectées par la sélection par tournoi restreint

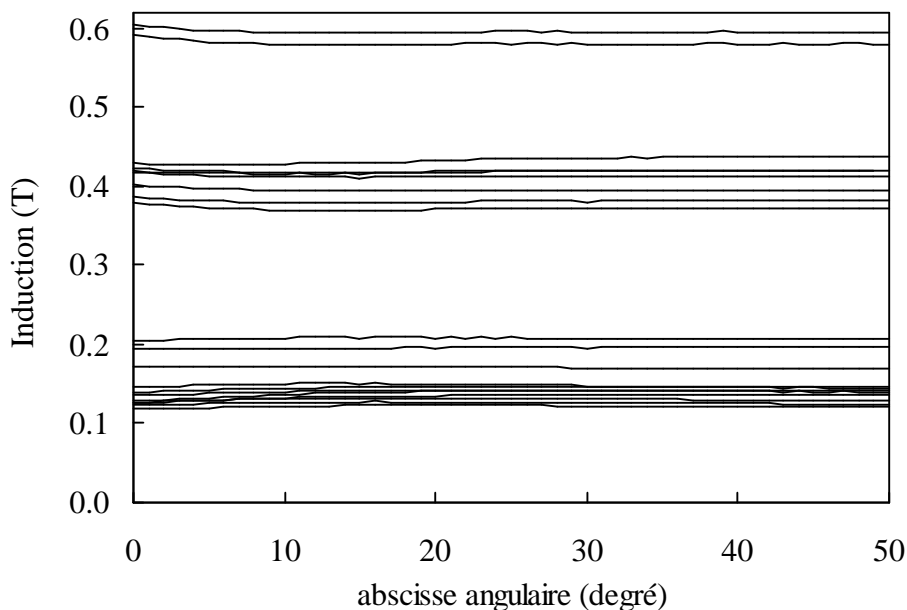


Figure 44. Valeur de l'induction sur l'arc AB pour les 20 profils optimaux

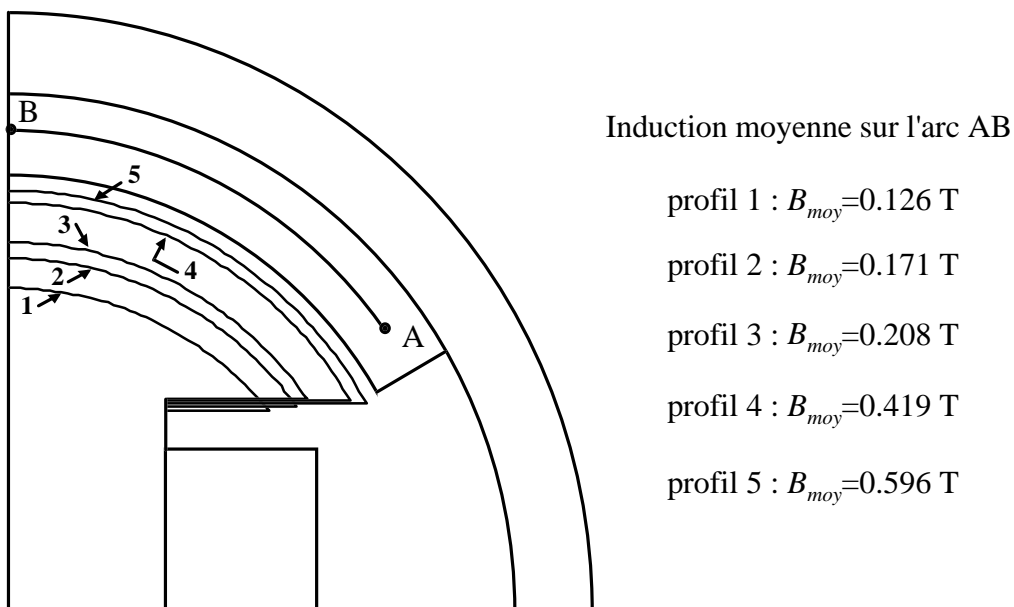


Figure 45. Exemples de profils optimaux

Comme nous pouvions nous y attendre, les profils détectés correspondent à des configurations de pôle à entrefer globalement constant. Avec le critère de distance lié aux paramètres géométriques, il n'est toutefois pas évident de répartir les solutions régulièrement

Chapitre IV

en fonction des niveaux d'induction uniformes sur l'arc AB. En effet, une variation linéaire des paramètres autour d'une solution optimale n'entraîne pas une variation linéaire de la valeur de l'induction sur l'arc AB (cette valeur est approximativement inversement proportionnelle à l'épaisseur de l'entrefer). Il est sans doute plus judicieux d'utiliser directement un critère de distance physique basé sur la valeur moyenne de l'induction sur l'arc AB pour définir la similarité entre individus au sein de la population. Nous avons testé le critère de distance suivant pour déterminer la similarité entre deux individus a_i^t et a_j^t :

$$d(a_i^t, a_j^t) = \left| \langle B(a_i^t) \rangle - \langle B(a_j^t) \rangle \right| \quad (122)$$

où $\langle B(a_k^t) \rangle$ désigne la moyenne de l'induction sur l'arc AB pour la configuration relative à l'individu a_k^t . Avec ce critère de distance et une méthode de nichage à voisinage explicite, nous pouvons espérer obtenir une répartition régulière des solutions optimales en fonction des niveaux d'induction uniformes.

Nous traitons le problème avec la méthode d'éclaircissement associée au critère de distance défini précédemment. Les paramètres de l'algorithme génétique sont identiques à ceux utilisés auparavant. Le rayon de niche est fixé à 0.05. De cette manière, nous nous attendons à trouver des profils de pôles donnant une induction uniforme sur l'arc AB et dont les niveaux varient au moins de 0.05 T de proche en proche.

La Figure 46 montre, dans l'espace des paramètres, la répartition des 7 configurations optimales obtenues avec le nouveau critère de distance employé. La fonction objectif f_{obj} de ces solutions est voisine de 0.003. La Figure 47 et la Figure 48 illustrent respectivement les valeurs de l'induction sur l'arc AB et la forme des profils correspondants. Nous constatons que nos prédictions sont bien vérifiées. L'utilisation d'une distance appropriée (pas nécessairement liée aux paramètres géométriques ou structurels) avec une méthode de nichage à voisinage explicite permet de contrôler la diversité et les caractéristiques des solutions finales.

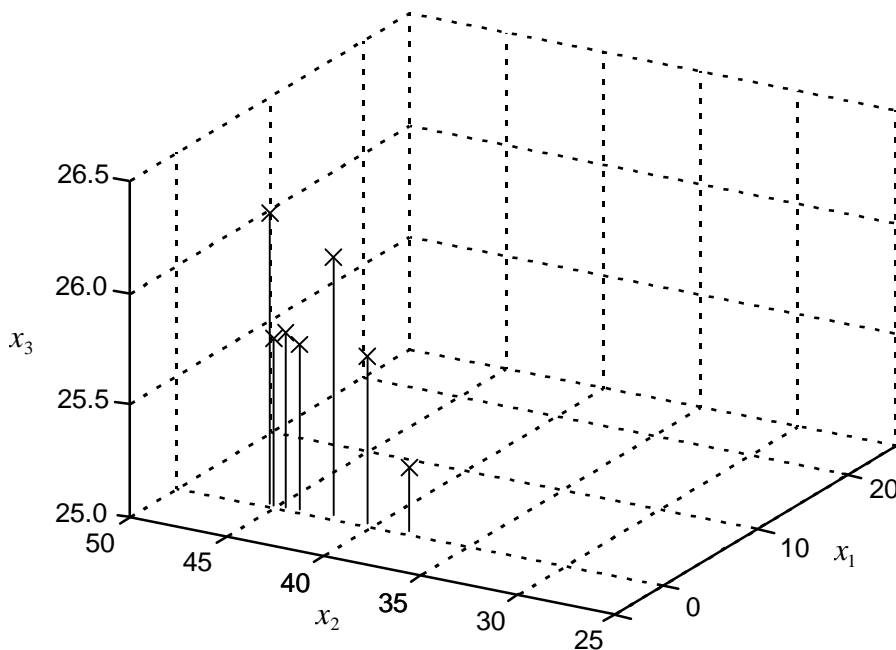


Figure 46. Distribution dans l'espace de recherche de 7 solutions optimales détectées par la méthode d'éclaircissement ($\sigma_s = 0.05$) avec un critère de distance physique liée à l'induction moyenne sur l'arc AB.

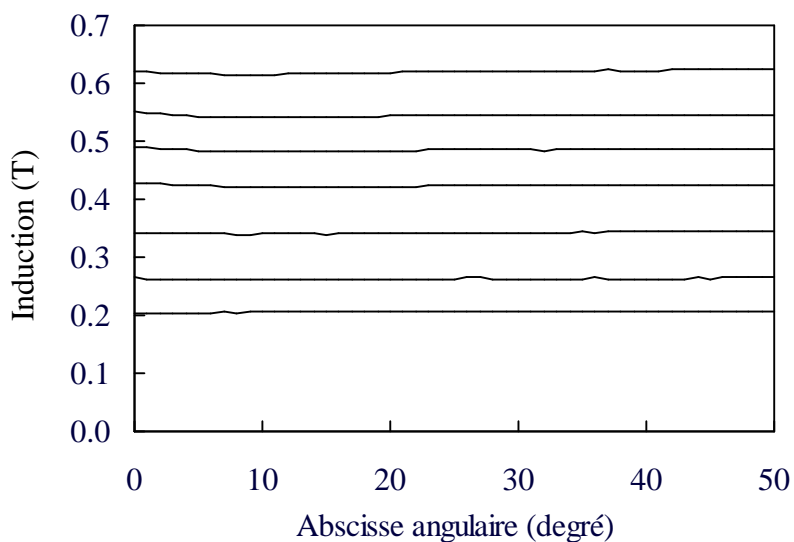


Figure 47. Valeur de l'induction sur l'arc AB pour les 7 profils optimaux

tel-00139202, version 1 - 29 Mar 2007

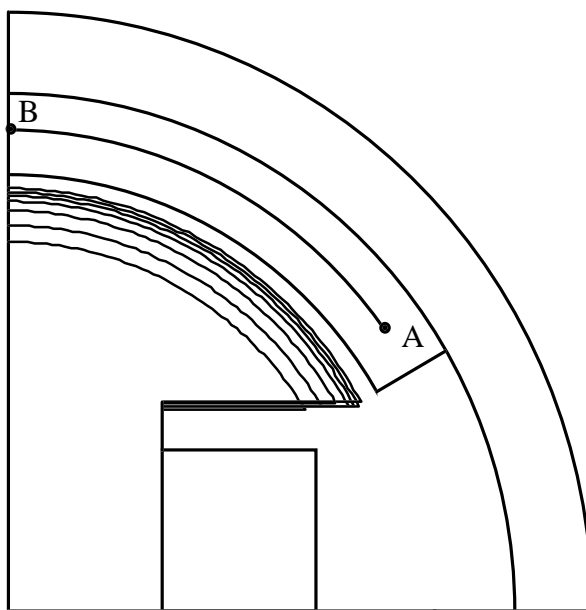


Figure 48. forme des 7 profils optimaux

IV.2.4. Induction sinusoïdale sur l'arc AB

Nous illustrons rapidement le cas de pôles donnant une induction sinusoïdale sur l'arc AB. Le problème est ici similaire à celui proposé par Üler et al. [Üler 95] mais nous recherchons une forme de pôle paramétrée de la même manière que précédemment. De plus, nous ne spécifions aucune valeur normative pour l'induction afin que le problème admette plusieurs solutions distinctes.

L'induction sur l'arc AB peut être écrite suivant une série de Fourier :

$$B(\theta) = \sum_{k=1}^{\infty} B_k \cos(k\theta + \varphi_k) \quad (123)$$

où B_k et φ_k désignent respectivement l'amplitude et la phase de l'harmonique k et θ l'abscisse angulaire sur l'arc AB.

Pour que l'induction soit sinusoïdale, il est nécessaire de réduire l'amplitude des harmoniques de rang supérieur à un. Plutôt, que de minimiser une somme pondérée d'harmoniques [Gizolme 97], nous avons choisi d'exprimer l'objectif sous la forme d'un critère flou :

$$f_{obj} = \max_{k=3,5,\dots,29} \frac{B_k}{B_1} \quad (124)$$

où B_1 représente l'amplitude du fondamental de l'induction. Le problème d'optimisation est formulé d'une façon analogue à (118) en considérant la fonction objectif définie précédemment et transformé en terme de maximisation suivant la relation (121).

Le problème est traité à l'aide du surpeuplement déterministe en utilisant une taille de population et un nombre de génération de 30. Les probabilités de croisement et de mutation sont identiques à celles employées précédemment. Le critère de distance est de type d^∞ normalisé.

La Figure 49 illustre, dans l'espace des paramètres, la répartition des individus de la population au bout de 30 générations. Le Tableau 25 présente six solutions extraites de la population finale. La valeur correspondante de l'induction sur l'arc AB est montrée à la Figure 50 en fonction du profil considéré.

Nous constatons qu'avec le critère d'optimisation et le paramétrage employés, nous pouvons identifier des profils de pôles donnant un induction sinusoïdale de phase opposée sur l'arc AB. L'amplitude des sinusoïdes et la composante continue correspondante sont variables suivant les profils.

Tableau 25. Exemple de profils optimaux

	x_1	x_2	x_3	f_{obj}
<i>profil 1</i>	4.02	38.59	26.26	0.238
<i>profil 2</i>	4.01	38.34	26.04	0.262
<i>profil 3</i>	6.35	35.60	25.82	0.324
<i>profil 4</i>	8.00	33.38	26.32	0.326
<i>profil 5</i>	-3.59	39.68	26.17	0.303
<i>profil 6</i>	-4.64	38.50	25.54	0.244

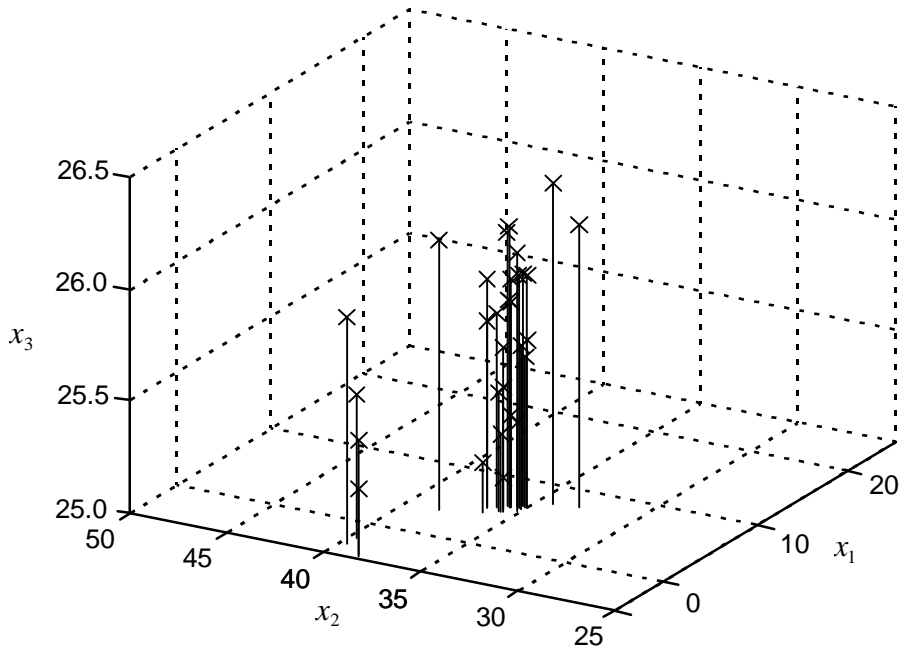


Figure 49. Répartition des individus de la population après 30 générations

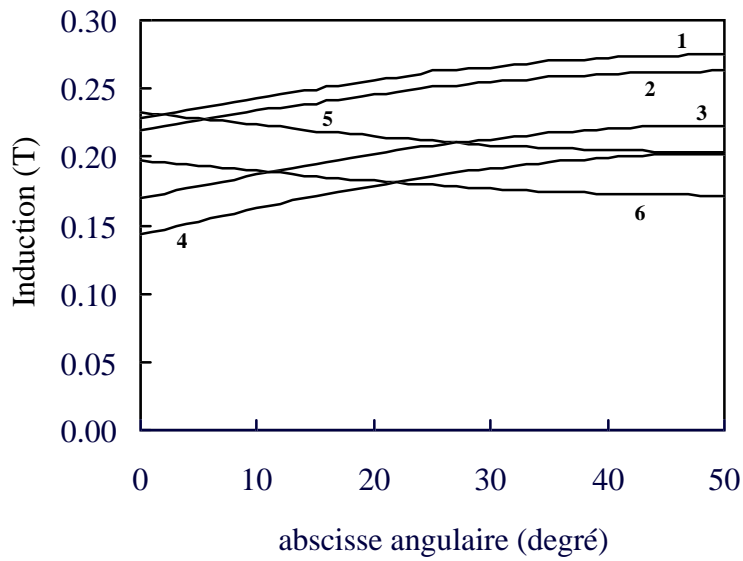


Figure 50. Valeur de l'induction sur l'arc AB pour les 6 solutions optimales

IV.3. CONCLUSION

Les tests effectués dans ce chapitre ont permis de caractériser le comportement des méthodes de nichage lorsque le problème d'optimisation admet un grand nombre de solutions distinctes. Les techniques de surpeuplement sont assez performantes quand elles sont utilisées

avec une faible taille de population. La méthode de partage est par contre très limitée dans ce cas de figure. La technique d'éclaircissement est beaucoup plus efficace surtout si le critère de distance est judicieusement choisi. Elle permet alors de répartir les individus de la population de façon régulière et quasi-uniforme dans les zones de l'espace où la fonction est optimale.

CHAPITRE V

OPTIMISATION DE FORME D'ELECTRODES

Dans ce chapitre, nous présentons une méthode originale de conception de formes optimales d'électrodes pour les systèmes 2D-plan ou axisymétriques. Le procédé s'appuie sur les idées qui avaient permis dès le XIX^{ème} siècle de trouver des solutions analytiques pour la répartition du potentiel entre électrodes dans de nombreuses configurations particulières. On parlait des répartitions générées dans le vide par des sources imposées (des charges ou des dipôles par exemple) ; toute courbe équipotentielle de ce système fictif pouvait alors représenter une électrode dans un dispositif équivalent, dont la solution était ainsi connue.

En C.A.O. traditionnelle, c'est la forme de l'électrode qui serait paramétrée de façon directe à l'aide de rayons et de centres de courbures. Avec ce procédé original, les paramètres sont les positions et les valeurs des sources. La conception d'une forme d'électrode se résume alors à la résolution d'un problème d'optimisation sous contraintes : les paramètres doivent être optimisés d'une part, de façon à ce que les formes générées respectent certaines contraintes géométriques (sous forme de limites à ne pas franchir), d'autre part pour qu'un objectif dépendant de la solution en champ électrique soit atteint.

V.1. EQUATIONS DE L'ELECTROSTATIQUE

Les phénomènes électrostatiques sont régis par les deux équations de Maxwell suivantes,

$$\nabla \times \mathbf{E} = 0 \quad (125)$$

$$\nabla \cdot \mathbf{D} = \rho \quad (126)$$

où \mathbf{E} est le champ électrique en $[\text{V.m}^{-1}]$

\mathbf{D} est l'induction électrique en $[\text{C.m}^{-2}]$

ρ est la densité de charge en $[\text{C.m}^{-3}]$

auxquelles il convient d'ajouter la relation caractérisant le milieu :

$$\mathbf{D} = \varepsilon \mathbf{E} \quad (127)$$

où ε désigne la permittivité en $[\text{F.m}^{-1}]$ du milieu.

Chapitre V

L'équation (125) implique que le champ électrique dérive d'une grandeur scalaire que l'on qualifie de potentiel électrique V ,

$$\mathbf{E} = -\nabla V \quad (128)$$

L'espace est alors caractérisé par une seule et unique équation obtenue en combinant (126), (127) et (128):

$$\nabla \cdot (\varepsilon \nabla V) = -\rho \quad (129)$$

Dans le cas où le milieu diélectrique est linéaire, isotrope et homogène, cette relation se réduit à l'équation de Poisson:

$$\Delta V = -\frac{\rho}{\varepsilon} \quad (130)$$

Si par ailleurs la densité de charge est nulle, le potentiel obéit en tout point de ce milieu à l'équation de Laplace:

$$\Delta V = 0 \quad (131)$$

Les solutions de ces équations différentielles s'obtiennent en fonction des conditions aux limites du domaine d'étude et aux frontières des milieux homogènes. Ces limites sont généralement :

- des surfaces où le potentiel est imposé (condition de Dirichlet)¹
- des surfaces où le potentiel est inconnu mais supposé constant (condition flottante)¹
- des plans de symétries où $\partial V / \partial n = 0$ (condition de Neuman homogène)
- le ou les plans de masse où $V=0$ (condition de Dirichlet homogène)

Les limites du domaine d'étude peuvent éventuellement être rejetées à l'infini où le potentiel V est supposé nul.

Pour les dispositifs comportant plusieurs milieux de constantes diélectriques différentes, les grandeurs électriques doivent satisfaire des conditions d'interface qui découlent de (125) à (128). Ainsi au passage d'un milieu 1 vers un milieu 2, nous aurons :

$$V_1 = V_2 \quad (132)$$

$$\varepsilon_1 \frac{\partial V_1}{\partial n} - \varepsilon_2 \frac{\partial V_2}{\partial n} = \rho_s \quad (133)$$

où ρ_s est la densité de charge surfacique sur l'interface, ε_1 et ε_2 désignant respectivement les permittivités des milieux 1 et 2. En l'absence de charge d'interface, l'équation (133) indique la conservation de la composante normale du déplacement électrique. L'équation (132) impose la continuité du potentiel au passage d'un milieu à l'autre.

V.2. DISTRIBUTIONS DE CHARGES DANS LE VIDE

Dans l'espace vide infini, le potentiel créé en un point P par une charge ponctuelle de valeur q située au point Q est donné par la loi de Coulomb :

$$V_q(P) = G(P, Q)q \quad (134)$$

où G est la fonction de Green

$$G(P, Q) = \frac{1}{4\pi\varepsilon_0 r_{PQ}} \quad (135)$$

r_{PQ} désignant la distance du point de calcul P à la charge q [en C].

Dans le cas d'un système de N charges ponctuelles, le potentiel s'obtient selon le principe de superposition en considérant chaque contribution élémentaire :

$$V(P) = \sum_{k=1}^N V_k(P) \quad (136)$$

Pour des distributions linéiques de charges constantes et de forme simple (segments de droite, anneaux ou lignes infinies), des expressions analytiques permettent de calculer le potentiel et le champ électrique en tout point de l'espace. L'équation (134) qui donne le potentiel créé en un point P par une charge ponctuelle Q peut être généralisée pour une distribution de charge de valeur q à la relation (137)²

$$V_q(P) = p \cdot q \quad (137)$$

où le facteur p est appelé coefficient de potentiel. Il dépend uniquement de la forme de la distribution et de sa position par rapport au point de calcul [Annexe IV]. Les composantes du champ électrique sont directement obtenues par dérivation des coefficients de potentiel :

$$\mathbf{E} = -\nabla p \cdot q \quad (138)$$

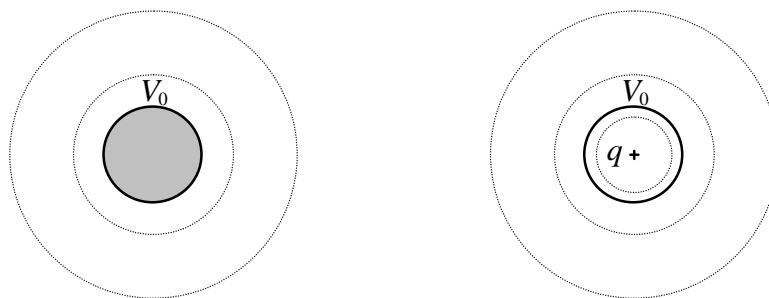
¹Ces surfaces correspondent aux conducteurs en électrostatique.

Pour un système de N distributions de charges, le potentiel est calculé de façon similaire à (136) par superposition en considérant chaque contribution élémentaire :

$$V(P) = \sum_{k=1}^N p_k \cdot q_k \tag{139}$$

V.3. SIMULATION DE CONDUCTEURS PAR DES CHARGES FICTIVES

La méthode d'optimisation de forme que nous avons mise au point est basée sur le principe des charges fictives utilisé dans la méthode des charges équivalentes [Annexe V]. Il s'agit, pour simuler les répartitions de potentiel et de champ électrique, de remplacer les conducteurs par des systèmes équivalents de charges fictives. Ainsi, par exemple, les répartitions de potentiel et de champ électrique, à l'extérieur d'une sphère conductrice de rayon R portée au potentiel V_0 , peuvent être obtenues en remplaçant cette sphère par une charge ponctuelle de valeur $q = 4\pi\epsilon_0 V_0 R$, placée en son centre (voir Figure 51). De la même manière, il est possible de simuler n'importe quelle électrode à l'aide d'un système équivalent de charges fictives, à condition que ces dernières soit situées à l'intérieur de l'électrode (voir Figure 52).



sphère conductrice de rayon R charge fictive équivalente $q = 4\pi\epsilon_0 V_0 R$

Figure 51. Simulation d'une sphère conductrice par une charge ponctuelle équivalente

² En 3D ou en 2D axisymétrique, q désigne la charge totale en [C] correspondant à la distribution linéique considérée (segment ou anneau chargé par exemple). En 2D plan, q représente la charge totale par unité de longueur en [C.m⁻¹].

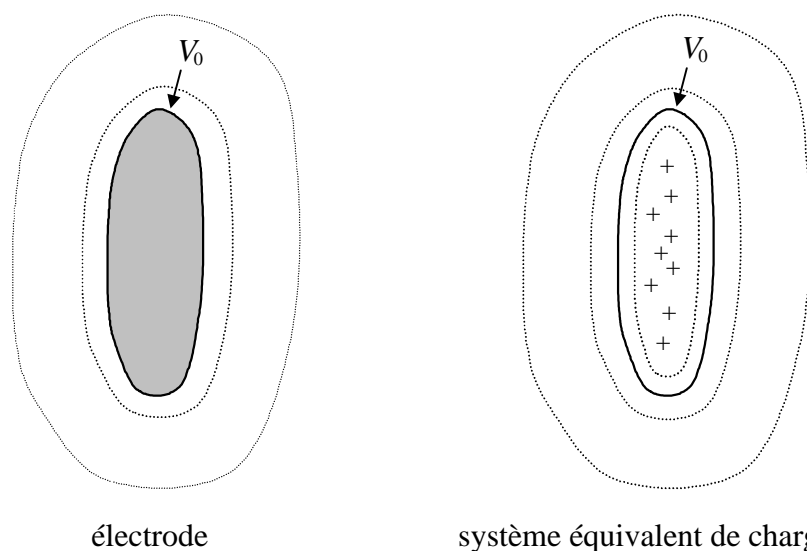


Figure 52. Simulation d'une électrode par un système équivalent de charges

V.4. METHODE DE GENERATION DE FORMES D'ELECTRODES OPTIMALES

V.4.1. Introduction

L'optimisation de la forme d'une électrode est habituellement réalisée en paramétrant directement son contour à l'aide de lignes prédéfinies (segments de droite, arcs, rayons de courbure, etc.) [Stih 86][Vasconcelos 94] [Mukherjee 96]. Dans l'exemple présenté par *Mukherjee* et al., l'extrémité de l'électrode est assimilée à une ellipse (voir Figure 53). Le profil est défini en fonction de deux paramètres x_1 et x_2 qui spécifient respectivement le demi-axe vertical et le centre de l'ellipse. Le demi-axe horizontal est fixé automatiquement lorsque le centre de l'ellipse est donné. Avec ce type de contour défini à l'avance, les degrés de liberté permettant la modification de la forme pour l'optimisation du champ électrique sont relativement limités. Nous présentons une approche originale qui permet de générer une forme quelconque à l'intérieur d'un gabarit géométrique donné à partir d'un système de charges ponctuelles. Le paramétrage indirect du profil de l'électrode permet d'explorer des formes plus variées que les méthodes classiques de paramétrage direct.

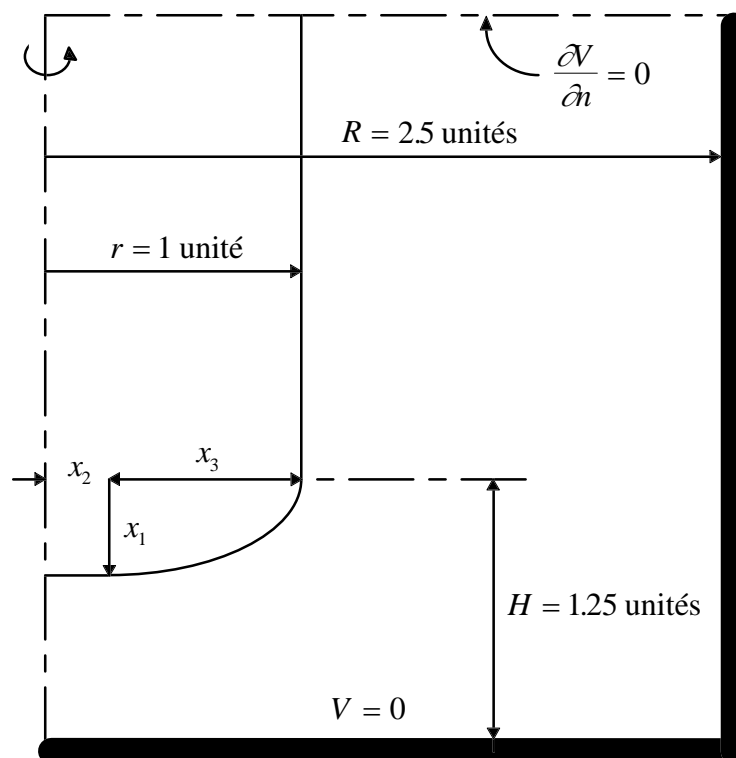


Figure 53. Exemple de forme d'électrode à paramétrage direct [Mukherjee 96]

V.4.2. Principe de la méthode

La méthode proposée consiste à définir le contour de l'électrode à partir d'un gabarit géométrique qui fixe les limites de la forme. Considérons par exemple le gabarit présenté à la Figure 54. Celui-ci est constitué de deux frontières, l'une interne et l'autre externe. Le profil de l'électrode recherchée peut prendre n'importe quelle forme à condition qu'elle "respecte" le gabarit géométrique, c'est à dire qu'elle ne sorte pas de l'aire délimitée par la frontière interne et la frontière externe (aire admissible hachurée).

Il faudra en plus que la forme choisie soit optimale en fonction de critères relatifs à son comportement électrostatique (contraintes en champ électrique minimales, uniformité du champ électrique,...). En définissant la forme à partir d'une équipotentielle d'un système de charges situées à l'intérieur de l'électrode, nous déduisons à la fois la *forme* de l'électrode et la *solution* exacte du problème d'électrostatique associé. Dès lors, la recherche de l'électrode optimale respectant un gabarit donné revient à résoudre un problème d'optimisation sous contraintes pour lequel les *paramètres* représentent les positions et les valeurs des charges. Les *contraintes* traduisent le respect du gabarit, l'*optimalité* étant liée pour sa part aux valeurs du champ électrique sur l'électrode.

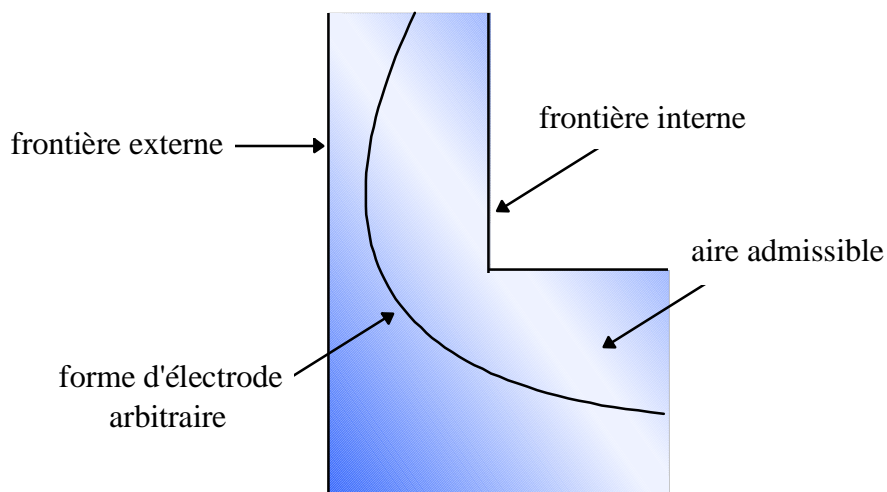


Figure 54. Gabarit géométrique d'une électrode

L'organigramme fonctionnel de la technique d'optimisation de forme développée est présenté à la Figure 55. Nous détaillerons chacune des étapes et analyserons les spécificités du problème par la suite.

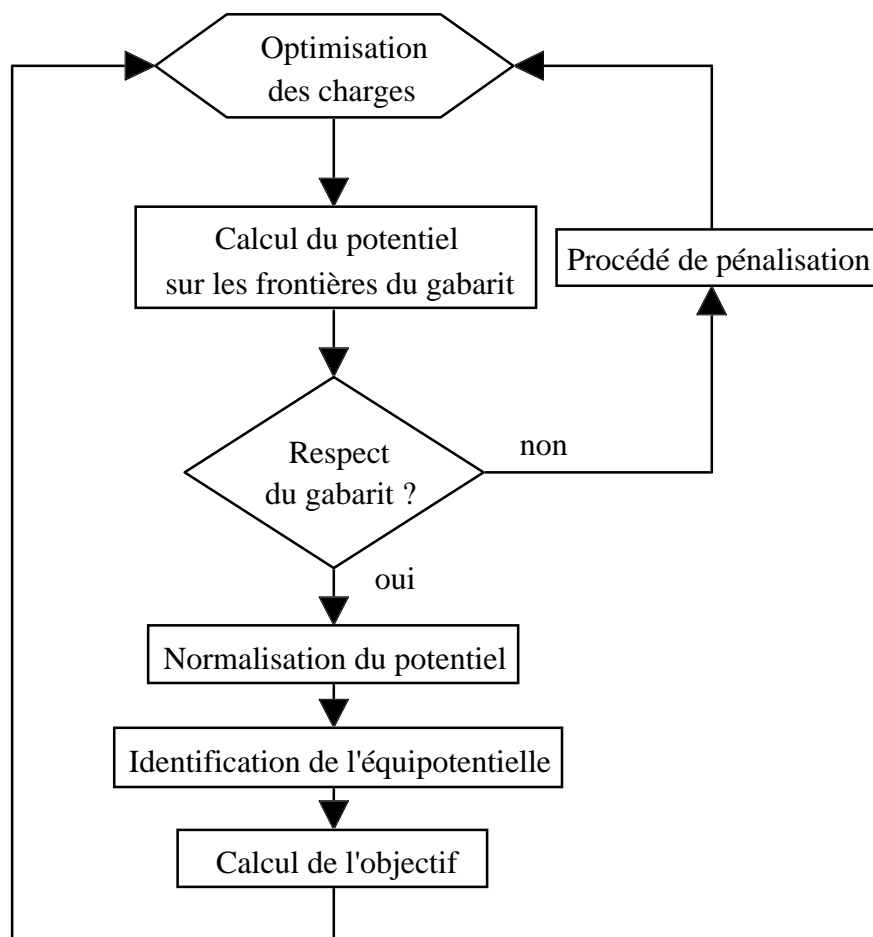


Figure 55. Organigramme de la méthode d'optimisation de forme

V.4.3. Optimisation des charges

La position et la valeur des charges sont donc déterminées par un algorithme d'optimisation de façon à respecter le gabarit géométrique et à remplir au mieux les critères d'optimalité. Outre le gros avantage apporté par les méthodes de nichage qui permettent d'explorer plusieurs profils optimaux distincts, nous recommandons l'utilisation d'algorithmes génétiques multimodaux en raison de la nature particulière du problème. Nous aborderons ce point plus précisément dans les paragraphes suivants.

V.4.4. Calcul des grandeurs électriques

Nous nous limiterons à l'utilisation d'anneaux chargés pour décrire des problèmes axisymétriques et de lignes de charges infinies dans le cas de domaines 2D plan. Le potentiel et le champ électriques peuvent alors être calculés à partir des distributions des charges fictives, par superposition (cf. paragraphe V.2).

V.4.5. Respect du gabarit géométrique

La Figure 56 présente l'ensemble des configurations possibles pour une équipotentielle de valeur V_{obj} par rapport au gabarit géométrique. Dans tous les cas, nous supposons que le potentiel diminue de la frontière interne vers la frontière externe. Les configurations (a), (b) et (c) correspondent à des profils interdits qui violent les contraintes géométriques tandis que les contours (d), (e), (f) et (g) représentent des formes d'électrodes respectant le gabarit.

Nous remarquons que chaque cas est défini en fonction des valeurs minimale et maximale du potentiel sur les frontières. Ceci nous permet de déduire une condition générale pour qu'une équipotentielle de valeur V_{obj} satisfasse les contraintes géométriques imposées par le gabarit :

$$V_{\max}^{\text{ext}} \leq V_{obj} \leq V_{\min}^{\text{int}} \quad (140)$$

où V_{\max}^{ext} représente la valeur maximale du potentiel sur la frontière externe et V_{\min}^{int} la valeur minimale du potentiel sur la frontière interne. Une relation duale de (140) est obtenue en échangeant les indices *ext* et *int* lorsque le potentiel augmente de la frontière interne vers la frontière externe.

Plus généralement la condition nécessaire et suffisante pour qu'il existe au moins une équipotentielle de valeur quelconque \tilde{V}_{obj} vérifiant les contraintes géométriques, s'écrit :

$$V_{\max}^{\text{ext}} \leq V_{\min}^{\text{int}} \quad (141)$$

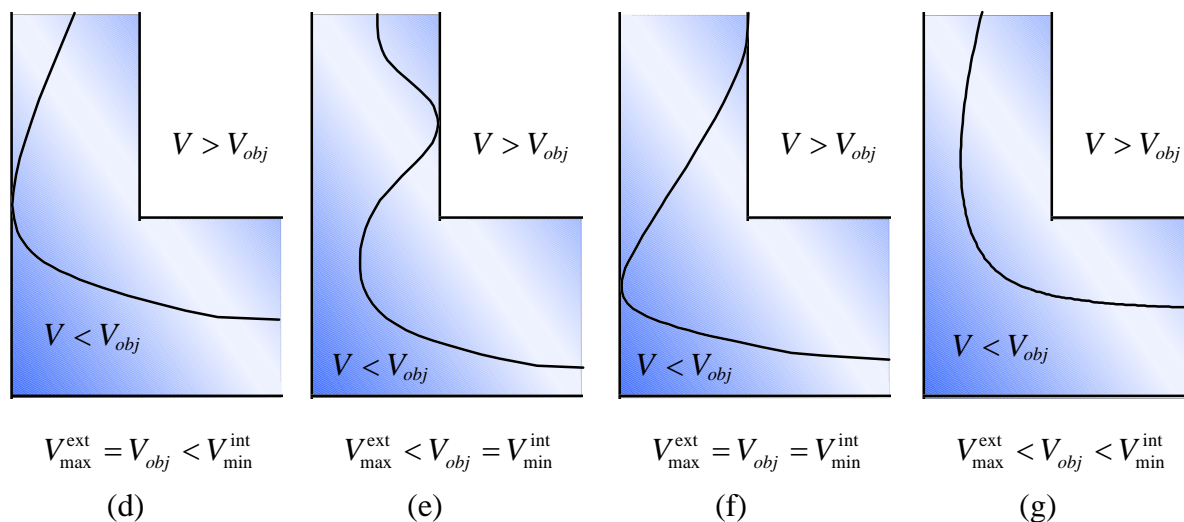
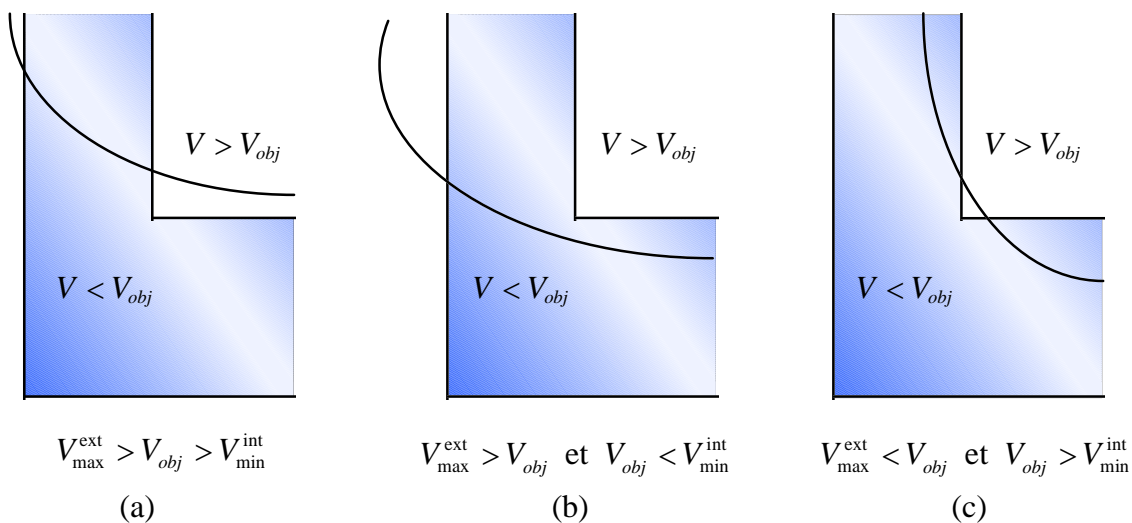


Figure 56. Exemples d'équipotentiels représentant des profils valides ou interdits

- (a) équipotentielle violant les deux frontières du gabarit
- (b) équipotentielle violant la frontière externe du gabarit
- (c) équipotentielle violant la frontière interne du gabarit
- (d) équipotentielle valide tangente à la frontière interne du gabarit
- (e) équipotentielle valide tangente à la frontière externe du gabarit
- (f) équipotentielle valide tangente aux deux frontières du gabarit
- (g) équipotentielle correspondant à un profil valide quelconque

V.4.6. Mise à l'échelle du potentiel

Trouver une équipotentielle de valeur fixée V_{obj} qui vérifie les contraintes géométriques n'est pas forcément évident, pour deux raisons. D'abord, parce qu'il est nécessaire d'avoir une

idée *a priori* sur les valeurs extrêmes des charges (q_{\min} et q_{\max}). On évite ainsi que l'espace irréalisable soit beaucoup plus grand que l'espace réalisable, ce qui pénaliserait fortement la recherche et augmenterait la difficulté du problème. Ensuite parce qu'il faut pour l'identification d'une équipotentielle donnée vérifier deux contraintes d'inégalité (Equation (140)) au lieu d'une seule (Equation (141)).

Il est donc plus judicieux de chercher une équipotentielle quelconque de valeur \tilde{V}_{obj} vérifiant l'équation (141) et de se ramener à une équipotentielle de référence $V = V_{obj}$ par une simple mise à l'échelle des charges et des grandeurs électriques en multipliant leur valeur par le terme $V_{obj} / \tilde{V}_{obj}$. Nous verrons par la suite que cette mise à l'échelle du potentiel, aussi appelée normalisation, facilite grandement la recherche des configurations réalisables.

Lorsque l'équation (141) est vérifiée, nous pouvons rechercher n'importe quelle équipotentielle valide comprise dans l'intervalle $[V_{\max}^{\text{ext}}, V_{\min}^{\text{int}}]$. L'ensemble des valeurs possibles pour \tilde{V}_{obj} est défini par l'équation (142)

$$\tilde{V}_{obj} = V_{\max}^{\text{ext}} + a(V_{\min}^{\text{int}} - V_{\max}^{\text{ext}}) \quad (142)$$

où $a \in [0,1]$ est appelé facteur de normalisation. Trois techniques de normalisation ont été développées pour fixer la valeur du coefficient a .

V.4.6.1. Normalisation intégrée

Le coefficient de normalisation est pris comme paramètre supplémentaire et est directement intégré à l'algorithme d'optimisation qui recherche la valeur et la position des charges.

V.4.6.2. Normalisation optimisée

Le coefficient de normalisation n'est pas intégré à l'algorithme de recherche global mais optimisé indépendamment à l'aide d'une méthode unidimensionnelle. Nous utiliserons la méthode du nombre d'or ou section dorée [Schwefel 95] pour déterminer une valeur optimale de a pour chaque configuration de charges respectant les contraintes.

Par rapport à la méthode de normalisation précédente, cette technique permet d'explorer l'ensemble des équipotentielles \tilde{V}_{obj} qui vérifient l'équation (141). La recherche des profils optimaux est donc plus précise mais plus coûteuse en temps de calcul puisqu'il est nécessaire d'explorer une équipotentielle par division d'intervalles du coefficient de normalisation.

V.4.6.3. Normalisation fixe

Ce mode de normalisation consiste à fixer la valeur du coefficient a lorsque les charges satisfont l'équation (141). Nous pouvons par exemple imposer que le profil recherché soit tangent à la frontière externe (normalisation externe) ou à la frontière interne (normalisation interne) en fixant la valeur de a à 0 ou à 1. Cette méthode, facile à mettre en œuvre, introduit une contrainte implicite supplémentaire et limite le nombre de profils réalisables par rapport aux techniques de normalisation précédentes.

V.4.7. Identification et exploration d'une équipotentielle valide

Lorsque la contrainte liée au gabarit géométrique est vérifiée, il faut évaluer la qualité du profil réalisable. Cela nécessite la détermination du champ électrique le long de l'électrode, donc l'identification de l'équipotentielle correspondante. Nous pourrions caractériser cette équipotentielle à partir d'un maillage comme dans une méthode d'éléments finis. Nous proposons ici une technique qui permet de profiter au maximum de la précision de la méthode de charges fictives utilisée.

V.4.7.1. Identification d'un point de l'équipotentielle

Un point d'une équipotentielle valide \tilde{V}_{obj} est recherché à l'aide de la méthode de *Newton-Raphson* [Schwefel 95] en parcourant une direction de l'espace. A partir d'un point initial x_k fixé, le point suivant x_{k+1} est calculé dans la direction x selon (143).

$$x_{k+1} = x_k + \frac{V(x_k) - \tilde{V}_{obj}}{E_x(x_k)} \quad (143)$$

où $E_x(x_k)$ est la composante du champ électrique dans la direction x au point x_k et $V(x_k)$ la valeur du potentiel en ce même point. La procédure est répétée jusqu'à ce que :

$$\left| V(x_{k+1}) - \tilde{V}_{obj} \right| \leq \varepsilon_v \quad (144)$$

avec ε_v précision requise sur le potentiel.

V.4.7.2. Exploration de l'ensemble des points de l'équipotentielle

Les autres points de l'équipotentielle sont détectés par exploration de voisinage à l'aide d'une grille à huit nœuds de centre C_k et de côté 2ξ (voir Figure 57).

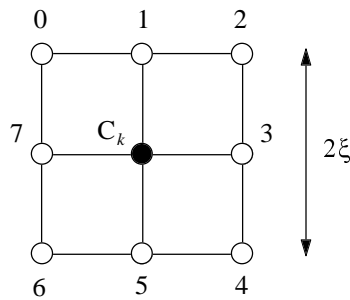
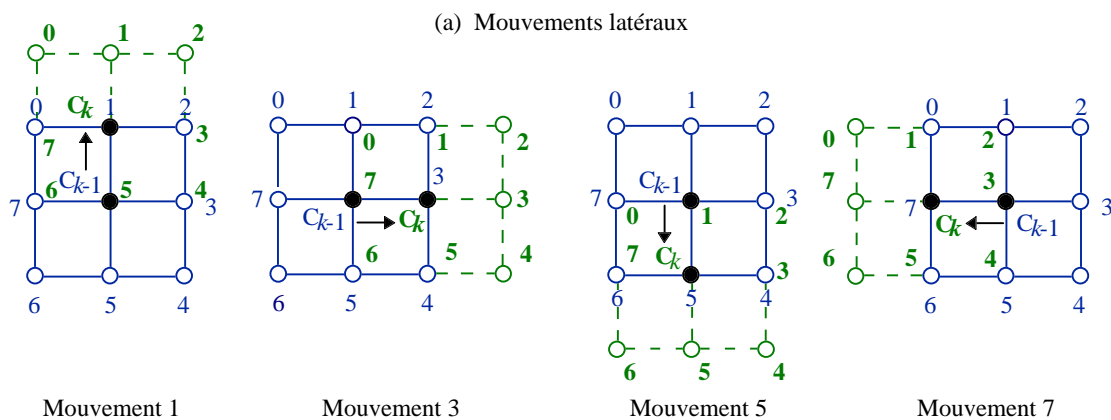


Figure 57. Grille élémentaire pour la détection des points de l'équipotentielle

Le centre C_0 de la grille initiale est placé en un point de l'équipotentielle (en l'occurrence au point identifié par la méthode de *Newton-Raphson*). La grille est ensuite déplacée séquentiellement, en tenant compte des mouvements précédents, et de façon à ce que le centre de la nouvelle grille soit situé au plus près de l'équipotentielle \tilde{V}_{obj} . Deux types de mouvements sont rendus possibles : les mouvements diagonaux et les mouvements latéraux. Dans chaque cas le nombre de points testés pour placer le centre de la grille suivante est limité pour ne pas réexplorer des points déjà identifiés. Le centre de la grille peut être déplacé en trois points après un mouvement latéral et en cinq après un mouvement diagonal. La Figure 58 illustre l'ensemble des mouvements possibles de la grille et les positions testées pour les centres suivants. Nous pouvons envisager des procédures plus complexes permettant des rotations ou des inclinaisons de la grille mais la méthode d'exploration proposée s'avère suffisamment fiable et précise quand la taille de la grille est faible par rapport aux dimensions du gabarit.



Nouveau centre pour un mouvement latéral i quelconque: $C_k = \text{noeud } k : \min_{k=i-1, i, i+1} |V_k - \tilde{V}_{obj}|$ $i = 1, 3, 5, 7$

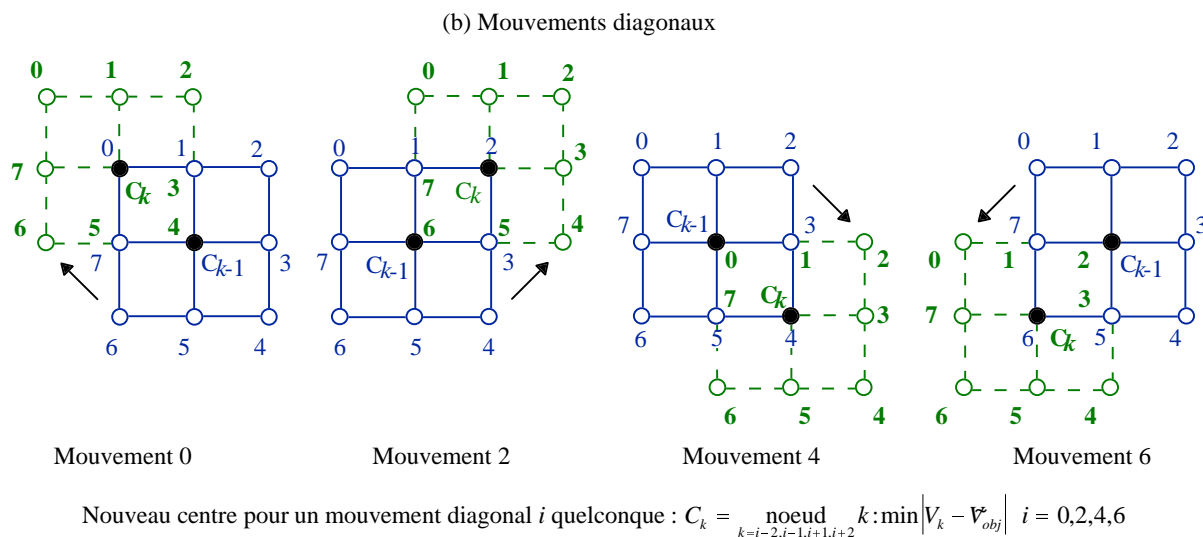


Figure 58. Mouvements de la grille. (a) latéraux (b) diagonaux

V.4.8. Critères d'optimisation et calcul de l'objectif

Lorsqu'une équipotentielle vérifie les contraintes géométriques, l'objectif correspondant à la forme d'électrode équivalente peut être calculé. Celui-ci dépend du critère d'optimalité recherché. L'optimisation de forme des éléments diélectriques des dispositifs haute tension est généralement réalisée avec les critères suivants [Vasconcelos 94] :

- l'uniformisation du champ tangentiel sur le contour
- l'uniformisation du champ total sur le contour
- la minimisation du champ total sur le contour

V.4.9. Traitement des contraintes et procédé de pénalisation

Lorsque au contraire les contraintes géométriques sont *violées*, le critère d'optimisation associé à la forme de l'électrode *n'est pas calculable*. En effet,

- une ligne équipotentielle ne respectant pas les contraintes géométriques peut avoir une forme quelconque et se situer très loin du gabarit rendant son identification difficile.
- des charges peuvent se trouver des "deux cotés" de l'équipotentielle. Il n'est alors plus possible d'assimiler l'équipotentielle à une électrode.

- Le choix du facteur normalisation et de la valeur de l'équipotentielle \tilde{V}_{obj} à identifier reste problématique.

Pour ces raisons, il n'existe aucun lien possible entre les valeurs de la fonction objectif dans le domaine réalisable et le domaine irréalisable. Il faut toujours s'assurer que la performance des configurations irréalisables reste inférieure à celles des configurations réalisables.

V.4.10. Spécificité du problème d'optimisation

Le problème d'optimisation, résultant de la technique décrite précédemment, présente divers modes, et possède ce caractère multimodal à plusieurs niveaux :

V.4.10.1. Modes géométriques recherchés

Plusieurs profils peuvent être optimaux au sens physique du terme, c'est à dire qu'ils présentent un même caractère optimal pour des contours géométriques différents. Ce sont bien évidemment ces solutions qui nous intéressent et que nous désirons détecter. Néanmoins, la définition particulière du problème introduit des modes supplémentaires à d'autres niveaux.

V.4.10.2. Modes dus aux possibles permutations de charges

Pour des charges variant dans un même intervalle tant en ce qui concerne leurs positions que leurs valeurs, nous pouvons trouver des solutions distinctes au sens mathématique mais qui conduisent au même profil. Par exemple, la solution constituée par deux charges $Q_1(x, y, q)$ et $Q_2(x', y', q')$ est équivalente à celle définie par les deux charges $Q_1(x', y', q')$ et $Q_2(x, y, q)$, les permutations de charges n'influençant pas la distribution des grandeurs électriques.

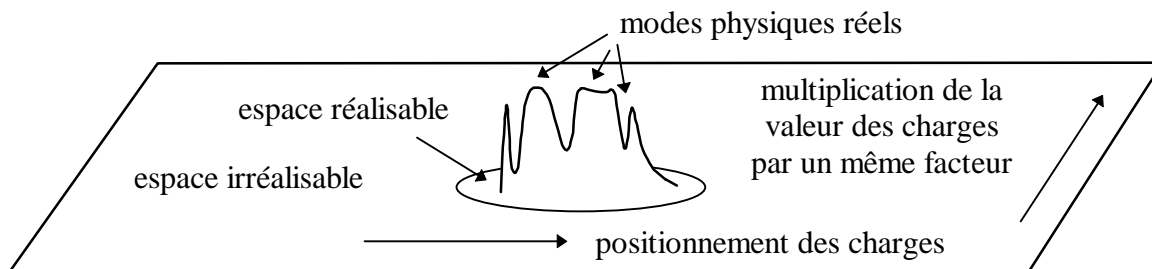
V.4.10.3. Modes dus à la mise à l'échelle du potentiel

De même, le procédé de normalisation du potentiel introduit au paragraphe V.4.6 implique qu'une solution représentée par des charges $Q_i(x_i, y_i, q_i)$ est équivalente à toute solution du type $Q_i(x_i, y_i, \alpha q_i)$ où α est une constante quelconque, identique pour toutes les charges i .

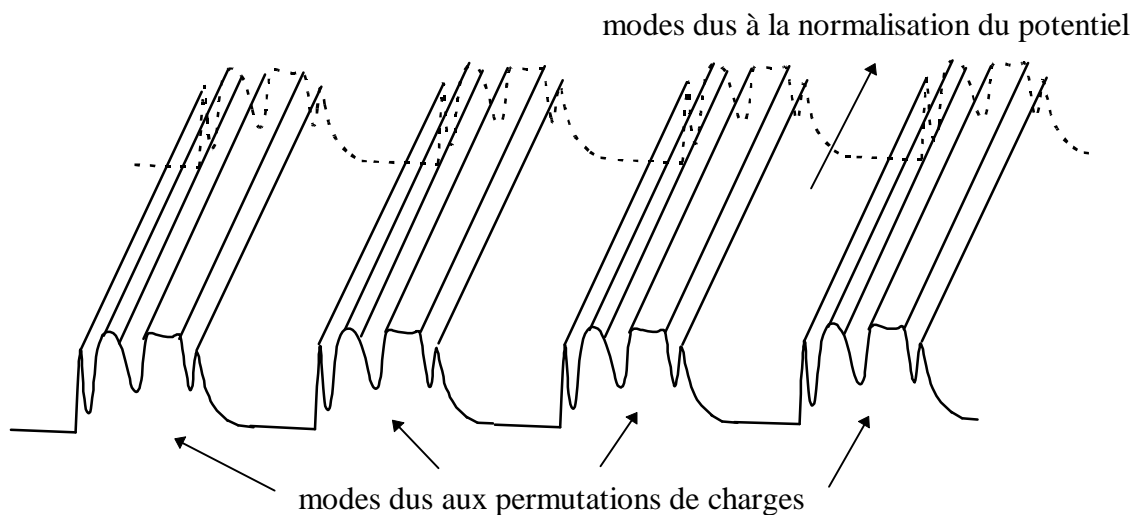
La Figure 59 schématise les modes résultant de la définition particulière du problème pour un critère formulé en terme de maximisation.

La méthode proposée conduit à un ensemble de solutions physiques équivalentes. Néanmoins, le paramétrage adopté simplifie grandement la définition du problème en évitant

l'adjonction de contraintes supplémentaires liées à la position et à la valeur des charges : ces contraintes pourraient limiter la diversité des formes et rendre la recherche des profils réalisables plus difficiles pour l'algorithme d'optimisation. Il sera intéressant de déterminer si les modes "parasites" introduits par le paramétrage de position et la normalisation du potentiel facilitent ou au contraire entravent la détection des profils optimaux réels. Nous tenterons de répondre à cette question dans les paragraphes suivants.



(a) problème contraint : pas de normalisation du potentiel, pas de permutations de charges
 ⇒ espace réalisable isolé



(b) paramétrage proposé : normalisation du potentiel et permutations de charges possibles
 ⇒ ensemble de modes équivalents (modes réels+modes "parasites")

Figure 59. Modes dus à la définition particulière du problème

tel-00139202, version 1 - 29 Mar 2007

V.5. EXEMPLES D'APPLICATIONS

V.5.1. Electrode simple

V.5.1.1. Définition du problème

Considérons le gabarit géométrique 2D d'une électrode infiniment longue dans la direction (Oz) représenté par la Figure 60a. A l'aide de la méthode décrite précédemment, nous recherchons une forme d'électrode qui respecte le gabarit géométrique et présente un champ électrique minimal sur son contour.

Le potentiel électrique est créé par quatre charges placées sur l'axe-y à des ordonnées variant entre 4.5 et 16. Pour tenir compte de la condition de symétrie par rapport à l'axe-x, quatre charges symétriques par rapport à cet axe et de même valeur respective sont ajoutées. Le potentiel en un point de l'espace de coordonnées x et y s'exprime à partir de (145) en considérant la contribution des 8 charges [Annexes IV]

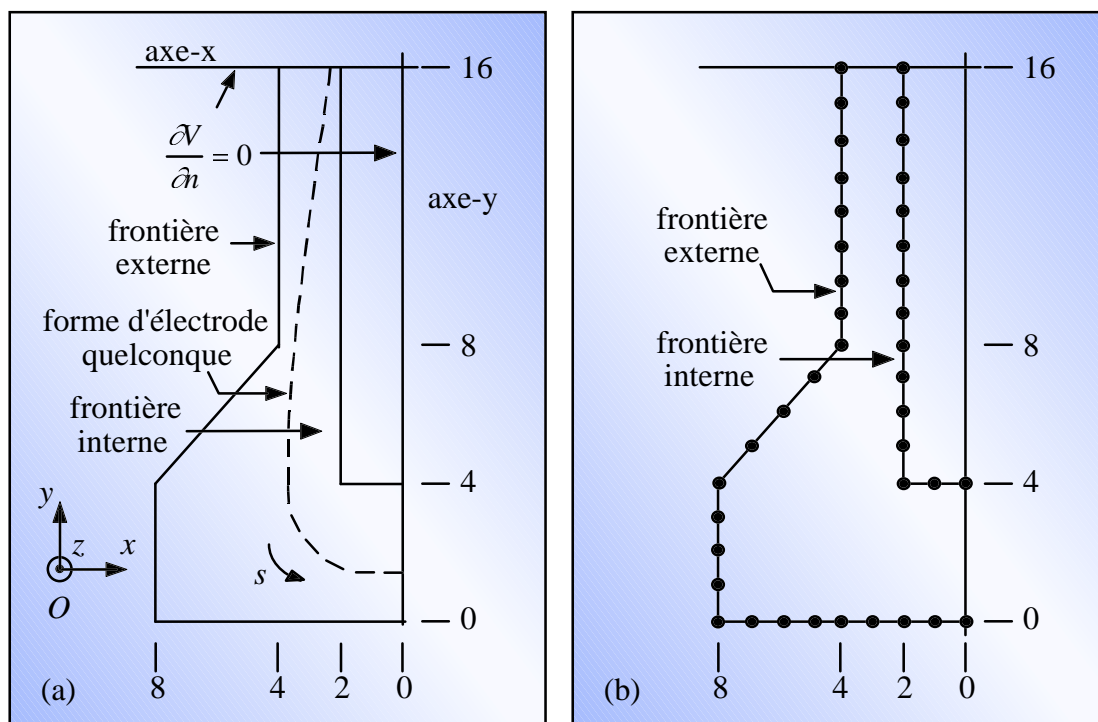


Figure 60. Gabarit géométrique de l'électrode

- (a) dimensions et conditions de symétrie
- (b) maillage des frontières

$$V(x, y) = \sum_{i=1}^8 q_i \ln r_i \quad (145)$$

où q_i désigne la valeur de la charge i et $r_i = \sqrt{(x_i - x)^2 + (y_i - y)^2}$ la distance entre le point de calcul et la charge correspondante. Notons que l'expression du potentiel a été simplifiée (abandon du facteur $-1/2\pi\epsilon_0$) et que nous considérons les longueurs sans dimensions car en définitive, seule la *forme* de l'électrode nous intéresse. La valeur des charges est comprise dans l'intervalle $[-1,1]$ et l'équipotentielle de référence recherchée est $V_{obj} = 1$.

Le potentiel logarithmique augmentant de la frontière intérieure à la frontière extérieure, la condition de respect du gabarit géométrique s'écrit sous une forme duale de (141) en échangeant les indices *ext* et *int*,

$$V_{\max}^{\text{int}} \leq V_{\min}^{\text{ext}} \quad (146)$$

donnant comme valeur possible pour l'équipotentielle recherchée :

$$\tilde{V}_{obj} = V_{\max}^{\text{int}} + a(V_{\min}^{\text{ext}} - V_{\max}^{\text{int}}) \quad a \in [0,1] \quad (147)$$

Pour savoir si une configuration de charges respecte la contrainte géométrique (146), les frontières internes et externes sont discrétisées en 15 et 25 points respectivement (voir Figure 60b). Le potentiel est calculé en chacun de ces points de contrôle et les valeurs extrêmes sur chaque frontière en sont déduites. Lorsqu'une solution correspond à un profil réalisable, l'équipotentielle \tilde{V}_{obj} est identifiée à partir du coefficient de normalisation a avec une grille de coté $2\xi=0.1$. Le champ électrique est calculé en chaque point de l'équipotentielle et sa valeur maximale \tilde{E}_{\max} permet de déduire par normalisation la valeur correspondante du champ électrique maximal sur l'équipotentielle de référence :

$$E_{\max} = \frac{V_{obj}}{\tilde{V}_{obj}} \tilde{E}_{\max} = \frac{1}{\tilde{V}_{obj}} \tilde{E}_{\max} \quad (148)$$

La fonction objectif adoptée pour caractériser la performance des configurations réalisables est la suivante :

$$F_{obj} = \frac{1}{1 + E_{\max}} \quad (149)$$

V.5.1.2. Résolution du problème par des algorithmes génétiques multimodaux.

Quatre algorithmes génétiques multimodaux ont été testés sur ce problème. La taille de la population est fixée dans chaque cas à 100 et le nombre de générations à 200. Les paramètres sont codés dans une chaîne binaire de 30 bits avec décodage dynamique. Les probabilités de

croisement et de mutation sont respectivement $p_c = 1$ et $p_m = 0.001$. Le critère de distance utilisé est de type d^∞ normalisé (voir chapitre III). Quatre exécutions sont effectuées avec une population initiale différente générée aléatoirement.

Les résultats obtenus avec les divers algorithmes génétiques dans le cas d'une normalisation intégrée et d'une pénalisation radicale sont résumés dans le Tableau 26. La moyenne (*moy*) et l'écart-type (σ) relatifs aux quatre exécutions effectuées sont précisés pour chaque grandeur présentée.

Tableau 26. Efficacité des divers algorithmes génétiques pour le problème proposé

Méthode d'optimisation	Nombre de profils optimaux distincts ($E_{\max} < 0.06$)			Meilleur profil détecté (E_{\max})		
	max	moy	σ	min	moy	σ
GA élitiste	1	1	0	0.0557	0.0564	0.0007
SH ($\sigma_s = 0.05$)	0	0	0	0.0663	0.0672	0.0011
SH ($\sigma_s = 0.2$)	0	0	0	0.0641	0.0668	0.0019
SH ($\sigma_s = 0.4$)	0	0	0	0.0722	0.0746	0.0045
CL ($\sigma_s = 0.05$)	77	60.50	22.13	0.0557	0.0561	0.0003
CL ($\sigma_s = 0.2$)	20	16.25	4.11	0.0558	0.0568	0.0010
CL ($\sigma_s = 0.4$)	12	7.50	3.11	0.0553	0.0559	0.0004
RTS ($w=30$)	29	18.50	11.09	0.0563	0.0567	0.0006
DC	3	2.25	0.96	0.0560	0.0569	0.0009

Ici encore, l'algorithme génétique standard (GA) converge vers une unique solution de l'espace mais les quatre exécutions donnent quatre solutions distinctes, de fonction objectif quasi-identique. Il semble en effet que l'espace réalisable correspondant aux optima physiques réels représente un plateau relativement large de très faible gradient. Ceci explique en partie l'échec de la méthode de partage (SH) qui, comme nous l'avons signalé, requiert généralement un réajustement en puissance de l'objectif pour localiser précisément les sommets.

L'éclaircissement (CL) détecte un grand nombre de solutions assez proches lorsqu'il est appliqué avec un faible rayon de niche ($\sigma_s = 0.05$) et quelques unes "plus distinctes" avec un rayon de niche plus grand ($\sigma_s = 0.2$ et $\sigma_s = 0.4$). Ceci démontre une fois de plus la nécessité de se doter d'une technique adaptative pour fixer le rayon de niche et s'affranchir du compromis lié au choix de ce facteur pour la recherche de solutions optimales distinctes.

La sélection par tournoi restreint (RTS) réalise assez bien ce compromis en détectant un assez grand nombre de solutions distinctes. Le surpeuplement déterministe (DC) n'a pu

localiser qu'un nombre très limité de profils avec cette normalisation et cette méthode de pénalisation.

V.5.1.3. Influence de la méthode de normalisation du potentiel

Le Tableau 27 résume les statistiques obtenues avec trois méthodes génétiques multimodales associées à une pénalisation radicale des solutions irréalisables pour les différentes techniques de mise à l'échelle du potentiel développées.

Tableau 27. Comparaison des techniques de normalisation du potentiel

Méthode d'optimisation Multimodale	Normalisation	Facteur de normalisation a		Nombre de profils optimaux distincts ($E_{\max} < 0.06$)		Moyenne des solutions optimales (E_{\max})	
		moy	σ	moy	σ	moy	σ
CL ($\sigma_s = 0.2$)	intégrée	0.745	0.162	16.25	4.11	0.0583	0.0008
	optimisée ³	0.994	0.002	12.50	4.51	0.0578	0.0012
	extérieure	1 (fixé)	0	17.00	2.45	0.0577	0.0007
RTS ($w=30$)	intégrée	0.663	0.130	18.50	11.09	0.0583	0.0003
	optimisée	0.994	0.001	9.00	3.56	0.0580	0.0006
	extérieure	1 (fixé)	0	17.00	2.45	0.0579	0.0002
DC	intégrée	0.712	0.204	2.25	0.96	0.0576	0.0008
	optimisée	0.995	0.000	2.25	1.50	0.0578	0.0013
	extérieure	1 (fixé)	0	2.50	1.29	0.0579	0.0005

Les résultats sont assez similaires quelle que soit la méthode de normalisation employée. Ils montrent que les meilleurs profils sont localisés sur la frontière externe du gabarit. La normalisation optimisée en particulier permet de trouver la valeur du coefficient de normalisation optimal. Bien qu'elle soit assez coûteuse en temps CPU en raison de l'exploration séquentielle des équipotentielles sur l'intervalle complet du coefficient de normalisation, elle est très intéressante car pour une configuration d'électrode quelconque, la valeur optimale du facteur a n'est pas connue *a priori*. La normalisation intégrée permet de détecter des solutions optimales de coefficient de normalisation quelconque mais la qualité de ces dernières est généralement inférieure à celles trouvées avec la méthode de normalisation optimisée. Nous mentionnons enfin qu'en l'absence du procédé normalisation, c'est à dire lorsque l'on utilise deux contraintes fixes duales de l'équation (140) et une pénalisation radicale des configurations irréalisables, aucune solution réalisable n'a pu être trouvée en

³ Pour la normalisation optimisée, le facteur de normalisation a est déterminé après neuf itérations de la méthode du nombre d'or.

20 000 évaluations du critère quel que soit l'algorithme d'optimisation [Sareni 98c]. Par une méthode de recherche purement aléatoire, 6 819 500 évaluations du critère ont été nécessaires pour détecter 10 solutions réalisables sur les intervalles de variation de position et de valeurs de charges précédemment définis. Ceci montre donc, l'intérêt et la nécessité du procédé de mise à l'échelle du potentiel pour la recherche des profils réalisables. La normalisation augmente en effet fortement le volume de l'espace réalisable dans l'espace de recherche complet, facilitant la détection des solutions optimales.

V.5.1.4. Influence de la méthode de pénalisation

Nous explorons enfin l'influence de la méthode de pénalisation sur le comportement des divers algorithmes génétiques multimodaux. Deux techniques de pénalisation ont été utilisées pour évaluer la performance des solutions irréalisables :

- la méthode de pénalisation radicale
- la méthode de pénalité extérieure

La fonction objectif caractérisant la performance des individus non réalisables est dans le cas d'une pénalisation radicale $F_{obj} = 0$, et pour la méthode de pénalité extérieure :

$$F_{obj} = \frac{1}{1 + \lambda_1 + \lambda_2 (V_{\max}^{\text{int}} - V_{\min}^{\text{ext}})} \quad (150)$$

où λ_1 et λ_2 sont des coefficients de pénalité. Il faut s'assurer que λ_1 est suffisamment grand pour éviter une éventuelle convergence de l'algorithme sur les frontières extérieures du domaine réalisable. D'un autre côté, prendre une valeur de λ_1 trop grande risquerait d'occasionner une forte discontinuité entre le domaine réalisable et le domaine irréalisable qui pourrait rendre plus délicate la recherche des solutions. Pour assurer une pénalisation suffisante et minimiser cette discontinuité, la valeur de λ_1 doit être la plus proche de celle de la plus "mauvaise" configuration de l'espace réalisable. Dans notre cas, la valeur λ_1 a été fixée à partir des tests effectués avec une méthode de recherche aléatoire dans l'espace réalisable [Sareni 98c]. Nous avons finalement pris $\lambda_1 = 0.1$ et $\lambda_2 = 100$.

Le Tableau 28 résume les résultats obtenus avec les différents algorithmes génétiques associés à une normalisation externe pour les deux techniques de pénalisation testées. Le taux d'exploration est défini comme le rapport du nombre total d'évaluations de configurations réalisables sur le nombre total d'évaluations dans l'espace de recherche complet.

Quel que soit l'algorithme génétique employé, une très faible proportion de la population se trouve dans l'espace réalisable lors des premières générations. Ceci est dû au fait que la taille de l'espace réalisable est très petite par rapport à celle de l'espace de recherche global, même avec la mise à l'échelle du potentiel. Le taux d'exploration d'une méthode purement aléatoire est d'ailleurs relativement faible, seulement de 0.8 % [Sareni 98c].

Tableau 28. Comparaisons des méthodes de pénalité

Algorithme d'optimisation	Pénalisation	Taux d'exploration (%)		Nombre de profils optimaux distincts ($E_{\max} < 0.06$)		Meilleur profil détecté (E_{\max})	
		moy	σ	moy	σ	moy	σ
GA élitiste	radicale	93.20	1.59	1	0	0.0558	0.0006
	extérieure	92.01	2.54	1	0	0.0564	0.0011
DC	radicale	2.97	0.56	2.50	1.29	0.0571	0.0006
	extérieure	19.43	2.99	17.50	3.11	0.0563	0.0004
RTS ($w=30$)	radicale	8.20	1.44	11.50	3.00	0.0561	0.0002
	extérieure	19.26	2.59	33.75	4.5	0.0561	0.0005
CL ($\sigma_s = 0.2$)	radicale	64.96	7.92	17	2.45	0.0564	0.0006
	extérieure	64.46	4.25	14	1.15	0.0563	0.0004

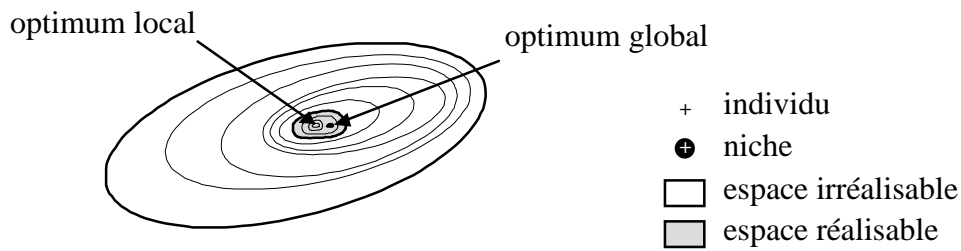
La Figure 61 résume le comportement des divers algorithmes face au problème contraint traité.

Un *algorithme génétique élitiste* converge très rapidement vers l'espace réalisable et concentre tous les individus en un même point entraînant un fort taux d'exploration. Néanmoins, c'est évidemment toujours le même individu qui est évalué par croisement et la recherche de nouvelles solutions réalisables repose alors nécessairement sur la mutation. Le risque de convergence prématuré est accru.

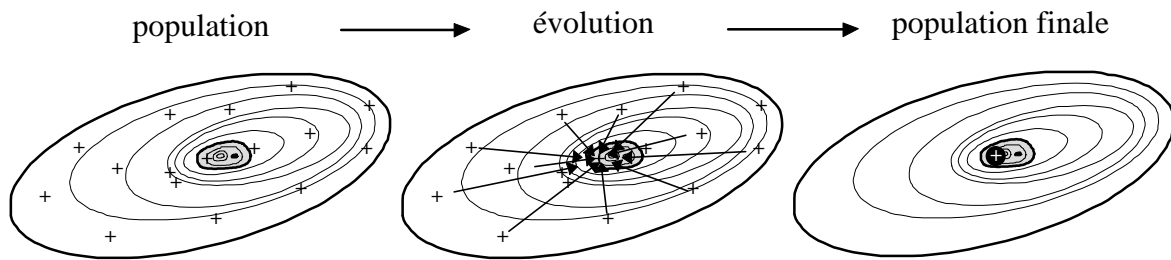
La *méthode d'éclaircissement* a un comportement assez similaire mais le procédé de réinitialisation et la préservation des individus dominants permet de maintenir des niches à la fois dans l'espace réalisable et dans l'espace irréalisable. Cette diversité combinée aux nombre important d'individus présents et reproduits dans l'espace réalisable offre une probabilité plus grande d'identifier les solutions optimales.

Par contre, les *méthodes de surpeuplement* souffrent d'un inconvénient majeur évoqué précédemment (voir chapitre III). Leurs mécanismes de remplacement minimisent les écarts de voisinage entre individus d'une génération à l'autre et ne permettent pas de concentrer très rapidement la population dans l'espace réalisable. Le taux d'exploration est assez faible et la distribution de la population finale proche de celle de la population initiale. Cette évolution lente pénalise fortement l'exploration de l'espace réalisable car la plupart des individus restent à l'extérieur de celui-ci. Or, la recombinaison d'individus non réalisables a peu de chance d'aboutir à des individus réalisables. Bien que les méthodes de surpeuplement soient assez performantes sur ce premier problème "facile", nous verrons que ce n'est plus le cas lorsque les contraintes à respecter sont plus sévères.

(a) représentation simplifiée de l'espace de recherche

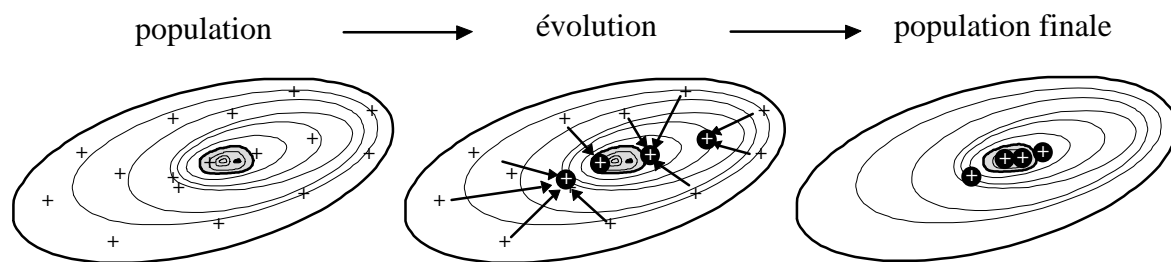


(b) AG standard élitiste



⇒ formation d'une seule niche - convergence rapide (parfois prématurée) due à la sélection

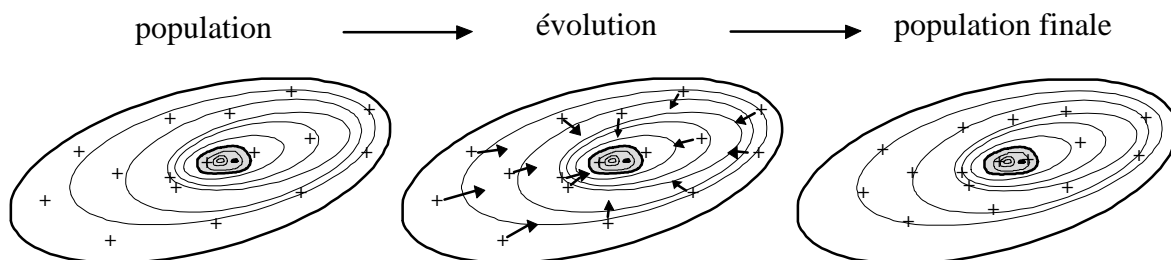
(c) Eclaircissement+élitisme



⇒ concentration d'individus dans le domaine réalisable grâce à la sélection

⇒ formation de niches dans l'espace réalisable et l'espace irréalisable grâce à la combinaison de l'éclaircissement et de l'élitisme

(d) Méthode de surpeuplement



⇒ remplacement des individus par voisinage - formation progressive mais lente des niches

⇒ population finale proche de la population initiale

Figure 61. Comportement des algorithmes génétiques multimodaux face à un problème contraint

Cette limitation est encore plus marquée si on emploie une méthode de pénalisation radicale car l'algorithme n'a aucun moyen d'améliorer les individus non réalisables. Le taux d'exploration et le nombre de solutions optimales détectées sont largement augmentés avec l'utilisation d'une méthode de pénalisation extérieure. Dans ce type de problème où l'espace réalisable est très réduit par rapport à l'espace de recherche global, la méthode de pénalisation radicale est à proscrire.

V.5.1.5. Localisation des profils optimaux dans l'espace de recherche

Nous présentons la dispersion des 4 charges de chaque solution optimale pour les divers algorithmes génétiques multimodaux (Figure 62 à Figure 65).

La position des charges, donnée par la coordonnée correspondante sur l'axe-y, est représentée en abscisse. La valeur des charges normalisée relative à l'équipotentielle de référence $V_{obj} = 1$ est illustrée en ordonnée.

Ces figures mettent en évidence les caractéristiques des diverses méthodes multimodales :

L'éclaircissement appliqué avec un rayon de niche faible ($\sigma_s = 0.05$) conduit à un ensemble de configurations optimales proches. La dispersion des charges est faible (voir Figure 62) et les profils correspondants sont similaires. L'éclaircissement est capable de localiser des modes équivalents dus aux permutations pour des charges proches (charges 1 et 4 sur la Figure 62). L'augmentation du rayon de niche permet d'identifier des configurations optimales "plus distinctes" et de détecter des modes équivalents dus aux permutations plus éloignés (Figure 63).

Les méthodes de surpeuplement à voisinage implicite s'affranchissent directement du compromis lié au choix du seuil de voisinage. Elles se révèlent assez efficaces sur ce problème en localisant un grand nombre de solutions optimales distinctes. (Figure 64 et Figure 65).

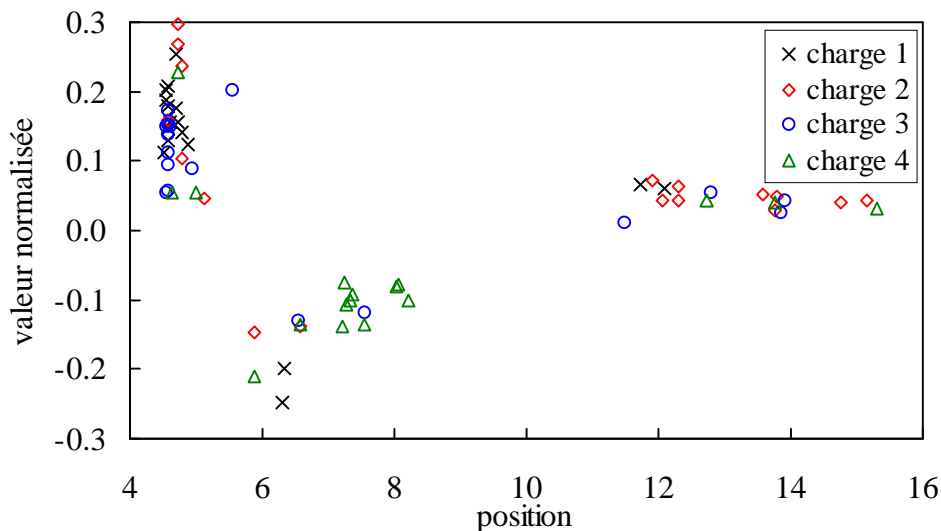


Figure 64. Répartition des charges de 17 solutions optimales distinctes déterminées par la méthode du surpeuplement déterministe.

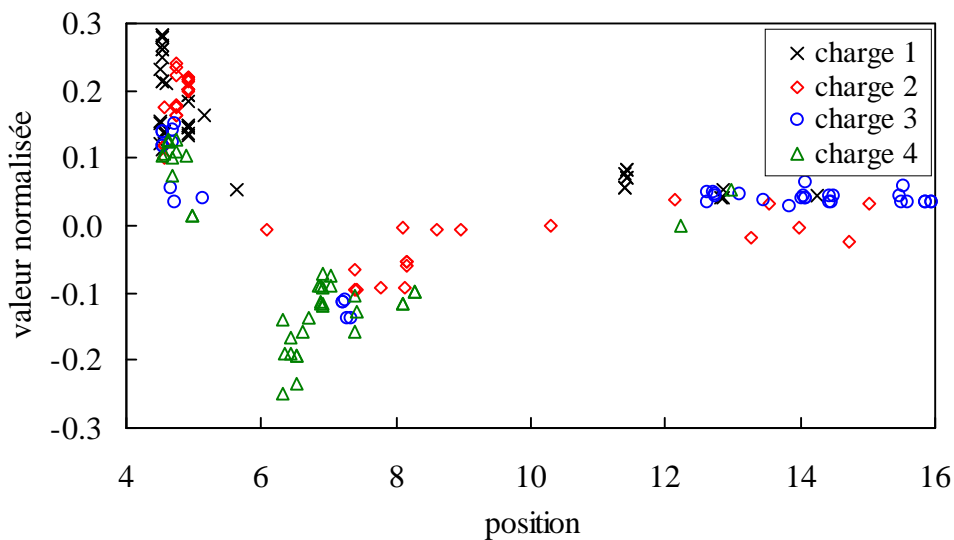


Figure 65. Répartition des charges de 40 solutions optimales distinctes déterminées par la sélection par tournoi restreint ($w=30$).

V.5.1.6. Exemple de profils optimaux

Nous présentons enfin six configurations de charges correspondant à six profils d'électrodes optimaux distincts. Le Tableau 29 donne la position et la valeur normalisée des charges conduisant à l'équipotentielle de référence $V_{obj} = 1$, ainsi que le champ électrique maximal sur le contour et la valeur de l'équipotentielle \tilde{V}_{obj} réellement détectée. Le facteur de normalisation a des solutions est égal à 1 dans chaque cas.

Tableau 29. Caractéristiques de six profils optimaux

<i>profil</i>	<i>charge 1</i>		<i>charge 2</i>		<i>charge 3</i>		<i>charge 4</i>		\tilde{V}_{obj}	E_{max}
	<i>position</i>	<i>valeur</i>	<i>position</i>	<i>valeur</i>	<i>position</i>	<i>valeur</i>	<i>position</i>	<i>valeur</i>		
1	4.52	0.279	6.90	-0.117	10.31	-0.001	12.73	0.047	3.45	0.0568
2	4.63	0.125	4.74	0.164	7.34	-0.135	11.40	0.056	4.37	0.0567
3	4.92	0.194	4.93	0.214	6.32	-0.250	12.60	0.050	3.31	0.0564
4	4.91	0.146	4.92	0.221	6.54	-0.192	15.48	0.035	4.52	0.0557
5	4.91	0.148	4.92	0.218	6.45	-0.191	15.93	0.035	4.54	0.0565
6	4.63	0.282	6.09	-0.005	6.89	-0.116	13.09	0.048	3.42	0.0565

La Figure 66 montre l'équipotentielle représentant les six configurations optimales précédentes en comparaison avec une équipotentielle quelconque. Le champ électrique dans le domaine 2D considéré est représenté en dégradé. Toutes ces formes sont identiques dans la zone où le champ est fort. Pour $y > 8$, nous pouvons choisir n'importe quelle forme pour l'électrode lors de la fabrication (par exemple un simple profil vertical). Enfin, la Figure 67 donne l'allure du champ électrique sur chacun des profils.

tel-00139202, version 1 - 29 Mar 2007

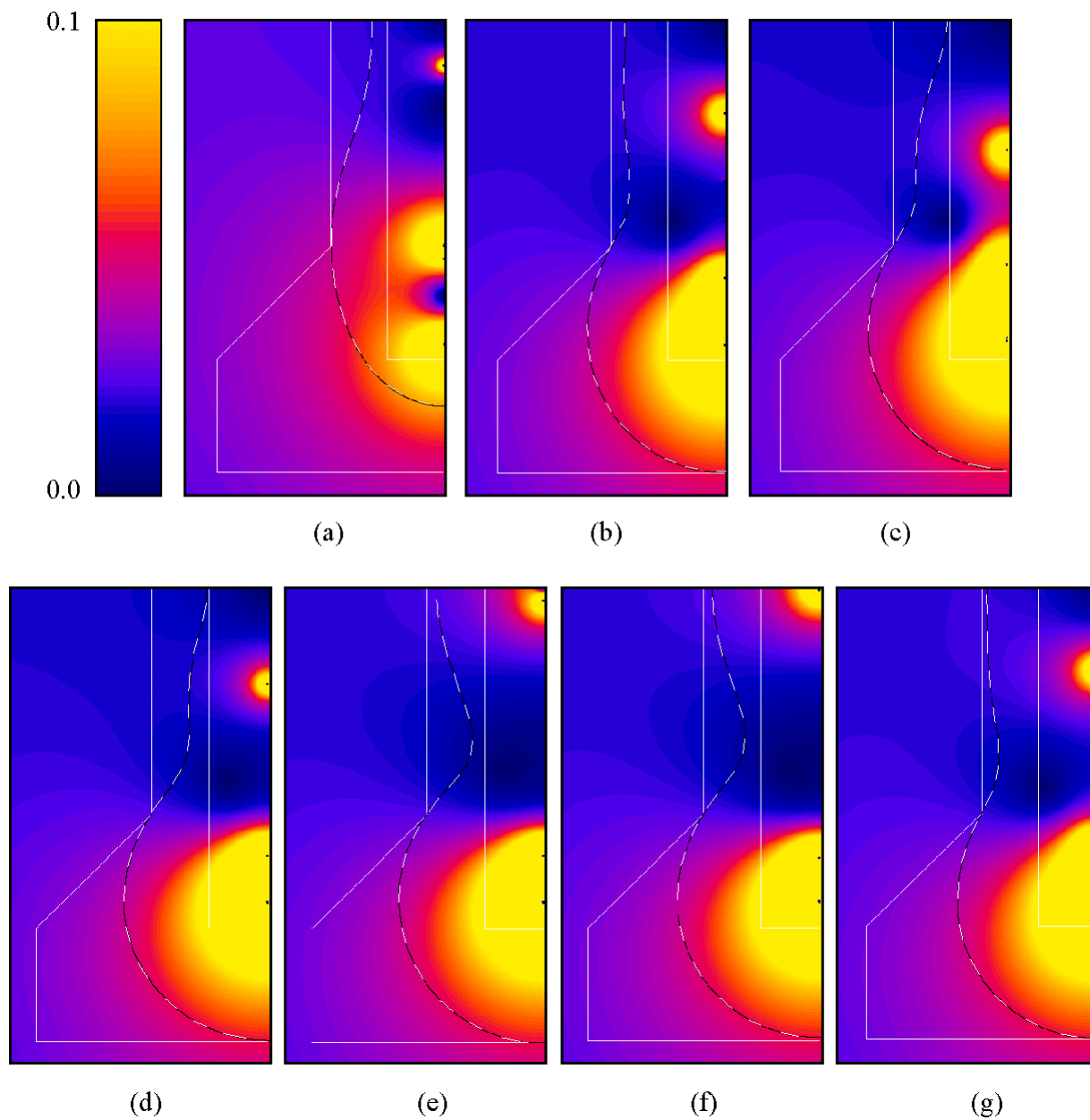


Figure 66. Exemples de profils
(a) quelconque - (b) à (f) optimaux

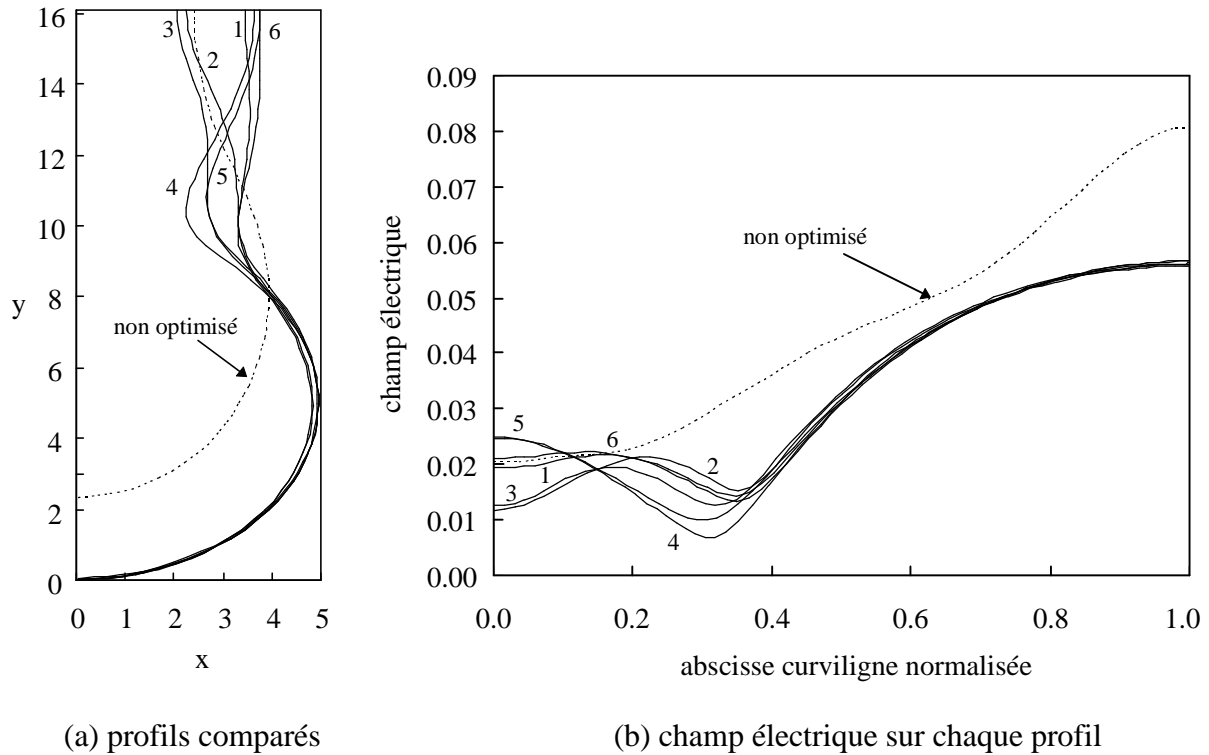


Figure 67. Champ électrique sur les profils

V.5.2. Condensateur plan à lames profilées

Il s'agit ici de traiter un problème d'optimisation de forme dont la solution exacte est connue analytiquement.

V.5.2.1. Définition du problème

On considère un plan de masse en regard avec une électrode infinie à gauche, d'épaisseur constante, qui se termine à droite par un arrondi (voir Figure 68).

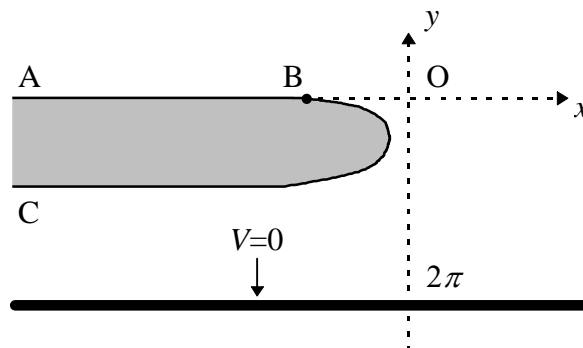


Figure 68. Profil de condensateur recherché

Le profil de l'électrode de telle sorte que le champ soit uniforme du point C au point B (défini par OB=1) peut être obtenu au moyen d'une transformation conforme [Durand 66 - Tome II, page 324]. Le contour à partir du point B est défini par l'équation suivante :

$$\begin{cases} x = 2 \ln \cos(\varphi / 2) - \cos \varphi \\ y = -\varphi + \sin \varphi \end{cases} \quad \varphi \in [0, \pi] \quad (151)$$

En complétant par symétrie, nous obtenons un condensateur plan à lames profilées présentant un champ uniforme sur l'extrémité inférieure des électrodes. La valeur du champ électrique uniforme sur le contour s'obtient en fonction de la différence de potentiel V_c aux bornes du condensateur :

$$E = \frac{V_c}{2\pi} \quad (152)$$

V.5.2.2. Résolution numérique

Nous nous proposons de retrouver ce résultat analytique à l'aide de la méthode d'optimisation de forme présentée précédemment. Le profil optimal de l'électrode est recherché à partir du gabarit de la Figure 69.

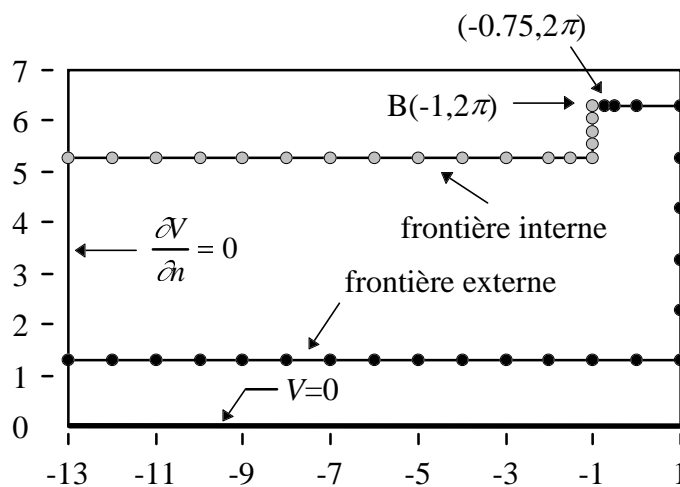


Figure 69. Gabarit géométrique de l'électrode et maillage associé

Onze charges⁴ placées à l'intérieur de la zone délimitée par la frontière interne du gabarit permettent de simuler l'équipotentielle représentant le profil de l'électrode. Comme

⁴ Onze charges supplémentaires sont considérées dans le calcul du potentiel et du champ électrique pour assurer sur la gauche la condition $\partial V / \partial n = 0$. Vingt deux autres charges sont ajoutées pour garantir la condition $V=0$ en bas.

précédemment, la valeur et la position des charges sont déterminées par un algorithme génétique. Le problème d'optimisation résultant est donc de dimension $n=33$ puisqu'il y a 3 degrés de liberté par charge (la valeur, l'abscisse et l'ordonnée).

Outre la contrainte géométrique relative au respect du gabarit exprimée par l'équation (141), nous imposons aussi que la composante du champ électrique suivant x soit nulle pour que l'électrode soit horizontale au point B. Cette contrainte de stricte égalité est formulée suivant deux contraintes d'inégalité :

$$\left| \frac{E_x(B)}{E(B)} \right| \leq 0.08 \quad (153)$$

qui garantissent un profil horizontal au point B avec une erreur maximale de ± 5 degrés.

Lorsqu'un individu ne satisfait pas la contrainte géométrique (141), sa fonction d'adaptation est calculée de façon similaire à (148) en prenant en compte la deuxième contrainte :

$$F_{obj} = \frac{1}{1 + \lambda_1 + \lambda_2 (V_{\max}^{\text{ext}} - V_{\min}^{\text{int}}) + \lambda_3 \max\left(0, \left| \frac{E_x(B)}{E(B)} \right| - 0.08\right)} \quad (154)$$

où $\lambda_2 \gg \lambda_3$. En effet, le coefficient de pénalité λ_2 associé à la contrainte géométrique doit être largement supérieur aux autres coefficients de pénalité de façon à ce que cette contrainte soit optimisée et satisfaite en premier lieu. La recherche des profils respectant le gabarit est prioritaire devant la contrainte liée à la courbure au point B afin d'évaluer le plus rapidement possible le critère d'optimalité.

Dans le cas où la contrainte géométrique est satisfaite, la normalisation du potentiel est effectuée par rapport à la valeur du potentiel au point B. L'équipotentielle à détecter est $\tilde{V}_{obj} = V(B)$. Après identification de l'équipotentielle avec une grille de coté $2\xi=0.1$, le critère d'optimalité est estimé. Nous recherchons cette fois à obtenir un champ uniforme sur l'équipotentielle à partir du point B jusqu'au point C (voir Figure 69). Ceci peut être réalisé en minimisant le facteur de dispersion des valeurs du champ électrique sur l'équipotentielle. En prenant en compte la contrainte liée à la courbure au point B, nous définissons la fonction d'adaptation des individus de la façon suivante :

$$F_{obj} = \frac{1}{1 + \frac{\sigma_E}{\mu_E} + \lambda_3 \max\left(0, \left| \frac{E_x(B)}{E(B)} \right| - 0.08\right)} \quad (155)$$

où μ_E représente la moyenne des valeurs du champ électrique aux points détectés sur l'équipotentielle et σ_E l'écart-type correspondant.

Trois algorithmes génétiques multimodaux ont été testés sur ce problème. La taille de la population est fixée à 100 et le nombre de générations à 200. Les paramètres sont codés chacun dans une chaîne binaire de 10 bits avec décodage dynamique. Les probabilités de croisement et de mutation sont respectivement $p_c = 1$ et $p_m = 0.001$. Le critère de distance utilisé est de type d^∞ normalisé (voir chapitre III). Les coefficients de pénalité ont été fixés suivant $\lambda_1 = \lambda_3 = 100$ et $\lambda_2 = 10^5$. Quatre exécutions sont effectuées avec une population initiale différente générée aléatoirement. Le Tableau 30 compare les résultats obtenus avec les algorithmes génétiques multimodaux et une méthode purement aléatoire. Les taux d'individus satisfaisant au cours des exécutions la contrainte géométrique uniquement, et les deux contraintes simultanément, sont désignés respectivement par T_1 et T_{12} .

Tableau 30. Comparaison d'algorithmes génétiques multimodaux sur le problème du condensateur à lames profilées - 200 individus - 200 générations - 4 exécutions

<i>Méthode d'optimisation</i>	T_1 (%)		T_{12} (%)		<i>Meilleure solution (F_{obj})</i>		
	moy	σ	moy	σ	moy	max	σ
aléatoire	0.26	0.03	0.016	0.011	0.869	0.895	0.026
RTS ($w = 30$)	14.45	5.08	3.17	2.62	0.918	0.943	0.020
DC	8.36	0.55	0.52	0.21	0.888	0.899	0.009
CL ($\sigma_s = 0.05$)	81.98	7.18	64.58	16.40	0.953	0.967	0.020

Les tests effectués avec une méthode purement aléatoire montrent que la difficulté du problème est liée au respect des contraintes. Sur 20 000 évaluations, seuls 104 individus respectent la contrainte géométrique tandis que 6 à 7 seulement satisfont les deux contraintes simultanément. Les *méthodes de surpeuplement* et en particulier le *surpeuplement déterministe* sont assez inefficaces. Comme nous l'avons déjà signalé au paragraphe V.5.1.4, la convergence lente qui les caractérise ne permet pas une exploration suffisante de l'espace des solutions réalisables. L'*éclaircissement* conduit à la meilleure solution mais cette dernière est encore loin de l'optimum réel. Le Tableau 31 montre que l'on peut améliorer la qualité du profil en choisissant un nombre de charges plus élevé. Ceci soulève un problème intéressant. En effet, l'accroissement du nombre de charges permet une augmentation de la diversité profils et des variétés de forme qui favorise la détection de l'optimum. Mais, la dimension du

problème d'optimisation résultant est aussi augmentée de 3 pour chaque charge ajoutée, rendant la résolution plus difficile. Il y a ici encore, un compromis à trouver dans le choix du nombre de charges pour augmenter la performance de la méthode.

Tableau 31. Influence du nombre de charges - 200 individus - 200 générations - 4 exécutions

Méthode d'optimisation	Nombre de charges	T ₁ (%)		T ₂ (%)		Meilleure solution (F _{obj})		
		moy	σ	moy	σ	moy	max	σ
CL (σ _s = 0.05)	11 charges	71.96	3.18	64.58	16.40	0.953	0.967	0.020
CL (σ _s = 0.05)	15 charges	74.80	27.62	66.10	32.13	0.972	0.986	0.012

V.5.2.3. Utilisation de charges fixes

Nous pouvons par ailleurs nous interroger sur l'intérêt du choix de charges mobiles dans la méthode développée. Plutôt que de choisir n_q charges *mobiles*, soit $3n_q$ degrés de liberté, ne serait-il pas plus judicieux de rechercher la forme du profil à partir de $3n_q$ charges à *position imposée* ?. Le Tableau 32 présente les résultats obtenus par la méthode d'éclaircissement avec l'utilisation de 45 charges fixes pour deux répartitions de charges différentes. Dans le premier cas, les charges sont distribuées de façon quasi-uniforme à l'intérieur de la zone délimitée par la frontière interne du gabarit à des abscisses $x=-1.05-0.8i$ ($i=0..14$) pour des ordonnées $y=2\pi-0.1, 2\pi-0.5, 2\pi-0.9$. Dans le deuxième cas, les charges sont placées le long de la frontière interne du gabarit à des coordonnées $x=-1.05-0.29i$ ($i=0..41$), $y=2\pi-0.9$ et $x=-1.05, y=2\pi-0.9+0.29j$ ($j=1..3$).

Tableau 32. Utilisation de 45 charge fixes - 200 générations - 200 individus - 4 exécutions

Méthode d'optimisation	Distribution des charges	T ₁ (%)		T ₁₂ (%)		Meilleure solution (F _{obj})		
		moy	σ	moy	σ	moy	max	σ
CL (σ _s = 0.05)	uniforme	65.69	14.18	61.20	11.84	0.919	0.967	0.033
CL (σ _s = 0.05)	sur la frontière	89.00	1.98	86.79	3.78	0.921	0.932	0.013

Nous constatons que pour les deux distributions de charges adoptées, l'utilisation d'un nombre plus élevé de charges fixes n'améliore pas la solution optimale trouvée précédemment. Il est donc plus intéressant de limiter le nombre de charges mais de rechercher à la fois leur position et leur valeur comme nous l'avons fait précédemment.

V.5.2.4. Adaptation du gabarit

L'autre moyen à envisager pour augmenter l'efficacité de la méthode est l'adaptation du gabarit géométrique de façon à donner plus de liberté de mouvement aux charges. La Figure 70 présente deux autres gabarits constitués à partir de la solution optimale trouvée au paragraphe V.5.2.2. La frontière interne de ces gabarits a été élargie par rapport à celle du gabarit initial en fonction de la position des charges et de la forme du profil correspondant à la solution optimale. Le gabarit 2 est défini pour $y \in [2\pi - 3, 2\pi]$ par l'équation suivante :

$$x = -1 - \frac{12}{3^6} (2\pi - y)^6 \tag{156}$$

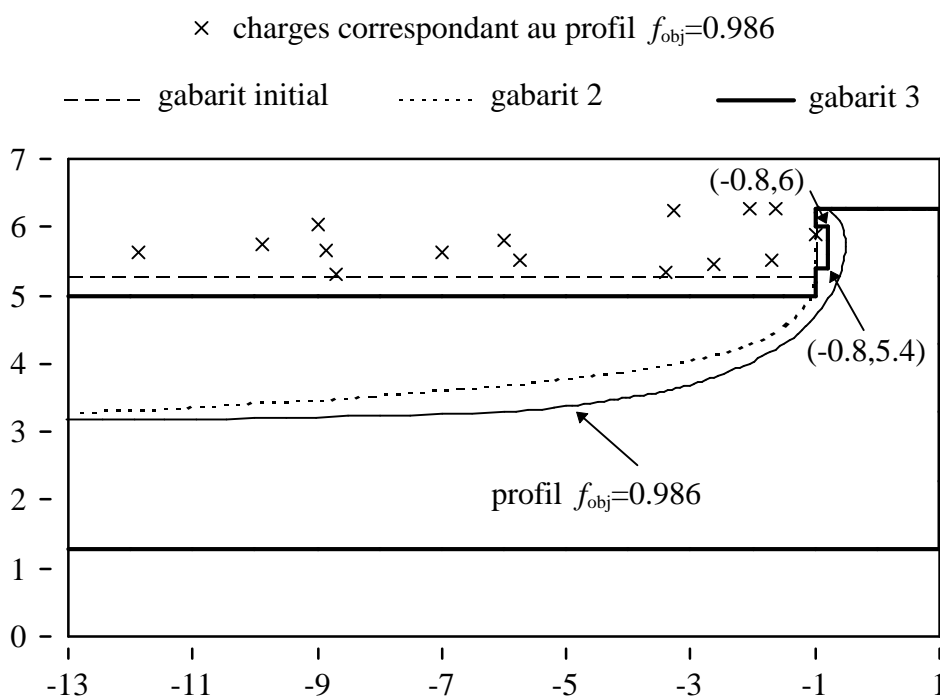


Figure 70. Adaptation du gabarit

- paramétrage du gabarit 2

Si nous utilisons un paramétrage linéaire de la position des charges comme nous l'avons fait précédemment, nous devons introduire une contrainte supplémentaire visant à assurer leur positionnement à l'intérieur de la zone délimitée par la frontière interne. Pour garder l'avantage des contraintes implicites de domaine respectées automatiquement par les algorithmes génétiques et éviter l'adjonction de cette contrainte, nous proposons un paramétrage non linéaire de la position des charges par rapport à la frontière interne du gabarit géométrique. Le principe du paramétrage proposé est illustré à la Figure 71.

tel-00139202, version 1 - 29 Mar 2007

Il s'agit de fixer la position de chaque charge à partir d'un point pris aléatoirement sur la frontière interne du gabarit $F(Y)$ et d'un coefficient de réduction α (ou rapport d'homothétie) compris entre 0 et 1.

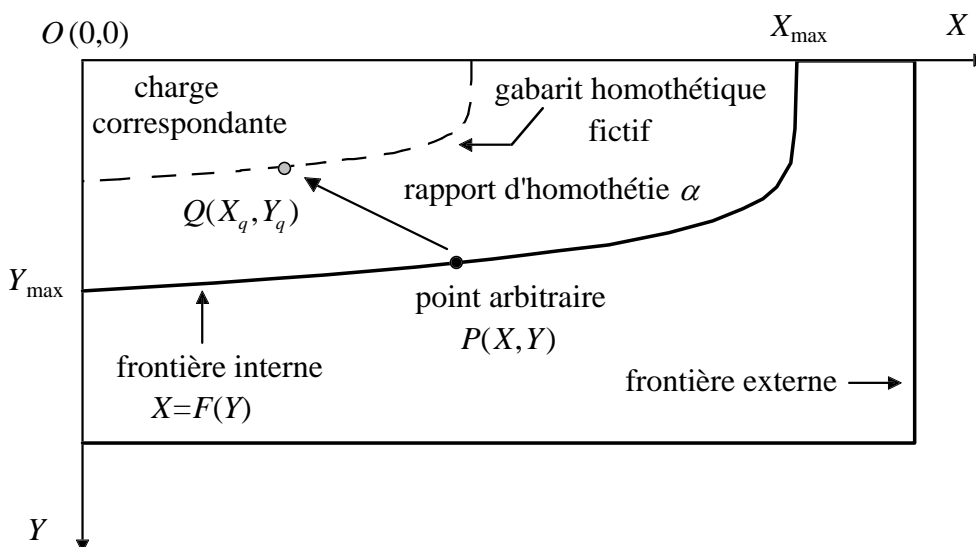


Figure 71. Paramétrage non linéaire de la position des charges

Pour chaque charge, un point arbitraire P du gabarit est déterminé au moyen d'une variable aléatoire u à densité uniforme sur l'intervalle $[0,1]$. Celle-ci fixe l'ordonnée du point $Y = uY_{\max}$ et par suite l'abscisse $X = F(uY_{\max})$. La position de la charge Q relative au point P , est obtenue par homothétie de centre O et de rapport α (α étant aussi une variable aléatoire comprise dans l'intervalle $[0,1]$). L'abscisse et l'ordonnée de la charge sont respectivement $X_q = \alpha X = \alpha F(uY_{\max})$ et $Y_q = \alpha Y = \alpha u Y_{\max}$. De cette manière, les charges restent toujours à l'intérieur de la zone délimitée par la frontière interne.

Si le coefficient de réduction α est une variable aléatoire à densité uniforme, les charges vont avoir tendance à se concentrer vers le point O , centre de l'homothétie. En effet, l'homothétie impose une réduction de l'aire du gabarit fictif d'un rapport α^2 par rapport au gabarit réel. Pour que les charges soient distribuées uniformément dans l'aire délimitée par la frontière interne, il est nécessaire de définir α de la manière suivante :

$$\alpha = \sqrt{v} \tag{157}$$

où v est une variable à densité uniforme sur l'intervalle $[0,1]$.

- paramétrage du gabarit 3

Le gabarit 3 est paramétré de façon standard en plaçant 14 charges dans l'aire définie par $x \in [-13, -1]$, $y \in [5, 2\pi]$ et une charge dans la région caractérisée par $x \in [-1.2, -0.8]$, $y \in [5.4, 6]$.

Le problème est traité comme précédemment avec la méthode d'éclaircissement en prenant une taille de population de 100 et un nombre de génération égale à 200. La solution identifiée au paragraphe V.5.2.2 avec le premier gabarit est insérée dans la population initiale pour le test des gabarits 2 et 3. Le Tableau 33 présente les résultats obtenus à la suite de quatre exécutions successives.

Tableau 33. Adaptation du gabarit - 200 générations - 200 individus - 4 exécutions

Méthode d'optimisation	Gabarit employé	T_1 (%)		T_2 (%)		Meilleure solution (F_{obj})		
		moy	σ	moy	σ	moy	max	σ
CL ($\sigma_s = 0.05$)	initial	74.80	27.62	66.10	32.13	0.972	0.986	0.012
CL ($\sigma_s = 0.05$)	gabarit 2	82.68	1.60	80.54	0.05	0.989	0.989	0.000
CL ($\sigma_s = 0.05$)	gabarit 3	91.96	1.27	89.64	0.75	0.991	0.991	0.000

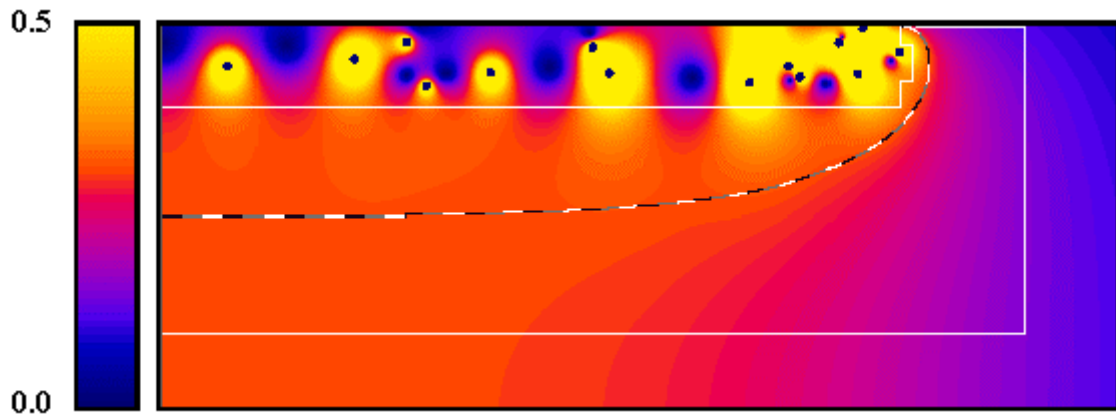
L'adaptation du gabarit géométrique n'a pas permis d'améliorer réellement la configuration trouvée avec le gabarit initial. Les charges restent concentrées dans une région incluse dans le gabarit initial quel que soit le gabarit testé. La solution a tout de même été affinée par les 20 000 évaluations supplémentaires. Le meilleur profil, détecté avec le gabarit 3, ne diffère que de 0.9% de l'optimum réel pour le critère d'optimisation considéré.

V.5.2.5. Profil optimal

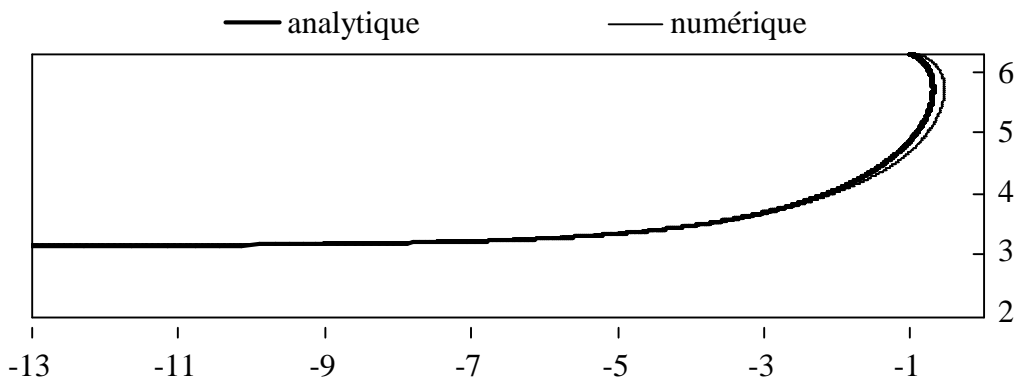
La Figure 72 présente les caractéristiques de la solution optimale trouvée par la méthode d'éclaircissement à savoir :

- la répartition de champ électrique dans l'ensemble du domaine (Figure 72a)
- le profil définitif obtenu en comparaison avec la solution analytique exacte (Figure 72b)
- le champ électrique sur le contour (Figure 72c)

(a) répartition de champ électrique dans l'ensemble du domaine



(b) profil optimal théorique (analytique) et expérimental (numérique)



(c) valeur du champ électrique sur le profil

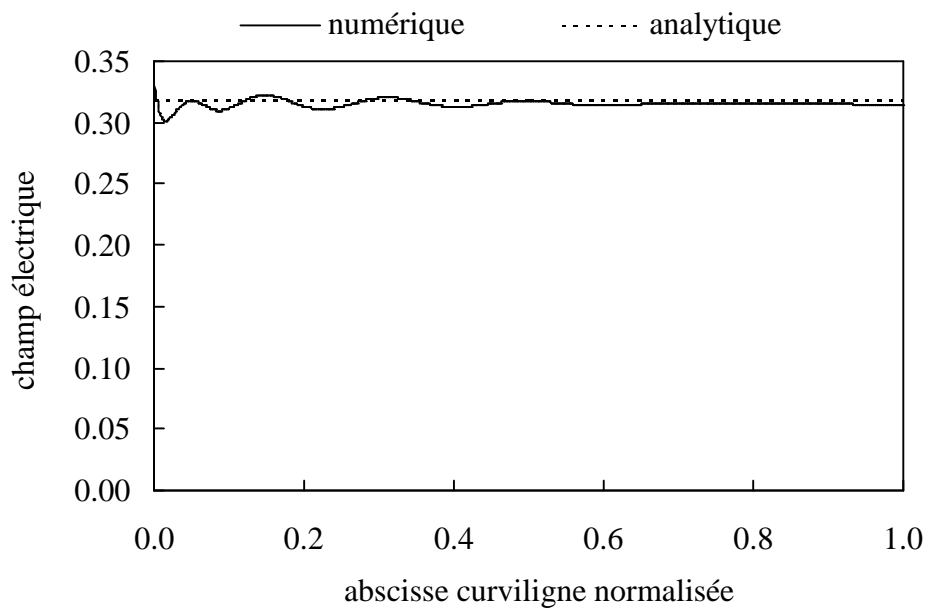


Figure 72. Caractéristiques du profil optimal détecté par algorithme génétique

V.5.3. Electrode avec blindage

Il s'agit d'un problème d'optimisation de forme analogue à celui présenté par [Mukherjee 96] ; l'auteur propose une solution à partir d'un paramétrage direct de la forme (voir Figure 53).

V.5.3.1. Définition du problème

Un problème équivalent est défini à partir du gabarit géométrique et des conditions aux limites de la Figure 73. Il s'agit de minimiser le champ électrique sur le contour d'une électrode satisfaisant le gabarit géométrique imposé. La difficulté du problème est liée à la présence de la condition de Dirichlet homogène $V=0$ pour $r=2.5$.

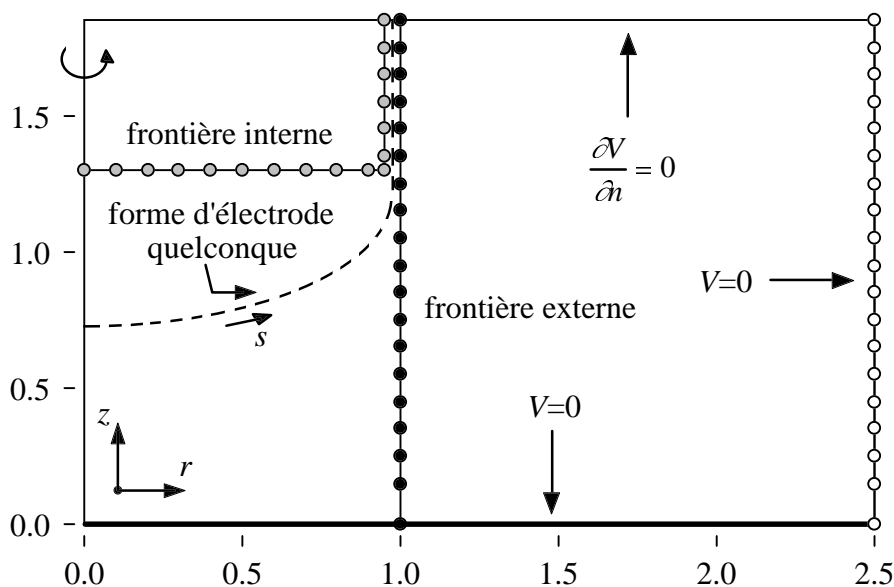


Figure 73. Electrode avec blindage

V.5.3.2. Résolution numérique

Six charges placées à l'intérieur de la zone délimitée par la frontière interne du gabarit simulent l'équipotentielle représentant le profil de l'électrode. Six autres charges caractérisées par $r \in [2.5, 5]$ et $z \in [0, 1.85]$ permettent d'assurer la condition de Dirichlet homogène en $r=2.5$. Comme précédemment, la valeur et la position des charges sont déterminées par un algorithme génétique. Le problème d'optimisation résultant est de dimension $n=36$. Le calcul des grandeurs électriques est réalisé en axisymétrie en considérant 12 anneaux chargés auxquels sont ajoutés 12 autres anneaux pour assurer approximativement la condition de Neuman homogène en $z=1.85$ et 24 autres pour garantir de façon exacte la condition de Dirichlet homogène en $z=0$.

Outre la contrainte géométrique relative au respect du gabarit exprimée par l'équation (141), le potentiel doit être nul en $r=2.5$. Cette condition est testée à partir des points de contrôle placés sur la ligne correspondante (voir Figure 73) et exprimée sous la forme d'une contrainte d'inégalité :

$$\max |V(r = 2.5)| \leq 1\% \quad (158)$$

qui garantit un potentiel nul sur l'ensemble des points de contrôle avec une erreur maximale de 1%.

Lorsqu'un individu ne satisfait pas la contrainte géométrique (141), sa fonction d'adaptation est calculée de façon similaire à (148) en prenant en compte la deuxième contrainte :

$$F_{obj} = \frac{1}{1 + \lambda_1 + \lambda_2 (V_{\max}^{\text{ext}} - V_{\min}^{\text{int}}) + \lambda_3 \max(0, \max |V(r = 2.5)| - 0.01)} \quad (159)$$

où $\lambda_2 \gg \lambda_3$ pour les raisons que nous avons évoquées au paragraphe V.5.2.2.

Dans le cas où la contrainte géométrique est satisfaite, la normalisation du potentiel est effectuée par rapport à la valeur du potentiel maximal sur la frontière externe $\tilde{V}_{obj} = V_{\max}^{\text{ext}}$. Après identification de l'équipotentielle avec une grille de coté $2\xi=0.01$, le champ électrique maximal sur le profil est estimé. En prenant en compte la contrainte liée au potentiel en $r=2.5$, nous définissons la fonction d'adaptation des individus de la façon suivante :

$$F_{obj} = \frac{1}{1 + E_{\max} + \lambda_3 \max(0, \max |V(r = 2.5)| - 0.01)} \quad (160)$$

où E_{\max} désigne le champ électrique maximal sur l'équipotentielle.

Trois algorithmes génétiques multimodaux sont testés sur ce problème. La taille de la population est fixée à 100 et le nombre de générations à 200. Les paramètres sont codés chacun dans une chaîne binaire de 10 bits avec décodage dynamique. Les probabilités de croisement et de mutation sont respectivement $p_c = 1$ et $p_m = 0.001$. Le critère de distance utilisé est de type d^∞ normalisé (voir chapitre III). Les coefficients de pénalité ont été fixés suivant $\lambda_1 = \lambda_3 = 100$ et $\lambda_2 = 10^5$. Quatre exécutions sont effectuées avec une population initiale différente générée aléatoirement. Le Tableau 34 compare les résultats obtenus avec les algorithmes génétiques multimodaux et une méthode purement aléatoire. Les taux d'individus satisfaisant au cours des exécutions, la contrainte géométrique uniquement, et les deux contraintes simultanément, sont désignés respectivement par T_1 et T_{12} .

Tableau 34. Comparaisons d'algorithmes génétiques multimodaux sur le problème de l'électrode à écran - 200 générations - 200 individus - 4 exécutions

Méthode d'optimisation	T_1 (%)		T_{12} (%)		Meilleure solution (E_{\max})		
	moy	σ	moy	σ	moy	min	σ
Aléatoire	0.17	0.02	0	0	-	-	-
RTS ($w = 30$)	3.69	0.42	0.02	0.02	2.02	1.99	0.05
DC	2.64	0.59	0.01	0.00	2.16	1.99	0.29
CL ($\sigma_s = 0.05$)	14.29	4.86	2.49	0.81	1.85	1.83	0.02

Les résultats sont assez similaires à ceux obtenus pour le problème du condensateur à lames profilées. Ici encore, la difficulté est liée au respect des contraintes et en particulier de la condition de Dirichlet homogène en $r=2.5$. Avec 20 000 évaluations aléatoires des paramètres, aucune configuration satisfaisant cette contrainte n'a pu être trouvée.

Les méthodes de surpeuplement sont à peine plus efficaces puisqu'elles ne détectent qu'au maximum une dizaine de configurations réalisables et ne permettent donc pas l'optimisation du champ électrique sur le profil.

Seule la méthode d'éclaircissement réussit là où les autres techniques échouent mais, les faibles pourcentages obtenus pour T_1 et T_{12} témoignent de la difficulté du problème. Avec cette approche, la prise en compte de conditions aux limites (ou de conditions d'interfaces) en un certain nombre de points est assez délicate en raison des contraintes sévères à respecter.

V.5.3.3. Profil optimal

La solution optimale obtenue avec la méthode d'éclaircissement après 20 000 évaluations est illustrée à la Figure 74 et comparée à deux profils quelconques. La solution donnée dans [Mukherjee 96] s'est avérée inexacte, comme l'a reconnu l'auteur, et ne peut donc malheureusement pas être présentée en comparaison!

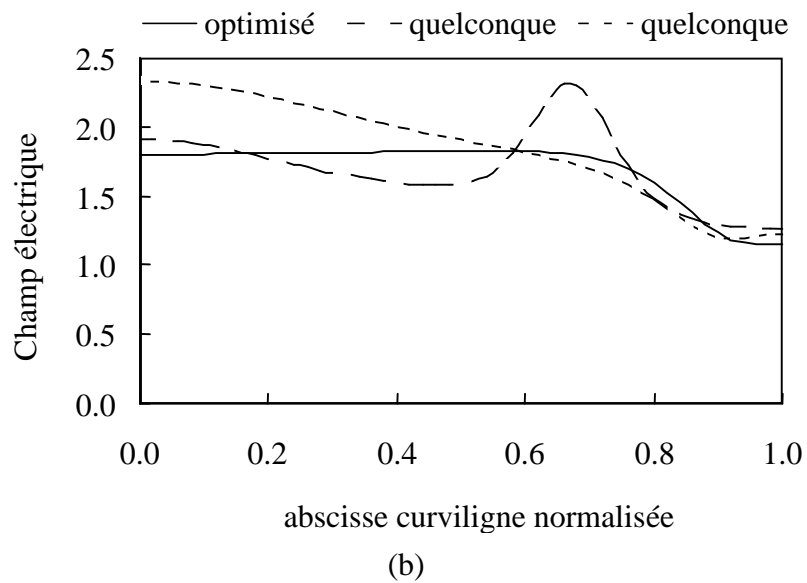
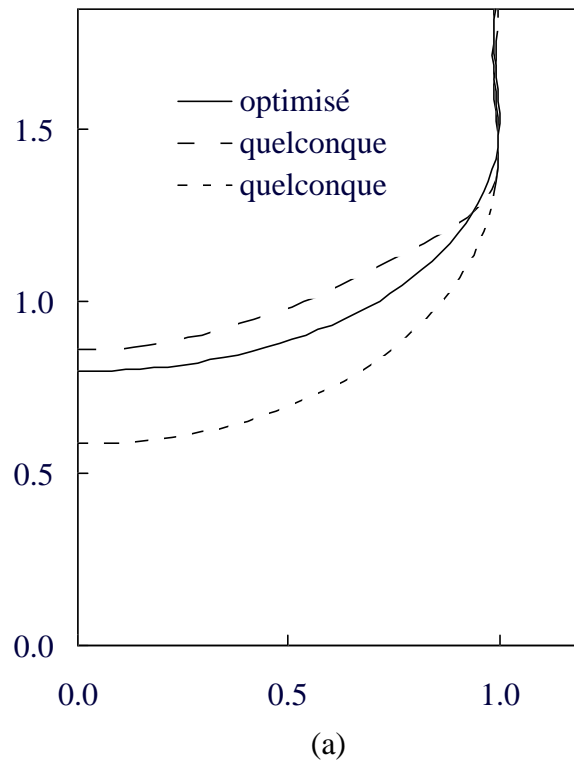


Figure 74. Electrode avec blindage

(a) profil optimal

(b) champ électrique sur le profil

V.6. CONCLUSION

Dans ce chapitre, nous avons présenté et validé sur des cas simples une méthode d'optimisation de forme d'électrode à paramétrage indirect basée sur des distributions de charges fictives. La technique développée consiste à déterminer le contour de l'électrode à

partir d'une équipotentielle comprise entre deux frontières d'un gabarit géométrique prédéfini. L'application d'algorithmes génétiques multimodaux pour la détermination de la position et de la valeur des charges offre la possibilité de détecter des formes multiples distinctes lorsque plusieurs solutions optimales existent pour un gabarit donné. La procédure d'optimisation est extensible à des électrodes en contact avec plusieurs milieux diélectriques ou soumises à diverses conditions aux limites. Néanmoins, elle nécessite alors l'introduction de contraintes supplémentaires qui augmentent considérablement la difficulté du problème et pénalisent la résolution.

Parmi l'ensemble des algorithmes génétiques multimodaux testés, la *méthode d'éclaircissement* s'avère la plus performante en particulier lorsque les contraintes liées à la géométrie ou aux conditions aux limites sont sévères. Les *techniques de surpeuplement* sont dans ce cas incapables de converger rapidement vers l'espace réalisable en raison de leur caractère stationnaire. Plusieurs pistes sont envisageables pour augmenter l'efficacité de cette méthode :

- l'introduction du *nombre* de charges fictives comme paramètre supplémentaire au procédé d'optimisation

Le nombre de charges fictives simulant l'équipotentielle relative à l'électrode pourrait être codé dans le chromosome des individus afin d'être adapté de façon optimale au cours de la recherche des profils. Cela nécessiterait la définition d'opérateurs spécifiques de croisement capables de recombinaison des configurations présentant un nombre de charges et des chromosomes différents. Le critère de distance employé par les algorithmes de nichage devrait par ailleurs être indépendant des charges mais lié directement à la forme de l'électrode pour la comparaison de profils distincts. A notre connaissance, l'optimisation à nombre variable de paramètres est un sujet encore vierge à ce jour.

- *l'adaptation du gabarit géométrique et du positionnement des charges* au cours de la recherche des profils optimaux

Il s'agirait d'adapter le gabarit et plus précisément l'espace de mobilité des charges pendant l'exécution de l'algorithme d'optimisation et non pas séquentiellement comme nous l'avons fait au paragraphe V.5.2.4.

- le *couplage* avec la *méthode des charges équivalentes*

Il faudrait pouvoir coupler la technique d'optimisation de forme à paramétrage indirect avec la méthode des charges équivalentes [Annexe V] pour traiter localement des conditions d'interface ou des conditions aux limites particulières. Cela permettrait d'envisager des problèmes plus complexes.

CONCLUSION GENERALE

Nous espérons avoir permis de mieux cerner l'efficacité des méthodes génétiques multimodales pour l'optimisation en électromagnétisme. Nous avons analysé les divers mécanismes qui conditionnent la recherche des optima et montré les forces et les faiblesses de chacune des méthodes pour des problèmes tests mathématiques et physiques.

Bien que la technique du partage soit la plus populaire dans la littérature, celle-ci a été mise en défaut dans la plupart des cas à cause de son instabilité et de son incapacité à converger lorsque le nombre d'individus de la population est limité. Les méthodes de surpeuplement sont intéressantes pour des problèmes d'optimisation réels car elles ne nécessitent pas, pour susciter la formation de niches, de définir un seuil de voisinage entre individus. Elles restent assez efficaces y compris lorsque la taille de la population est très limitée mais peuvent échouer si la fonction objectif est trop sélective (ou si les contraintes de conception sont sévères) en raison de leur caractère stationnaire. La méthode d'éclaircissement est dans ce cas plus robuste, à condition qu'elle soit utilisée avec une taille de population suffisante (au moins quatre individus par nombre de solutions optimales que l'on veut identifier) et un rayon de niche approprié. La définition *a priori* de ce rayon est problématique. Néanmoins, le développement de procédures de tri et d'identification des individus dominants devrait permettre de s'affranchir de cette difficulté et d'appliquer l'éclaircissement indépendamment d'un seuil de similarité fixé de façon hasardeuse.

Le nichage offre alors des perspectives séduisantes notamment pour une meilleure caractérisation de la sensibilité et de la stabilité des solutions optimales. Les méthodes multimodales devraient par ailleurs permettre de traiter plus aisément des problèmes d'optimisation à objectifs multiples, en assurant une identification globale de la frontière de *Pareto*.

Du point de vue purement algorithmique, il nous semble intéressant d'envisager le couplage entre les méthodes de nichage et d'autres techniques d'optimisation évolutionnistes. L'algorithme génétique à codage binaire semble aujourd'hui montrer quelques limitations pour l'optimisation continue de paramètres réels. L'association des méthodes de nichage avec les stratégies d'évolution, l'évolution différentielle ou les algorithmes génétiques à codage réel, devrait être plus efficace, tant en ce qui concerne la précision de la localisation des optima, que la vitesse de convergence elle-même.

Conclusion générale

Un autre aspect fondamental du nichage, peu abordé dans ce travail, concerne la définition et l'étude de critères de distance appropriés pour l'optimisation de forme ou de structures (choix d'un matériau, homogénéité des propriétés physiques, isotropie, ...). Le choix de la distance qui caractérise la similarité entre configurations à optimiser n'est jamais unique. La caractérisation de critères de distance ne dépendant pas explicitement des paramètres du problème est à envisager.

Nous avons aussi présenté une méthode originale de conception de forme d'électrodes optimales basée sur une technique de distribution de charges fictives qui s'avère être un très bon exemple pour tester et faire progresser les algorithmes d'optimisation sous contraintes. En outre, la notion de paramétrage indirecte des contours que nous avons introduite constitue une nouvelle approche de l'optimisation de forme. Bien que la technique mise en œuvre au moyen de charges fictives soit spécifique à l'électrostatique, il pourrait être intéressant de rechercher des procédés de conception automatique analogues dans un cas plus général.

Enfin, ce travail devrait aboutir, dans les mois qui viennent, à l'étude d'applications plus concrètes concernant l'optimisation de parafoudres H.T., la résolution de problèmes inverses en C.E.M. et en C.N.D., la conception automatique de circuits électriques, et l'étude de matériaux à propriétés physiques spécifiques.

VALIDATION DU GENERATEUR DE NOMBRES ALEATOIRES

AI.1. INTRODUCTION

Le bon fonctionnement d'un algorithme stochastique ne peut être déterminé que si le générateur de nombres aléatoires ou de probabilité est fiable. Il est donc impératif de s'assurer que ce dernier se rapproche d'un générateur idéal et qu'il ne présente aucune tendance ou périodicité.

AI.2. GENERATEUR DE NOMBRE ALEATOIRE EMPLOYE

Les générateurs de nombres aléatoires utilisés par les algorithmes stochastiques qui ont été développés, sont dérivés du générateur de nombres pseudo-aléatoires du langage ANSY-C. Dans ce langage, la fonction **rand()** renvoie un nombre entier compris entre 0 et 2^{32} avec une densité de probabilité uniforme sur l'intervalle complet. Le générateur de nombres réels à densité uniforme sur l'intervalle [min,max] est défini à partir de la fonction **rand()** de la façon suivante :

$$\mathbf{Rrand}(\mathbf{min},\mathbf{max}) = \mathbf{min} + (\mathbf{max} - \mathbf{min}) * \mathbf{rand}() / 2^{32}$$

Le générateur de probabilité à densité uniforme est simplement **Rrand(0,1)**. La période de ce générateur pseudo-aléatoire est $T=2^{32}$

AI.3. PERFORMANCE D'UN GENERATEUR DE NOMBRES ALEATOIRES

La densité de probabilité caractérisant un générateur de nombre aléatoire doit se rapprocher le plus possible de celle d'une distribution idéale. Le générateur ne doit pas présenter de biais, de tendances ou de périodicités. Nous nous intéresserons par la suite à la validation du générateur de probabilité **Rrand(0,1)** utilisé par tous les algorithmes génétiques développés. Deux critères de performance sont généralement employés pour tester l'efficacité d'un générateur de nombres aléatoires à densité uniforme. Il s'agit du test d'écart de distribution χ^2 et du test de corrélation retardée.

AI.3.1. Le test du khi-deux

Le test du khi-deux χ^2 mesure l'écart de distribution du générateur de nombres aléatoires par rapport à une distribution supposée idéale. L'intervalle complet du générateur est divisé en un nombre fini de sous-intervalles. Il n'est pas nécessaire que chacun de ces intervalles soit de même longueur. Néanmoins, nous l'imposerons par raison de simplicité. Le test consiste, à partir d'un certain nombre N de valeurs extraites du générateur, à compter le nombre d'échantillons présents dans chacun des sous-intervalles et à comparer ce dernier avec le nombre d'échantillons attendu dans le cas d'une distribution idéale. L'écart de distribution alors exprimé de la façon suivante :

$$\chi^2 = \sum_{i=1}^m \frac{(X_i - \mu_i)^2}{\mu_i} \quad (\text{AI.1})$$

où m désigne le nombre de sous-intervalles, X_i le nombre d'échantillons dans l'intervalle i et μ_i le nombre d'occurrences dans l'intervalle correspondant pour une distribution supposée idéale. Dans le cas d'un générateur de probabilité à densité uniforme, le nombre d'occurrences attendu dans chaque sous-intervalle est simplement le nombre total de valeurs extraites du générateur divisé par le nombre total d'intervalles :

$$\mu_i = \frac{N}{m} \quad (\text{AI.2})$$

Un bon générateur doit présenter une valeur de χ^2 inférieure au nombre d'intervalle. Plus la valeur du critère est petite, meilleur est le générateur. Des tables de statistiques précisent la valeur limite du critère χ^2 à partir de laquelle nous pouvons considérer que les deux distributions diffèrent avec un niveau de confiance γ [Pearson 54]. Cette valeur est fonction du nombre de degrés de liberté égal à $m-1$. Nous pouvons de même déduire de ces tables, la valeur limite à partir de laquelle les deux distributions peuvent être considérées comme identiques avec un niveau de confiance γ' . Le Tableau AI.1 donne les valeurs limites du khi-deux pour 29 degrés de liberté et différents niveaux de confiance. Si par exemple, le khi-deux d'un générateur de probabilité à densité uniforme est estimé à 21, le générateur peut être considéré uniforme avec un niveau de confiance $\gamma'=80\%$ et non uniforme avec un niveau de confiance $\gamma=10\%$.

Tableau AI.1. Valeur critique du khi-deux pour 29 degrés de liberté

γ (%)	0.5	1	2.5	5	10	20	30	50	70	80	90	95	97.5	99	99.5
γ' (%)	99.5	99	97.5	95	90	80	70	50	30	20	10	5	2.5	1	0.5
χ^2	13.1	14.3	16.0	17.7	19.8	22.5	24.6	28.3	32.5	35.1	39.1	42.6	45.7	49.6	52.3

AI.3.2. La corrélation retardée

La corrélation retardée souligne les dépendances entre deux nombres générés à deux instants différents. Elle révèle les éventuelles tendances ou périodicités. La corrélation au retard τ est calculée de la manière suivante :

$$\rho_{xy}(\tau) = \frac{\sum_{i=1}^{N-\tau} x_i y_{i+\tau} - \frac{1}{N-\tau} \sum_{i=1}^{N-\tau} x_i \sum_{i=1}^{N-\tau} y_{i+\tau}}{\sqrt{\sum_{i=1}^{N-\tau} x_i^2 \sum_{i=1}^{N-\tau} y_{i+\tau}^2}} \quad (\text{AI.3})$$

où N est le nombre d'échantillons. Cette quantité représente l'intercorrélation lorsque x et y désignent deux séries de données différentes et l'autocorrélation lorsque celles-ci sont identiques. Pour évaluer la performance du générateur de probabilité à densité uniforme, nous nous intéressons uniquement à l'autocorrélation :

$$\rho_{xx}(\tau) = \frac{\sum_{i=1}^{N-\tau} x_i x_{i+\tau} - \frac{1}{N-\tau} \sum_{i=1}^{N-\tau} x_i \sum_{i=1}^{N-\tau} x_{i+\tau}}{\sqrt{\sum_{i=1}^{N-\tau} x_i^2 \sum_{i=1}^{N-\tau} x_{i+\tau}^2}} \quad (\text{AI.4})$$

$$\rho_{xx}(0) = \frac{\sum_{i=1}^N x_i^2 - \frac{1}{N} \left(\sum_{i=1}^N x_i \right)^2}{\sum_{i=1}^N x_i^2} \quad (\text{AI.5})$$

La valeur de l'autocorrélation d'un générateur de nombre aléatoire idéal est égale à un lorsque $\tau = 0$ et zéro pour toute autre valeur de τ . Une valeur importante de l'autocorrélation pour des valeurs non nulles de τ indique une périodicité dans le temps. Si l'autocorrélation chute très lentement vers zéro alors que τ augmente, cela signifie que les nombres générés ne sont pas vraiment indépendants les uns des autres.

AI.4. CARACTERISATION DU GENERATEUR DE PROBABILITES

AI.4.1. Test qualitatif

Ce premier test consiste à faire figurer dans le domaine 2D un nuage de points correspondant à un ensemble de couples de valeurs aléatoires indépendantes. Ce type de tracé permet de se rendre compte directement de la qualité du générateur. Les regroupements de points et les chemins constitués dans le nuage caractérisent directement les tendances, les

biais et les périodicités du générateur. La Figure AI.1 présente un exemple de tracé obtenu pour 2000 couples de points. Nous constatons que la distribution des points est quasi-uniforme dans les deux directions. Il n'apparaît pas de chemins ou de regroupements de points marqués. Il n'y a pas non plus de zones délaissées par les points. Le générateur semble à première vue avoir un comportement satisfaisant.

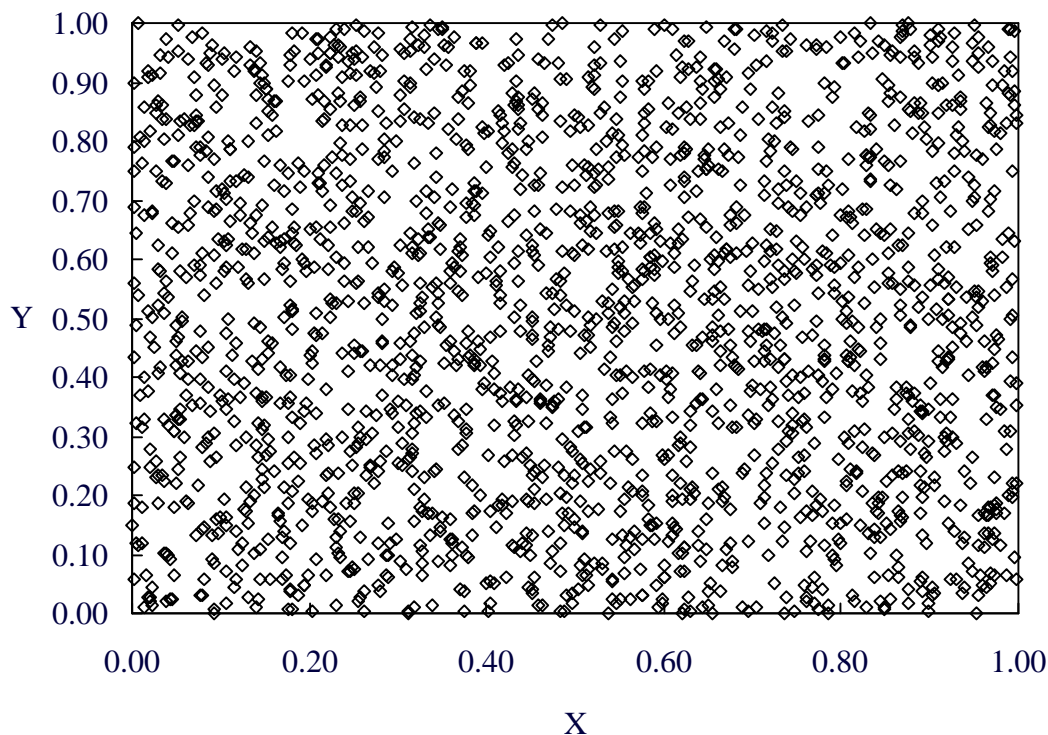


Figure AI.1. Tracé de 2000 couples de points aléatoires indépendants

AI.4.2. Test du khi-deux

L'écart de distribution χ^2 est calculé à partir d'un échantillon de 3000 valeurs et de 30 intervalles de divisions. Le nombre attendu de valeurs dans chaque intervalle est par conséquent égal à 100. 100 évaluations du critère sont effectuées pour différents échantillons de 3000 valeurs. Le khi-deux moyen obtenu à partir des 100 tests est égal à 19.10 et l'écart type correspondant vaut 6.97. Nous pouvons donc conclure d'après les données du Tableau AI.1 que la distribution du générateur de nombres aléatoires est uniforme avec un niveau de confiance d'au moins 90%

La Figure AI.2 présente le nombre d'occurrences dans les 30 intervalles de longueur identique obtenu à partir d'un échantillon de 3000 valeurs. Le critère khi-deux correspondant est de 19.

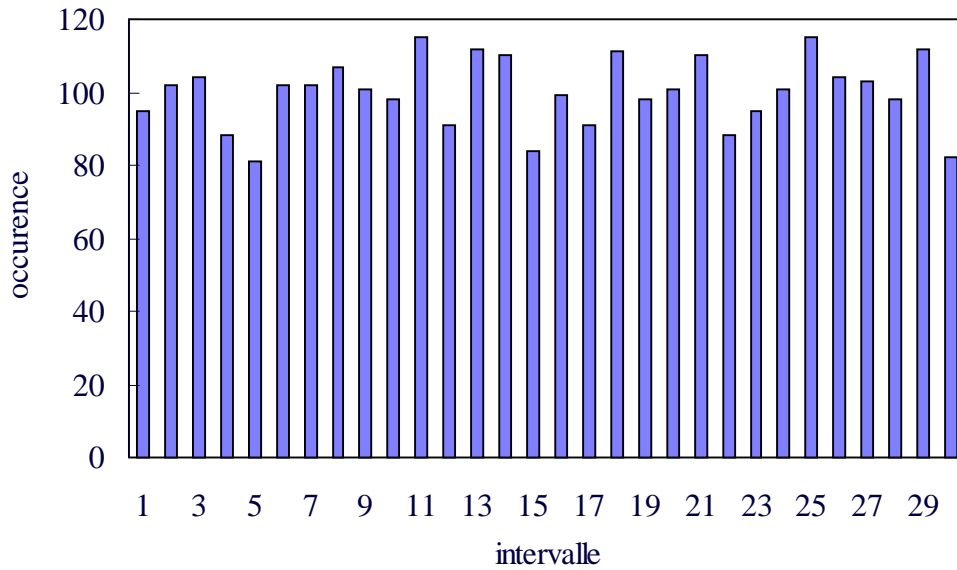


Figure AI.2. Nombre d'occurrences typique dans chacun des 30 intervalles de même longueur pour un échantillon de 3000 valeurs ($\chi^2 = 19$).

AI.4.3. Test d'autocorrélation

La valeur de l'autocorrélation est calculée à partir d'un échantillon de 3000 valeurs pour un retard variant de 0 à 2500. Les résultats obtenus sont présentés à la Figure AI.3.

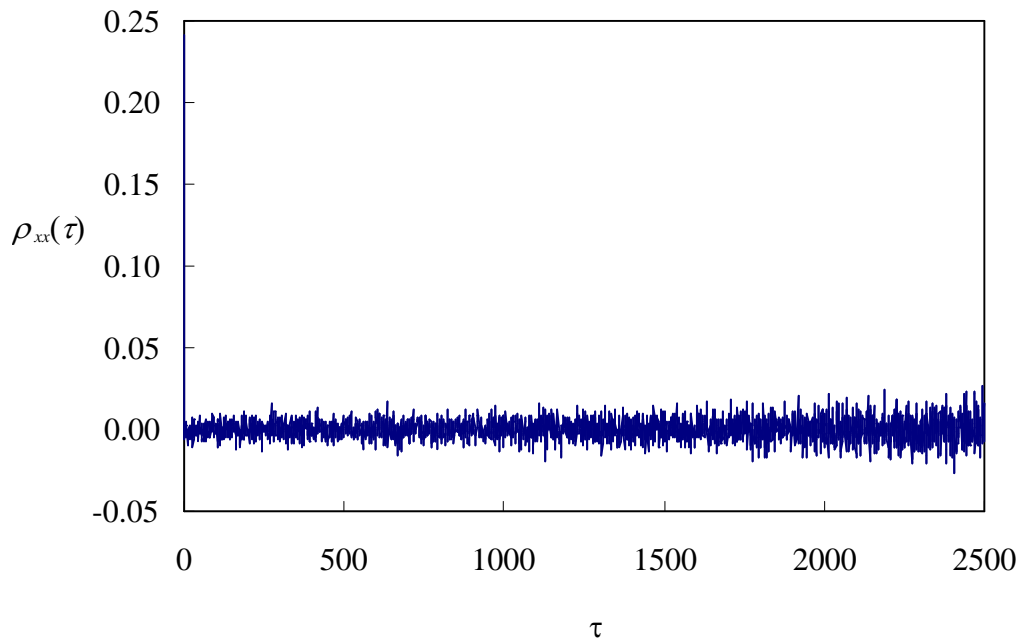


Figure AI.3. Autocorrélation retardée calculée à partir d'un échantillon de 3000 valeurs.

Annexe I

Nous constatons bien l'absence de dépendances ou de périodicités entre les nombres générés pour les 2500 valeurs testées. L'autocorrélation reste très faible lorsque τ est supérieur à zéro validant le bon fonctionnement du générateur.

ANNEXE II

FONCTIONS TESTS MATHÉMATIQUES

Dans cette partie, nous présentons des fonctions tests classiques de la littérature. La première série est généralement utilisée pour tester des algorithmes évolutionnistes unimodaux. Il s'agit des cinq fonctions populaires de *De Jong* [De Jong 75] auxquelles nous avons ajouté la fonction fortement multimodale de *Rastrigin* [Mühlenbein 93]. La deuxième série de tests est employée pour évaluer la performance des algorithmes génétiques multimodaux.

AII.1. FONCTIONS TESTS POUR ALGORITHME GENETIQUE UNIMODAL

Cette série de problèmes spécifiquement créés pour tester la performance des algorithmes d'optimisation, inclue six fonctions mathématiques :

- continues ou discontinues
- convexes ou concaves
- unimodales ou multimodales
- quadratiques ou non quadratiques
- de faible dimension ou de dimension élevée
- déterministes ou stochastiques

AII.1.1. Fonction sphère

La fonction *sphère* ou *paraboloïde* est un problème unimodal, convexe, continu et symétrique :

$$f_1(x) = \sum_{i=1}^n x_i^2 \quad (\text{AII.1})$$

avec $-5.12 \leq x_i \leq 5.12$

Annexe II

Cette fonction possède un minimum $f_1(x) = 0$ en $x = (0,0,\dots,0)$. Les paramètres x_i sont habituellement codés sur 10 bits. La dimension adoptée lors des tests est généralement $n=3$. Le problème équivalent, formulé en terme de maximisation, s'écrit :

$$\max F_1(x) = 30n - f_1(x) \quad (\text{AII.2})$$

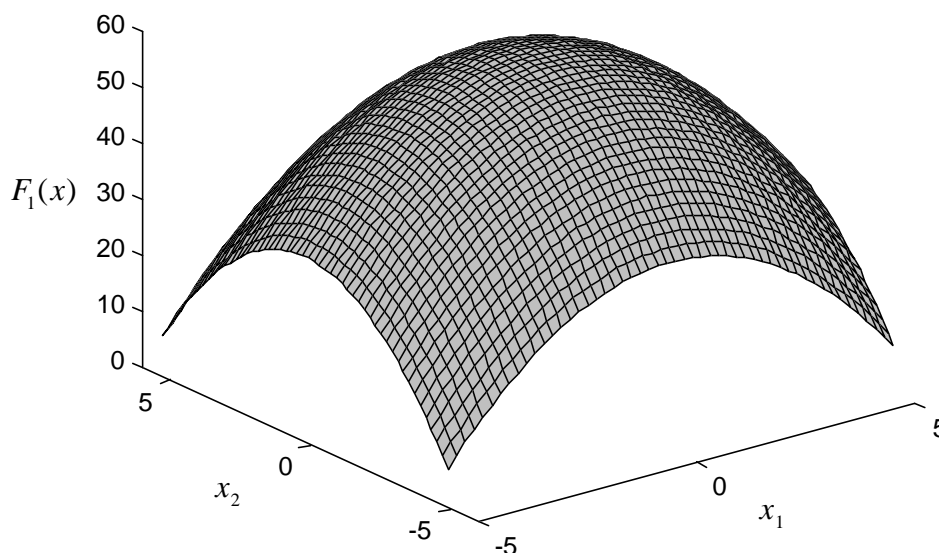


Figure AII.1. Représentation de $F_1(x)$ pour $n=2$

AII.1.2. Fonction de Rosenbrock

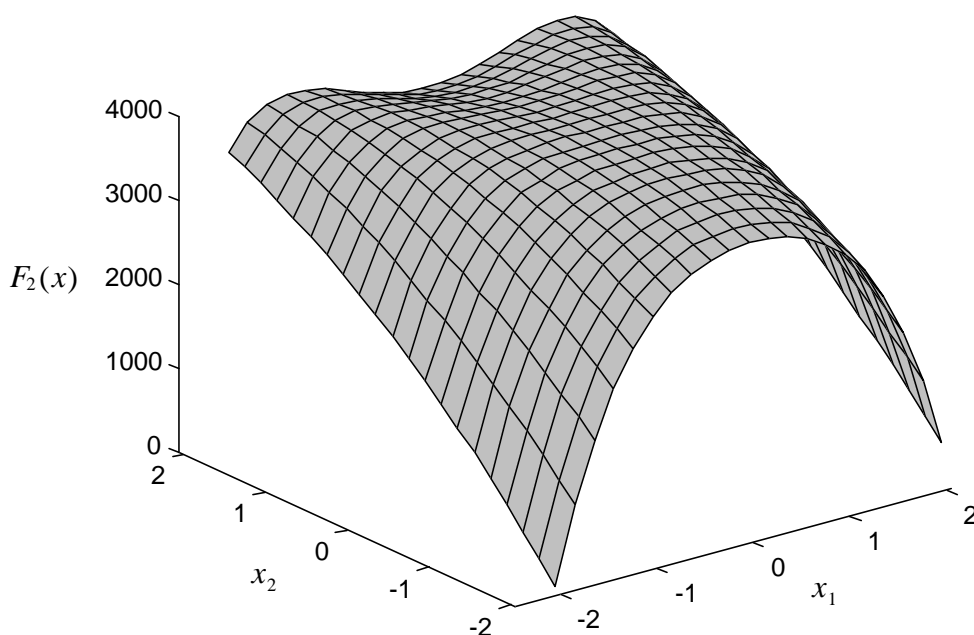
La fonction de *Rosenbrock* est unimodale mais le minimum se trouve dans une région très étroite à l'intersection de vallées très peu pentues :

$$f_2(x) = \sum_{i=1}^{n-1} 100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2 \quad (\text{AII.3})$$

$$\text{avec } -2.048 \leq x_i \leq 2.048$$

Le minimum de la fonction est localisé en $x = (1,1,\dots,1)$ et vaut $f_2(x) = 0$. Les paramètres x_i sont habituellement codés sur 12 bits. La dimension adoptée lors des tests est généralement $n=2$. Le problème équivalent, formulé en terme de maximisation, s'écrit :

$$\max F_2(x) = 4000(n - 1) - f_2(x) \quad (\text{AII.4})$$

Figure AII.2. Représentation de $F_2(x)$ pour $n=2$

AII.1.3. Fonction paliers

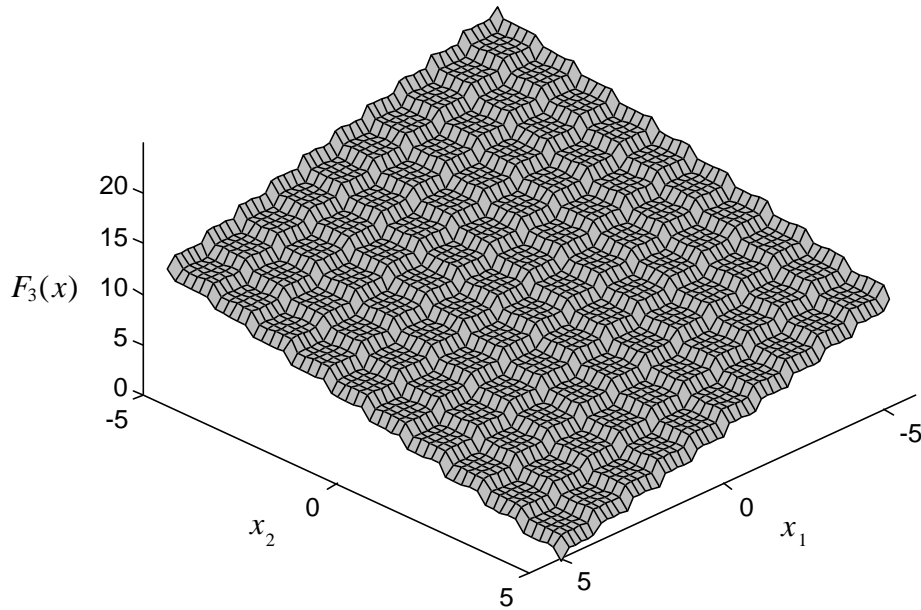
La fonction *paliers* est un problème discontinu à multiple optima.

$$f_3(x) = \sum_{i=1}^n \text{partie entière}(x_i) \quad (\text{AII.5})$$

avec $-5.12 \leq x_i \leq 5.12$

Les minima de la fonction sont localisés aux points de coordonnées $x = ([-5.12, -5], [-5.12, -5], \dots, [-5.12, -5])$ et ont pour valeur $f_3(x) = -6n$. Les paramètres x_i sont habituellement codés sur 10 bits. La dimension adoptée lors des tests est généralement $n=5$. Le problème équivalent, formulé en terme de maximisation, s'écrit :

$$\max F_3(x) = 5n - f_3(x) \quad (\text{AII.6})$$

Figure AII.3. Représentation de $F_3(x)$ pour $n=2$

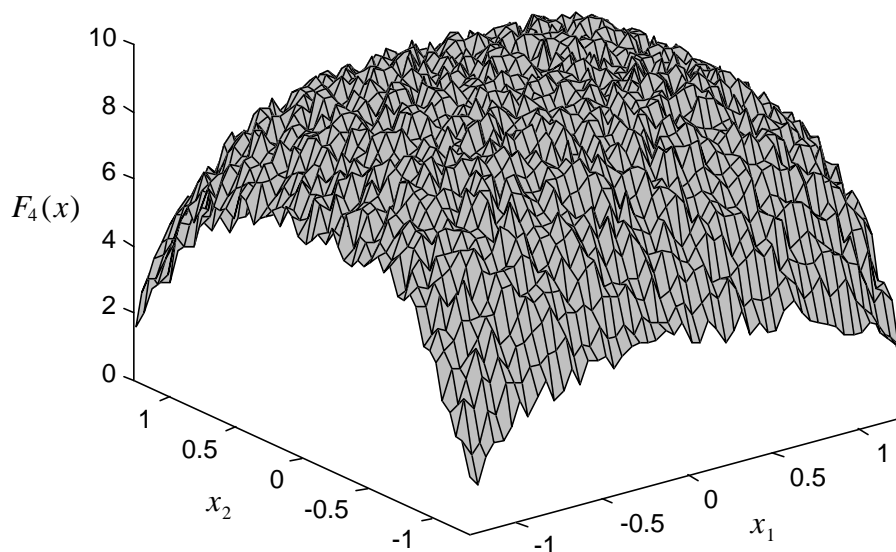
AII.1.4. Fonction bicarée bruitée

$$f_4(x) = \sum_{i=1}^n ix_i^4 + \eta \quad (\text{AII.7})$$

avec $-1.28 \leq x_i \leq 1.28$

La difficulté de cette fonction est liée au bruit stochastique ajouté. Dans l'étude originale de De Jong [De Jong 75], η est une variable gaussienne aléatoire de moyenne nulle et d'écart-type unitaire. Bien que cette fonction ait été conçue pour tester la performance d'une méthode d'optimisation en présence de bruit, elle n'est pas nécessairement appropriée car aucun minimum n'existe réellement. Une fois qu'un algorithme a réussi à minimiser le premier terme de la fonction, il peut continuer à rechercher indéfiniment des valeurs de η faibles. Pour assurer la présence d'un unique optimum, il est préférable d'utiliser une variable aléatoire à densité uniforme sur l'intervalle $[0,1]$ [Ingber 92]. Le minimum de la fonction est alors localisé en $x = (0,0,\dots,0)$ et vaut $f_4(x) = 0$. Les paramètres x_i sont habituellement codés sur 8 bits. La dimension adoptée lors des tests est généralement $n=30$. Le problème équivalent, formulé en terme de maximisation, s'écrit :

$$\max F_4(x) = \frac{3}{2}n(n+1) - f_4(x) \quad (\text{AII.8})$$

Figure AII.4. Représentation de $F_4(x)$ pour $n=2$

AII.1.5. Fonction "trous de renard" de Shekel

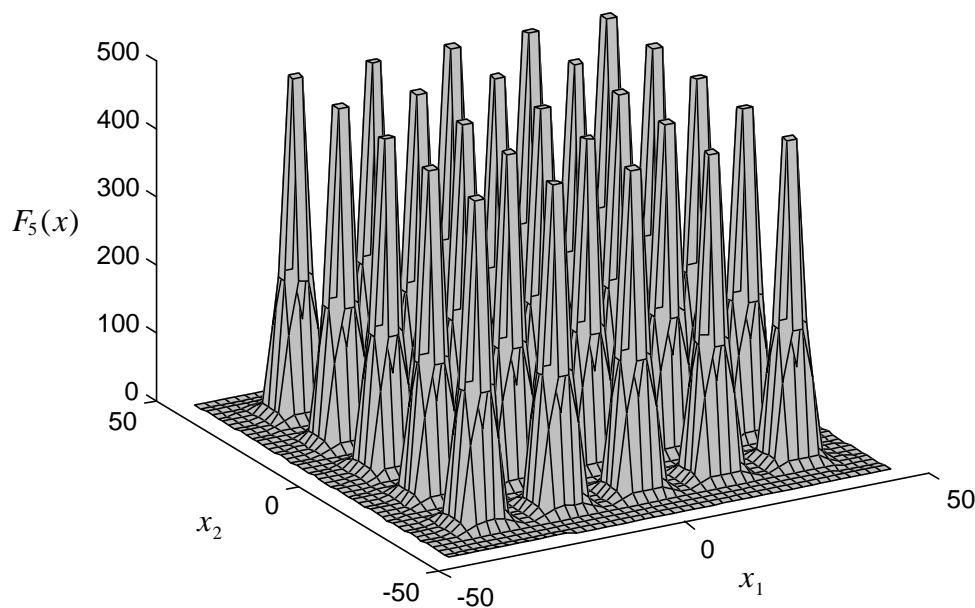
$$f_5(x) = \frac{1}{0.002 + \sum_{i=0}^{24} \frac{1}{1+i + (x_1 - X(i))^6 + (x_2 - Y(i))^6}} \quad (\text{AII.9})$$

avec $-65.536 \leq x_i \leq 65.536$

où $X(i) = [16(i \bmod 5) - 2]$ et $Y(i) = 16[\text{int}(i/5) - 2]$, $a \bmod b$ et $\text{int}(a/b)$ désignant respectivement le reste de la division de a par b et la partie entière de a/b .

Cette fonction de deux variables possède 25 minima de valeur quasi-identique situés aux coordonnées $(32j, 32k)$, $j = -1, 0, 1$ et $k = -1, 0, 1$. Le minimum global, centré sur le point de coordonnées $(-32, -32)$, vaut approximativement 0.998004. Les paramètres x_i sont habituellement codés sur 17 bits. Le problème équivalent, formulé en terme de maximisation, s'écrit :

$$\max F_5(x) = 500 - f_5(x) \quad (\text{AII.10})$$

Figure AII.5. Représentation de $F_5(x)$

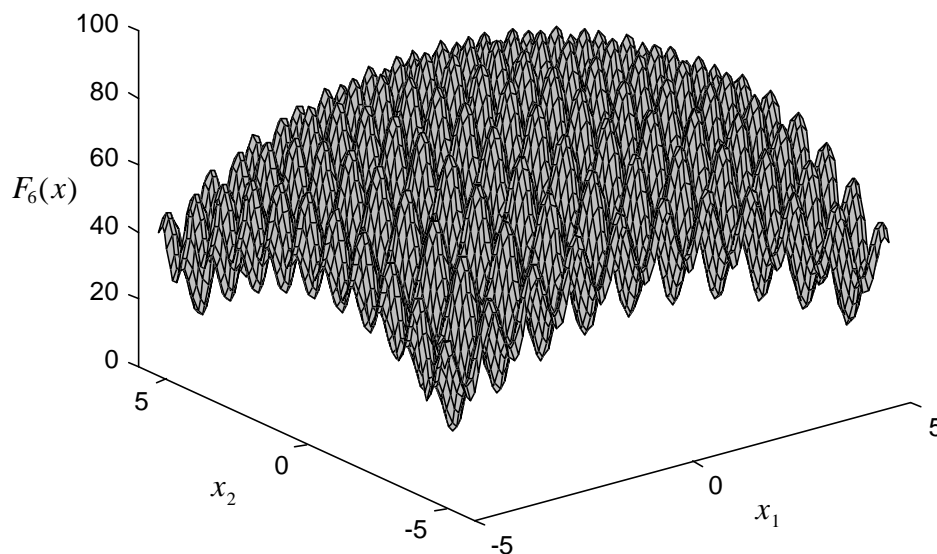
AII.1.6. Fonction multimodale de Rastrigin

$$f_5(x) = 10n + \sum_{i=1}^n x_i^2 - 10 \cos(2\pi x_i) \quad (\text{AII.11})$$

$$\text{avec } -5.12 \leq x_i \leq 5.12$$

Cette fonction possède 10^n minima locaux. Le minimum global, localisé en $x = (0,0,\dots,0)$, vaut $f_6(x) = 0$. Les paramètres x_i sont habituellement codés sur 10 bits. La dimension adoptée lors des tests est généralement $n=10$. Le problème équivalent, formulé en terme de maximisation, s'écrit :

$$\max F_6(x) = 50n - f_6(x) \quad (\text{AII.12})$$

Figure AII.6. Représentation de $F_6(x)$ pour $n=2$

Le Tableau AII.1 résume les caractéristiques des six fonctions utilisées pour évaluer la performance d'un algorithme génétique unimodal.

Tableau AII.1. Fonctions tests pour algorithme génétique unimodal (n est la dimension du problème, l la longueur de codage conseillée pour chaque gène)

fonction	nom	n	l	paramètres	caractéristiques
f_1	sphère	3	10	$-5.12 \leq x_i \leq 5.12$	unimodale, continue, convexe
f_2	Rosenbrock	2	12	$-2.048 \leq x_i \leq 2.048$	concave
f_3	palier	5	10	$-5.12 \leq x_i \leq 5.12$	discontinue
f_4	bicarrée bruitée	30	8	$-1.28 \leq x_i \leq 1.28$	bruitée, dimension élevée
f_5	trous de renard	2	17	$-65.536 \leq x_i \leq 65.536$	multimodale
f_6	Rastrigin	10	10	$-5.12 \leq x_i \leq 5.12$	multimodale, dimension élevée

AII.2. FONCTION TESTS POUR ALGORITHME GENETIQUE MULTIMODAL

Parmi les problèmes tests multimodaux les plus populaires présentés dans [Mahfoud 95], nous avons retenu les six fonctions suivantes :

AII.2.1. Fonction sinus

$$M_3(x) = \sin^6(5\pi[x^{3/4} - 0.05]) \quad (\text{AII.13})$$

Cette fonction définie dans l'intervalle $[0,1]$ est constituée de cinq sommets de même hauteur irrégulièrement espacés. Les maxima de valeur 1 sont situés aux abscisses $x = 0.080, 0.247, 0.451, 0.681, 0.934$. Le paramètre x est généralement codé sur 30 bits.

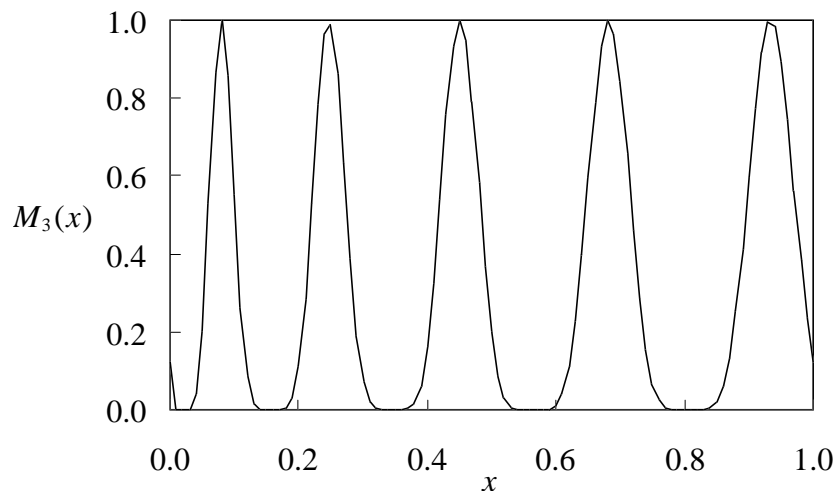
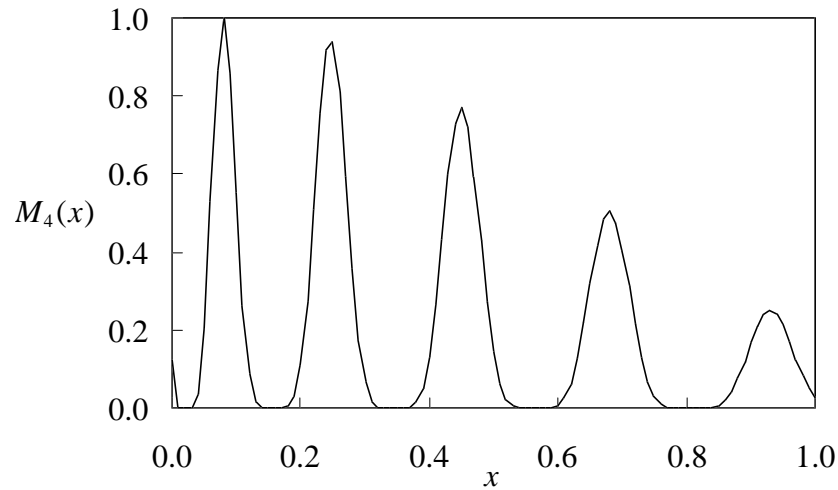


Figure AII.7. Représentation de $M_3(x)$

AII.2.2. Fonction sinus amorti

$$M_4(x) = e^{-2(\ln 2)\left(\frac{x-0.08}{0.854}\right)^2} \sin^6(5\pi[x^{3/4} - 0.05]) \quad (\text{AII.14})$$

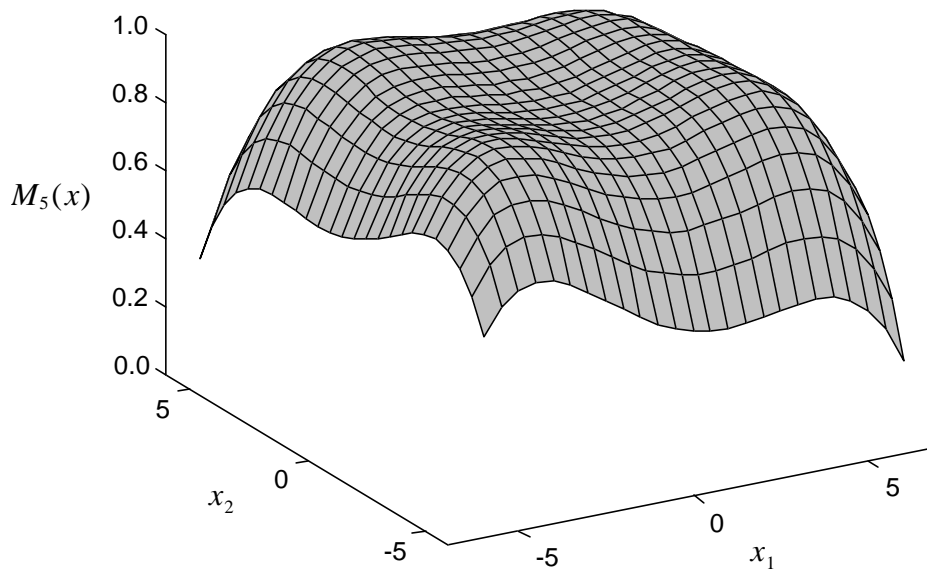
Cette fonction définie dans l'intervalle $[0,1]$ est constituée de cinq sommets de hauteur variable irrégulièrement espacés. Les maxima situés aux abscisses $x = 0.080, 0.247, 0.451, 0.681, 0.934$ ont pour hauteur respective 1.000, 0.948, 0.770, 0.503, 0.250. Le paramètre x est généralement codé sur 30 bits.

Figure AII.8. Représentation de $M_4(x)$

AII.2.3. Fonction de Himmelbau

$$M_5(x) = \frac{2186 - (x_1^2 + x_2 - 11)^2 - (x_1 + x_2^2 - 7)^2}{2186} \quad (\text{AII.15})$$

M_5 est une fonction de deux variables définies dans l'intervalle $[-6,6]$. Elle est constituée de quatre sommets de hauteur 1 situés aux coordonnées suivantes $(3,2)$, $(-3.584,-1.848)$, $(-3.779,-3.283)$, $(-2.779,3.131)$. Les paramètres x_1 et x_2 sont généralement codés sur 15 bits.

Figure AII.9. Représentation de $M_5(x)$

AII.2.4. Fonction "trous de renard" de Shekel

$$M_6(x) = F_5(x) \tag{AII.16}$$

M_6 est la fonction de *Shekel* modifiée présentée précédemment (voir paragraphe AII.1.5). Elle possède 25 sommets de différente hauteur variant entre 476.191 et 499.002. Les optima sont situés aux coordonnées $(32j, 32k)$, $j = -1,0,1$ et $k = -1,0,1$. Le maximum global est centré sur le point de coordonnées $(-32,-32)$.

AII.2.5. Fonction massivement multimodale et trompeuse

$$M_7(x_0, x_1, \dots, x_{29}) = \sum_{i=0}^4 u \left(\sum_{j=0}^5 x_{6i+j} \right) \tag{AII.17}$$

où $\forall k \in [0,29], x_k \in \{0,1\}$

M_7 est la somme de 5 sous-fonctions bimodales trompeuses définies à la Figure AII.10. Elle possède 32 optima globaux de hauteur 5 et plusieurs millions de maxima locaux variant entre 3.203 et 4.641. Les optima globaux correspondent aux séquences binaires de 30 bits constituées de groupe de six "0" ou de six "1" adjacents tels que :

```
00000000000000000000000000000000
00000000000000000000000000000111111
0000000000000000000000000111111000000
0000000000000000000000000111111111111
etc...
```

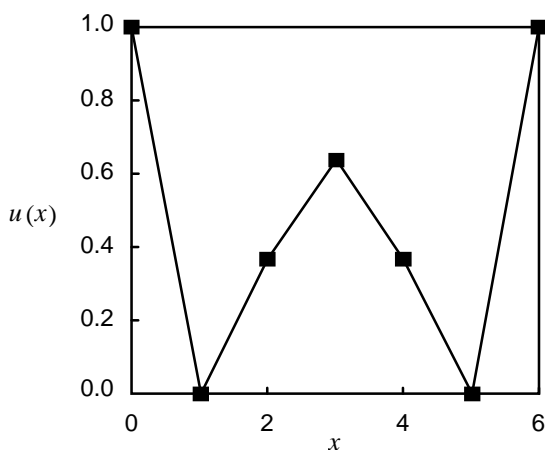


Figure AII.10. Fonction bimodale trompeuse

Le Tableau AII.2 résume les caractéristiques des cinq fonctions employées pour évaluer la performance des algorithmes génétiques multimodaux.

Tableau AII.2. Fonctions tests pour algorithme génétique multimodal (n est la dimension du problème, l la longueur de codage conseillée pour chaque gène)

<i>fonction</i>	<i>nom</i>	<i>n</i>	<i>l</i>	<i>paramètres</i>	<i>caractéristiques</i>
M_3	<i>sinus</i>	1	30	$0 \leq x_i \leq 1$	5 optima globaux
M_4	<i>sinus amorti</i>	1	30	$0 \leq x_i \leq 1$	1 optimum global 4 optima locaux
M_5	<i>Himmelbau</i>	2	15	$-6 \leq x_i \leq 6$	4 optima globaux
M_6	<i>trous de renard</i>	2	17	$-65.536 \leq x_i \leq 65.536$	1 optimum global 24 optima locaux
M_7	<i>massivement multimodale</i>	30	1	binaire 0-1	32 optima globaux, 5 153 600 optima locaux

ANNEXE III

EFFICACITE DES OPERATEURS GENETIQUES

Dans cette partie, nous caractérisons l'efficacité et la sensibilité des différents opérateurs de l'algorithme génétique standard à partir des fonctions mathématiques de test définies à l'Annexe II. Nous nous intéressons en particulier à l'influence du type de codage, de sélection et de recombinaison sur la performance de l'algorithme.

AIII.1. SIMULATIONS NUMERIQUES

Les caractéristiques de l'algorithme génétique utilisé lors des différents tests sont les suivantes :

Taille de la population : $N = 30$ ¹

Probabilité de croisement : $p_c = 1.0$

Probabilité de mutation : $p_m = 0.001$

Elitisme standard (préservation du meilleur individu)

Dans chaque cas étudié, l'algorithme est exécuté pendant au maximum 1000 générations et arrêté en cas de convergence à 10^{-3} près par rapport à l'optimum réel. 100 tests sont effectués pour tenir compte de l'aspect stochastique de la recherche. A l'issue des tests, quatre grandeurs sont présentées:

- Le *taux de réussite* ou taux de succès qui correspond au nombre de fois où l'algorithme a convergé.
- Le *nombre d'évaluation de fonctions* ou nombre d'appels à la fonction en cas de convergence (la moyenne est calculée uniquement à partir des tests ayant convergé ; l'écart-type correspondant est précisé dans chaque tableau entre parenthèse)
- La fonction objectif de la *solution optimale* moyennée sur l'ensemble des 100 tests (y compris ceux où l'algorithme n'a pas convergé). L'écart-type correspondant est aussi précisé entre parenthèse.
- La fonction objectif de la *meilleure solution trouvée* au cours des 100 tests.

¹ La faible taille de population a été choisie de façon à se placer dans des conditions de risque maximal concernant la convergence prématurée. L'ordre de grandeur est semblable à celui imposé lorsque l'évaluation de la fonction à optimiser est coûteux en temps CPU.

AIII.2. INFLUENCE DU TYPE DE CODAGE

Les performances de l'algorithme génétique obtenues avec le code Gray et le code binaire sont comparées dans les cas suivants :

- longueur de chromosome courte ($l=4$ bits) associée à un codage dynamique (*dyn*)
- longueur de chromosome habituelle (valeurs définies par *De Jong*) associée à un codage standard discret.
- longueur de chromosome habituelle associée à un codage dynamique (*dyn*).

L'algorithme génétique employé utilise une sélection de type proportionnelle (roulette classique : RWS) et une procédure de croisement à un site.

Les Tableaux AIII.1 à AIII.6 présentent les résultats obtenus pour chaque fonction testée.

Le codage binaire s'est révélé plus efficace que le code Gray pour la plupart des fonctions tests. A l'opposé du code Gray qui augmente l'exploration locale, le codage binaire semble favoriser la recherche globale. Ceci expliquerait les faibles taux de réussite et la lenteur de convergence obtenus avec le code Gray (en particulier pour la fonction F_3). Pour une taille de population réduite, il est impératif d'avoir une exploration globale forte pour éviter la convergence prématurée vers un optimum local.

Contrairement à ce qui est indiqué dans [Schraudolph 92], la réduction de la longueur des chromosomes au profit du codage dynamique n'améliore pas nécessairement la qualité des solutions (voir Tableau AIII.1, Tableau AIII.4 et Tableau AIII.6). Le codage dynamique est très intéressant car il permet d'explorer l'espace complet des paramètres et augmente le pouvoir de recherche local de l'algorithme génétique. Néanmoins, il n'est pas judicieux de réduire la longueur des chromosomes et de laisser l'exploration de l'espace au seul opérateur de codage dynamique, car celui-ci se comporte comme une méthode de recherche aléatoire. On perd dans ce cas, toute l'efficacité apportée par les mécanismes génétiques. Par ailleurs, lorsque la taille des chromosomes est limitée, les génotypes les plus proches de l'optimum peuvent être faiblement adaptés. La sélection risque alors de ralentir la convergence en orientant la recherche vers des configurations éloignées de l'optimum.

Tableau AIII.1. Influence du type de codage - Fonction F_1

	Binaire			Gray		
	$l=4$ dyn	$l=10$ std	$l=10$ dyn	$l=4$ dyn	$l=10$ std	$l=10$ dyn
<i>Taux de réussite</i>	45%	100%	100%	60%	100%	100%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	3223 (1916)	2970 (1362)	2778 (1302)	6359 (7400)	5515 (2345)	6162 (2960)
<i>Solution optimale (écart-type)</i>	0.08665 (0.09127)	0.00037 (0.00029)	0.00053 (0.00030)	0.07512 (0.17353)	0.00048 (0.00032)	0.00073 (0.00027)
<i>Meilleure solution trouvée</i>	0.00007	0.00008	0.00004	0.00002	0.00008	0.000003

Tableau AIII.2. Influence du type de codage - Fonction F_2

	Binaire			Gray		
	$l=4$ dyn	$l=12$ std	$l=12$ dyn	$l=4$ dyn	$l=12$ std	$l=12$ dyn
<i>Taux de réussite</i>	23%	9%	11%	31%	4%	9%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	1900 (2477)	6617 (10845)	622 (361)	3104 (150)	1290 (1056)	2350 (3019)
<i>Solution optimale (écart-type)</i>	0.07467 (0.22675)	0.14611 (0.20898)	0.15980 (0.33989)	0.04217 (0.06661)	0.15831 (0.19126)	0.17389 (0.36330)
<i>Meilleure solution trouvée</i>	0.00003	0.00006	0.00008	0.00003	0.00016	0.00006

Tableau AIII.3. Influence du type de codage - Fonction F_3

	Binaire			Gray		
	$l=4$ dyn	$l=10$ std	$l=10$ dyn	$l=4$ dyn	$l=10$ std	$l=10$ dyn
<i>Taux de réussite</i>	95%	93%	93%	3%	0%	0%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	12386 (7563)	12455 (7250)	12016 (7273)	14300 (2637)	- -	- -
<i>Solution optimale (écart-type)</i>	0.05000 (0.21904)	0.07000 (0.25643)	0.07000 (0.25643)	4.81000 (2.25941)	5.20000 (1.78659)	5.03000 (1.59833)
<i>Meilleure solution trouvée</i>	0	0	0	0	1	1

Tableau AIII.4. Influence du type de codage - Fonction F_4

	<i>Binaire</i>			<i>Gray</i>		
	<i>l=4 dyn</i>	<i>l=8 std</i>	<i>l=8 dyn</i>	<i>l=4 dyn</i>	<i>l=8 std</i>	<i>l=8 dyn</i>
<i>Taux de réussite</i>	0%	0%	0%	0%	0%	0%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	- -	- -	- -	- -	- -	- -
<i>Solution optimale (écart-type)</i>	0.27802 (0.11145)	0.04022 (0.01681)	0.04563 (0.02107)	9.07334 (6.04083)	0.10518 (0.07057)	0.16252 (0.13628)
<i>Meilleure solution trouvée</i>	0.09809	0.00980	0.01007	0.64122	0.03078	0.03744

Tableau AIII.5. Influence du type de codage - Fonction F_5

	<i>Binaire</i>			<i>Gray</i>		
	<i>l=4 dyn</i>	<i>l=17 std</i>	<i>l=17 dyn</i>	<i>l=4 dyn</i>	<i>l=17 std</i>	<i>l=17 dyn</i>
<i>Taux de réussite</i>	27%	28%	22%	10%	26%	22%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	2682 (3893)	2291 (5052)	1968 (1543)	648 (352)	4823 (7780)	5671 (7647)
<i>Solution optimale (écart-type)</i>	4.44508 (4.12723)	4.63592 (4.80217)	5.12425 (5.17216)	7.31439 (5.22669)	2.97117 (2.28827)	2.97117 (2.28827)
<i>Meilleure solution trouvée</i>	0.99802	0.99800	0.99802	0.99800	0.99800	0.99800

Tableau AIII.6. Influence du type de codage - Fonction F_6

	<i>Binaire</i>			<i>Gray</i>		
	<i>l=4 dyn</i>	<i>l=10 std</i>	<i>l=10 dyn</i>	<i>l=4 dyn</i>	<i>l=10 std</i>	<i>l=10 dyn</i>
<i>Taux de réussite</i>	0%	0%	0%	0%	0%	0%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	- -	- -	- -	- -	- -	- -
<i>Solution optimale (écart-type)</i>	28.5907 (7.25459)	17.4852 (5.93134)	18.0385 (6.47940)	34.5373 (8.39591)	17.9267 (4.46379)	17.84706 (6.98434)
<i>Meilleure solution trouvée</i>	10.22717	2.84375	6.66268	20.41106	5.00066	6.98434

AIII.3. INFLUENCE DU MODE DE SÉLECTION

Les performances de l'algorithme génétique sont évaluées en fonction des modes de sélection suivants :

- la sélection proportionnelle standard (RWS)²
- la sélection à reste stochastique (SRS)
- la sélection par tournoi de deux individus (TS)
- la sélection stochastique universelle (SUS)²
- la sélection linéaire par rapport au rang (LRS)³
- la sélection uniforme par rapport au rang (URS)⁴

L'algorithme génétique utilise un codage binaire dynamique et une procédure de croisement à un site. La longueur des chromosomes est celle recommandée par *De Jong*.

Les Tableaux AIII.7 à AIII.12 présentent les résultats obtenus pour les différentes fonctions testées.

Il n'y a pas de différence flagrante entre les divers modes de sélection. Les méthodes de sélection dynamiques ont toutefois surpassé les méthodes de sélection statiques. Lorsque les sélection de type proportionnelle sont utilisées avec une procédure de réajustement de l'objectif, elles permettent à l'algorithme génétique de converger rapidement vers l'optimum sans dégrader le taux de réussite. La sélection stochastique universelle et la sélection à reste stochastiques ont été les plus efficaces. Ces conclusions sont en accord avec celles tirées des travaux de Baker [Baker 87].

² Une procédure de réajustement de la fonction est effectuée avant la sélection, le facteur de mise à l'échelle étant fixé à 2.

³ L'intensité de la sélection est caractérisée par $\eta=2$.

⁴ La sélection est effectuée à partir des neuf premiers individus ($\mu=9$).

Tableau AIII.7. Influence du mode de sélection - Fonction F_1

	SUS	SRS	RWS	TS	LRS	URS
<i>Taux de réussite</i>	100%	100%	100%	100%	100%	100%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	2008 (1178)	2356 (1145)	2778 (1302)	2700 (1517)	2651 (1329)	2642 (1468)
<i>Solution optimale (écart-type)</i>	0.00051 (0.00032)	0.00051 (0.00029)	0.00053 (0.00030)	0.00044 (0.00031)	0.00046 (0.00030)	0.00047 (0.00033)
<i>Meilleure solution trouvée</i>	0.00001	0.00001	0.00004	0.00002	0.00002	0.00000

Tableau AIII.8. Influence du mode de sélection - Fonction F_2

	SUS	SRS	RWS	TS	LRS	URS
<i>Taux de réussite</i>	13%	18%	11%	1%	6%	3%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	2961 (7530)	2265 (5034)	622 (361)	11310 -	1235 (1451)	1910 (1022)
<i>Solution optimale (écart-type)</i>	0.08159 (0.17793)	0.05210 (0.08735)	0.15980 (0.33989)	0.36844 (0.67777)	0.19946 (0.30956)	0.30946 (0.50884)
<i>Meilleure solution trouvée</i>	0.00001	0.00017	0.00008	0.00069	0.00027	0.00030

Tableau AIII.9. Influence du mode de sélection - Fonction F_3

	SUS	SRS	RWS	TS	LRS	URS
<i>Taux de réussite</i>	100%	100%	93%	73%	76%	62%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	5851 (3549)	7051 (4000)	12016 (7273)	14479 (6911)	13172 (8044)	14249 (7162)
<i>Solution optimale (écart-type)</i>	0 (0)	0 (0)	0.07000 (0.25643)	0.28000 (0.47312)	0.24000 (0.42923)	0.45000 (0.62563)
<i>Meilleure solution trouvée</i>	0	0	0	0	0	0

Tableau AIII.10. Influence du mode de sélection - Fonction F_4

	SUS	SRS	RWS	TS	LRS	URS
<i>Taux de réussite</i>	0%	0%	0%	0%	0%	0%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	- -	- -	- -	- -	- -	- -
<i>Solution optimale (écart-type)</i>	0.02631 (0.00980)	0.02980 (0.01307)	0.04563 (0.02107)	0.04696 (0.02014)	0.03350 (0.01420)	0.04197 (0.01879)
<i>Meilleure solution trouvée</i>	0.00955	0.00885	0.01007	0.01237	0.01126	0.00832

Tableau AIII.11. Influence du mode de sélection - Fonction F_5

	SUS	SRS	RWS	TS	LRS	URS
<i>Taux de réussite</i>	38%	30%	22%	24%	28%	24%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	1122 (870)	1177 (1017)	1968 (1543)	1806 (1178)	2324 (3347)	2263 (2826)
<i>Solution optimale (écart-type)</i>	3.36820 (3.76748)	4.49289 (4.75167)	5.12425 (5.17216)	5.11254 (5.42066)	4.49038 (5.04902)	5.96827 (5.70368)
<i>Meilleure solution trouvée</i>	0.99800	0.99800	0.99802	0.99801	0.99800	0.99800

Tableau AIII.12. Influence du mode de sélection - Fonction F_6

	SUS	SRS	RWS	TS	LRS	URS
<i>Taux de réussite</i>	0%	0%	0%	0%	0%	0%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	- -	- -	- -	- -	- -	- -
<i>Solution optimale (écart-type)</i>	14.6121 (5.46355)	16.7690 (5.20092)	18.0385 (6.47940)	20.8167 (7.32702)	18.7544 (6.68277)	19.6815 (7.23276)
<i>Meilleure solution trouvée</i>	2.02292	5.77291	6.66268	4.17984	4.22981	6.37216

AIII.4. INFLUENCE DU MODE DE RECOMBINAISON

Les performances de l'algorithme génétique sont évaluées en fonction des modes de recombinaison suivants :

- le croisement standard à un site.
- le croisement à deux sites.
- le croisement uniforme paramétré de probabilité d'échange des gènes $p_0 = 0.5$ (croisement uniforme standard, désigné par la suite par unif 0.5), $p_0 = 0.1$ (unif 0.1) et $p_0 = 0.01$ (unif 0.01).

L'algorithme génétique utilise un codage binaire dynamique et une procédure de sélection stochastique universelle. La longueur des chromosomes est celle recommandée par *De Jong*.

Les Tableaux AIII.13 à AIII.18 présentent les résultats obtenus pour les différentes fonctions testées.

Il n'y a pas de différence majeure entre les divers modes de recombinaison si l'on excepte les résultats médiocres du croisement uniforme lorsque la probabilité d'échange des allèles est faible. Dans ce cas, l'opérateur de recombinaison ne permet pas de produire de solutions suffisamment distinctes avec la faible taille de population utilisée. L'exploration globale de l'espace est insuffisante, ce qui entraîne une convergence lente vers l'optimum. Les procédures de croisement à 1 et 2 sites, ainsi que la technique de croisement uniforme standard, donnent des résultats globalement équivalents.

Tableau AIII.13. Influence du mode de recombinaison - Fonction F_1

	1 site	2 sites	unif 0.5	unif 0.1	unif 0.01
<i>Taux de réussite</i>	100%	100%	100%	100%	100%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	2008 (1178)	1752 (885)	1460 (856)	2436 (1456)	3850 (1510)
<i>Solution optimale (écart-type)</i>	0.00051 (0.00032)	0.00051 (0.00032)	0.00057 (0.00032)	0.00056 (0.00033)	0.00051 (0.00033)
<i>Meilleure solution trouvée</i>	0.00001	0.00003	0.00002	0.00000	0.00002

Tableau AIII.14. Influence du mode de recombinaison - Fonction F_2

	1 site	2 sites	unif 0.5	unif 0.1	unif 0.01
<i>Taux de réussite</i>	13%	23%	44%	19%	5%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	2961 (7530)	1195 (1223)	2320 (4873)	711 (398)	2364 (1541)
<i>Solution optimale (écart-type)</i>	0.08159 (0.17793)	0.04405 (0.14341)	0.01384 (0.02836)	0.04847 (0.11661)	0.043692 (0.95728)
<i>Meilleure solution trouvée</i>	0.00001	0.00008	0.00002	0.00001	0.00046

Tableau AIII.15. Influence du mode de recombinaison - Fonction F_3

	1 site	2 sites	unif 0.5	unif 0.1	unif 0.01
<i>Taux de réussite</i>	100%	99%	100%	100%	100%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	5851 (3549)	3879 (3215)	1022 (953)	4798 (2663)	12320 (5041)
<i>Solution optimale (écart-type)</i>	0 (0)	0.01 (0.1)	0 (0)	0 (0)	0 (0)
<i>Meilleure solution trouvée</i>	0	0	0	0	0

Tableau AIII.16. Influence du mode de recombinaison - Fonction F_4

	1 site	2 sites	unif 0.5	unif 0.1	unif 0.01
<i>Taux de réussite</i>	0%	0%	0%	0%	0%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	- -	- -	- -	- -	- -
<i>Solution optimale (écart-type)</i>	0.02631 (0.00980)	0.02420 (0.01056)	0.01892 (0.00715)	0.02559 (0.01124)	0.04416 (0.02028)
<i>Meilleure solution trouvée</i>	0.00955	0.00748	0.00449	0.00750	0.00952

Tableau AIII.17. Influence du mode de recombinaison - Fonction F_5

	1 site	2 sites	unif 0.5	unif 0.1	unif 0.01
<i>Taux de réussite</i>	38%	38%	26%	15%	12%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	1122 (870)	1328 (2397)	843 (1506)	2810 (3315)	2583 (1737)
<i>Solution optimale (écart-type)</i>	3.36820 (3.76748)	3.15708 (3.76326)	2.37706 (2.27600)	5.75303 (5.31836)	6.42204 (5.81758)
<i>Meilleure solution trouvée</i>	0.99800	0.99800	0.99800	0.99801	0.99800

Tableau AIII.18. Influence du mode de recombinaison - Fonction F_6

	1 site	2 sites	unif 0.5	unif 0.1	unif 0.01
<i>Taux de réussite</i>	0%	0%	0%	0%	0%
<i>Nombre d'évaluations de fonctions (écart-type)</i>	- -	- -	- -	- -	- -
<i>Solution optimale (écart-type)</i>	14.6121 (5.46355)	13.2020 (6.13202)	15.0549 (4.96895)	19.5901 (6.55825)	22.2886 (7.26028)
<i>Meilleure solution trouvée</i>	2.02292	3.38329	4.49136	6.28435	8.48129

AIII.5. CONCLUSIONS

Pour avoir une robustesse satisfaisante, nous recommandons surtout l'utilisation des opérateurs suivant :

- codage binaire standard (sur une longueur fonction de la précision souhaitée) associé à un décodage dynamique pour l'exploration de l'ensemble de l'espace.
- sélection proportionnelle de type SUS ou SRS avec réajustement préalable de l'objectif.
- croisement uniforme ou croisement à un ou deux sites.
- élitisme classique (préservation du meilleur individu)

ANNEXE IV

COEFFICIENTS DE POTENTIEL ET DE CHAMP

Nous présentons ici les expressions des coefficients de potentiel et de champ pour des distributions de charge linéiques de forme simple. Nous définissons ces coefficients en fonction de la charge totale q_i des distributions données (excepté en 2D plan, pour le cas d'une ligne chargée infiniment longue où q_i désigne une densité linéique de charge). Le coefficient de potentiel d'une distribution quelconque est défini à partir du potentiel V au point considéré :

$$p = \frac{V}{q_i} \quad (\text{AIV.1})$$

Le vecteur \mathbf{C} des coefficients de champ est obtenu par dérivation des coefficients de potentiel :

$$\mathbf{C} = -\nabla p \quad (\text{AIV.2})$$

ce qui donne pour le champ électrique $\mathbf{E} = \mathbf{C}q_i$

AIV.1. COEFFICIENTS DE POTENTIEL ET DE CHAMP POUR UNE CHARGE PONCTUELLE

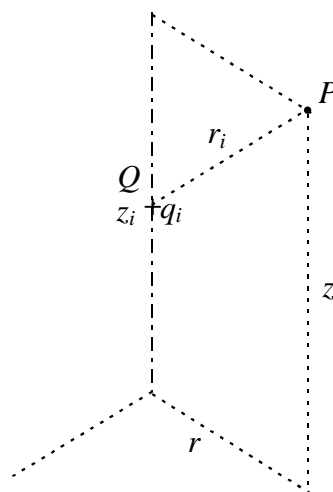


Figure AIV.1. Charge ponctuelle

Annexes IV

Le coefficient de potentiel d'une charge ponctuelle est :

$$p = \frac{1}{4\pi\epsilon_0} \frac{1}{r_i} \quad (\text{AIV.3})$$

avec $r_i = \sqrt{r^2 + (z - z_i)^2}$.

Les coefficients de champ s'expriment alors selon (AIV.4) et (AIV.5) :

$$C_r = -\frac{\partial p}{\partial r} = \frac{1}{4\pi\epsilon_0} \frac{r}{r_i^3} \quad (\text{AIV.4})$$

$$C_z = -\frac{\partial p}{\partial z} = \frac{1}{4\pi\epsilon_0} \frac{z - z_i}{r_i^3} \quad (\text{AIV.5})$$

AIV.2. COEFFICIENTS DE POTENTIEL ET DE CHAMP POUR UN SEGMENT UNIFORMEMENT CHARGÉ

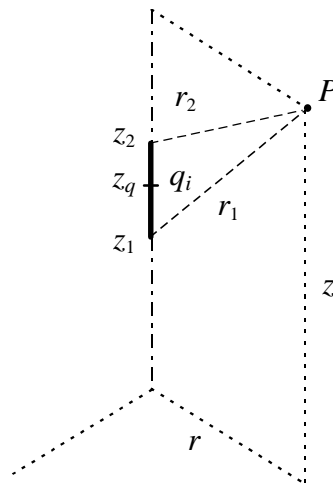


Figure AIV.2. Segment uniformément chargé

Le coefficient de potentiel d'un segment uniformément chargé de longueur $\Delta z = z_2 - z_1$ s'écrit :

$$p = \frac{1}{4\pi\epsilon_0} \frac{1}{\Delta z} \ln \left(\frac{z_2 - z + r_2}{z_1 - z + r_1} \right) \quad (\text{AIV.6})$$

avec $r_1 = \sqrt{(z_1 - z)^2 + r^2}$ et $r_2 = \sqrt{(z_2 - z)^2 + r^2}$.

Les coefficients de champ sont donnés par :

$$C_r = -\frac{\partial p}{\partial r} = \frac{1}{4\pi\epsilon_0} \frac{1}{\Delta z} \left(\frac{z_2 - z}{r_2} - \frac{z_1 - z}{r_1} \right) \quad (\text{AIV.7})$$

$$C_z = -\frac{\partial p}{\partial z} = \frac{1}{4\pi\epsilon_0} \frac{1}{\Delta z} \left(\frac{1}{r_2} - \frac{1}{r_1} \right) \quad (\text{AIV.8})$$

AIV.3. COEFFICIENTS DE POTENTIEL ET DE CHAMP POUR UN ANNEAU UNIFORMEMENT CHARGÉ

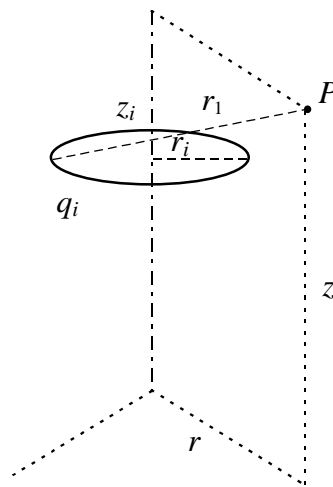


Figure AIV.3. Anneau uniformément chargé

Les coefficients de champ et de potentiel s'expriment en fonction des intégrales de Legendre de première et de deuxième espèce :

$$K(m) = \int_0^{\pi/2} \frac{d\varphi}{\sqrt{1 - m^2 \sin^2 \varphi}} \quad \text{Intégrale de Legendre de première espèce} \quad (\text{AIV.9})$$

$$E(m) = \int_0^{\pi/2} \sqrt{1 - m^2 \sin^2 \varphi} d\varphi \quad \text{Intégrale de Legendre de deuxième espèce} \quad (\text{AIV.10})$$

avec $m = \frac{2\sqrt{r_i r}}{r_1}$, $r_1 = \sqrt{(r + r_i)^2 + (z - z_i)^2}$ et $r_i = \sqrt{r^2 + (z - z_i)^2}$.

Le coefficient de potentiel vaut :

$$p = \frac{r_i}{\pi\epsilon_0 r_1} \frac{1}{2\pi r_i} K(m) \quad (\text{AIV.11})$$

Les coefficients de champ sont exprimés par :

$$C_r = -\frac{\partial p}{\partial r} = \frac{r_i}{2\pi\epsilon_0 r_1 r} \frac{1}{2\pi r_i} \left(K(m) - \frac{r_i^2 - r^2 + (z - z_i)^2}{r_1^2} \frac{E(m)}{1 - m^2} \right) \quad (\text{AIV.12})$$

$$C_z = -\frac{\partial p}{\partial z} = \frac{(z - z_i) r_i}{\pi\epsilon_0 r_1^3} \frac{1}{2\pi r_i} \frac{E(m)}{1 - m^2} \quad (\text{AIV.13})$$

AIV.4. COEFFICIENTS DE POTENTIEL ET DE CHAMP POUR UNE LIGNE DE CHARGE INFINIE

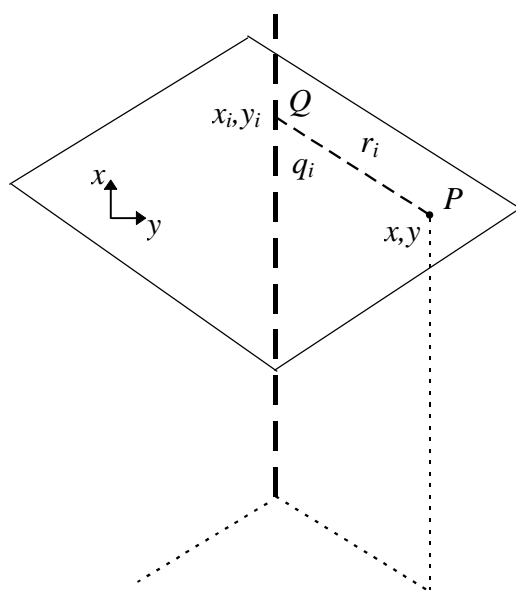


Figure AIV.4. Ligne de charge infinie

Le coefficient de potentiel est défini par la relation suivante :

$$p = -\frac{1}{2\pi\epsilon_0} \ln r_i \quad (\text{AIV.14})$$

avec $r_i = \sqrt{(x - x_i)^2 + (y - y_i)^2}$.

Les coefficients de champ correspondants sont :

$$C_x = -\frac{\partial p}{\partial x} = \frac{1}{2\pi\epsilon_0} \frac{x - x_i}{r_i^2} \quad (\text{AIV.15})$$

$$C_y = -\frac{\partial p}{\partial y} = \frac{1}{2\pi\epsilon_0} \frac{y - y_i}{r_i^2} \quad (\text{AIV.16})$$

LA METHODE DES CHARGES EQUIVALENTES

AV.1. PRINCIPE

La méthode des charges équivalentes ou méthode de simulation de charges (*Charge Simulation Method*) consiste à simuler les répartitions de champ et de potentiel dans des objets diélectriques ou conducteurs à partir d'un ensemble de charges fictives [Singer 74] [Abdel-Salam 87][Guillen 92][Błaszczuk 94]. Le principe de la méthode est illustré par la Figure AV.1. Le domaine physique à caractériser se compose de plusieurs régions. Une région homogène est constituée par un seul et unique diélectrique ou conducteur. Une série de charges fictives entoure chaque région séparément. Le champ et le potentiel à l'intérieur d'une région sont calculés par superposition en prenant en compte les contributions des charges associées à la région considérée.

Les charges représentant une région diélectrique homogène ne sont pas placées à l'intérieur de celle-ci mais à l'extérieur et à une certaine distance des frontières. Les régions pour lesquelles le champ électrique est déjà connu (les conducteurs par exemple) ne requièrent aucune charge.

Il est à noter que les charges discrètes utilisées pour la description des régions n'ont pas de signification physique. La charge "réelle" des objets doit être calculée à partir du flux traversant leurs frontières.

La mise en œuvre de la méthode nécessite par ailleurs le choix d'un ensemble de points de référence situés sur les frontières des régions. Ces derniers appelés points de collocation sont des points vérifiant des conditions aux limites de type Dirichlet ou flottante, ou des conditions d'interface.

La méthode des charges équivalentes réside dans la détermination des valeurs des charges de façon à ce que le potentiel et le champ électrique résultant de leur distribution vérifient les conditions aux limites fixées et les conditions d'interface pour tous les points de collocation.

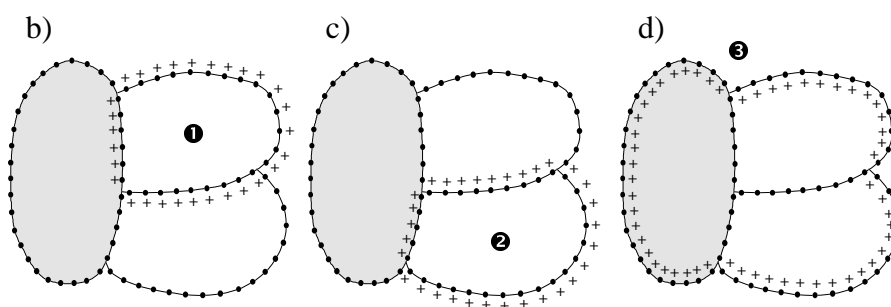
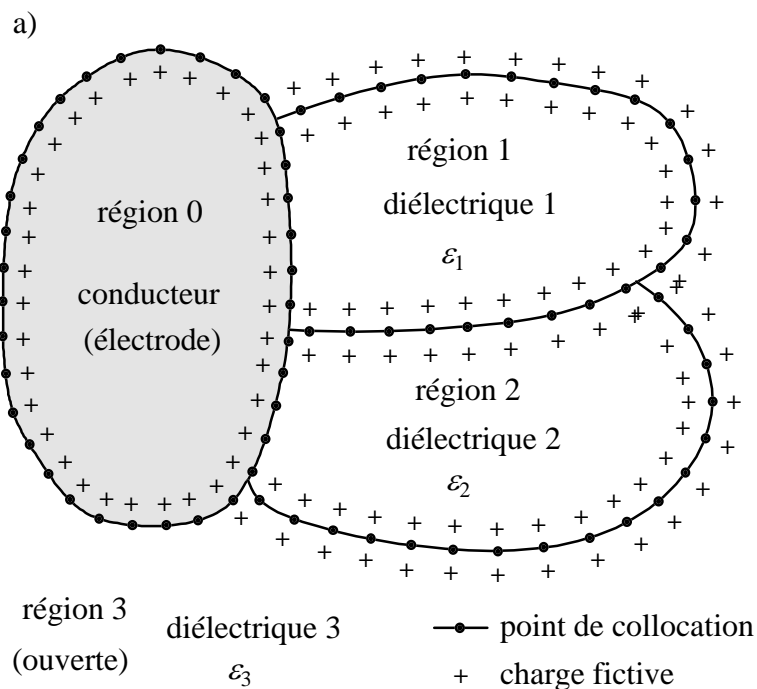


Figure AV.1. Principe de la méthode des charges équivalentes

- (a) Charges simulant les grandeurs électriques dans l'ensemble du domaine.
- (b) Charges utilisées pour le calcul des grandeurs électriques dans la région 1.
- (c) Charges utilisées pour le calcul des grandeurs électriques dans la région 2.
- (d) Charges utilisées pour le calcul des grandeurs électriques dans la région 3.

AV.2. FORMULATION

Les équations aux points de collocation sont formulées en fonction du type de condition de frontière. Nous distinguons habituellement quatre types de conditions possibles pour les différents points de collocation:

AV.2.1. Points de collocation à potentiel imposé de valeur V_0

Ce sont les points de référence situés aux frontières des conducteurs parfaits et notamment sur les surfaces des électrodes. Pour un point i situé à la frontière d'une région diélectrique M_1 , nous avons:

$$\sum_{k \in M_1} p_{ik} \cdot q_k = V_0 \quad (\text{AV.1})$$

où la somme est effectuée sur l'ensemble des charges simulant la région M_1 . Nous rappelons que ces charges ne sont pas situées à l'intérieur de M_1 mais qu'elles se trouvent dans les régions diélectriques M_j ($j \neq 1$) et les conducteurs C_k en contact avec M_1 au niveau des interfaces M_1/C_k et M_1/M_j .

AV.2.2. Points de collocation à potentiel flottant inconnu V_F

De la même manière, pour un point i de potentiel flottant inconnu situé à la frontière d'une région diélectrique M_1 , nous avons:

$$\sum_{k \in M_1} p_{ik} \cdot q_k - V_F = 0 \quad (\text{AV.2})$$

AV.2.3. Points de collocation à densité de flux imposée ψ_0

La relation caractérisant un point i de densité de flux imposée situé à la frontière d'une région diélectrique M_1 est la suivante:

$$-\varepsilon_0 \varepsilon_1 \sum_{k \in M_1} \bar{n} \times \nabla p_{ik} \cdot q_k = \psi_0 \quad (\text{AV.3})$$

où ε_1 désigne la permittivité relative du milieu M_1 et \bar{n} représente le vecteur normal à la frontière au point de collocation.

AV.2.4. Points de collocation à l'interface de deux milieux diélectriques

Un point de collocation situé à la l'interface des milieux diélectriques M_1 et M_2 doit vérifier les deux équations suivantes:

$$\sum_{k \in M_1} p_{ik} \cdot q_k - \sum_{k \in M_2} p_{ik} \cdot q_k = 0 \quad (\text{AV.4})$$

$$-\varepsilon_1 \sum_{k \in M_1} \bar{n} \times \nabla p_{ik} \cdot q_k + \varepsilon_2 \sum_{k \in M_2} \bar{n} \times \nabla p_{ik} \cdot q_k = \rho_s \quad (\text{AV.5})$$

où ρ_s est la densité de charge surfacique à l'interface entre les milieux M_1 et M_2 , ε_1 et ε_2 désignant respectivement les permittivités des milieux M_1 et M_2 .

La prise en compte des conditions de Neuman homogène ($\partial V/\partial n = 0$) et de Dirichlet homogène ($V=0$) est réalisée implicitement par la méthode des images en ajoutant aux charges des régions considérées les contributions de charges fictives symétriques de même valeur et respectivement de valeur opposée.

Une équation supplémentaire caractérise les conducteurs à potentiel flottant. En effet, la charge d'un conducteur F à potentiel flottant a une valeur Q_F qui peut être obtenue par intégration du flux sur toute la surface du conducteur:

$$Q_F = \varepsilon_0 \sum_{i \in F} S_i \cdot \varepsilon_i \sum_{k \in M_i} \bar{n} \times \nabla p_{ik} \cdot q_k \quad (\text{AV.6})$$

où S_i désigne la surface d'un élément de frontière de point de collocation i , M_i le milieu diélectrique externe au conducteur en ce même point de collocation et ε_i la permittivité correspondante du milieu M_i .

AV.3. DISCRETISATION ET POSITIONNEMENT DES CHARGES

Les logiciels basés sur la méthode des charges équivalentes utilisent généralement quatre types de distributions de charges : des segments de droite, des lignes infinies et des anneaux chargés en 2D et des charges ponctuelles en 3D. La position des charges est déterminée en fonction des paramètres des éléments de frontières correspondants. La Figure AV.2 illustre le positionnement des charges pour des éléments 2D et 3D. Les charges sont placées dans la direction de la normale à l'élément de frontière à une distance d_i du centre de l'élément. Cette distance est fonction de la taille de l'élément (la longueur l_i en 2D et la surface S_i en 3D) et d'un facteur d'échelle f . Ce facteur est constant pour tous les éléments et doit prendre une valeur comprise dans l'intervalle [1.0 2.0].

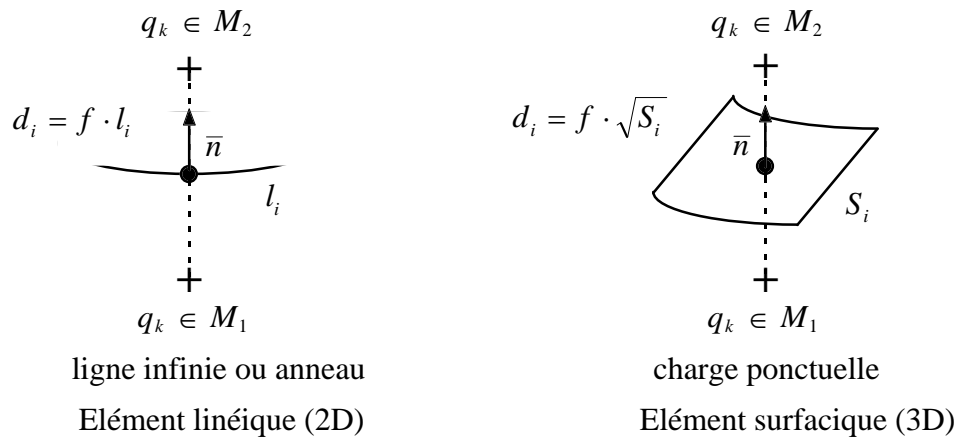


Figure AV.2. Positionnement des charges pour des éléments de frontière linéiques et surfaciques.

Une fois le positionnement des charges établi, la distance entre une charge et le centre son élément de frontière doit être inférieure aux distances entre cette même charge et les centres des autres éléments de frontière de la région à laquelle la charge appartient. Cette règle est toujours vérifiée pour les régions convexes. Pour des régions non convexes, la règle de distance peut être violée notamment dans les deux cas suivants. Le premier cas concerne les angles aigus. Pour éviter le chevauchement de charges discrètes au voisinage d'un angle aigu, la valeur du facteur d'échelle doit être réduite tant que la règle précédente n'est pas vérifiée (voir Figure AV.3). Le second cas concerne les couches minces diélectriques ou les barrières conductrices d'épaisseur quasi-nulle. Pour ces structures particulières, la règle peut être satisfaite en réduisant la taille des éléments de frontière. Néanmoins, comme cette solution conduit à un plus grand nombre d'inconnues, il est plutôt recommandé d'introduire des régions fictives [Blaszczyk 94].

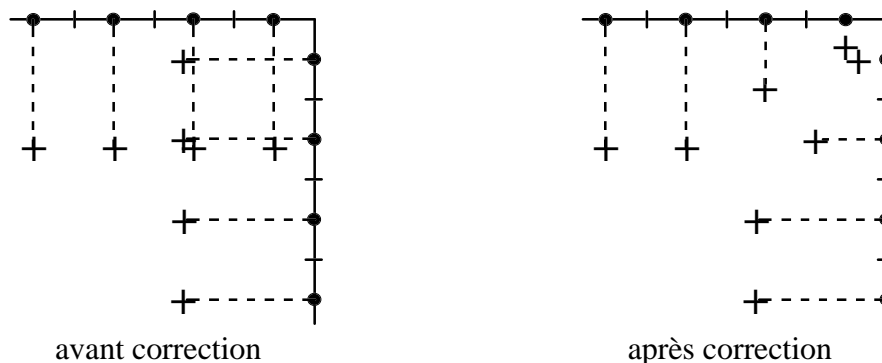


Figure AV.3. Correction de la position des charges pour une électrode à angle droit

AV.4. MISE EN EQUATION MATRICIELLE ET RESOLUTION

La discrétisation conduit à placer n_{kj} charges à l'intérieur d'un conducteur C_k à la frontière d'une région diélectrique M_j et n'_{ij} charges à l'intérieur d'une région diélectrique M_i à l'interface d'un milieu diélectrique M_j . Le nombre total de charges internes aux régions conductrices est :

$$N_c = \sum_k \sum_j n_{kj} \quad (\text{AV.7})$$

où les indices k et j désignent respectivement les indentificateurs des régions conductrices C_k et des régions diélectriques M_j . Ce nombre est évidemment identique au nombre total de points de collocation sur les frontières des conducteurs.

De même, le nombre total de charges internes aux régions diélectriques est :

$$N'_d = \sum_i \sum_j n'_{ij} \quad (i \neq j) \quad (\text{AV.8})$$

Comme il y a autant de charges de part et d'autre d'une interface ($n'_{ij} = n'_{ji}$), la relation précédente peut être réécrite en fonction du nombre total de points de collocation N_d à l'interface des milieux diélectriques :

$$N'_d = 2 \sum_i \sum_j n'_{ij} = 2N_d \quad (i > j) \quad (\text{AV.9})$$

Nous vérifions ici qu'il y a bien deux charges associées à un point de collocation situé sur une interface diélectrique. Il y a autant d'équations que d'inconnues car une seule équation est nécessaire pour caractériser un point de collocation à la frontière d'un conducteur (l'équation (AV.1) ou (AV.3)) mais deux équations de passage doivent être vérifiées à l'interface de deux milieux diélectriques (les équations (AV.4) et (AV.5)). La méthode des charges équivalentes conduit à un système linéaire de $N_c + 2N_d$ équations à $N_c + 2N_d$ inconnues de la forme :

$$[A][Q_k] = [B] \quad (\text{AV.10})$$

auxquels il est éventuellement nécessaire d'ajouter les équations relatives aux potentiels flottants. Par exemple, la forme de la matrice $[A]$ correspondant au problème de la Figure AV.1 est décrite à la Figure AV.4. Cette matrice contient de nombreux termes nuls. En réordonnant les équations et les indentificateurs des régions, ces termes peuvent être concentrés le long de la diagonale de la matrice. Le système matriciel est par la suite résolu de façon plus efficace au moyen de solveurs basés sur la méthode d'élimination de Gauss [Blaszczyk 94].

charges fictives (q_k) Point de collocation i		interne au conducteur C_0			interne au milieu M_1		interne au milieu M_2		interne au milieu M_3	
		n_{01}	n_{02}	n_{03}	n'_{12}	n'_{13}	n'_{21}	n'_{23}	n'_{32}	n'_{31}
Frontière C_0/M_1	Equation (AV.1)	-	0	0	0	0	-	0	0	-
Frontière C_0/M_2	Equation (AV.1)	0	-	0	-	0	0	0	-	0
Frontière C_0/M_3	Equation (AV.1)	0	0	-	0	-	0	-	0	0
Interface M_1/M_2	Equation (AV.4)	-	-	0	-	0	-	0	-	-
	Equation (AV.5)	-	-	0	-	0	-	0	-	-
Interface M_1/M_3	Equation (AV.4)	-	0	-	0	-	-	-	0	-
	Equation (AV.5)	-	0	-	0	-	-	-	0	-
Interface M_2/M_3	Equation (AV.4)	0	-	-	-	-	0	-	-	0
	Equation (AV.5)	0	-	-	-	-	0	-	-	0

n_{0j} = charges appartenant à l'électrode C_0 simulant les grandeurs physiques dans la région M_j

n'_{ij} = charges appartenant à la région M_i simulant les grandeurs physiques dans la région M_j

- = termes non nuls de la matrice

0 = termes nuls de la matrice

Figure AV.4. Exemple de forme de matrice obtenue par la méthode des charges équivalentes.

BIBLIOGRAPHIE

- [Aarts 90] E.H.L. Aarts, J. Korst, - *Simulated Annealing and Boltzmann Machines* - John Wiley & Sons, 1990.
- [Abdel-Salam 87] M. Abdel-Salam, E.K. Stanek, "Optimizing Field Stress on High-Voltage Insulators", *IEEE Trans. on Electrical Insulation*, Vol. 22, No. 1, pp. 47-56, 1987.
- [Alotto 96] P. Alotto, A.V. Kuntsevitch, Ch. Magele, G. Molinari, C. Paul, K. Preis, M. Repetto, K.R. Richter, "Multiobjective Optimization in Magnetostatics: A Proposal for Benchmark Problems", *IEEE Trans. on Magnetics*, Vol. 32, No. 3, 1238-1241, 1996.
- [Bäck 91] Th. Bäck, F. Hoffmeister, "Extended selection mechanisms in genetic algorithms", R.K. Belew and L.B. Booker (Eds.), *Proceedings of the 4th International Conference on Genetic Algorithms and Their Applications*, University of California, San Diego, Morgan Kauffmann Publishers, pp. 92-99, 1991.
- [Bäck 92a] Th. Bäck, "The interaction of mutation rate, selection and self-adaptation within a genetic algorithm", R. Männer and B. Manderick (Eds.), *Parallel Problem Solving from Nature 2 - PPSN II*, Amsterdam, Elsevier, pp. 85-94, 1992.
- [Bäck 92b] Th. Bäck, "Self-adaptation in genetic algorithms", F.J. Varela and P. Bourguine (Eds.), *Proceedings of the First European Conference of Artificial Life*, Cambridge, MA, The MIT Press, pp. 263-271, 1992.
- [Bäck 93] Th. Bäck, "Optimal mutation rates in genetic search", S. Forrest (Ed.), *Proceedings of the 5th International Conference on Genetic Algorithms and Their Applications*, Morgan Kauffmann Publishers, pp. 2-9, 1993.
- [Bäck 96] Th. Bäck, - *Evolutionary Algorithms in Theory and Practice* -, Oxford University Press, N.Y., 1996
- [Baker 85] J.E. Baker, "Adaptative selection methods for genetic algorithms", J.J. Grefenstette (Ed.), *Proceedings of the 1st International Conference on Genetic Algorithms and Their Applications*, Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp. 101-111, 1985.

Bibliographie

- [Baker 87] J.E. Baker, "Reducing bias and inefficiency in the selection algorithm", J.J. Grefenstette (Ed.), *Proceedings of the 2nd International Conference on Genetic Algorithms and Their Applications*, Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp. 14-21, 1987.
- [Battiti 94] R. Battiti, G. Tecchiolli, "The Reactive Tabu Search", *ORSA Journal of Computing*, Vol. 6, No. 2, pp. 126-140, 1994.
- [Beasley 93] D. Beasley, D.R. Bull, R. Martin, "A sequential niche technique for multimodal function optimization", *Evolutionary Computation*, Vol. 1, No. 2, pp 101-125, 1993.
- [Bertoni 92] A. Bertoni, M. Dorigo, - *Implicit parallelism in genetic algorithms* -, Technical Report 93-012, Politecnico di Milano, 1992.
- [Blaszczyk 94] A. Blaszczyk, H. Steinbigler, "Region Oriented Charge Simulation", *IEEE Trans. on Magnetics*, Vol. 30, No. 5, pp. 2924-2927, 1994.
- [Bucci 90] O.M. Bucci, G. D'Elia, G. Leone, R. Pierri, "Far-Field Pattern Determination from the Near-Field Amplitude on Two Surfaces", *IEEE Trans. Antennas Propagat.*, Vol. 38, No. 11, pp. 1772-1779, 1990.
- [Caruana 88] R.A. Caruana, J.D. Schaffer, "Representation and hidden bias: Gray vs. binary coding for genetic algorithms", J. Laird (Ed.), *Proceedings of the 5th International Conference on Machine Learning*, San Mateo, CA, Morgan Kauffmann Publishers, pp. 153-161, 1988.
- [Cavicchio 70] D.J. Cavicchio, - *Adaptative search using simulated evolution* -, Doctoral Dissertation, University Microfilms, No. 25.0199, 1970
- [Chat-uthai 96] C. Chat-uthai, J.A. Ramirez, E.M. Freeman, "An Improved Constrained Quasi-Newton Method for the Solution of Inverse Electromagnetic Problems", *IEEE Trans. on Magnetics*, Vol. 32, No. 3, pp. 1318-1321, 1996.
- [Chiampi 96] M. Chiampi, C. Ragusa, M. Repetto, "Fuzzy Approach for Multiobjective Optimization in Magnetics", *IEEE Trans. on Magnetics*, Vol. 32, No. 3, pp. 1234-1237, 1996.
- [Chun 98] J.S. Chun, H.K. Jung, S.Y. Hahn, "A Study on Comparison of Optimization Performances between Immune Algorithm and other Heuristic Algorithms", *IEEE Trans. on Magnetics*, Vol. 34, No. 5, pp. 2972-2975, 1998.

- [Coulomb 81] J.L. Coulomb, - *Analyse tridimensionnelle des champs électriques et magnétiques par la méthode des éléments finis* - Thèse de Docteur Es-Sciences Physiques, Institut National Polytechnique de Grenoble, France, 1981.
- [Darwen 95] P. Darwen, X. Yao, "A dilemma for fitness sharing with a scaling function", *IEEE International Conference on Evolutionary Computing (ICEC'95)*, IEEE Press, pp. 166-171, 1995.
- [Darwin 59] Ch. Darwin, - *The origin of species by means of natural selection* -, 1859.
- [Davidor 91] Y. Davidor, "A naturally occurring niche and species phenomenon: the model and first results", *Proceedings of the 4th International Conference on Genetic Algorithms*, Morgan Kaufmann, pp. 257-263, 1991.
- [Davis 85] L. Davis, "Applying adaptative algorithms to epistatic domains", *Proceedings of the 9th International Joint Conference on Artificial Intelligence*, pp. 162-164, 1985.
- [Deb 89] K. Deb, D.E. Goldberg, "An investigation of niche and species formation in genetic function optimization", *Proceedings of the 3rd International Conference on Genetic Algorithms*, Morgan Kaufmann, , pp. 42-50, 1989.
- [De Jong 75] K.A. De Jong, - *An Analysis of the behavior of a class of genetic adaptative systems* -, PhD thesis, University of Michigan, *Dissertation Abstracts International*, 36(10), 5140B, University Microfilms No. 76-9381, 1975.
- [De Jong 92] K.A. De Jong, W.M. Spears, "A formal analysis of the role of multi-point crossover in genetic algorithms", *Annals of Mathematics and Artificial Intelligence Journal*, Vol. 5, No. 1, pp. 1-26, 1992.
- [Driankov 93] D. Driankov, H. Hellendoorn, M. Reinfrank, - *An introduction to fuzzy logic control* -, Springer Verlag, 1993.
- [Durand 66] E. Durand, *Electrostatique*, Tome I, II, III, Masson, Paris, 1966.
- [Dyck 96] D.N. Dyck, D.A. Lowther, "Automated Design of Magnetic Devices by Optimizing Material Distribution", *IEEE Trans. on Magnetics*, Vol 32, No. 3, pp. 1188-1193, 1996.
- [Ebner 98] Th. Ebner, Ch. Magele, B.R. Brandstätter, K.R. Richter, "Utilizing Feed Forward Neural Networks for Acceleration of Global Optimization Procedures", *IEEE Trans. on Magnetics*, Vol. 34, No. 5, pp. 2928-2931, 1998.

Bibliographie

- [Flux2D] Flux2D : Analyse des dispositifs électriques, magnétiques et thermiques par la méthode des éléments finis, Cedrat (CEE), Magsoft (USA), 1994.
- [Fogarty 89] T.C. Fogarty, "Varying the probability of mutation in the genetic algorithm", J.D. Schaffer (Ed.), *Proceedings of the 3rd International Conference on Genetic Algorithms and Their Applications*, San Mateo, CA, Morgan Kaufmann Publishers, pp. 104-109, 1989.
- [Formisano 96] A. Formisano, R. Martone, F. Villone, "A Lagrangian Approach to Shape Inverse Electromagnetic Problems", *IEEE Trans. on Magnetics*, Vol 32, No. 3, pp. 1314-1317, 1996.
- [Gizolme 97] O. Gizolme, - *Modélisation et optimisation d'une machine synchrone et de son alimentation pour la traction électrique* - Thèse de Doctorat, Ecole Centrale de Lyon, France, 1997.
- [Goldberg 87] D.E. Goldberg, J. Richardson, "Genetic algorithms with sharing for multimodal function optimization", *Proceedings of the 2nd International Conference on Genetic Algorithms*, pp. 41-49, 1987.
- [Goldberg 89] D.E. Goldberg, - *Genetic Algorithms in Search Optimization & Machine Learning* -, Addison Wesley, 1989.
- [Goldberg 91] D.E. Goldberg, "Construction of high-order deceptive functions using low-order Walsh coefficients", *Annals of Mathematics and Artificial Intelligence*, Vol. 5, pp. 35-48, 1991.
- [Goldberg 92] D.E. Goldberg, K. Deb, J. Horn, "Massive multimodality, deception, and genetic algorithms", R. Männer and B. Manderick (Eds.), *Parallel Problem Solving from Nature 2 - PPSN II*, Amsterdam, Elsevier, pp. 37-46, 1992.
- [Goldberg 93] D.E. Goldberg, K. Deb, D. Thierens, - *Toward a better understanding of mixing in genetic algorithms* - Technical Report No. 93004, Urbana: University of Illinois at Urbana-Champaign, Illinois Genetic Algorithms Laboratory, 1993.
- [Gottvald 92] A. Gottvald, K. Preis, Ch. Magele, O. Biro, A. Savini, "Global Optimization Methods for Computational Electromagnetics", *IEEE Trans. on Magnetics*, Vol. 28, No. 2, 1537-1540, 1992.
- [Guarnieri 90] M. Guarnieri, A. Stella, F. Trevisan, "A methodological analysis of different formulations for solving inverse electromagnetic problems", *IEEE Trans. on Magnetics*, Vol. 26, No. 2, pp. 622-625, 1990.

- [Guillen 93] M. Guillen, - *Réalisation et validation d'un modèle hybride basé sur la méthode des équations intégrales de frontière et la méthode des charges équivalentes pour le calcul de champs électriques tridimensionnels* -, Thèse de Doctorat, Ecole Centrale de Lyon, France, 1993.
- [Hesser 91] J. Hesser, R. Männer, "Towards an optimal mutation probability in genetic algorithms", H.P. Schwefel and R. Männer (Eds.), *Parallel Problem Solving from Nature - PPSN I*, Volume 496 of *Lecture Notes in Computer Science*, Berlin, Springer Verlag, pp. 23-32, 1991.
- [Holland 75] J. Holland, - *Adaptation in Natural and Artificial Systems* -, The Univ. of Michigan Press, Ann Arbor, 1975.
- [Hoole 91] S.R.H. Hoole, S. Subramaniam, R. Saldanha, J.L. Coulomb, J.C. Sabonnadière, "Inverse Problem Methodology and Finite Elements in the Identification of Cracks, Sources, Materials, and their Geometry in Inaccessible locations", *IEEE Trans. on Magnetics*, Vol. 27, No. 3, pp. 3433-3443, 1991.
- [Horn 93] J. Horn, N. Nafpliotis, - *Multiobjective Optimization Using The Niche Pareto Genetic Algorithm* -, Technical Report No. 93005, Urbana: University of Illinois at Urbana-Champaign, Illinois Genetic Algorithms Laboratory, 1993.
- [Ingber 92] L. Ingber, B. Rosen, "Genetic algorithms and very fast simulated reannealing : a comparison", *Mathematical and Computer Modelling*, Vol. 16, No. 11, pp. 87-100, 1992.
- [Kallel 97] L. Kallel, M. Schoenauer, "Alternative Random Initialization in Genetic Algorithms", *Proceedings of the 7th International Conference on Genetic Algorithms*, Morgan Kaufmann, pp. 268-275, 1997.
- [Kargupta 92] H. Kargupta, K. Deb, D.E. Goldberg, "Ordering genetic algorithms and deception", R. Männer and B. Manderick (Eds.), *Parallel Problem Solving from Nature 2 - PPSN II*, Amsterdam, Elsevier, pp. 47-56, 1992.
- [Kirkpatrick 83] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, "Optimization by Simulated Annealing", *Science*, Vol. 220, No. 4598, pp. 671-680, 1983.
- [Krähenbühl 83] L. Krähenbühl,- *La méthode des équation intégrales de frontière pour la résolution des problèmes de potentiel en électrotechnique, et sa formulation axisymétrique* - Thèse de Docteur-Ingénieur, Ecole Centrale de Lyon, France, 1983.

Bibliographie

- [Leschi 94] D. Leschi, N. Burais, J.Y. Gaspard, "Inductors Modelling and Optimization in Cooking Induction Heating Systems", *2nd International Workshop on Electric and Magnetic Fields*, Leuven, Belgique, 1994.
- [Metropolis 53] N. Metropolis, A.W. Rosenbluth, M.W. Rosenbluth, A.H. Teller, E. Teller, "Equation of state calculations by fast computing machines", *J. Chem. Phys.*, Vol. 21, pp. 1087-1092, 1953.
- [Mahfoud 95] S.W. Mahfoud, - *Niching methods for genetic algorithms* - , Doctoral dissertation/Technical report No. 95001, University of Illinois at Urbana Champaign, Illinois Genetic Algorithm Laboratory, 1995.
- [Messine 98] F. Messine, B. Nogarède, J.L. Lagounaelle, "Optimal Design of Electro-Mechanical Actuators: A New Method Based on Global Optimization", *IEEE Trans. on Magnetics*, Vol. 34, No. 1, pp. 299-308, 1998.
- [Michalewicz 96] Z. Michalewicz, D. Dasgupta, R. Leriche, M. Schoenauer, "Evolutionary Algorithms for Constrained Engineering Problems", *Computers & Industrial Engineering Journal*, Vol. 30, No. 2, 1996
- [Miller 95a] B.L. Miller, D.E. Goldberg, - *Genetic algorithms, tournament selection and the effect of noise-*, Technical Report No. 95006, Urbana: University of Illinois at Urbana-Champaign, Illinois Genetic Algorithms Laboratory, 1995.
- [Miller 95b] B.L. Miller, M.J. Shaw, - *Genetic algoritms with dynamic niche sharing for multimodal function optimization* -, Technical report No. 95010, University of Illinois at Urbana Champaign, Illinois Genetic Algorithm Laboratory, 1995.
- [Mühlenbein 91] H. Mühlenbein, "Evolution in time and space - The parallel genetic algorithm", J.G.E. Rawlings (Ed), *Foundations of Genetic Algorithms*, San Mateo, CA, Morgan Kaufmann Publishers, pp. 316-337, 1991.
- [Mühlenbein 92] H. Mühlenbein, "How genetic algorithm really work: I. mutation and hillclimbing", R. Männer and B. Manderick (Eds.), *Parrallel Problem Solving from Nature 2 - PPSN II*, Amsterdam, Elsevier, pp. 15-25, 1992.
- [Mühlenbein 93] H. Mühlenbein, H. Schlierkamp-Voosen, "Predictive models for the breeder genetic algorithm: I. Continuous parameter optimization", *Evolutionary computation*, Vol. 1, No. 1, pp. 25-49, 1993.

- [Mühlenbein 94] H. Mühlenbein, - *The optimal population size for uniform crossover and truncation selection* -, Technical Report GMD-AS-TR-94-11, GMD, St Augustin, Germany, 1994.
- [Mukherjee 96] P.K. Mukherjee, C. Trinitis, H. Steinbigler, "Optimization of HV Electrode Systems by Neural Networks using a New Learning Method", *IEEE Trans. on Dielectrics and Electrical Insulation*, Vol. 3, No. 6, pp.737-742, 1996.
- [Nicolas 83] A. Nicolas, - *Application de la méthode des équations intégrales de frontière à la modélisation des phénomènes d'induction* - Thèse d'état, Lyon, France, 1983.
- [Oei 91] C.K. Oei, D.E. Goldberg, S. Chang, - *Tournament selection, niching and the preservation of diversity* -, Technical report No. 91011, University of Illinois at Urbana Champaign, Illinois Genetic Algorithm Laboratory, 1991.
- [Petrowski 96] A. Petrowski, "A clearing procedure as a niching method for genetic algorithms", *Proceedings of the IEEE International Conference on Evolutionary Computation*, pp. 798-803, 1996.
- [Petrowski 97a] A. Petrowski, "A new selection operator dedicated to selection", *Proceedings of the 7th International Conference on Genetic Algorithms*, Morgan Kaufmann, San Francisco, CA, 1997.
- [Petrowski 97b] A. Petrowski, "An Efficient Hierarchical Clustering Technique for Speciation", *Proceeding of the 3rd Conference on Artificial Evolution*, AE'97, Nimes, France, 1997.
- [Rao 96] S. S. Rao, - *Engineering Optimization, Theory and Practice* -, Wiley, 1996.
- [Russenschuck 94] S. Russenschuck, T. Tortschanoff, "Mathematical Optimization of Superconducting Accelerator Magnets", *IEEE Trans. on Magnetics*, Vol. 30, No. 5, pp. 3419-3422, 1994.
- [Sado 92] G. Sado, M.C. Sado, - *Les plans d'expériences* -, afnor technique, 1992.
- [Saludjian 97] L. Saludjian, - *Optimisations en Electrotechnique par Algorithmes Génétiques* -, Thèse de Doctorat, I.N.P.G., No.0073, 1997.
- [Sareni 96] B. Sareni, L. Krähenbühl, A. Beroual, C. Brosseau, "Effective Dielectric Constant of Periodic Composite Materials", *Journal of Applied Physics*, Vol. 80, No. 3, pp. 1688-1696, 1996.

Bibliographie

- [Sareni 98a] B. Sareni, L. Krähenbühl, "Fitness Sharing and Niching Methods Revisited", *IEEE Trans. on Evolutionary Computation*, Vol. 2, No. 3, pp. 97-106, 1998.
- [Sareni 98b] B. Sareni, L. Krähenbühl, A. Nicolas, "Niching Genetic Algorithms for Optimization in Electromagnetics, I. Fundamentals", *IEEE Trans. on Magnetics*, Vol. 34, No. 5, pp. 2984-2987, 1998.
- [Sareni 98c] B. Sareni, L. Krähenbühl, D. Muller, "Niching Genetic Algorithms for Optimization in Electromagnetics, II. Shape Optimization of Electrodes using the CSM", *IEEE Trans. on Magnetics*, Vol. 34, No. 5, pp. 2988-2991, 1998.
- [Schaffer 89] J.D. Schaffer, R.A. Caruana, L.J. Eshelman, R. Das, "A study of control parameters affecting online performance of genetic algorithms for function optimization", J.D. Schaffer (Ed.), *Proceedings of the 3rd International Conference on Genetic Algorithms and Their Applications*, San Mateo, CA, Morgan Kaufmann Publishers, pp. 51-60, 1989.
- [Schlierkamp 94] D. Schlierkamp-Voosen, H. Mühlenbein, "Strategy Adaptation by Competing Subpopulations", *Parallel Problem Solving from Nature 3 - PPSN III*, Jerusalem, Springer, pp. 199-208, 1994.
- [Schoenauer 93] M. Schoenauer, S. Xanthakis, "Constrained GA optimization", S. Forrest (Ed.), *Proceedings of the 5th International Conference on Genetic Algorithms and Their Applications*, Morgan Kauffmann Publishers, 1993.
- [Schraudolph 92] N.N. Schraudolph, R.K. Belew, - *Dynamic parameter encoding for genetic algorithms* - Technical Report CS 90-175, La Jolla, University of California, Computer Science and Engineering Department, 1992.
- [Schwefel 95] H.P. Schwefel, - *Evolution and Optimum Seeking* - Sixth-Generation Computer Technology Series, John Wiley & Sons, N.Y., 1996.
- [Simkin 92] J. Simkin, C.W. Trowbridge, "Optimizing electromagnetic devices combining direct search methods and simulated annealing", *IEEE Trans. on Magnetics*, Vol 28, No. 2, 1545-1548, 1992.
- [Singer 74] H. Singer, H. Steinbigler, P. Weiss, "A Charge Simulation Method for the Calculation of High Voltage Fields", *IEEE Trans. on PAS*, Vol. 93, No. 5, pp. 1660-1668, 1974.

- [Smith 93a] R.E. Smith, S. Forrest, A.S. Perelson, "Searching for diverse cooperative populations with genetic algorithms", *Evolutionary Computation*, Vol. 1, No. 2, pp. 127-149, 1993.
- [Smith 93b] R.E. Smith, S. Forrest, A. Perelson, "Population diversity in a immune system model: implications for genetic search", *Foundations of Genetic Algorithms 2*, San Mateo, pp. 153-165, 1993.
- [Spears 90] W.M. Spears, K.A. De Jong, "An analysis of multi-point crossover", G. Rawlings (Ed.), *Proceedings of the Foundations of Genetic Algorithms Workshop*, Morgan Kaufmann Publishers, pp. 221-241, 1990.
- [Spears 91] W.M. Spears, K.A. De Jong, "On the virtues of parametrized uniform crossover", R.K. Belew and L.B. Booker (Eds.), *Proceedings of the 4th International Conference on Genetic Algorithms and Their Applications*, University of California, San Diego, Morgan Kauffmann Publishers, pp. 230-236, 1991.
- [Spears 94] W.M. Spears, "Simple subpopulations schemes", *Proceedings of the 3rd Annual Conference on Evolutionary Programming*, World Scientific, pp. 296-307, 1994.
- [Spears 95] W.M. Spears, "Adaptating crossover in evolutionary algorithms", *Proceedings of the 5th Annual Conference on Evolutionary Programming*, San Diego, CA, Morgan Kaufmann Publishers, 1995.
- [Storn 96] R. Storn, K. Price, "Minimizing the real functions of the ICEC'96 contest by Differential Evolution", *International Conference on Evolutionary Computation*, Nagoya, Japan, 1996.
- [Syswerda 89] G. Syswerda, "Uniform crossover in genetic algorithms", J.D. Schaffer (Ed.), *Proceedings of the 3rd International Conference on Genetic Algorithms and Their Applications*, San Mateo, CA, Morgan Kaufmann Publishers, pp. 2-8, 1989.
- [Thollon 95] F. Thollon, - *Conception et optimisation des capteurs à courants de Foucault pour la détection de défauts profonds dans les matériaux amagnétiques* - Thèse de Doctorat, Ecole Centrale de Lyon, France, 1995.
- [Üler 95] G.F. Üler, O.A. Mohammed, C.S. Koh, "Design Optimization of Electrical Machines using Genetic Algorithms", *IEEE Tran. on Magnetics*, Vol. 31, No. 3, 1995.

Bibliographie

- [Vasconcelos 94] J.A. Vasconcelos, - *Optimisation de forme des structures électromagnétiques* - Thèse de Doctorat, No. 94-32, Ecole Centrale de Lyon, France, 1994.
- [Vasconcelos 97] J.A. Vasconcelos, R.R. Saldanha, L. Krähenbühl, A. Nicolas, "Genetic algorithm coupled with a deterministic method for optimization in electromagnetics", *IEEE Trans. on Magnetics*, Vol. 33, No. 2, pp. 1860-1863, 1997.
- [Yao 91] X. Yao, "Simulated Annealing with Extended Neighbourhood", *International Journal of Computer Mathematics*, Vol. 40, pp. 169-189, 1991.
- [Yin 93] X. Yin, N. Gerday, "A fast genetic algorithm with sharing scheme using cluster analysis methods in multimodal function optimization", *Proceedings of the International Conference on Artificial Neural Nets and Genetic Algorithms*, pp. 450-457, 1993.
- [Zadeh 65] L.A. Zadeh, "Fuzzy Sets", *Information and Control*, Vol. 8, pp. 338-353, 1965.

Titre :

Méthodes d'optimisation multimodales associées à la modélisation numérique en électromagnétisme.

Title :

Multimodal optimization methods related to numerical modeling in electromagnetics.

Résumé :

L'essor de l'informatique et des techniques d'intelligence artificielle a conduit ces dernières années à un développement sans précédent des procédés d'optimisation automatique qui peuvent aujourd'hui prendre en compte des dizaines de paramètres de conception. En particulier, les méthodes évolutionnistes ont connu depuis le début des années soixante une croissance exponentielle et s'affirment peu à peu comme les techniques les plus robustes : d'une part, elles permettent de localiser l'optimum d'une fonction dans l'espace des paramètres sans avoir recours aux dérivées de la fonction par rapport à ces paramètres ; d'autre part, elles ne se laissent pas piéger par un optimum local et réussissent le plus souvent à déterminer l'optimum global de la fonction considérée. Cependant, la traduction d'un problème d'optimisation réel avec tous ses aspects (performance "pure" mais aussi sensibilité, facilité de fabrication, prix de revient, ...) sous forme d'une fonction à optimiser n'est pas toujours une chose simple. Dès lors, le concepteur apprécie lorsqu'il étudie l'un des aspects d'être conduit à plusieurs possibilités (plus ou moins parfaites suivant cet aspect là) plutôt qu'à une solution unique. Les méthodes génétiques multimodales ou méthodes de nichage offrent des perspectives intéressantes en permettant la localisation de solutions optimales multiples, aussi bien locales que globales. Notre travail est centré sur la caractérisation de ces nouvelles techniques d'optimisation numériques. Chaque méthode a été testée de façon classique à partir de fonctions mathématiques ainsi que sur des problèmes d'électromagnétisme et sur un procédé très original de conception de formes optimales d'électrodes. Nous décrivons une nouvelle approche pour des systèmes 2D-plan ou axisymétriques, où la forme de l'électrode est identifiée à une ligne équipotentielle obtenue par optimisation du positionnement et de la valeur d'un certain nombre de charges fictives.

Mots clés :

Conception automatique optimale, algorithmes génétiques, méthodes de nichage, optimisation de forme, électrostatique, électrode, charges équivalentes.

Keywords :

Automated optimal design, genetic algorithms, niching methods, shape optimization, electrostatics, electrode, charge simulation method.

Direction de recherche

Laurent KRÄHENBÜHL, Directeur de recherche CNRS (*Laurent.Krahenbuhl@ec-lyon.fr*)

Alain NICOLAS, Professeur des Universités

Centre de Génie Electrique de Lyon (CEGELY) - UPRESA CNRS n° 5005

Ecole Centrale de Lyon - B.P. 163 - 69131 Ecully Cedex (France)