HAL
open science

Analyse spatiale et spectrale des motifs d'échantillonnage pour l'intégration Monte Carlo

Adrien Pilleboue

To cite this version:

Adrien Pilleboue. Analyse spatiale et spectrale des motifs d'échantillonnage pour l'intégration Monte Carlo. Traitement des images [eess.IV]. Université Claude Bernard - Lyon I, 2015. Français. NNT : 2015LYO10225 . tel-01286224v2

HAL Id: tel-01286224

https://theses.hal.science/tel-01286224v2

Submitted on 23 Nov 2017

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

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

Université Claude Bernard Lyon I
LIRIS

Thèse

présentée en première version en vu d'obtenir le grade de Docteur
spécialité « Informatique »

par
Adrien Pilleboue

Directeurs de thèse
Victor Ostromoukhov
David Coeurjolly

Analyse spatiale et spectrale des motifs d'échantillonnage pour l'intégration Monte Carlo

Thèse soutenue le 19 novembre 2015 devant le jury composé de :

M. Kadi Bouatouch – Université de Rennes, IRISA	Rapporteur
M. Mathias Paulin – Université Paul Sabatier de Toulouse, IRIT	Rapporteur
M. Frédo Durand – Massachusetts Institute of Technology, CSAIL	Examinateur
M. Nicolas Holzschuch – INRIA Grenoble Rhône-Alpes	Examinateur
M. Victor Ostromoukhov – Université de Lyon, LIRIS	Directeur
M. David Coeurjolly – CNRS, LIRIS	Directeur

الى رادا ...

Table des matières

Table des matières v

Liste des figures viii

Résumé 1

Abstract 3

1 Introduction 5
- 1.1 Critère de qualité en informatique graphique . . . 6
- 1.2 Contributions et plan de thèse . . . 7

2 Prérequis mathématiques 11
- 2.1 Notations . . . 11
- 2.2 Théorie de l'intégration . . . 11
  - 2.2.1 Topologie . . . 12
  - 2.2.2 Mesure . . . 12
  - 2.2.3 Intégrale au sens de Lebesgue . . . 13
  - 2.2.4 Intégrale d'un espace discret . . . 14
- 2.3 Fonction δ de Dirac . . . 15
- 2.4 Autres fonctions . . . 15
- 2.5 Notations en «grand O » . . . 16

3 Ensemble de points et espace d'échantillonnage 17
- 3.1 Ensemble de points . . . 17
- 3.2 Espace d'échantillonnage . . . 19
  - 3.2.1 Espace euclidien $\mathcal{E}^d$ . . . 22
  - 3.2.2 Espace hyper-cubique périodique $\mathcal{T}^d$ . . . 22
  - 3.2.3 Espace hexagonal periodique $\mathcal{H}^2$ . . . 23
  - 3.2.4 Espace sphérique $\mathcal{S}^2$ . . . 23
  - 3.2.5 Espace discret régulier $\mathcal{L}^d$ . . . 24
- 3.3 Espace d'échantillonnage restreint . . . 24
  - 3.3.1 Espace hyper-cubique unitaire $\mathcal{U}^d$ . . . 24
  - 3.3.2 Espace hémisphérique $\mathcal{HS}^2$ . . . 25
- 3.4 Intégration et espace d'échantillonnage . . . 25
- 3.5 Homogénéité statistique . . . 26
- Conclusion . . . 28

4 Méthodes d'échantillonnage 29
- 4.1 Méthodes d'échantillonnage bruit blanc . . . 29
- 4.2 Disques de Poisson . . . 30

4.3 Méthodes d'échantillonnage par pavage . . . 32
4.3.1 Grille régulière . . . 33
4.3.2 *Jittered sampling* . . . 34
4.3.3 Hypercube Latin . . . 35
4.4 Optimisation d'un ensemble de points . . . 36
4.4.1 Relaxation de Lloyd . . . 36
4.4.2 Optimisation de la capacité . . . 38
4.4.3 Optimisation de la distance minimale . . . 39
4.4.4 Contrôle spectral . . . 39
4.5 Échantillonnage par pavage hiérarchique . . . 40
4.6 Séquences à basse discrépance . . . 42
Conclusion . . . 43

5 Analyse spatiale **45**
5.1 Densité . . . 45
5.1.1 Densité de paires . . . 47
5.1.2 Densité de n-uplets . . . 48
5.2 Densité différentielle . . . 49
5.2.1 Fonction de distribution de distance . . . 50
5.3 Autres fonctions liées à la densité . . . 52
5.3.1 Fonction de corrélation totale . . . 52
5.3.2 Auto-corrélation . . . 53
5.3.3 Distribution de voisinage . . . 54
5.4 Diagramme de Voronoi . . . 54
5.5 Uniformité . . . 55
5.5.1 Test de Kolmogorov-Smirnov . . . 56
5.5.2 Discrépance . . . 56
5.5.3 Discrépance stochastique . . . 58
Conclusion . . . 60

6 Analyse spectrale **61**
6.1 Transformée de Fourier . . . 61
6.1.1 Transformée de Fourier et intégrales . . . 63
6.1.2 Spectre de puissance . . . 63
6.2 Généralisation de la transformée de Fourier . . . 63
6.2.1 Espace de Hilbert . . . 64
6.2.2 Décomposition d'un espace pré-hilbertien . . . 65
6.2.3 Similitude avec la transformée de Fourier . . . 66
6.2.4 Théorie des représentations . . . 68
6.3 Transformée de Fourier de l'espace d'échantillonnage . . . 71
6.3.1 Espace euclidien $\mathcal{E}^d$ . . . 72
6.3.2 Espace hyper-cubique periodique $\mathcal{T}^d$ . . . 72
6.3.3 Espace sphérique $\mathcal{S}^2$ . . . 73
6.4 Transformée de Fourier et ensemble de points . . . 73
6.4.1 Ensembles de points homogènes . . . 75
6.5 Lien avec l'analyse spatiale . . . 77
Conclusion . . . 77

7 Intégration Monte Carlo **79**

7.1 Définition formelle . . . 79
7.2 Erreur d'intégration . . . 80
7.3 Ensemble de points homogène et biais . . . 81
7.4 Lien avec la discrépance à l'origine . . . 82
7.4.1 Variation de Hardy-Krause . . . 83
7.4.2 Inégalité de Koksma-Hlawka multi-dimensionnelle . . . 84
7.5 Lien avec l'analyse spatiale . . . 85
7.6 Lien avec l'analyse spectrale . . . 86
7.7 Vitesse de convergence . . . 89
7.7.1 Spectre de puissance et densité dans l'espace euclidien . . . 90
7.7.2 Spectre de puissance et densité dans les espaces à mesure finie . . . 91
7.7.3 Borne supérieur de la variance . . . 92
7.7.4 Spectre de puissance de l'intégrande . . . 93
7.7.5 Vitesse de convergence par classe de fonctions . . . 95
Conclusion . . . 98

8 Analyse de l'état de l'art **99**
8.1 Méthodologie expérimentale . . . 99
8.1.1 Catalogue d'ensembles de points . . . 99
8.1.2 Fonction de distribution différentielle . . . 100
8.1.3 Spectre de puissance . . . 100
8.1.4 Intégrandes . . . 100
8.2 Bruit blanc . . . 101
8.2.1 Fonction de distribution différentielle . . . 102
8.2.2 Spectre de puissance . . . 103
8.2.3 Variance de l'intégration Monté Carlo . . . 103
8.3 Disques de Poisson . . . 104
8.3.1 Fonction de distribution différentielle . . . 104
8.3.2 Spectre de puissance . . . 104
8.3.3 Variance de l'intégration Monte Carlo . . . 104
8.4 Échantillonnage *jittered* . . . 105
8.4.1 Densité locale . . . 105
8.4.2 Fonction de distribution différentielle . . . 105
8.4.3 Spectre de puissance . . . 106
8.4.4 Variance de l'intégration Monte Carlo . . . 107
8.5 *Farthest-Point Optimization* . . . 108
8.5.1 Variance de l'intégration Monte Carlo . . . 108
8.6 *Blue Noise through Optimal Transport* . . . 109
8.6.1 Variance de l'intégration Monte Carlo . . . 109
8.7 *Step Blue Noise* . . . 110
8.7.1 Variance de l'intégration Monte Carlo . . . 110
8.8 Séquence de Halton . . . 110
8.9 Intégration d'une texture . . . 111
8.10 Test *zoneplate* . . . 111

Conclusion **129**
8.11 Limitations et perspectives . . . 130

Index terminologique **133**

Table des symboles 137

Bibliographie 141

Liste des figures

1.1 Illustration du lancé de rayon 6
1.2 Rôle de la position des échantillons dans la couleur d'un pixel 7

3.1 Différents espaces d'échantillonnage 21
3.2 L'espace carré périodique 22

4.1 Exemple d'ensemble de points bruit blanc dans differents espaces 30
4.2 Progression de l'algorithme générant un ensemble de disques de Poisson 31
4.3 Exemple de grilles régulières dans differents espaces 33
4.4 Exemple d'ensembles de points *jittered* dans différents espaces d'échantillonnage 34
4.5 Illustration d'une minimisation d'énergie par une méthode d'optimisation d'un ensemble de points. 36
4.6 Relaxation de Lloyd 37
4.7 Blue Noise through Optimal Transport 39
4.8 Différents pavages et les ensembles de points qui y sont associés 41
4.9 Échantillonnage par séquences à basse discrépance dans différentes dimensions 42
4.10 Résultat de plusieurs méthodes d'échantillonnage 44

5.1 Évolution de l'estimation de la densité locale du bruit blanc pour un nombre de réalisations croissant 46
5.2 Exemple de fonction de distribution différentielle et de distribution de distance d'un ensemble de points optimisé par BNOT 51
5.3 Utilisation des diagrames de Voronoi dans l'analyse spatiale des ensembles de points 55
5.4 Trois méthodes de calcul de la discrépance 57

6.1 La fonction Π dans le domaine spatial et spectral 62
6.2 Illustration d'une representation de G vers $GL(H)$ 68
6.3 Exemple d'harmoniques sphériques pour différentes fréquences 72
6.4 Exemple d'harmoniques sphériques pour différentes fréquences 73
6.5 Exemple de spectre de puissance 74

7.1 Exemple de profils spectraux de l'ensemble de points pour différentes valeurs de b 92
7.2 Exemples de profils spectraux de l'intégrande pour différentes valeurs de s 93

7.3 Évolution du produit de deux spectres de puissance en fonction de la densité de l'ensemble de points 97

8.1 Analyse spatiale et spectrale des ensembles de points bruit blanc 114
8.2 Test d'intégration avec les ensembles de points bruit blanc 115
8.3 Analyse spatiale et spectrale des ensembles de disques de Poisson 116
8.4 Test d'intégration avec les ensembles de disques de Poisson 117
8.5 Analyse spatiale et spectrale des ensembles de points *jittered* 118
8.6 Test d'intégration avec les ensembles de points *jittered* 119
8.7 Analyse spatiale et spectrale des ensembles de points optimisés par FPO 120
8.8 Test d'intégration avec les ensembles de points optimisés par FPO 121
8.9 Analyse spatiale et spectrale des ensembles de points optimisés par BNOT 122
8.10 Test d'intégration avec les ensembles de points optimisés par BNOT . . 123
8.11 Analyse spatiale et spectrale des ensembles de points *step blue noise* . . . 124
8.12 Test d'intégration avec les ensembles de points *step blue noise* 125
8.13 Analyse spatiale et spectrale de la séquence de points de Halton 126
8.14 Test d'intégration avec la sequence de Halton 127
8.15 Variance de l'intégration pour différentes méthodes et différentes intégrandes 128

Résumé

L'échantillonnage est une étape clé dans le rendu graphique. Il permet d'intégrer la lumière arrivant en un point de la scène pour en calculer sa couleur. Généralement, la méthode utilisée est l'intégration Monte Carlo qui approxime cette intégrale en choisissant un nombre fini d'échantillons. La réduction du biais et de la variance de l'intégration Monte Carlo est devenue une des grandes problématiques en rendu réaliste. Les techniques trouvées consistent à placer les points d'échantillonnage avec intelligence de façon à rendre la distribution la plus uniforme possible tout en évitant les régularités. Les années 80 ont été de ce point de vue un tournant dans ce domaine, avec l'apparition de nouvelles méthodes stochastiques. Ces méthodes ont, grâce à une meilleure compréhension des liens entre intégration Monte Carlo et échantillonnage, permis de réduire le bruit et la variance des images générées, et donc d'améliorer leur qualité. En parallèle, la complexité des méthodes d'échantillonnage s'est considérablement améliorée, permettant d'obtenir des méthodes à la fois rapides et efficaces en termes de qualité. Cependant, ces avancées ont jusqu'à là été faites par tâtonnement et se sont axées sur deux points majeurs : l'amélioration de l'uniformité du motif d'échantillonnage et la suppression des régularités. Bien que des théories permettant de borner l'erreur d'intégration existent, elles sont souvent limitées, voire inapplicables dans le domaine de l'informatique graphique.

Cette thèse propose de rassembler les outils d'analyse des motifs d'échantillonnages et de les mettre en relation. Ces outils peuvent caractériser des propriétés spatiales, comme la distribution des distances entre points, ou bien spectrales à l'aide de la transformée de Fourier. Nous avons ensuite utilisé ces outils afin de donner une expression simple de la variance et du biais dans l'intégration Monte Carlo, en utilisant des prérequis compatibles avec le rendu d'image. Finalement, nous présentons une boite à outils théorique permettant de déterminer la vitesse de convergence d'une méthode d'échantillonnage à partir de son profil spectral. Cette boite à outils est notamment utilisée afin de classifier les méthodes d'échantillonnage existantes, mais aussi pour donner des indications sur les principes fondamentaux nécessaires à la conception de nouveaux algorithmes d'échantillonnage.

Abstract

Sampling is a key step in rendering pipeline. It allows the integration of light arriving to a point of the scene in order to calculate its color. Monte Carlo integration is generally the most used method to approximate that integral by choosing a finite number of samples. Reducing the bias and the variance of Monte Carlo integration has become one of the most important issues in realistic rendering. The solutions found are based on smartly positioning the samples points in a way that maximizes the uniformity of the distribution while avoiding the regularities. From this point of view, the 80s were a turning point in this domain, as new stochastic methods appeared. With a better comprehension of links between Monte Carlo integration and sampling, these methods allow the reduction of noise and of variance in rendered images. In parallel, the complexity of sampling methods has considerably enhanced, enabling to have fast as well as good quality methods.

However, these improvements have been done by trial and error focusing on two major points : the improvement of sampling pattern uniformity, and the suppression of regularities. Even though there exists some theories allowing to bound the error of the integration, they are usually limited, and even inapplicable in computer graphics.

This thesis proposes to gather the analysis tools of sampling patterns and to connect them together. These tools can characterize spatial properties such as the distribution of distances between points, as well as spectral properties via Fourier transformation. Secondly, we have used these tools in order to give a simple expression of the bias and the variance for Monte Carlo integration ; this is done by using prerequisites compatible with image rendering. Finally, we present a theoretical toolbox allowing to determine the convergence speed of a sampling method from its spectral profile. This toolbox is used specifically to give indications about the design principles necessary for new sampling algorithms.

Introduction 1

Le travail présenté dans ce manuscrit a été effectué dans le domaine de la synthèse d'image, au sein de l'équipe R3AM du laboratoire LIRIS. Il présente donc une approche orientée sur cette thématique. La synthèse d'image vise à générer des images artificielles à partir d'une description numérique d'une scène. À la manière d'un appareil photo évoluant dans un environnement réel pour immortaliser un instant et un point de vue précis du monde, le moteur de rendu se déplace dans une scène numérique afin d'en créer une image. Les objets de ces scènes peuvent provenir de capteur numérique ou tout simplement être créés par des artistes. Le fait de prendre des "photos" dans un monde numérique plutôt que réel impose des contraintes totalement différentes, puisqu'un élément crucial est manquant, la lumière. En effet, la lumière, qui permet de voir dans le monde réel, est totalement absente dans les scènes numériques et doit donc être simulée. De nombreuses méthodes de rendu actuelles cherchent donc à recréer la trajectoire des photons, de la source de lumière jusqu'a l'objectif de la camera, ou bien de la camera jusqu'à la source de lumière. La figure (1.1) illustre ce concept. Les deux grandes problématiques actuelles se résument donc de cette façon : comment simuler correctement la lumière et comment calculer correctement cette simulation ?

En général, l'image finale est composée de pixels dont le moteur de rendu cherche à calculer la couleur (au sens large). On souhaite donc simuler la quantité de lumière, pour chaque longueur d'onde, arrivant à la surface du pixel. Mathématiquement, la quantité de lumière arrivant en un point est modélisée par une fonction et la quantité le lumière atteignant la surface, et donc la couleur du pixel, est modélisée par l'intégrale de cette fonction. Malheureusement, le processus de simulation est souvent trop complexe pour permettre de connaître le résultat de cette intégrale de façon analytique. Les chercheurs en synthèse d'images ont donc recours à des méthodes d'approximation d'intégrale, à l'origine créées pour des raisons tout autres [1]. Le choix de la méthode d'approximation dépend notamment de ce que l'on est capable de savoir sur la fonction à intégrer. Dans les algorithmes de rendu d'image, il n'est possible de calculer que la trajectoire d'un photon à la fois. Il est donc possible de connaître l'intensité lumineuse en un point de la surface du pixel, mais pas sur l'ensemble de la surface. Le choix de la méthode d'approximation de l'intégrale s'est donc naturellement porté vers l'échantillonnage. L'échantillonnage consiste à prélever la fonction en un certain nombre de points et de calculer la moyenne de ces valeurs. À mesure que le nombre d'échantillons augmente, la moyenne devient de plus en plus précise jusqu'à atteindre un seuil d'acceptabilité. On parle d'intégration Monte Carlo quand la position des points est aléatoire, et d'intégration Quasi Monte Carlo quand la position

1. On notera par exemple l'impact du Projet Manhattan sur les méthodes Monte Carlo.

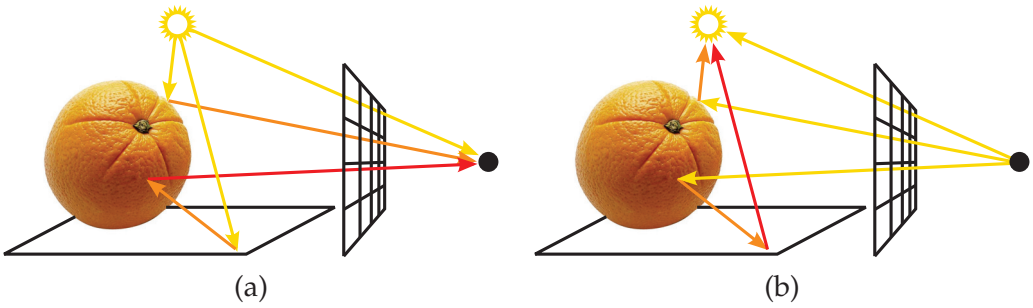


FIGURE 1.1 – *Illustration du modèle de lumière utilisé en rendu d'images.* ***(a) :*** *Les rayons simulant les photons vont d'une source de lumière vers l'objectif.* ***(b) :*** *Les rayons partent cette fois de l'objectif pour atteindre les sources lumineuses. Dans les deux images, les flèches jaunes indiquent la trajectoire d'un rayon après son émission, les flèches oranges une des trajectoires possibles d'un rayon après un «rebond » sur une surface (ici, sur le fruit et sur le sol) et les flèches rouges une des trajectoires possibles d'un rayon après un deuxième rebond. L'image finale est constituée par les rayons qui traversent le plan image, dont les longueurs d'ondes sont déterminées par leurs interactions avec la scène.*

des points est déterministe. Toutefois, nous utilisons abusivement dans ce manuscrit le terme Monte Carlo pour designer ces deux méthodes.

1.1 Critère de qualité en informatique graphique

L'intégration Monte Carlo étant une approximation, on parle d'erreur d'intégration pour designer la différence entre cette approximation et l'intégrale que l'on cherche à estimer. Il est rapidement apparu que la position des points d'échantillonnage sur la fonction affecte grandement la qualité de cette estimation. La position des échantillons doit par exemple être la plus uniforme possible, c'est à dire éviter que certaines parties de la fonction soient plus échantillonnées que d'autres. Partant de ce constat, les échantillons ont donc été disposés de façon régulière sur le domaine d'intégration. Cependant, cette disposition n'est pas sans créer des difficultés. Comme le montre la figure (1.2), lorsque la position des échantillons est trop régulière, l'intégration Monte Carlo devient très sensible à certains détails de la fonction, comme les discontinuités. Les alignements des échantillons, et de manière plus générale les structures régulières dans leur disposition, sont donc devenues une des principales causes d'erreur dans le rendu graphique. Le phénomène le plus connu, l'*aliasing*[2] , résulte de ces défauts dans la position des échantillons. Pour supprimer ces structures, des processus stochastiques et quasi-stochastiques ont été introduits dans les échantillonneurs. La sensibilité aux discontinuités dans le signal est alors réduite, mais au prix d'une augmentation de la variance. En effet, puisque la position des échantillons change à chaque estimation de l'intégrale, du fait de l'utilisation de processus stochastique, le résultat de l'estimation varie d'une fois à l'autre.

L'objectif actuel des méthodes d'échantillonnage est donc de calculer la position des échantillons d'une façon non-régulière mais uniforme. En parallèle, certains cher-

2. Le terme d'*aliasing*, issue de l'anglais, se traduit généralement en français par crénelage, terme qui correspond en fait à une des manifestations de l'*aliasing*. Nous utilisons donc dans ce manuscrit le terme anglais pour éviter toute confusion.

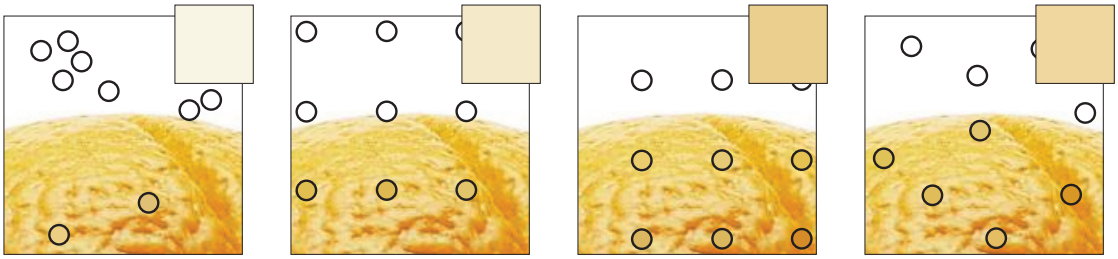


FIGURE 1.2 – *Rôle de la position des échantillons dans la couleur d'un pixel. La portion de scène visible depuis le pixel est constituée d'un fond blanc et d'une partie de l'orange. Les cercles représentent des échantillons. Le carré en haut à droite de chaque image montre la couleur du pixel, calculée en moyennant les échantillons.* ***(a) :*** *Les échantillons sont mal répartis sur la surface du pixel. Le fond de la scène est trop représenté, aboutissant à un pixel trop blanc.* ***(b) et (c) :*** *Les échantillons sont placés de façon régulière. On remarque que la couleur peut varier fortement lorsque l'on décale la position des points. L'alignement des échantillons rend l'intégration Monte Carlo très sensible aux contours présents dans la scène* ***(d) :*** *Les échantillons sont placés aléatoirement, tout en restant répartis sur la surface du pixel.*

cheurs se sont employés à déterminer le lien théorique entre l'erreur d'intégration et certaines propriétés dans la disposition des points d'échantillonnage. Cette recherche, plus fondamentale, permet de déterminer la position souhaitée des échantillons et ainsi de donner un objectif clairement défini en terme de placement de points pour les méthodes d'échantillonnage. Dans ce domaine, Ulichney [1987] a proposé un modèle basé sur les propriétés spectrales de l'ensemble de points d'échantillonnage[3]. Bien qu'empirique, ce modèle a servi de boussole à une partie de la communauté scientifique pour le développement de nouveaux algorithmes d'échantillonnage. Nous reviendrons dans le dernier chapitre de ce manuscrit sur la pertinence de ce modèle. Depuis les années 2010, de nouveaux travaux [Durand, 2011; Subr and Kautz, 2013] portant sur les liens entre intégration Monte Carlo et position des échantillons ont permis de mieux comprendre comment positionner les échantillons afin de limiter la variance de l'intégration. Le travail présenté dans ce manuscrit est dans la continuité de ces travaux.

1.2 Contributions et plan de thèse

Ce manuscrit, organisé en huit chapitres, contient plusieurs contributions scientifiques dans le domaine de la synthèse d'image et de l'échantillonnage.

Formalisme mathématique Dans le chapitre 3, nous détaillons une formalisation de l'espace d'échantillonnage et des ensembles de points dans la perspective de l'intégration Monte Carlo pour l'informatique graphique. Les définitions énoncées ainsi que les propriétés qui en découlent, forment un tout permettant de mieux comprendre les ensembles de points et les relations qui les lient à l'intégration Monte Carlo. Ce formalisme est orienté vers une conception particulière de l'échantillonnage, où aucunes informations sur l'intégrande n'est disponible et où l' *aliasing* est une contrainte supplémentaire. Dans ce contexte, nous avons donc axé notre vision mathématique

3. Pour une définition de l'analyse spectrale, voir le chapitre 6.

de l'échantillonnage sur les processus stochastiques et sur l'homogénéité. Les définitions et lemmes présentés ici ne doivent pas être vu comme des contributions. C'est la réunion de ces définitions et propriétés, dans une perspective et un formalisme communs qui peut être considérée comme tel.

État de l'art Le chapitre 4 détaille les méthodes d'échantillonnage mises en place pour l'intégration Monte Carlo dans la synthèse d'images. Cette vue d'ensemble permettra d'avoir une idée concrète de ce que sont les ensembles de points, et des enjeux mis en place pour les améliorer. Nous verrons en particulier les méthodes d'échantillonnage utilisant des processus stochastiques, celles utilisant des partitions de l'espace, ainsi que des méthodes par optimisation d'énergie.

Outils théoriques d'analyse Les chapitres 5 et 6 rassemblent un ensemble d'outils théoriques pour analyser les propriétés des ensembles de points. Bien que ces outils soient connus, nous les avons regroupés et liés ensemble à travers le formalisme mathématique que nous avons introduit. Nous détaillons également les propriétés des ensembles de points homogènes pour chacun de ces outils.

Intégration Monte Carlo En poursuivant les travaux de Durand [2011] et Subr [Subr and Kautz, 2013], nous montrons, dans le chapitre 7, comment le spectre de puissance est relié à la variance de l'intégration Monte Carlo, dans le cas des ensembles de points homogènes. Ce lien permet de mieux diriger la recherche portant sur les méthodes d'échantillonnage stochastiques en décrivant explicitement des conditions suffisantes pour qu'un ensemble de points soit efficace pour l'intégration. À l'heure où les méthodes de contrôle spectral commencent à apparaître, tel que le contenu spectral de l'ensemble de points prend une importance de plus en plus grande, il nous a semblé important de formaliser ce le lien entre le domaine spectral et l'intégration Monte Carlo.

Analyse de l'état de l'art Dans le dernier chapitre, nous nous sommes attelés à évaluer, à l'aide des mêmes outils et des mêmes normalisations, les méthodes d'échantillonnages stochastiques qui nous semblent importantes. Les résultats produits permettent de mieux comprendre ce que chaque méthode a réussi à accomplir, tout en soulignant les différences fondamentales de leurs caractéristiques. L'ensemble des outils d'analyse que nous avons utilisés est mis à la disposition de la communauté scientifique.

Publications Dans [Pilleboue *et al.*, 2015], publié à la conférence SIGGRAPH 2015, nous détaillons comment exprimer la variance de l'intégration en fonction du spectre de puissance de l'ensemble de points et de l'intégrande. Nous montrons également comment la vitesse de convergence de l'intégration est affectée par la forme du spectre de puissance dans les basses fréquences. Ces résultats théoriques sont montrés pour l'espace hyper-cubique périodique ainsi que l'espace sphérique. Dans [Kazhdan *et al.*, 2015], nous montrons que l'expression de la variance en fonction du spectre de puissance peut être généralisée à d'autres espaces en utilisant la théorie des représentations. Ce manuscrit reprend l'essentiel des contributions présentées dans ces deux

articles, mais présenté dans un cadre théorique unifiée. L'étude de la vitesse de convergence n'est ainsi plus limitée à l'espace hyper-cubique périodique et à l'espace sphérique. Finalement, dans [Wachtel *et al.*, 2014], nous pressentons une méthode d'échantillonnage utilisant des pavages hiérarchiques. Bien que nous mentionnons cette méthode dans l'état de l'art, chapitre 4, ce manuscrit ne détaille pas les contributions de cet article.

A la fin de ce manuscrit, une table des notations, ainsi qu'un index terminologique, sont disponibles pour faciliter la lecture. Maintenant que nous avons présenté le contexte de recherche et la structure de ce manuscrit, il est temps d'entrer concrètement dans le sujet. Cependant, il est important de souligner que nos travaux étant principalement théoriques, un prérequis mathématique est nécessaire pour une pleine compréhension de nos résultats. Aussi, avant de définir formellement l'espace d'échantillonnage dans le chapitre 3, nous présentons dans le chapitre suivant les notations et les concepts nécessaires pour appréhender le reste de ce manuscrit.

Prérequis mathématiques 2

Ce chapitre à pour vocation d'introduire les notations mathématiques et de définir les concepts fondamentaux nécessaires à la compréhension de nos travaux. En particulier, nous présentons ici la définition d'une topologie, la définition de l'intégrale au sens de Lebesgue et la notation en «grand O ».

2.1 Notations

On note $\mathbb{N}$ l'ensemble des nombres entiers naturels, $\mathbb{Z}$ l'ensemble des nombres entiers relatifs, $\mathbb{R}$ l'ensemble des nombres réels et $\mathbb{C}$ l'ensemble des nombres complexes. Pour les nombres complexes, on utilisera la notation i tel que $i^2 = -1$, et $\bar{c}$ pour le conjugué d'un nombre complexe c. Soit x une variable représentant un élément d'un ensemble X, nous utilisons la notation x' et x'' pour désigner deux autres variables représentant des éléments de X. On retrouve également cette notation dans la littérature pour désigner la dérivée d'une fonction. Pour éviter toutes confusions, nous utilisons la notation de Leibniz pour les dérivées, telle que $\mathrm{d}f(x)/\mathrm{d}x$ signifie la dérivée de la fonction f par rapport à x. Il faudra donc interpréter f' comme une autre fonction d'un ensemble de fonctions et non pas comme la dérivée d'une fonction f. Soit X un ensemble, nous utilisons la notation

$$\{x : x \in X, C(x)\} \ , \tag{2.1}$$

pour designer l'ensemble des x appartenant à X respectant la condition C. L'ensemble des nombres rationnel s'écrit ainsi comme

$$\mathbb{Q} := \left\{ \frac{m}{n} : m \in \mathbb{Z}, n \in \mathbb{Z} \setminus \{0\} \right\} \ . \tag{2.2}$$

Le symbole $\emptyset$ correspond à l'ensemble vide et le symbole $:=$ correspond à la définition d'un terme.

2.2 Théorie de l'intégration

Dans ce manuscrit, les intégrales jouent un rôle central, à la fois comme objets d'étude avec l'intégration Monte Carlo, et comme outils mathématiques dans la plupart de nos dérivations. Nous utilisons ce concept dans sa forme classique au sens de Lebesgue et à l'aide des mesures usuelles des surfaces et des groupes. Pour les lecteurs curieux souhaitant savoir ce qu'est une intégrale, nous définissons ici l'intégrale

au sens de Lebesgue, en introduisant les concepts nécessaires à sa définition tels que les topologies, les σ-algèbres et les mesures. La définition exacte de ces concepts, présentée ici, n'est pas indispensable à la compréhension du travail que nous présentons.

2.2.1 Topologie

Définition 2.1 *[Rudin, 1987, def. 1.2.a] Soit X un ensemble et P un ensemble de sous-parties $p \subseteq X$, on dit que (X, P) est un* ***espace topologique*** *si*

- $\emptyset \in P$ *et* $X \in P$,
- $p_1, ... p_n \in P \Rightarrow p_1 \cap ... \cap p_n \in P$,
- *Soit Q un sous-ensemble fini ou infini de P, on a* $\left(\bigcup_{p \in Q} p \right) \in P$.

Par souci de concision, on dira que X est un espace topologique sans mentionner P. X reste donc un ensemble, mais pour lequel nous assumons que P existe et est connu. On dit que les sous-parties de X appartenant à P sont des ouverts et on appelle une union d'ouverts, un borélien. Le formalisme des espaces topologiques permet de définir les notions de continuité et de limite. Pour plus d'informations sur les topologies, voir [Lee, 2000].

Définition 2.2 *[Barut and Rączka, 1986, chap. 2] Soit (X, P) un espace topologique, on dit que (X, P) est un* ***espace de Hausdorff*** *si la propriété suivante est respectée*

$$\forall x, x' \in X, x \neq x', \exists p, p' \in P \text{ tel que } x \in p, x' \in p' \text{ et } p \cap p' = \emptyset . \tag{2.3}$$

Dans un espace de Hausdorff, lorsque l'on se rapproche d'un point, on s'éloigne des autres. Ces espaces sont importants puisqu'ils garantissent que les limites n'admettent qu'une seule valeur. On appelle recouvrement d'un ensemble X, un ensemble de sous-parties de X tel que X est un sous ensemble de l'union de ces sous-parties. On appelle sous-recouvrement un sous-ensemble d'un recouvrement de X qui recouvre toujours X.

Définition 2.3 *[Lee, 2000, lemma 4.27] Soit (X, P) un espace topologique, et p un sous ensemble de X, on dit que p est* ***compact*** *si tout recouvrement de p par des ouverts admet un sous-recouvrement fini.*

Définition 2.4 *[Barut and Rączka, 1986, chap. 2] Soit (X, P) un espace de Hausdorff, on dit que (X, P) est un* ***espace topologique localement compact*** *si pour tout $x \in X$, il existe $p \in P$ tel que p est compact et $x \in p$.*

Les espaces topologiques localement compact ne sont pas directement utilisés dans la définition de l'intégrale au sens de Lebesgue. Cependant, ils possèdent de nombreuses propriétés, dont plusieurs seront utilisées dans ce manuscrit.

2.2.2 Mesure

Définition 2.5 *[Rudin, 1987, def. 1.3.a] Soit X un ensemble et P un ensemble de sous-parties $p \subseteq X$, on dit que P est une* ***σ-algèbre*** *sur X si P vérifie les propriétés suivantes :*

- $X \in P$,
- $\forall p \in P, \quad X \setminus p \in P$,

— *Soit* $\{p_i\}_{i\in\mathbb{N}}$ *vérifiant pour tout* $i, p_i \in P$, *et* p *un ensemble tel que* $p = \left(\bigcup_{i\in\mathbb{N}} p_i\right)$, *alors* $p \in P$.

Lorsque X possède une σ-algèbre, on dit que (X, P) est un **espace mesurable**, ou tout simplement que X est un espace mesurable.

Définition 2.6 *[Rudin, 1987, def. 1.3.c] Soit* (X, P_X) *un espace mesurable et* (Y, P_Y) *un espace topologique, on dit qu'une fonction* $f : X \to Y$ *est* ***mesurable*** *si pour tout* $p \in P_Y$, *on a*

$$\{x : x \in X, f(x) \in p\} \in P_X\,. \tag{2.4}$$

Définition 2.7 *[Rudin, 1987, def. 1.18.a] Soit* X *un ensemble et* P *une* σ*-algèbre sur* X, *on dit que la fonction* $\mu : P \to \mathbb{R}^+ \cup \{\infty\}$ *est une* ***mesure positive*** *si elle est* σ*-additive, c'est à dire si pour tout ensemble dénombrable* p_i *d'éléments disjoints deux à deux, on a*

$$\mu\left(\bigcup_{i\in\mathbb{N}} p_i\right) = \sum_{i\in\mathbb{N}} \mu\left(p_i\right). \tag{2.5}$$

Dans ce formalisme, la σ-algèbre P sur X permet de définir l'ensemble des sous-parties de X que l'on peut mesurer, et la mesure est la fonction permettant de quantifier une "aire" associée à chaque sous-partie mesurable.

2.2.3 Intégrale au sens de Lebesgue

Définition 2.8 *[Rudin, 1987, def. 1.16] Soit* X *un espace mesurable,* $s : X \to \mathbb{C}$ *une fonction, on dit que* s *est une* ***fonction étagée*** *si elle peut être exprimée sous la forme*

$$s(x) = \sum_{i\in I} \alpha_i \chi_{A_i}(x)\,, \tag{2.6}$$

où I *est un ensemble d'indices fini,* $\alpha_i \in \mathbb{C}$, $A_i := \{x : x \in X, s(x) = \alpha_i\}$, *et* χ_{A_i} *est la fonction caractéristique de* A_i *telle que*

$$\chi_{A_i}(x) = \begin{cases} 1 & si\ x \in A_i \\ 0 & sinon\,. \end{cases} \tag{2.7}$$

Définition 2.9 *[Rudin, 1987, def. 1.23] Soit* X *un espace mesurable,* μ *une mesure de* X, f *une fonction définie dans* $X \to \mathbb{R}^+ \cup \{\infty\}$ *et* E *un sous ensemble de* X *appartenant à la* σ*-algèbre de* X *on appelle* ***intégrale au sens de Lebesgue*** *sur* E, *ou tout simplement intégrale sur* E, *l'opération suivante*

$$\int_E f(x)\mu(\mathrm{d}x) = \sup_{s\in S} \sum_{i\in I} \alpha_i \mu A_i \cap E\,, \tag{2.8}$$

où S *est l'ensemble des fonctions étagées réelles mesurables telles que pour tout* $x, s(x) \leq f(x)$.

On étend cette définition aux fonctions à valeurs dans $\mathbb{C}$ de la façon suivante

$$\begin{aligned}\int_E f(x)\mu(\mathrm{d}x) = &\int_E Re\left(f(x)\right)^+ \mu(\mathrm{d}x) - \int_E Re\left(f(x)\right)^- \mu(\mathrm{d}x) \\ &+ i\left(\int_E Im\left(f(x)\right)^+ \mu(\mathrm{d}x) - \int_E Im\left(f(x)\right)^- \mu(\mathrm{d}x)\right)\,,\end{aligned} \tag{2.9}$$

où $Re\,(f(x))^+$ et $Im\,(f(x))^+$ sont les parties positives réelles et imaginaires et $Re\,(f(x))^-$ et $Im\,(f(x))^-$ sont les parties négatives réelles et imaginaires de f. Lorsque la mesure est implicitement donnée par l'espace mesurable, on écrira l'intégrale simplement comme

$$\int_E f(x)\mathrm{d}x\,, \tag{2.10}$$

avec $\mathrm{d}x$ à la place de $\mu(\mathrm{d}x)$.

2.2.4 Intégrale d'un espace discret

Tout le formalisme que nous venons de définir est compatible avec les espaces discrets. Ainsi, si X est discret, on peut définir une σ-algèbre simple en prenant toutes les sous-parties possibles de X, et utiliser une mesure de comptage μ_C telle que

$$\mu_C\,(p) = Card(p)\,, \tag{2.11}$$

c'est à dire le nombre d'éléments contenus dans p.

Lemme 2.1 *Soit X un espace mesurable discret, f une fonction définie dans $X \to \mathbb{R}^+ \cup \{\infty\}$ et E un sous ensemble de X appartenant à la σ-algèbre de X*

$$\int_E f(x)\mu_C(\mathrm{d}x) = \sum_{x\in E} f(x)\,, \tag{2.12}$$

où la mesure utilisée pour l'intégrale est une mesure de comptage.

Preuve 2.1 *On pose Y égal à l'ensemble des valeurs possibles de la fonction f et A_y l'ensemble des éléments de X tel que*

$$A_i = \{x : f(x) = y\}\ . \tag{2.13}$$

Comme X est dénombrable, alors Y l'est aussi. On peut alors écrire que

$$f(x) = \sum_{y\in Y} y\chi_{A_y}(x)\,. \tag{2.14}$$

Cependant, cette somme ne représente pas une fonction étagée car Y peut être infini. On note Y_i l'ensemble des i premières valeurs de Y et on ré-écrit $f(x)$ sous la forme

$$f(x) = \lim_{i\to\infty} \sum_{y\in Y_i} y\chi_{A_y}(x)\,. \tag{2.15}$$

Sachant que pour tout $y \in Y_i$, $y \geq 0$, on peut dire que la suite à l'intérieur de la limite dans Eq. (2.15) est croissante, par rapport à i, car

$$\left(\sum_{y\in Y_{i+1}} y\chi_{A_y}(x)\right) - \left(\sum_{y\in Y_i} y\chi_{A_y}(x)\right) = y_{i+1}\chi_{A_y}(x)\,, \tag{2.16}$$

où y_{i+1} est la dernière valeur de Y_{i+1}. Ainsi, la fonction étagée réelle la plus proche de f, mais qui soit toujours inférieure ou égale à f, est donnée par cette limite. La preuve est terminée en reprenant la définition de l'intégrale donnée dans Def. (2.9) et en remarquant que tout les A_i sont mesurables. □

Toute somme d'éléments d'une fonction f peut s'écrire dans une forme intégrale en assumant que l'on utilise une mesure de comptage. Il ne faut donc pas voir l'intégrale comme la version continue de la somme, mais plutôt comme une généralisation de la somme aux espaces discrets et continus. Dans ce manuscrit, nous utiliserons donc parfois des intégrales à la place des sommes pour garder une expression aussi indépendante que possible du domaine de définition de la fonction.

2.3 Fonction δ de Dirac

La fonction δ de Dirac est une fonction $\mathbb{R} \to \mathbb{R}$ permettant de représenter une impulsion. Elle est définie par les deux propriétés suivantes :

$$\begin{cases} \delta(x) = 0 \text{ si } x \neq 0 \\ \displaystyle\int_{-\infty}^{\infty} \delta(x)\mathrm{d}x = 1\,. \end{cases} \tag{2.17}$$

Comme on le voit, cette fonction n'est pas complètement définie par ces propriétés puisque quand $x = 0$, le résultat est inconnu. En échantillonnage, on ajoute une propriété supplémentaire :

$$\int_{-\infty}^{\infty} \delta(x)f(x)\mathrm{d}x = f(0)\,, \tag{2.18}$$

où f est une fonction $\mathbb{R} \to \mathbb{R}$. Cette propriété permet de modéliser le prélèvement d'une valeur à l'aide d'une intégrale. Lorsque l'on veut prélever une valeur de f en x_0, on utilise une fonction δ de Dirac décalée en écrivant $\delta(x - x_0)$.

Dans ce manuscrit, nous serons amené à utiliser les fonctions δ de Dirac dans d'autres domaines que $\mathbb{R}$. Ces domaines pourront être multidimensionnels et non-euclidiens. On utilisera donc une version alternative de la fonction δ de Dirac, donnée par

$$\begin{cases} \delta(x, x_0) = 0 \text{ si } x \neq x_0 \\ \displaystyle\int \delta(x, x_0)\mathrm{d}x = 1 \\ \displaystyle\int \delta(x, x_0)f(x)\mathrm{d}x = f(x_0)\,, \end{cases} \tag{2.19}$$

où $\delta : X \times X \to \mathbb{C}$, $f : X \to \mathbb{C}$, et $x, x_0 \in X$. L'intégrale de la fonction δ de Dirac étant égale à 1, cette fonction est utilisée pour représenter des lois de probabilité. On parle alors de distribution de Dirac.

2.4 Autres fonctions

Nous utiliserons dans ce manuscrit certaines fonctions afin de simplifier les expressions. La fonction $\Pi : \mathbb{R} \to \mathbb{R}$ est définie par la relation suivante

$$\Pi(x) = \begin{cases} 1 & \text{si } -\dfrac{1}{2} \leq x \leq \dfrac{1}{2} \\ 0 & \text{sinon}\,. \end{cases} \tag{2.20}$$

La fonction $sinc : \mathbb{R} \to \mathbb{R}$ est définie par la relation suivante [Bracewell, 1999, chap. 4]

$$sinc(x) = \begin{cases} 1 & \text{si } x = 0 \\ \dfrac{sin(\pi x)}{\pi x} & \text{sinon}\,. \end{cases} \tag{2.21}$$

2.5 Notations en «grand O »

Nous utiliserons les notations dites en «grand O » pour caractériser le comportement asymptotique des fonctions. Ce type de notation est courant dans l'informatique, notamment pour étudier le temps de calcul d'un algorithme. On trouve donc des définitions pour ces notations pour les suites et les fonctions $\mathbb{N} \to \mathbb{R}$. Dans notre cas, les fonctions que nous allons étudier sont définies dans des espaces topologiques multidimensionnels. Il nous semble donc nécessaire de redéfinir ici cette notation et d'en donner quelques propriétés.

Définition 2.10 *Soit (X, P) un espace topologique, f, g deux fonctions $X \to \mathbb{C}$ et $a \in X$, on dit que f est en grand O de g quand x tend vers a, noté $f \in O_{x \to a}(g(x))$, s'il existe un entourage $A \in P$ tel que $a \in A$ et un réel M tel que*

$$\forall x \in A, |f(x)| \leq M|g(x)|. \tag{2.22}$$

Définition 2.11 *Soit (X, P) un espace topologique possédant une origine notée x_0 et f, g deux fonctions $X \to \mathbb{C}$, on dit que f est en grand O de g quand x tend vers l'infini, noté $f \in O_{x \to \infty}(g(x))$, s'il existe un entourage $X_0 \in P$ tel que $x_0 \in X_0$ et un réel M tel que*

$$\forall x \notin X_0, |f(x)| \leq M|g(x)|. \tag{2.23}$$

Ici, les entourages permettent de définir des seuils à partir desquels on vérifie l'équation $|f(x)| \leq M|g(x)|$. Dans le cas des séries ou des fonctions de $\mathbb{R}^+$, ces entourage sont simplement des intervalles.

Propriété 2.1 *Soit X un espace topologique, f, f', g, g' quatre fonctions $X \to \mathbb{C}$ et $k \in \mathbb{C}$,*

$$f \in O(g) \quad \Rightarrow \quad k \cdot f \in O(g) \tag{2.24}$$

$$f \in O(g), f' \in O(g') \quad \Rightarrow \quad f + f' \in O(|g| + |g'|), \tag{2.25}$$

dans le cas où x tend vers un élément de X ou vers l'infini.

Nous venons de présenter les prérequis mathématiques nécessaires, sous forme de notations, définitions et lemmes. Les bases mathématiques étant posées, nous pouvons maintenant entrer dans le vif du sujet de ce manuscrit : l'échantillonnage. Dans le chapitre qui suit, nous définissons ce que sont les ensembles de points, c'est à dire les positions des échantillons, et l'espace d'échantillonnage dans lequel ils évoluent.

Ensemble de points et espace d'échantillonnage

3

Dans ce chapitre, nous définissons ce que sont les ensembles de points, l'espace d'échantillonnage et l'échantillonneur. La définition formelle que nous proposons sert de base à l'ensemble de ce manuscrit et permet d'offrir un cadre général à la plupart des outils et des résultats que nous présentons. Une attention particulière est portée sur la définition de l'espace d'échantillonnage, laquelle peut être vue comme une contribution étant donné qu'une grande partie de nos résultats en dépende. Ce chapitre introduit aussi l'utilisation de variables aléatoires pour décrire le comportement des échantillonneurs stochastiques. Nous pressentons en complément, une liste des espaces d'échantillonnage couramment utilisés.

3.1 Ensemble de points

Définition 3.1 *Soit $\mathbb{D}_{\mathcal{E}}$ un domaine, on appelle* ***ensemble de points****, noté $\mathcal{S}$, l'ensemble indexé*

$$\mathcal{S} := \{\boldsymbol{s}_1, ..., \boldsymbol{s}_N\} \ , \tag{3.1}$$

où chaque $\boldsymbol{s}_k$ est un élément de $\mathbb{D}_{\mathcal{E}}$ et N un entier naturel strictement positif appelé ***nombre de points****.*

Chaque point étant indicé, on peut considérer aussi l'ensemble de points comme un vecteur de dimension N. On acceptera donc que des points d'indice différent soient égaux. L'ensemble de points, quand il est utilisé pour l'échantillonnage, peut aussi être appelé «motif d'échantillonnage ».

Définition 3.2 *Soit $\mathcal{S}$ un ensemble de points dans $\mathbb{D}_{\mathcal{E}}$, on appelle* ***fonction représentative*** *de $\mathcal{S}$, la fonction $S : \mathbb{D}_{\mathcal{E}} \to \mathbb{R}$ définie par*

$$S(x) := \sum_{k=1}^{N} \delta(x, \boldsymbol{s}_k) \, . \tag{3.2}$$

Cette définition de la fonction représentative de $\mathcal{S}$ est liée à l'utilisation de l'ensemble de points dans l'intégration Monte Carlo. Puisque le but des ensembles de points dans ce contexte est de prélever une valeur, il est naturel d'utiliser la fonction δ de Dirac.

L'échantillonneur est un algorithme générant l'ensemble de points $\mathcal{S}$. Il existe de multiples façons de générer un ensemble de points que nous détaillons plus loin dans ce chapitre. Avant cela, nous allons introduire les deux grandes classes d'échantillonneurs : les échantillonneurs déterministes et les échantillonneurs stochastiques.

Un **échantillonneur déterministe** produit un ensemble de points pour un jeu de paramètres donné. Ces paramètres peuvent être le nombre de points de $\mathcal{S}$, mais aussi d'autres paramètres propres à la méthode d'échantillonnage. Nous excluons ici les algorithmes utilisant des générateurs de nombres aléatoires, car bien qu'ils soient techniquement déterministes, leur objectif est de simuler un processus stochastique. On note $\mathfrak{M}_{\text{Determ.}}$ la classe des ensembles de points générés par un échantillonneur déterministe, et on utilisera la notation $\mathcal{S} \in \mathfrak{M}_{\text{Determ.}}$ pour dire que $\mathcal{S}$ est un ensemble de points déterministe.

À l'opposé des échantillonneurs déterministes, on trouve les **échantillonneurs stochastiques**. Ces échantillonneurs utilisent des processus stochastiques pour générer la position des points de $\mathcal{S}$. Chaque fois que l'échantillonneur est exécuté, un nouvel ensemble de points est créé. On appelle chaque ensemble de points créé par un échantillonneur une **réalisation** et on note la m-ième réalisation $\mathcal{S}[m]$. Afin d'étudier les propriétés des ensembles de points stochastiques, on utilise un formalisme faisant intervenir des variables aléatoires. L'ensemble de points pouvant être vu comme un vecteur de taille N, l'espace $\mathbb{D}_{\mathcal{E}}^{N}$ représente alors toutes combinaisons possibles de points formant l'ensemble $\mathcal{S}$. On redéfinit donc l'ensemble de points $\mathcal{S}$ comme une variable aléatoire à valeur dans $\mathbb{D}_{\mathcal{E}}^{N}$ à laquelle on associe une loi de probabilité $\mathbb{P}_{\mathcal{S}}$.

Cette loi de probabilité dépend de l'échantillonneur et de ses paramètres. On dit que deux ensembles de points sont statistiquement équivalents s'ils suivent la même loi de probabilité. Soit f une fonction applicable sur un ensemble de points, on définit **l'espérance** de cette fonction appliquée à $\mathcal{S}$ par

$$\langle f(\mathcal{S})\rangle := \int_{\mathbb{D}_{\mathcal{E}}^{N}} f(s^N) PDF_{\mathcal{S}}(s^N) \mathrm{d}s^N \,, \tag{3.3}$$

où l'opérateur $\langle ... \rangle$ désigne l'espérance sur l'ensemble des réalisations de $\mathcal{S}$ et $PDF_{\mathcal{S}}$ désigne la fonction de densité de probabilité associée à la loi $\mathbb{P}_{\mathcal{S}}$. On peut également exprimer cette espérance de façon empirique en utilisant chaque réalisation de l'échantillonneur.

$$\langle f(\mathcal{S})\rangle = \lim_{M\to\infty} \frac{1}{M} \sum_{m=1}^{M} f(\mathcal{S}[m]) \,, \tag{3.4}$$

où M est le nombre de réalisations de $\mathcal{S}$ pris en compte dans la moyenne empirique. La fonction f peut représenter toute opération sur un ensemble de points, ou utilisant un ensemble de points. Cela peut être la fonction représentative S, la transformée de Fourier de S (voir Chapitre 6), ou le résultat de l'intégration Monte-Carlo avec $\mathcal{S}$ comme motif d'échantillonnage (voir Chapitre 7). Cela peut-être aussi des opérations simples comme «retourner le k-ième point de l'ensemble », c'est à dire $\mathbf{s}_k$. Tout comme $\mathcal{S}$, $\mathbf{s}_k$ devient lui aussi une variable aléatoire et possède donc sa propre loi de probabilité, qui découle de $\mathbb{P}_{\mathcal{S}}$. La fonction de densité de probabilité de $\mathbf{s}_k$, notée $PDF_{\mathbf{s}_k}$, est donnée par

$$PDF_{\mathbf{s}_k}(x) = \langle \delta(x, \mathbf{s}_k)\rangle \,. \tag{3.5}$$

De l'expression de l'espérance donnée par Eq. (3.3), nous pouvons définir la covariance entre deux fonctions f et g appliquées sur l'ensemble de points :

$$Cov(f(\mathcal{S}), g(\mathcal{S})) = \langle f(\mathcal{S})g(\mathcal{S})\rangle - \langle f(\mathcal{S})\rangle \langle g(\mathcal{S})\rangle \,. \tag{3.6}$$

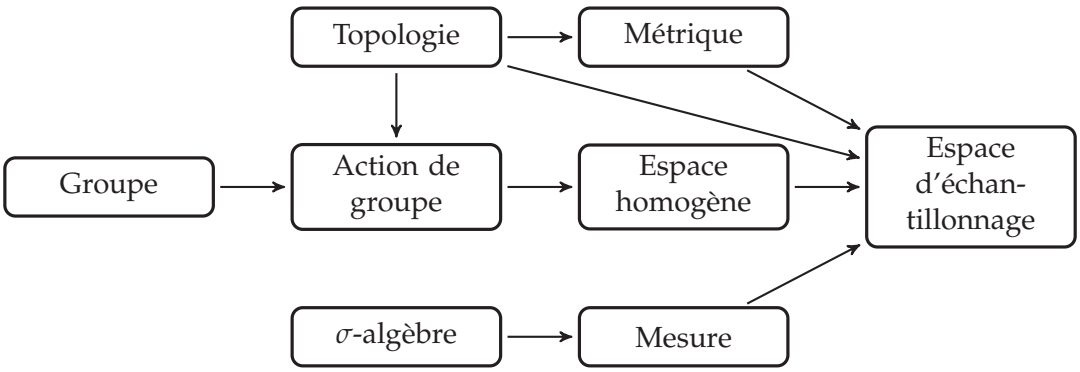
La variance d'une fonction f appliquée sur $\mathcal{S}$ est donnée par la covariance de cette fonction avec elle même :

$$Var(f(\mathcal{S})) = \langle f(\mathcal{S})^2 \rangle - \langle f(\mathcal{S}) \rangle^2 \ . \tag{3.7}$$

Lorsqu'un ensemble de points est stochastique, on note que $\mathcal{S} \in \mathfrak{M}_{\text{Stocha.}}$, où $\mathfrak{M}_{\text{Stocha.}}$ est la classe des ensembles de points générés par un échantillonneur stochastique. Pour ce manuscrit, nous utiliserons également le formalisme des ensembles de points stochastiques pour les ensembles de points déterministes. Dans ce cas, on considérera simplement que toutes les réalisations de l'échantillonneur sont identiques.

3.2 Espace d'échantillonnage

L'espace d'échantillonnage $\mathcal{E}$ est l'espace dans lequel évolue l'ensemble de points $\mathcal{S}$. Nous apportons ici une attention particulière à sa définition, puisque la plupart des résultats présentés dans ce manuscrit dépendent étroitement des propriétés de l'espace utilisé. L'espace d'échantillonnage tel que nous le définissons est un concept abstrait permettant de généraliser les propriétés des ensembles de points généralement utilisées en informatique graphique. Nous commencerons donc par introduire les concepts nécessaires, puis nous définirons clairement l'espace d'échantillonnage avant de finalement donner quelques exemples. La figure ci-dessous résume les dépendances entre chaque concept mathématique. Dans cette figure, les flèches indiquent qu'un concept mathématique (queue de la flèche) est utilisé pour en définir un autre (tête de la flèche).



Définition 3.3 *[Rudin, 1987, chap. 1] Soit X un ensemble, on dit que la fonction dist : $X \times X \to \mathbb{R}^+$ est une* ***métrique*** *si elle respecte les conditions suivantes*

— $\forall x, x' \in X, dist(x, x') = 0 \Rightarrow x = x'$,
— $\forall x, x' \in X, dist(x, x') = dist(x', x)$,
— $\forall x, x', x'' \in X, dist(x, x'') \leq dist(x, x') + dist(x', x'')$.

Lorsqu'un ensemble X est muni d'une métrique, on dit que l'ensemble est métrisable. On appelle également *dist* une **distance** et on utilisera indistinctement les deux termes. Soit $x \in X$ et $r \in \mathbb{R}^+$, on appelle boule ouverte centrée en x et de rayon r, l'ensemble $\{x' : x' \in X, dist(x, x') < r\}$. Une façon simple de créer une topologie à partir d'un ensemble métrisable X est de considérer l'ensemble des boules ouvertes centrées en chaque élément de X et pour tout les rayons possibles.

Définition 3.4 *[Cornwell, 1997, chap. 1] Soit G un ensemble et $\star$ un opérateur, on dit que $(G, \star)$, ou implicitement G, est un* ***groupe*** *si les conditions suivantes sont respectées*
- *l'opérateur $\star$ est une loi de composition interne : $\forall g, g' \in G, \quad g \star g' \in G$,*
- *l'opérateur est associatif sur G : $\forall g, g', g'' \in G, \quad g \star (g' \star g'') = (g \star g') \star g''$,*
- *il existe un élément neutre : $\exists g_0 \in G$ tel que $\forall g \in G, \quad g_0 \star g = g \star g_0 = g$,*
- *il existe un inverse : $\forall g \in G, \exists g^{-1} \in G$ tel que $g \star g^{-1} = g^{-1} \star g = g_0$.*

Dans le contexte de cette thèse, le groupe représentera généralement un ensemble de translations ou de rotations qu'il est possible de combiner ensemble. On dit que G est un groupe topologique si G est un espace topologique et que les fonctions $f_{\text{inv.}}(g) = g^{-1}$ et $f_\star(g, g') = g \star g'$ sont continues.

Définition 3.5 *[Bredon, 1972, sec. 1.1] Soit G un groupe topologique, X un espace topologique et $*$ un opérateur $X \times G \to X$, on dit que G agit sur X, et que G est une* ***action de groupe****, si les conditions suivantes sont respectées*
- *$\forall x \in X, \quad x * g_0 = x$,*
- *$\forall g, g' \in G, x \in X, \quad x * (g \star g') = (x * g) * g'$.*

On a donc un groupe, généralement un ensemble de translations ou de rotations, qui agit sur un ensemble, par exemple des points à déplacer.

Définition 3.6 *[Bredon, 1972] Soit X un espace topologique et G un groupe qui agit sur X, on dit que X est un* ***espace homogène*** *si l'action de G sur X est transitive, c'est à dire si $\forall x, x' \in X, \exists g \in G$ telle que $x * g = x'$.*

Avec cet ensemble de concepts, nous sommes maintenant capable de définir l'espace d'échantillonnage.

Définition 3.7 *Soit $\mathbb{D}_\mathcal{E}$ un espace topologique muni d'une métrique $\text{dist}_\mathcal{E}$ et $\mathbb{T}_\mathcal{E}$ un groupe topologique localement compact, on dit que l'espace $\mathcal{E} := (\mathbb{D}_\mathcal{E}, \mathbb{T}_\mathcal{E})$ est un* ***espace d'échantillonnage*** *si les conditions suivantes sont respectées :*
- *le groupe $\mathbb{T}_\mathcal{E}$ agit sur $\mathbb{D}_\mathcal{E}$ de façon transitive,*
- *le groupe $\mathbb{T}_\mathcal{E}$ agit sur $\mathbb{D}_\mathcal{E}$ en préservant les distances.*

La preservation des distances implique que

$$\forall x, x' \in \mathbb{D}_\mathcal{E}, \forall \tau \in \mathbb{T}_\mathcal{E}, \text{dist}_\mathcal{E}(x, x') = \text{dist}_\mathcal{E}(x * \tau, x' * \tau)\,. \tag{3.8}$$

D'un point de vue concret, le domaine d'échantillonnage $\mathbb{D}_\mathcal{E}$ est le domaine auquel appartiennent les points de l'ensemble $\mathcal{S}$, et le groupe $\mathbb{T}_\mathcal{E}$ est un ensemble d'actions sur le domaine, généralement des translations ou des rotations. On appelle les éléments de $\mathbb{T}_\mathcal{E}$ des **actions**. Dans le reste de ce manuscrit, x sera toujours un élément de $\mathbb{D}_\mathcal{E}$ et τ un élément de $\mathbb{T}_\mathcal{E}$. Par ailleurs, on note $\tau(\mathcal{S})$ l'ensemble de points dont tous les points ont été déplacé par l'action τ, et τ_0 l'action neutre de $\mathbb{T}_\mathcal{E}$.

Dans la section suivante, nous détaillons un ensemble d'espaces d'échantillonnage qui nous semblent importants pour l'informatique graphique, en précisant à chaque fois le domaine de définition, la métrique et l'ensemble des actions. Nous précisons aussi la valeur de la mesure appliquée sur l'ensemble du domaine, noté $\mu_\mathcal{E}(\mathbb{D}_\mathcal{E})$ et la dimension topologique de cet espace. Enfin, une courte explication de l'utilisation de chaque espace dans l'informatique graphique est donnée.

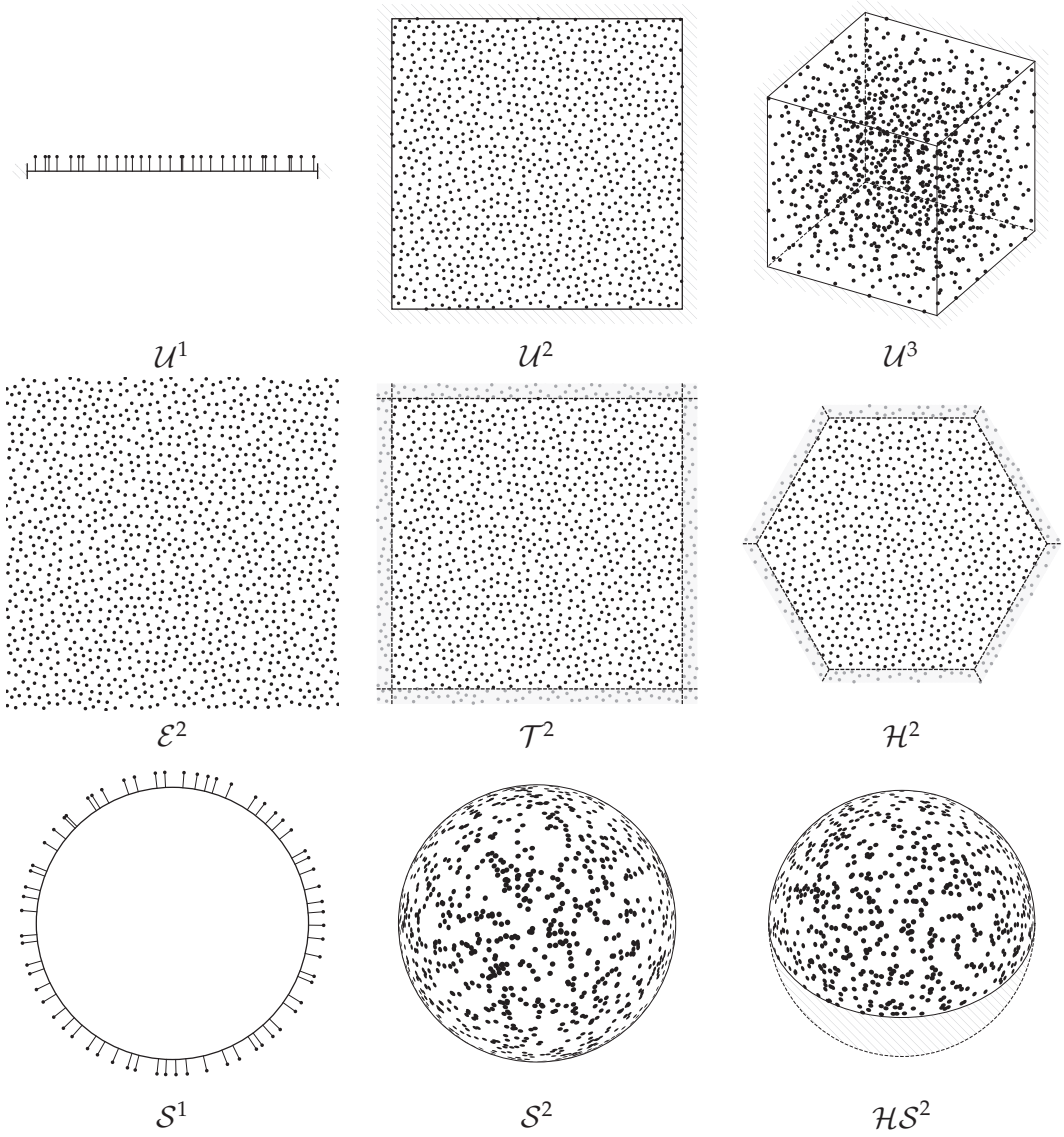
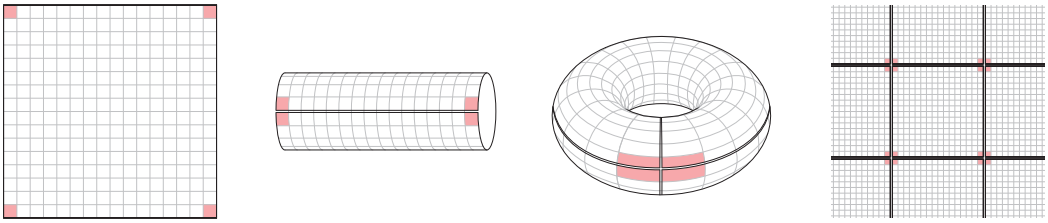


FIGURE 3.1 – *Différents espaces d'échantillonnage. De gauche à droite et de haut en bas, les espaces linéaire unitaire, carré unitaire, cubique unitaire, euclidien en dimension deux, carré périodique, hexagonal périodique, circulaire, sphérique et hémisphérique.*



(a) (b) (c) (d)

FIGURE 3.2 – *Différentes représentations de l'espace carré périodique.* ***(a) :*** *L'espace carré périodique dans une représentation classique. Les quatre carrés rouges ont l'air éloignés, pourtant, ils sont proches les uns des autres.* ***(b) et (c) :*** *On peut représenter l'espace carré périodique comme un espace replié sur lui-même dans chaque dimension. Dans cette représentation, on voit pourquoi les quatre carrés rouges sont proches.* ***(d) :*** *On peut aussi représenter cet espace comme un domaine périodiquement répété. Ici, les quatre carrés rouges sont également proches.*

3.2.1 Espace euclidien $\mathcal{E}^d$

L'espace euclidien $\mathcal{E}^d$ est un espace infini dont le domaine de définition est $\mathbb{D}_{\mathcal{E}^d} := \mathbb{R}^d$. Sa mesure $\mu_{\mathcal{E}}(\mathbb{D}_{\mathcal{E}^d})$ est infini et sa dimension topologique est égale à d. La métrique de cet espace est la norme euclidienne

$$\mathrm{dist}_{\mathcal{E}^d}(x, x') := \|x - x'\| . \tag{3.9}$$

L'action correspond à la translation classique dans l'espace euclidien. Dans cet espace, chaque point est un vecteur de coordonnées. Les translations sont quand à elle représentable par des vecteurs. L'ensemble des actions $\mathbb{T}_{\mathcal{E}}$, représentées par des vecteurs, est égale à $\mathbb{R}^d$. L'espace $\mathcal{E}^d$ étant infini, il est rarement utilisé pour l'intégration Monte Carlo. En revanche, on trouve quelques applications pour le placement d'objet où l'ensemble de points est généré au fur et à mesure que l'on explore l'espace. Dans le cadre de ce manuscrit, nous aborderons tout de même ce domaine dans nos développements, puisqu'il est relativement simple à étudier et qu'il fournit un bon point de départ à l'étude des autres espaces. La figure (3.1) montre un exemple d'un ensemble de points dans l'espace $\mathcal{E}^2$.

3.2.2 Espace hyper-cubique périodique $\mathcal{T}^d$

L'espace hyper-cubique périodique est un espace d'échantillonnage dont le domaine de définition est $\mathbb{D}_{\mathcal{T}^d} := \mathbb{R}^d/\mathbb{Z}^d$, c'est à dire le quotient entre l'espace des réels et l'espace des entiers relatifs. Cette notation veut dire que les points ne peuvent appartenir qu'au sous-domaine $[0, 1[^d$. La mesure de cet espace est donc égale à 1 et sa dimension topologique égale à d. On peut voir cet espace comme un domaine dont les éléments sont répétés périodiquement à l'aide des vecteurs de la base canonique. Cet espace peut aussi être vu comme un espace replié sur lui-même tel que si l'on se déplace dans une direction, on fini par retomber au même endroit. La figure (3.2) illustre ces deux représentations. Pour des raisons pratiques, nous avons choisi d'utiliser la représentation basée sur la périodicité. La figure (3.1) montre un exemple d'un ensemble de points évoluant dans cet espace. La métrique est définie par

$$\mathrm{dist}_{\mathcal{T}^d}(x, x') := \min_{p \in \mathbb{Z}^d} \|x - x' + p\| , \tag{3.10}$$

prenant donc en compte la distance la plus petite entre toutes les copies possibles de chaque point répliqué. Comme l'espace est périodique, seul un sous-ensemble des translations euclidiennes est nécessaire pour couvrir tout le domaine d'échantillonnage. On écrira donc que $\mathbb{T}_{\mathcal{T}^d} := [0,1[^d$.

3.2.3 Espace hexagonal periodique $\mathcal{H}^2$

La périodicité décrite dans l'espace hyper-cubique periodique n'est pas la seule solution pour obtenir un espace periodique. En dimension deux, plusieurs types de périodicités existent dont la périodicité hexagonale [Grunbaum and Shephard, 1986, sec. 1.3]. Cette périodicité est caractérisée par le couple de vecteur $b_0 := (0,1)$ et $b_1 := (\sqrt{3}/2, 1/2)$. Le domaine de définition est un hexagone régulier où chaque coté à une longueur de $1/\sqrt{3}$. La mesure du domaine $\mu_{\mathcal{E}}(\mathbb{D}_{\mathcal{H}^2})$ est de $\sqrt{3}/2$ et sa métrique est donnée par

$$\mathrm{dist}_{\mathcal{H}^2}(x, x') := \min_{(i,j)\in\mathbb{Z}^d} \|x - x' + i\, b_0 + j\, b_1\| \,. \tag{3.11}$$

Comme pour l'espace hyper-cubique périodique, l'ensemble des actions de l'espace hexagonale périodique est un sous-ensemble de l'espace euclidien défini par $\mathbb{T}_{\mathcal{H}^2} = \mathbb{D}_{\mathcal{H}^2}$. Il est utile de préciser aussi que cet espace peut être représenté sous la forme d'un parallélogramme periodique en utilisant les mêmes vecteurs b_0 et b_1. Nous avons choisi ici, pour des raisons pédagogiques, de conserver la représentation hexagonale. La figure (3.1) montre un exemple d'ensemble de points dans l'espace périodique hexagonale.

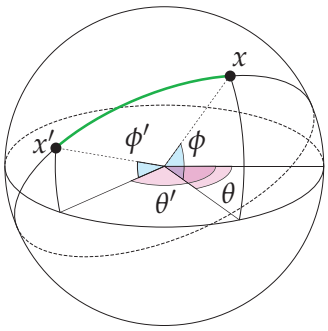
Cet espace n'a pas d'application directe en rendu graphique. Cependant, il est intéressant à étudier car il permet de mieux rendre compte du comportement de certains algorithmes d'échantillonnage. En effet, cet espace peut être pavé par un pavage hexagonale. Comme ce pavage est le plus compact possible dans l'espace euclidien de dimension 2, beaucoup de méthodes de minimisation d'énergie convergent vers ce type de structure. Or comme nous l'avons vu dans l'introduction, les structures régulières sont néfastes pour le rendu graphique. Il est donc utile de tester les méthodes d'échantillonnage dans ce type d'espace pour être sûr qu'elles ne créent pas de structures régulières. Cet espace est au cœur de la méthode d'échantillonnage que nous avons présenté dans [Wachtel *et al.*, 2014].

3.2.4 Espace sphérique $\mathcal{S}^2$

L'espace sphérique est le premier espace représentant des surfaces non-euclidiennes que nous voyons dans ce chapitre. Son domaine de définition est donné par

$$\mathbb{D}_{\mathcal{S}^2} := \left\{ \|x\| = 1 \quad | \quad x \in \mathbb{R}^3 \right\} , \tag{3.12}$$

et sa mesure est donnée par $\mu_{\mathcal{E}}(\mathbb{D}_{\mathcal{S}^2}) = 4\pi$. En informatique graphique, ce domaine sert à représenter des directions dans un espace euclidien de dimension trois. Il est particulièrement utilisé pour intégrer la lumière arrivant en un point de l'espace. L'espace sphérique est de dimension topologique deux, il est donc nécessaire d'avoir deux coordonnées pour définir la position d'un point. La façon usuelle est d'utiliser le couple longitude, θ (l'équivalent du fuseau horaire) et latitude ϕ (la distance par rapport à



l'équateur), comme le montre la figure ci-contre. La métrique de cet espace entre les points x et x' est donnée par la longueur du plus petit arc du grand cercle passant par ces deux points. Dans la figure ci-contre, cet arc du grand cercle est montré en vert. Sur la sphère, les déplacements se font par rotation. L'ensemble $\mathbb{T}_{\mathcal{S}^2}$ est donc le groupe des rotations en 3D, souvent appelé $SO(3)$. La figure (3.1) montre un exemple d'ensemble de points dans l'espace sphérique.

3.2.5 Espace discret régulier $\mathcal{L}^d$

Les espaces discrets réguliers sont défini par le domaine $\mathbb{D}_{\mathcal{L}^d} := \mathbb{Z}^d$. La mesure utilisée dans ce domaine est la mesure de comptage et cette mesure appliquée sur l'ensemble du domaine est infinie. Le groupe des actions correspond lui aussi à $\mathbb{T}_{\mathcal{L}^d} := \mathbb{Z}^d$. La métrique utilisée est celle de l'espace euclidien, bien que d'autres soient possibles. En informatique graphique, cet espace sert en particulier à représenter des images, où l'information de chaque pixel est rattachée à un élément discret de l'espace discret régulier.

3.3 Espace d'échantillonnage restreint

Il est en général difficile d'utiliser numériquement les espaces dont la mesure est infinie. Pour cela, on a recours aux espaces d'échantillonnage restreints.

Définition 3.8 *Soit $\mathcal{E}$ un espace d'échantillonnage et A un sous-ensemble de $\mathbb{D}_{\mathcal{E}}$, on dit que $\mathcal{E}_{\backslash A}$ est un* ***espace d'échantillonnage restreint*** *sur A si la mesure, la métrique et l'ensemble des actions de $\mathcal{E}_{\backslash A}$ sont les mêmes que ceux de $\mathcal{E}$, et que $\mathbb{D}_{\mathcal{E}_{\backslash A}} := A \subset \mathbb{D}_{\mathcal{E}}$.*

La plupart des propriétés des espaces d'échantillonnage ne s'appliquent pas aux espaces d'échantillonnage restreints. Par exemple, l'application d'une action sur un point peut faire sortir ce point du sous-domaine $\mathbb{D}_{\mathcal{E}_{\backslash A}}$. Pour s'en convaincre, il suffit d'imaginer l'espace euclidien $\mathcal{E}^2$ restreint au sous-domaine $[0,1[^2$ et d'appliquer une translation par un vecteur $(2,0)$. Tout point du sous-domaine $[0,1[^2$ se trouvera alors en dehors du domaine restreint après application de cette translation. En conséquence, l'ensemble $\mathbb{T}_{\mathcal{E}_{\backslash A}}$ n'est plus une action de groupe sur l'ensemble $\mathbb{D}_{\mathcal{E}_{\backslash A}}$ et $\mathbb{D}_{\mathcal{E}_{\backslash A}}$ n'est plus un espace homogène. L'intérêt de l'espace d'échantillonnage réside simplement dans le fait que les points de l'ensemble $\mathcal{S}$ ne peuvent appartenir qu'au sous-ensemble $\mathbb{D}_{\mathcal{E}_{\backslash A}}$. Les règles géométriques restent, elles, dictées par l'espace d'échantillonnage $\mathcal{E}$.

3.3.1 Espace hyper-cubique unitaire $\mathcal{U}^d$

L'espace hyper-cubique unitaire $\mathcal{U}^d$ est un espace similaire à l'espace euclidien, mais cette fois-ci restreint au domaine de définition $\mathbb{D}_{\mathcal{U}^d} := [0,1]^d$. Sa mesure $\mu_{\mathcal{E}}\left(\mathcal{U}^d\right)$ est égale à 1 et sa métrique est la même que celle de l'espace euclidien. Cet espace est un des plus utilisés puisqu'il permet de représenter une texture ou un pixel à échantillonner. En dimension 1, on parlera d'espace linéaire unitaire, en dimension 2 d'espace carré unitaire et en dimension 3 d'espace cubique unitaire. La figure (3.1) montre des exemples d'ensembles de points dans ces trois espaces. Dans nos figures, nous représenterons toujours ce domaine entouré de hachures pour symboliser la restriction du domaine.

3.3.2 Espace hémisphérique $\mathcal{HS}^2$

Comme nous l'avons dit, l'espace d'échantillonnage sphérique sert à représenter des directions. Cependant, lorsque le point dont on veut calculer la couleur se trouve sur une surface, il n'est pas nécessaire d'échantillonner les directions pointant vers la surface [1]. On a donc recours à un espace d'échantillonnage $\mathcal{S}^2$ restreint à une hémisphère. Par convention, on définira cette hémisphère par

$$\mathbb{D}_{\mathcal{HS}^2} := \{\phi > 0 : (\phi, \theta) \in \mathbb{D}_{\mathcal{S}^2}\} , \tag{3.13}$$

où (ϕ, θ) sont des coordonnées sphériques telles que $\phi \in [-\pi/2, \pi/2]$ correspond à la latitude. La mesure $\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{HS}^2}\right)$ du domaine restreint est égale à 2π. La figure (3.1) montre un exemple d'ensembles de points dans l'espace d'échantillonnage hémisphérique.

3.4 Intégration et espace d'échantillonnage

Étant donné que dans le contexte de cette thèse, le but de l'échantillonnage est d'approximer une intégrale, il est nécessaire de définir ce qu'est une intégrale dans le domaine d'échantillonnage. Comme nous l'avons vu dans l'introduction, une intégrale nécessite une mesure. Or nos espaces d'échantillonnage sont tous dotés d'une mesure, qui sera précisément celle utilisée pour le calcul d'intégrale. On écrit donc une intégrale d'une fonction $f : \mathbb{D}_{\mathcal{E}} \to \mathbb{C}$ sur le domaine d'échantillonnage comme

$$\int_{\mathbb{D}_{\mathcal{E}}} f(x)\mu_{\mathcal{E}}(\mathrm{d}x) , \tag{3.14}$$

ou tout simplement

$$\int_{\mathbb{D}_{\mathcal{E}}} f(x)\mathrm{d}x , \tag{3.15}$$

puisque l'on considère que la mesure est implicitement donnée par le domaine de définition de la fonction f. Parfois, nous aurons besoin d'intégrer sur le groupe des actions $\mathbb{T}_{\mathcal{E}}$. On se pose alors la question de savoir quelle mesure l'on peut utiliser.

Définition 3.9 *[Barut and Rączka, 1986, sec. 2.3] Soit G un groupe topologique localement compact, on appelle* ***mesures de Haar*** *les mesures d'un groupe topologique localement compact qui sont invariantes, c'est à dire les mesures μ_H qui vérifient pour tout borélien A et pour tout $g \in G$*

$$\mu_H\left(\{g \star g' : g' \in A\}\right) = \mu_H\left(A\right) . \tag{3.16}$$

La mesure de Haar est aussi invariante à droite. La construction de la mesure de Haar n'est pas explicite. En revanche, on sait que tout groupe topologique localement compact admet au moins une mesure de Haar, et que toutes les mesures de Haar non nulles d'un groupe topologique localement compact sont proportionnelles [Barut and Rączka, 1986, sec. 2.3]. Dans ce manuscrit, nous utilisons la mesure de Haar à chaque fois que l'on intégrera sur le groupe des actions $\mathbb{T}_{\mathcal{E}}$. L'invariance de la mesure de Haar nous permet d'écrire pour tout τ' la relation suivante

$$\int_{\mathbb{T}_{\mathcal{E}}} f(\tau)\mu_H(\mathrm{d}\tau) = \int_{\mathbb{T}_{\mathcal{E}}} f(\tau' \star \tau)\mu_H(\mathrm{d}\tau) . \tag{3.17}$$

1. Nous assumons ici que la surface n'est ni luminescente ni translucide.

Dans ce manuscrit, les intégrales sur $\mathbb{T}_{\mathcal{E}}$ utiliseront toujours la mesure de Haar de façon implicite. Nous ne préciserons donc plus la mesure de Haar dans ces intégrales.

Lemme 3.1 *Soit $\mathcal{E}$ un espace d'échantillonnage, on a, pour tout $x_0 \in \mathbb{D}_{\mathcal{E}}$, la relation suivante*

$$\int_{\mathbb{D}_{\mathcal{E}}} f(x)\mathrm{d}x = \frac{\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)}{\mu_H\left(\mathbb{T}_{\mathcal{E}}\right)} \int_{\mathbb{T}_{\mathcal{E}}} f(x_0 * \tau)\mathrm{d}\tau\,. \tag{3.18}$$

Preuve 3.1 *Comme l'action de $\mathbb{T}_{\mathcal{E}}$ sur $\mathbb{D}_{\mathcal{E}}$ préserve les distances et les mesures, on a, pour tout $\tau \in \mathbb{T}_{\mathcal{E}}$, l'égalité suivante*

$$\int_{\mathbb{D}_{\mathcal{E}}} f(x)\mathrm{d}x = \int_{\mathbb{D}_{\mathcal{E}}} f(x * \tau)\mathrm{d}x\,. \tag{3.19}$$

On peut par le même raisonnement écrire que l'intégrale d'une fonction sur le domaine est égale à la moyenne de cette intégrale sur le groupe des actions

$$\int_{\mathbb{D}_{\mathcal{E}}} f(x)\mathrm{d}x = \frac{1}{\mu_H\left(\mathbb{T}_{\mathcal{E}}\right)} \int_{\mathbb{T}_{\mathcal{E}}} \int_{\mathbb{D}_{\mathcal{E}}} f(x * \tau)\mathrm{d}x\mathrm{d}\tau\,. \tag{3.20}$$

Comme $\mathbb{T}_{\mathcal{E}}$ est transitive, on sait que pour tout $x, x' \in \mathbb{D}_{\mathcal{E}}$, il existe τ' tel que

$$\begin{aligned}\int_{\mathbb{T}_{\mathcal{E}}} f(x * \tau)\mathrm{d}\tau &= \int_{\mathbb{T}_{\mathcal{E}}} f((x' * \tau') * \tau)\mathrm{d}\tau \\ &= \int_{\mathbb{T}_{\mathcal{E}}} f((x' * (\tau' \star \tau))\mathrm{d}\tau\,.\end{aligned} \tag{3.21}$$

Le fait que la mesure de Haar soit invariante sur $\mathbb{T}_{\mathcal{E}}$ permet alors d'écrire que

$$\int_{\mathbb{T}_{\mathcal{E}}} f(x * \tau)\mathrm{d}\tau = \int_{\mathbb{T}_{\mathcal{E}}} f(x' * \tau)\mathrm{d}\tau\,. \tag{3.22}$$

En en déduit alors que le point x peut être substitué par tout autre point de $\mathbb{D}_{\mathcal{E}}$. Soit x_0 un point arbitraire de $\mathbb{D}_{\mathcal{E}}$, on peut écrire à partir de Eq. (3.20) que

$$\begin{aligned}\int_{\mathbb{D}_{\mathcal{E}}} f(x)\mathrm{d}x &= \frac{1}{\mu_H\left(\mathbb{T}_{\mathcal{E}}\right)} \int_{\mathbb{D}_{\mathcal{E}}} \int_{\mathbb{T}_{\mathcal{E}}} f(x_0 * \tau)\mathrm{d}\tau\mathrm{d}x \\ &= \frac{\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)}{\mu_H\left(\mathbb{T}_{\mathcal{E}}\right)} \int_{\mathbb{T}_{\mathcal{E}}} f(x_0 * \tau)\mathrm{d}x\,.\end{aligned} \tag{3.23}$$

□

L'intérêt de ce lemme est qu'il montre qu'il est équivalent d'intégrer une fonction sur le domaine d'échantillonnage ou sur le groupe des actions. Cette propriété sera utilisée notamment dans les derniers chapitres.

3.5 Homogénéité statistique

Définition 3.10 *Soit $\mathcal{E}$ un espace d'échantillonnage, et $\mathcal{S}$ un ensemble de points stochastique dans $\mathcal{E}$, on dit que $\mathcal{S}$ est* ***homogène*** *s'il est invariant par action, c'est à dire si pour tout τ, $\tau(\mathcal{S})$ est statistiquement équivalent à $\mathcal{S}$.*

On note $\mathfrak{M}_{\text{Homo.}}$ la classe des ensembles de points homogènes. On pourra écrire lorsque $\mathcal{S} \in \mathfrak{M}_{\text{Homo.}}$ que

$$\forall \tau \in \mathbb{T}_{\mathcal{E}}, \quad f(\mathcal{S}) = f(\tau(\mathcal{S}))\,, \tag{3.24}$$

ou bien encore que

$$f(\mathcal{S}) = \frac{1}{\mu_H\left(\mathbb{T}_{\mathcal{E}}\right)} \int_{\mathbb{T}_{\mathcal{E}}} f(\tau(\mathcal{S}))\mathrm{d}\tau\,. \tag{3.25}$$

Cette équation sera souvent utilisée pour démontrer certaines propriétés des ensembles de points homogènes. Il est important de bien voir que $\mathcal{S}$ est ici une variable aléatoire et donc $f(\mathcal{S})$ l'est également. L'équation (3.24) montre donc une équivalence statistique entre les deux termes.

Lemme 3.2 *Soit $\mathcal{S}$ un ensemble de points homogène dans un espace d'échantillonnage $\mathcal{E}$, la densité de probabilité du k-ième point de $\mathcal{S}$ est uniforme sur le domaine d'échantillonnage :*

$$PDF_{s_k} = \frac{1}{\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)}\,. \tag{3.26}$$

Preuve 3.2 *D'après Eq. (3.25), on a*

$$PDF_{s_k}(x) = \frac{1}{\mu_H\left(\mathbb{T}_{\mathcal{E}}\right)} \int_{\mathbb{T}_{\mathcal{E}}} PDF_{s_k * \tau}(x)\mathrm{d}\tau\,. \tag{3.27}$$

Évaluer $PDF_{s_k*\tau}(x)$ revient à évaluer $PDF_{s_k}(x * \tau^{-1})$, on peut donc écrire que

$$PDF_{s_k}(x) = \frac{1}{\mu_H\left(\mathbb{T}_{\mathcal{E}}\right)} \int_{\mathbb{T}_{\mathcal{E}}} PDF_{s_k}(x * \tau)\mathrm{d}\tau\,, \tag{3.28}$$

où l'on a remplacé τ^{-1} par τ par changement de variable. En utilisant Lemme (3.1), on peut transformer l'intégrale sur le groupe des actions en intégrale sur le domaine d'échantillonnage. On obtient alors

$$PDF_{s_k}(x) = \frac{1}{\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)} \int_{\mathbb{D}_{\mathcal{E}}} PDF_{s_k}(x')\mathrm{d}x'\,. \tag{3.29}$$

La preuve est terminée en remarquant que l'intégrale d'une densité de probabilité est par définition égale à 1. □

L'homogénéité est une propriété des ensembles de points régulièrement étudiée dans la littérature. On la retrouve par exemple en physique dans l'étude des gazes [Hansen and McDonald, 1990]. Elle est aussi utilisée en informatique graphique, cette fois sous le nom de *widesense stationary process* [Dippé and Wold, 1985]. La version que nous proposons dans ce manuscrit est une généralisation permettant de parler d'ensemble de points homogène dans des espaces d'échantillonnage arbitraires. Toutefois, il est important de garder à l'esprit que l'homogénéité d'un ensemble de points dépend du groupe des actions utilisé. Ainsi, un ensemble de points sera homogène pour une action de groupe donnée, mais ne le sera pas forcement pour une autre. C'est une des raisons pour lesquelles nous avons intégré l'action de groupe dans l'espace d'échantillonnage. Ainsi, il n'y a aucune ambiguïté quant au groupe utilisé, du moment que l'on sait dans quel espace d'échantillonnage un ensemble de points évolue. Nous ferons souvent référence à la notion d'homogénéité afin de démontrer un certain nombre de propriétés, étant donné, comme nous le verrons dans le chapitre suivant,

que beaucoup de méthodes d'échantillonnage stochastique produisent des ensembles de points homogènes. De plus, il est extrêmement simple de créer un ensemble de points homogène à partir d'un ensemble de points quelconque, comme le montre le lemme suivant.

Lemme 3.3 *Soit $\mathcal{S}$ un ensemble de points dans $\mathcal{E}$ et ν une variable aléatoire représentant une action dans $\mathbb{T}_{\mathcal{E}}$ telle que la densité de probabilité de ν soit uniforme, alors l'ensemble de points*

$$\mathcal{S}' := \{s_k * \nu : s_k \in \mathcal{S}\} \ , \tag{3.30}$$

est un ensemble de points homogène de $\mathcal{E}$.

Preuve 3.3 *On veut montrer que $\nu(\mathcal{S})$ est homogène, c'est à dire que la fonction de densité de probabilité de $\tau'(\nu\mathcal{S})$ est la même que celle de $\nu\mathcal{S}$. On peut écrire l'égalité suivante*

$$PDF_{\tau'(\nu(\mathcal{S}))}(s^N) = PDF_{\nu(\mathcal{S})}(\tau'^{-1}(s^N)) \, . \tag{3.31}$$

Sachant que les variables aléatoires $\mathcal{S}$ et ν sont indépendante, la fonction de densité de probabilité de $\nu\mathcal{S}$ est égale à l'espérance de la fonction de densité de probabilité de $\mathcal{S}$ par rapport à ν. On écrit alors que

$$PDF_{\tau'(\nu(\mathcal{S}))}(s^N) = \int_{\mathbb{T}_{\mathcal{E}}} PDF_{\tau(\mathcal{S})}(\tau'^{-1}(s^N)) PDF_{\nu}(\tau) \mathrm{d}\tau' \, . \tag{3.32}$$

ν étant uniforme, l'équation se simplifie de la façon suivante,

$$PDF_{\tau'(\nu(\mathcal{S}))}(s^N) = \frac{1}{\mu_H(\mathbb{T}_{\mathcal{E}})} \int_{\mathbb{T}_{\mathcal{E}}} PDF_{\tau(\mathcal{S})}(\tau'^{-1}(s^N)) \mathrm{d}\tau \, . \tag{3.33}$$

Comme la mesure de Haar est invariante, on à l'égalité suivante,

$$\begin{aligned} PDF_{\tau'(\nu(\mathcal{S}))}(s^N) &= \frac{1}{\mu_H(\mathbb{T}_{\mathcal{E}})} \int_{\mathbb{T}_{\mathcal{E}}} PDF_{\mathcal{S}}(\tau^{-1}(\tau'^{-1}(s^N))) \mathrm{d}\tau \\ &= \frac{1}{\mu_H(\mathbb{T}_{\mathcal{E}})} \int_{\mathbb{T}_{\mathcal{E}}} PDF_{\mathcal{S}}((\tau'^{-1} \star \tau^{-1})(s^N))) \mathrm{d}\tau \\ &= \frac{1}{\mu_H(\mathbb{T}_{\mathcal{E}})} \int_{\mathbb{T}_{\mathcal{E}}} PDF_{\mathcal{S}}(\tau^{-1}(s^N)) \mathrm{d}\tau \\ &= \frac{1}{\mu_H(\mathbb{T}_{\mathcal{E}})} \int_{\mathbb{T}_{\mathcal{E}}} PDF_{\tau(\mathcal{S})}(s^N) \mathrm{d}\tau \, . \end{aligned} \tag{3.34}$$

Finalement, on peut repasser de l'intégrale à la fonction de densité de probabilité de $\nu(\mathcal{S})$ et obtenir

$$PDF_{\tau'(\nu(\mathcal{S}))}(s^N) = PDF_{\nu(\mathcal{S})}(s^N) \, . \tag{3.35}$$

On en conclut que $\mathcal{S}'$ est statistiquement équivalent à $\tau'(\mathcal{S}')$ et donc que $\mathcal{S}'$ est homogène. □

Conclusion

Nous avons défini dans ce chapitre l'environnement théorique nécessaire à l'analyse des ensembles de points pour l'intégration Monte Carlo. Les choix que nous avons fait dans la modélisation mathématique de l'ensemble de points et de l'espace d'échantillonnage sont, à notre connaissance, nouveaux au sein de la communauté du rendu d'images et peuvent donc par conséquent être considérés comme des contributions. Le chapitre suivant offre une vision plus concrète de ce que peuvent être les ensembles de points dédiés à l'intégration Monte Carlo, en parcourant les techniques d'échantillonnage les plus populaires.

Méthodes d'échantillonnage 4

Dans ce chapitre, nous détaillons une liste non-exhaustive de méthodes d'échantillonnage, regroupées en cinq catégories. Tout d'abord, nous présentons les méthodes d'échantillonnage stochastiques les plus simples (bruit blanc et disques de Poisson). S'ensuivent, les méthodes utilisant une partition de l'espace à l'aide d'un pavage (grilles régulières et ensemble de points *jittered*), puis les algorithmes d'optimisations des ensembles de points, de la relaxation de Lloyd jusqu'au contrôle spectral. Finalement, nous abordons les échantillonneurs à base de pavages hiérarchiques et les séquences à basse discrépance. Cette vue d'ensemble permet de donner un aperçu des stratégies mises en place pour obtenir un bon motif d'échantillonnage pour l'intégration Monte Carlo. Pour chaque méthode, nous précisons les raisons derrières chaque approche et les bases des algorithmes utilisés. L'analyse de la qualité de chaque méthode est réservée au chapitre 8, puisqu'il nous est nécessaire avant d'introduire chaque outil d'analyse.

4.1 Méthodes d'échantillonnage bruit blanc

Le bruit blanc, parfois appelé distribution de Poisson, consiste à choisir aléatoirement et uniformément les points dans le domaine d'échantillonnage, indépendamment les uns des autres. On dit qu'un ensemble de points est bruit blanc, noté $\mathcal{S} \in \mathfrak{M}_{\text{WN}}$, s'il respecte les conditions suivantes :

$$\begin{cases} \forall k, \quad PDF_{\mathbf{s}_k}(x) = \dfrac{1}{\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)} \\ \forall i, k, \quad i \neq k \Rightarrow \mathbf{s}_i \text{ et } \mathbf{s}_k \text{ sont indépendants}\,. \end{cases} \tag{4.1}$$

L'intérêt de cette méthode est qu'elle ne génère pas d'*aliasing* quand on l'utilise pour l'intégration. Mais cet avantage est compensé par le niveau de bruit important qu'elle produit. La figure (4.1) montre des exemples d'ensembles des points de type bruit blanc dans plusieurs espaces d'échantillonnage. On remarque tout de suite que les ensembles de points ne sont pas bien répartis dans le domaine, c'est à dire que de grandes régions du domaines ne sont pas échantillonnées.

Par définition, ces ensembles de points sont stochastiques. Ils sont aussi homogènes, puisque qu'ils ne sont définis que par une loi de probabilité uniforme, qui est donc invariante par translation. On peut donc écrire que

$$\begin{cases} \mathfrak{M}_{\text{WN}} \subset \mathfrak{M}_{\text{Stocha.}} \\ \mathfrak{M}_{\text{WN}} \subset \mathfrak{M}_{\text{Homo.}}\,. \end{cases} \tag{4.2}$$

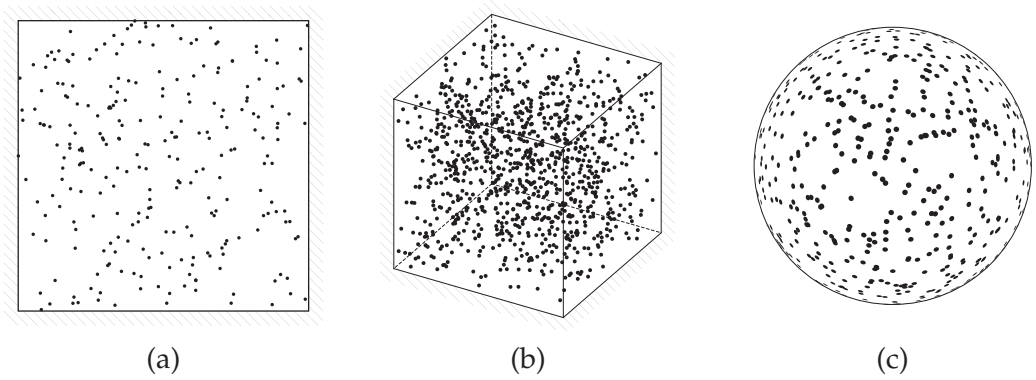


FIGURE 4.1 – *Exemple d'ensemble de points bruit blanc dans différents espaces d'échantillonnage :* ***(a)*** *dans l'espace* $\mathcal{U}^2$, ***(b)*** *dans l'espace* $\mathcal{U}^3$ *et* ***(c)*** *dans l'espace* $\mathcal{S}^2$.

Dans les espaces hyper-cubiques, les coordonnées des points sont générées en choisissant aléatoirement dans chaque dimension un nombre entre 0 et 1. Dans l'espace sphérique, le problème est un peu plus compliqué si l'on ne dispose que d'un générateur de nombre aléatoire uniforme sur l'intervalle $[0, 1[$, ce qui est souvent le cas. Dans ce cas, plusieurs méthodes existent [Knuth, 1968; Marsaglia and others, 1972] et reposent, soit sur un système de rejet de points, soit sur des transformations de variables aléatoires à l'aide des fonctions trigonométriques. Une façon simple à implémenter consiste à prendre deux variables aléatoires X_0 et X_1 sur l'intervalle $[0, 1[$ et à appliquer les transformations suivantes

$$\phi = 2sin^{-1}(\sqrt{X_0}) - \pi/2 \tag{4.3}$$
$$\theta = 2\pi X_1 \,, \tag{4.4}$$

où (ϕ, θ) correspondent à la latitude et la longitude du point à générer.

4.2 Disques de Poisson

Les méthodes des disques de Poisson[1] sont une alternative à l'échantillonnage bruit blanc. Dans la littérature, on peut aussi les trouver sous le terme de *random sequential addition of hard spheres* (RSA) ou *dart throwing*. Ces ensembles de points sont étudiés dans de nombreux domaines, à la fois en informatique, mais aussi en biologie (distribution de cellules ou d'arbres) et en physique [Torquato, 2002]. En informatique graphique, cette méthode d'échantillonnage a été introduite par Cook [1986] pour sa ressemblance avec la distribution des cônes et des bâtonnets dans la rétine humaine. Le principe de la méthode est une extension de l'échantillonnage bruit blanc ajoutant une condition de distance minimale entre les points.

$$\forall i, j \text{ avec } i \neq j, \mathrm{dist}_{\mathcal{E}}(\mathbf{s}_i, \mathbf{s}_j) \geq 2r \,, \tag{4.5}$$

où r est le rayon des disques centrés en chaque point de l'ensemble. Pour générer l'ensemble de points, la méthode classique consiste à ajouter un point avec une probabilité uniforme sur $\mathbb{D}_{\mathcal{E}}$ à chaque itération. Si la distance entre ce nouveau point et les autres points déjà ajoutés est inférieure à $2r$, c'est à dire si les disques associés à chaque point s'interpénètrent, alors ce nouveau point est rejeté. L'algorithme est terminé quand il

1. À ne pas confondre avec les distributions de Poisson

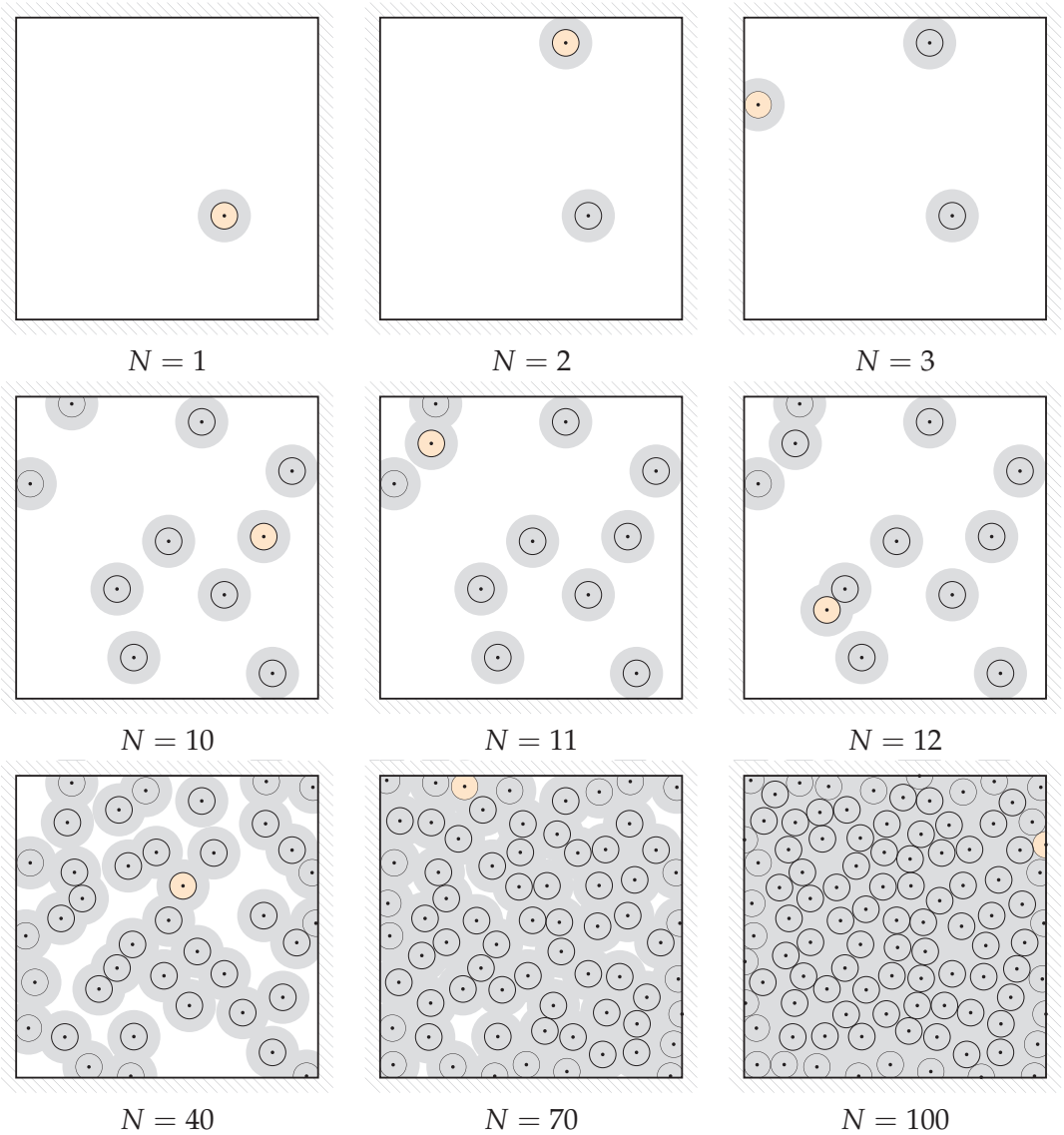


FIGURE 4.2 – *Progression de l'algorithme générant un ensemble de disques de Poisson avec un rayon de* 0.0879694 *dans* $\mathcal{U}^2$. *L'ensemble de points final comporte* 100 *points. Le cercle noir associé à chaque point représente le disque de rayon* r, *et la zone grise représente l'espace où ne peut pas être ajouté de nouveaux points. Pour chaque étape, le dernier disque inséré est affiché en orange. Les étapes de rejet ne sont pas montrées.*

n'est plus possible d'ajouter un nouveau point. On dit alors que l'ensemble de points est maximal. La figure (4.2) montre plusieurs itérations pour générer un ensemble de disques de Poisson.

L'un des problèmes de cette méthode est qu'elle nécessite un nombre d'itérations extrêmement grand pour arriver à un ensemble maximal. Une façon de contourner ce problème est de directement ajouter un point dans les zones valides et donc d'éviter les phases de rejet. Soit N' le nombre de disques déjà acceptés dans la distribution, le domaine $\mathcal{P}_{N'+1}$ dans lequel peut être inséré le $(N'+1)$-ième point est donné par

$$\mathcal{P}_{N'+1} := \mathbb{D}_{\mathcal{E}} \setminus \left(\bigcup_{k=1}^{N'} \{ \|x - \mathbf{s}_k\| < 2r \quad | \quad x \in \mathbb{D}_{\mathcal{E}} \} \right), \tag{4.6}$$

c'est à dire toutes les zones qui sont à plus de $2r$ de distance d'un point. Pour insérer un nouveau point, il faut alors tirer une valeur uniformément dans $\mathcal{P}_{N'+1}$. Toutefois, bien que cette méthode ne nécessitent que N itérations pour arriver à un ensemble de disques de Poisson maximal, la complexité nécessaire pour tirer uniformément un point dans le domaine $\mathcal{P}_{N'+1}$ rend cet algorithme difficile à mettre en place. Les détails d'implémentation de cette approche peuvent être trouvés dans [Gamito and Maddock, 2009] et [Dunbar and Humphreys, 2006].

Un autre problème concernant l'algorithme de génération est que l'on ne sait pas quel sera le nombre de points final de la distribution. Il a été montré que dans l'espace euclidien $\mathcal{E}^d$, le nombre de points dans un ensemble de disques de Poisson tend à suivre une loi normale à mesure que le rayon des disques diminue [Penrose and Yukich, 2002]. Cela signifie que pour deux réalisations utilisant un même rayon r, le nombre de points dans chaque ensemble pourra être très différent. On cherche alors une fonction qui donnerait $\langle N \rangle$, le nombre de points se trouvant en moyenne dans l'ensemble, pour un rayon donné. En dimension 1, ce résultat est connu, puisque la méthode des disques de Poisson est similaire au problème du parking de Rényi [Dvoretzky and Robbins, 1985]. La proportion de l'espace compris dans les disques associés à l'ensemble de points est égale à la constante de Rényi, $R := 0.74759...$, [Solomon, 1967, p. 6]. En revanche, pour les espaces euclidiens de dimension supérieure, il n'est pas possible d'obtenir analytiquement un résultat similaire [Torquato *et al.*, 2006, p. 4]. La seule solution restante est donc d'utiliser des approximations expérimentales (voir [Torquato *et al.*, 2006] pour les dimensions de 1 à 6).

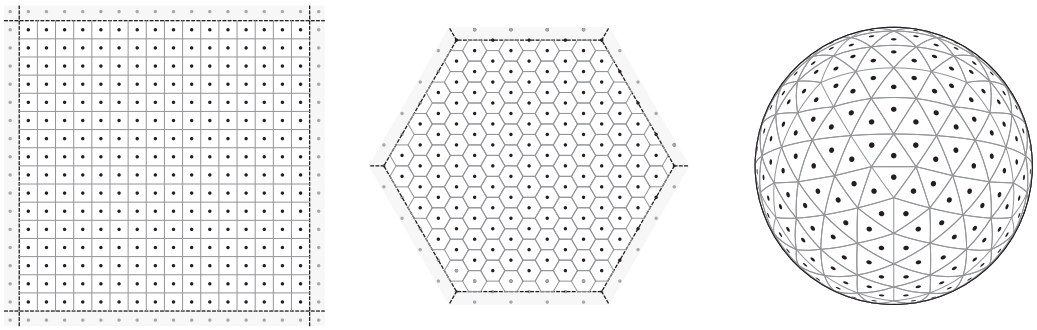
Étant donné que la méthode choisit uniformément les points à ajouter, les ensembles de disques de Poisson sont homogènes. On écrit alors,

$$\begin{cases} \mathfrak{M}_{\text{P.Disk}} \subset \mathfrak{M}_{\text{Stocha.}} \\ \mathfrak{M}_{\text{P.Disk}} \subset \mathfrak{M}_{\text{Homo.}}, \end{cases} \tag{4.7}$$

où $\mathfrak{M}_{\text{P.Disk}}$ est la classe des ensembles de disques de Poisson.

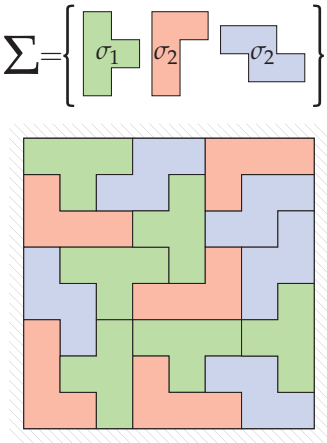
Dans l'espace sphérique, il est également possible de générer des ensembles de disques de Poisson, puisque l'espace d'échantillonnage possède une métrique. Dans ce cas-là, les disques sont remplacés par des calottes. Plusieurs implémentations des disques de Poisson existent sur la sphère [Cline *et al.*, 2009; Gamito and Maddock, 2009; Peyrot *et al.*, 2013].

4.3 Méthodes d'échantillonnage par pavage



(a) (b) (c)

FIGURE 4.3 – *Exemple de grilles régulières dans différents espaces d'échantillonnage :* ***(a)*** *une grille carrée dans l'espace* $\mathcal{T}^2$*,* ***(b)*** *une grille hexagonale dans l'espace* $\mathcal{H}^2$ *et* ***(c)*** *un icosaèdre subdivisé dans l'espace* $\mathcal{S}^2$*.*

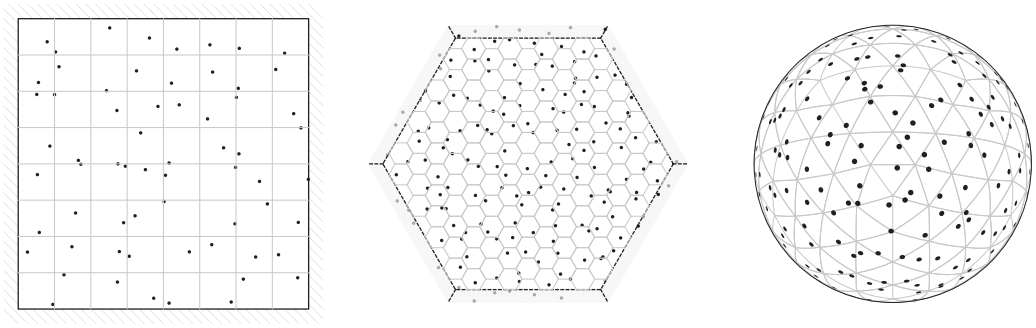


Lorsque l'on échantillonne une fonction, on souhaite que la position des points recouvre l'ensemble du domaine de la façon la plus uniforme possible. Pour cela, une des approches consiste à partitionner le domaine d'échantillonnage en sous-domaines. On appelle **tuile**, notée $\hat{\sigma}$, un sous-domaine de $\mathbb{D}_{\mathcal{E}}$ et **pavage** un ensemble de tuiles recouvrant tout le domaine d'échantillonnage. En général, lorsque l'on veut paver un domaine, on dispose d'un ensemble de **proto-tuiles** $\Sigma := \{\sigma_1, \sigma_2, ...\}$ que l'on va "placer" dans le domaine à l'aide de translation, rotation et réflexion. Une tuile est donc aussi définie comme un couple formé d'une proto-tuile et d'une transformation. La figure ci-contre montre un exemple d'ensemble de proto-tuiles, et un pavage du domaine $\mathbb{D}_{\mathcal{U}^2}$ à partir de cet ensemble de proto-tuiles. Il est possible de construire un ensemble de points à partir d'un pavage en ajoutant un ou plusieurs points par tuile. Les propriétés de l'ensemble de points dépendront du type de pavage et de la façon dont les points ont été ajoutés dans les tuiles.

On notera $\hat{\sigma}_k$ la tuile contenant $\mathbf{s}_k$, et σ_k la proto-tuile de $\hat{\sigma}_k$. La classe des ensembles de points reposant sur un pavage est noté $\mathfrak{M}_{\text{Pavage}}$.

4.3.1 Grille régulière

Les grilles régulières sont des ensembles de points reposant sur des pavages réguliers. Une définition précise des pavages réguliers peut être trouvée dans [Grunbaum and Shephard, 1986, sec. 1.3]. De façon informelle, on définira un pavage régulier comme étant un pavage ne possédant qu'une seule proto-tuile et dont toutes les tuiles ont la même aire. Dans l'espace euclidien de dimension 1, la proto-tuile est un segment. En dimension deux, il existe trois types de pavages réguliers utilisant respectivement le triangle isocèle, le carré et l'hexagone régulier [Grunbaum and Shephard, 1986, sec. 1.3]. En dimension supérieure, le problème est plus difficile, mais des solutions existent. La méthode la plus simple consiste à subdiviser l'espace avec des hyper-cubes dont chaque coté a une longueur de $1/\sqrt[d]{N}$. Pour l'espace sphérique, le problème est plus complexe car il n'existe pas de méthode connue pour générer des pavages régulier à des densités arbitraires. En relâchant les conditions sur le nombre



(a) (b) (c)

FIGURE 4.4 – *Exemple d'ensembles de points* jittered *dans différents espaces d'échantillonnage :* ***(a)*** *dans l'espace $\mathcal{U}^2$ à partir d'une grille carrée,* ***(b)*** *dans l'espace $\mathcal{H}^2$ à partir d'une grille hexagonale et* ***(c)*** *dans l'espace $\mathcal{S}^2$ à partir d'un icosaèdre subdivisé.*

de proto-tuiles ou sur l'aire des tuiles, ont peut trouver des solutions acceptables. On peut par exemple utiliser un icosaèdre subdivisé et re-projeté sur la surface de la sphère, mais à condition d'accepter d'avoir un pavage dont les aires des tuiles sont différentes. On peut également utiliser la méthode de projection HEALPix [Gorski *et al.*, 2005] qui donne de bon résultat, mais en acceptant cette fois d'avoir plusieurs proto-tuiles. Une dernière solution consiste à utiliser les solutions au problème de Thomson [Thomson, 1904], donnant ainsi la distribution de points qui minimise le plus l'énergie de distance. La figure (4.3) montre plusieurs exemples de grilles régulières dans des espaces différents.

Pour passer du pavage à l'ensemble de points, les grilles régulières placent un point au barycentre de chaque tuile. L'ensemble de points obtenu est ainsi déterministe et on écrira donc que

$$\begin{cases} \mathfrak{M}_{\text{Regular}} \subset \mathfrak{M}_{\text{Pavage}} \\ \mathfrak{M}_{\text{Regular}} \subset \mathfrak{M}_{\text{Determ.}} \end{cases} . \tag{4.8}$$

où $\mathfrak{M}_{\text{Regular}}$ est la classe des motifs d'échantillonnage réguliers.

4.3.2 *Jittered sampling*

Les grilles régulières ont l'avantage de partitionner l'ensemble du domaine de façon simple, mais la forte régularité des points génère de nombreux défauts dans le rendu graphique. Pour casser cette régularité, tout en gardant un pavage simple à générer, les méthodes de *jittering* placent un point d'échantillonnage aléatoirement et uniformément dans chaque tuile. Ainsi, la disposition globale des points reste uniforme, mais localement, la position des points est aléatoire. On appelle ces ensembles de points les ensembles *jittered* et on notera leurs classe $\mathfrak{M}_{\text{Jittered}}$. Les ensembles *jittered* ont les propriétés suivantes :

$$\begin{cases} \forall k \in \{1, ..., N\}\,, \quad PDF_{\mathbf{s}_k}(x) = \dfrac{\chi_{\hat{\sigma}_k}(x)}{\mu_{\mathcal{E}}\left(\hat{\sigma}_k\right)} \\ \forall i, k \in \{1, ..., N\}\,, \quad i \neq k \Rightarrow \mathbf{s}_i \text{ et } \mathbf{s}_k \text{ sont indépendants}\,. \end{cases} \tag{4.9}$$

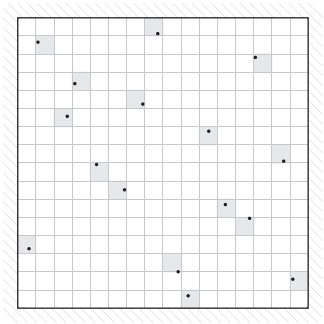
En informatique graphique, ils ont été introduits dans la deuxième moitié des années 80 [Dippé and Wold, 1985; Cook, 1986; Mitchell, 1991]. Jusqu'à aujourd'hui, ils sont très utilisés car ils donnent de bons résultats tout en restant très simples à générer.

Les ensembles de points *jittered* appartiennent à la classe des ensembles stochastiques et des ensembles reposant sur des pavages. En revanche, ils ne sont pas homogènes. En effet, les propriétés des ensembles de points *jittered* ne sont pas les mêmes sur tout le domaine et dépendent du pavage utilisé. Par exemple, il est possible que deux points soient proches à la frontière entre deux tuiles, mais il est impossible de trouver deux points proches au centre d'une tuile puisqu'il n'y a qu'un point par tuile. On a donc

$$\begin{cases} \mathfrak{M}_{\text{Jittered}} \subset \mathfrak{M}_{\text{Stocha.}} \\ \mathfrak{M}_{\text{Jittered}} \subset \mathfrak{M}_{\text{Pavage}} \\ \mathfrak{M}_{\text{Jittered}} \not\subset \mathfrak{M}_{\text{Homo.}} \,. \end{cases} \tag{4.10}$$

La figure (4.4) montre plusieurs exemples d'ensemble de points *jittered* dans des espaces différents.

4.3.3 Hypercube Latin



Introduits par McKay [1979] dans la fin des années 70, les ensembles de points hypercubes Latin sont une alternative à l'échantillonnage *jittered*. On dit d'une matrice carrée qu'elle est un carré Latin si chaque élément de la matrice n'est présent qu'une fois par ligne et par colonne. En échantillonnage, on peut utiliser ce principe en partitionnant l'espace d'échantillonnage (par exemple $\mathcal{U}^2$ ou $\mathcal{T}^2$) par un pavage régulier formé de $N \times N$ tuiles et en ajoutant N points de façon à ce qu'il n'y ait qu'un point par ligne et par colonne. Les ensembles de points hypercubes Latin sont une généralisation de cette idée pour des dimensions arbitraires. C'est un cas particulier d'échantillonnage par pavage, puisque certaines tuiles ne contiennent aucun point. La figure ci-contre montre un exemple d'ensemble de points hypercube Latin dans $\mathcal{U}^2$. Il existe plusieurs façons de générer des ensembles de points hypercubes Latin. La plus simple, consiste à choisir aléatoirement les lignes et les colonnes associées aux cases contenant un point, et à placer le point uniformément au sein de chaque case. D'autres méthodes reprennent l'idée des ensembles de points *jittered* à l'aide de deux niveaux de grille [Shirley and Wang, 1994; Kensler, 2013]. La première grille permet de placer un point par tuile, et de s'approcher de cette façon des ensembles de points *jittered*. La deuxième grille, plus fine, sert de matrice pour vérifier les propriétés des ensembles de points hypercubes Latin. Dans tous les cas, la méthode n'est pas homogène, pour les mêmes raisons que les ensembles de points *jittered*, mais reste tout de même stochastique. On peut donc écrire que

$$\begin{cases} \mathfrak{M}_{\text{LHC}} \subset \mathfrak{M}_{\text{Stocha.}} \\ \mathfrak{M}_{\text{LHC}} \subset \mathfrak{M}_{\text{Pavage}} \\ \mathfrak{M}_{\text{LHC}} \not\subset \mathfrak{M}_{\text{Homo.}} \,, \end{cases} \tag{4.11}$$

où $\mathfrak{M}_{\text{LHC}}$ est la classe des ensembles de points hypercubes Latin. Dans les espaces $\mathcal{U}^1$ et $\mathcal{T}^1$, cette classe de méthodes d'échantillonnage produit les mêmes ensembles de points que les ensembles de points *jittered*.

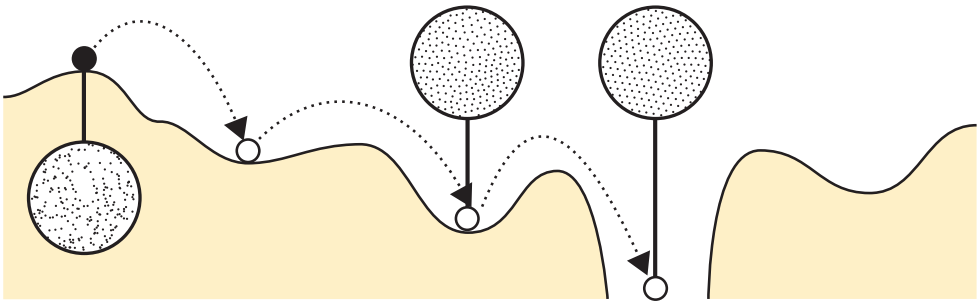


FIGURE 4.5 – *Illustration d'une minimisation d'énergie par une méthode d'optimisation d'un ensemble de points. L'ensemble de points initial, à gauche, n'est pas uniforme et possède une forte énergie. À chaque itération, symbolisée par une flèche, l'ensemble de points devient plus uniforme. Le minimum global de l'énergie correspond à la dernière itération. Bien que ce minimum soit très uniforme, il est aussi trop régulier. Il faut donc s'arrêter à la deuxième itération qui est un bon compromis.*

4.4 Optimisation d'un ensemble de points

Nous venons de voir comment générer des ensembles de points de façon stochastique ou à l'aide d'un pavage. Pour améliorer la qualité de l'ensemble de points, on a recourt à des techniques d'optimisation. Le critère de qualité visé est traduit en énergie, et une technique de minimisation d'énergie est ensuite appliquée sur l'ensemble de points. Bien que ces méthodes soient efficaces, elles présentent en générale plusieurs défauts. Tout d'abord, elles sont souvent lentes car le processus de minimisation d'énergie n'est jamais triviale. Ensuite, le critère d'énergie utilisé est généralement choisi de façon ad-hoc, sans lien théorique clair avec l'erreur d'intégration. Lorsque le minimum global est atteint, l'ensemble de points obtenu est souvent loin d'être de bonne qualité, car «trop optimisé ». Par exemple, si l'on cherche à maximiser l'uniformité de l'ensemble, on obtiendra un ensemble de points régulier, ce qui n'est pas satisfaisant. On considère alors que l'objectif de ces méthodes d'optimisation n'est pas d'atteindre un minimum global, mais un minimum local. Ce minimum local est atteint en arrêtant le processus d'optimisation au bon moment. La figure (4.5) illustre ce principe. La qualité finale d'un ensemble de points optimisé dépend donc de l'ensemble de points initial, de l'expression de l'énergie formulée, de la méthode d'optimisation et du critère d'arrêt. Certains optimiseurs sont compatibles avec l'homogénéité, c'est à dire que si un ensemble de points est homogène, alors l'ensemble de points optimisé l'est aussi.

4.4.1 Relaxation de Lloyd

La relaxation de Lloyd [Lloyd, 1983] est une méthode d'optimisation basée sur les diagrammes de Voronoi. Pour une description détaillée des diagrammes de Voronoi, voir le chapitre suivant, section 5.4. L'objectif de cette méthode est que chaque point de l'ensemble corresponde au barycentre de sa cellule de Voronoi. On dit alors que le diagramme de Voronoi de l'ensemble forme un **pavage barycentrique de Voronoi**. En terme d'optimisation, la fonction d'énergie à minimiser, $E_{CVT}(\mathcal{S})$, est donnée par

$$E_{CVT}(\mathcal{S}) := \sum_{k=1}^{N} \int_{\nu_k} \|x - \mathbf{s}_k\|^2 \mathrm{d}x \,, \tag{4.12}$$

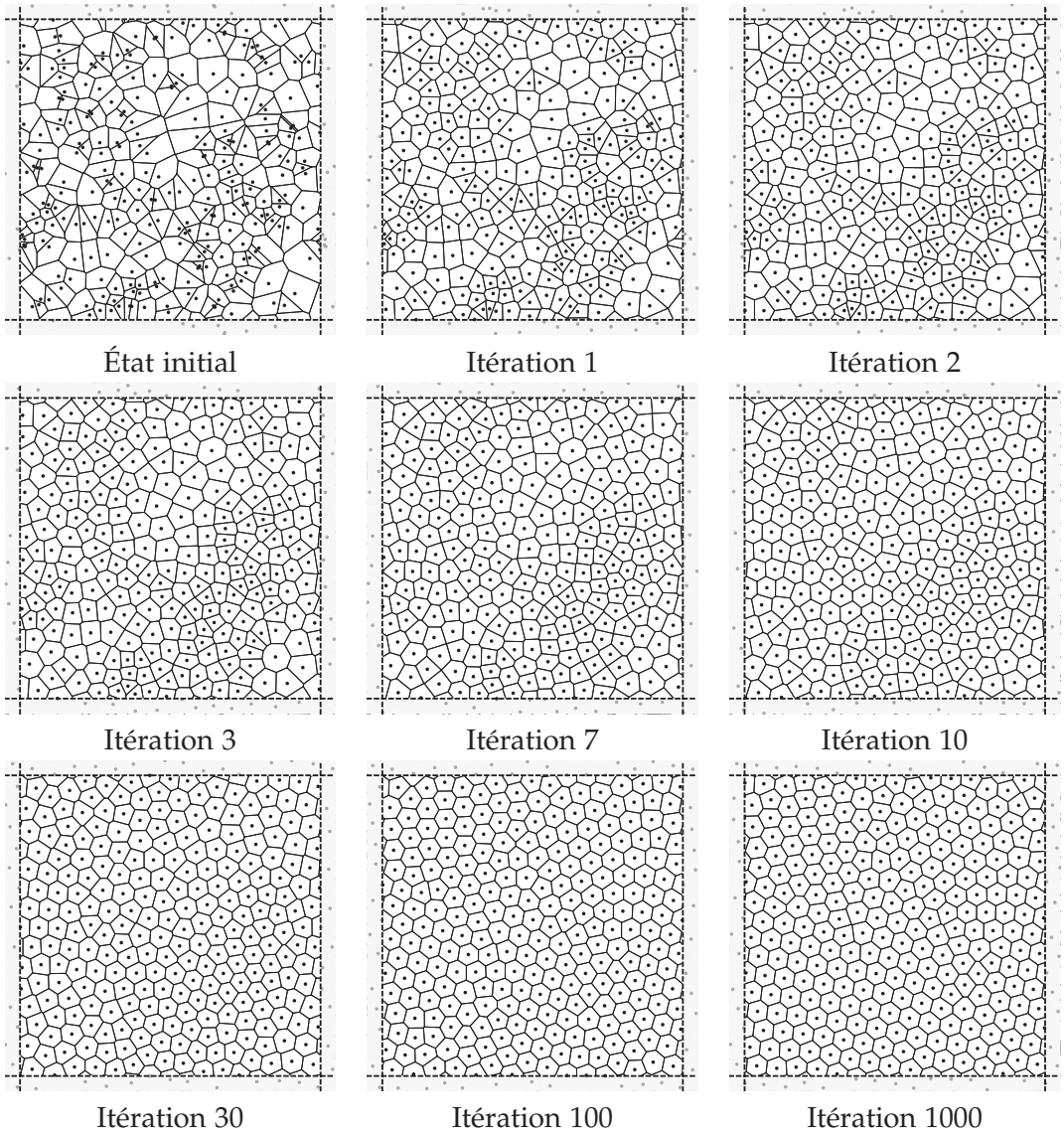
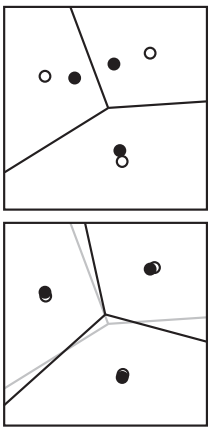


FIGURE 4.6 – *Progression d'une relaxation de Lloyd de* 256 *points dans* $\mathcal{T}^2$. *L'état initial correspond à un ensemble de points bruit blanc. Au fur et à mesure des itérations, l'ensemble de points devient de plus en plus uniforme. À l'itération* 1000, *l'ensemble de points est uniforme mais aussi très régulier car il comporte de nombreuses configurations hexagonales.*

où v_k est la cellule de Voronoi du point $\mathbf{s}_k$ [Du *et al.*, 1999]. Le but de cette fonction d'énergie est de quantifier à quel point les cellules de Voronoi de $\mathcal{S}$ sont compactes.

Pour minimiser cette énergie, la technique la plus populaire est la relaxation de Lloyd [1983]. À partir d'un ensemble de points déjà existant, en général un bruit blanc, on déplace chaque point vers le barycentre de sa cellule de Voronoi. Comme les points ont bougé, les cellules de Voronoi sont différentes et le barycentre s'est déplacé. On applique donc la relaxation de Lloyd plusieurs fois jusqu'a ce que le point converge vers le barycentre de sa cellule de Voronoi. La figure ci-contre illustre une itération appliquée sur trois points. Après une relaxation de Lloyd, On remarque que les cellules de Voronoi sont différentes et que les points sont maintenant plus proches de leur barycentre. Des méthodes ont depuis été proposées pour améliorer la convergence de cet algorithme, en utilisant par exemple une optimisation de Newton [Liu *et al.*, 2009].



Il n'existe pas de critère clair sur le nombre d'itérations nécessaire pour obtenir un ensemble de points créant un pavage barycentrique de Voronoi. Cette question est d'autant plus délicate que si un trop grand nombre de relaxations est appliqué, l'ensemble de points devient trop régulier [Du *et al.*, 1999]. En informatique graphique, on applique généralement seulement une ou deux itérations sur des ensembles de points qui sont déjà de bonne qualité. Un exemple de convergence de relaxation de Lloyd est montré dans la figure (4.6).

Si l'ensemble de points est homogène, alors les barycentres de ses cellules de Voronoi le sont aussi. La relaxation de Lloyd conserve donc l'homogénéité.

4.4.2 Optimisation de la capacité

Depuis les travaux de Balzer et al. [2009], une nouvelle génération d'échantillonneurs basée sur la relaxation de Lloyd a fait son apparition. En plus de l'expression énergétique $E_{CVT}(\mathcal{S})$ utilisée par la relaxation de Lloyd, une contrainte de capacité est ajoutée

$$\forall k \in \{1, ..., N\}, \quad \mu_{\mathcal{E}}(v_k) = \frac{\mu_{\mathcal{E}}(\mathbb{D}_{\mathcal{E}})}{N}. \tag{4.13}$$

Ici, la capacité d'un point represente l'aire de la cellule de Voronoi. Une contrainte de non régularité est aussi ajoutée, pour éviter que l'ensemble de points ne converge vers une grille régulière.

La méthode proposée par Balzer et al. [2009], appelée *Capacity-Constrained Voronoi Tesselation* (CCVT), consiste à subdiviser le domaine d'échantillonnage en petits fragments. Chaque fragment représente un petit élément de capacité à associer à un point de l'ensemble de points. L'algorithme est initialisé en associant les fragments aux points de façon aléatoire, en faisant attention que chaque point ait le même nombre de fragments. A ce stade, la capacité de chaque point est la même, mais la surface représentée par l'union des fragments associés à un point ne correspond pas encore à sa cellule de Voronoi. Pour faire converger les unions de fragments vers le diagramme de Voronoi, on déplace chaque point vers le barycentre de son ensemble de fragments. Puis on cherche à minimiser l'énérgie (la somme des distances entre points et fragments) en modifiant l'association entre points et fragments par des permutations. Ces deux étapes sont repétées jusqu'à ce qu'il ne soient plus possible de minimiser l'énérgie par des permutations de fragments. Cette méthode est proche de la relaxation de

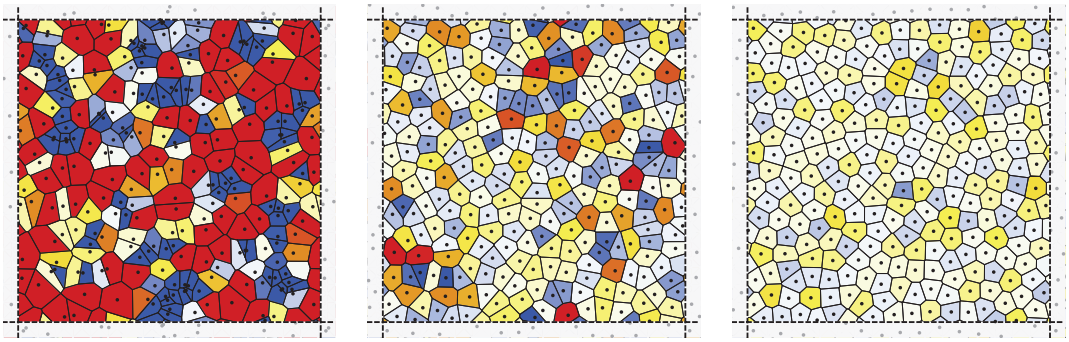


FIGURE 4.7 – *Évolution de la capacité de chaque point d'un ensemble de* 256 *points dans l'algorithme de transport optimal, à partir d'un ensemble de points bruit blanc. En rouge, les cellules de Voronoi qui ont une capacité trop grande et en bleu, celles qui ont une capacité trop faible.*

Lloyd au sens où les points poursuivent leur barycentre au fil des iterations, mais l'utilisation des fragments permet de contraindre la capacité de chaque point. Les ensembles de points optimisés par CCVT sont de bonne qualité pour le rendu graphique, mais le processus d'optimisation n'est pas efficace sur le plan algorithmique. De plus, le fait que cette méthode converge ou non vers une grille régulière dépend fortement de l'initialisation et du nombre de fragments utilisés.

Depuis l'apparition de la méthode CCVT, plusieurs articles sont apparus pour proposer des méthodes n'utilisant pas de discrétisation de l'espace par des fragments. [Chen *et al.*, 2012; de Goes *et al.*, 2012]. De Goes et al. [2012] ont proposé une solution, *Blue Noise through Optimal Transport* (BNOT), utilisant des diagrammes de puissance avec poids et un algorithme de transport optimal. La figure (4.7) montre l'évolution d'un ensemble de points bruit blanc optimisé par BNOT. Ce nouvel algorithme contrairement à CCVT préserve le caractère homogène d'un ensemble de points.

4.4.3 Optimisation de la distance minimale

Schlömer et al.[2011] ont proposé une méthode, *Farthest-Point Optimization* (FPO), pour optimiser un ensemble de points afin d'augmenter la distance minimale entre points. L'algorithme consiste, pour chaque point de l'ensemble, à le supprimer puis à l'insérer à l'endroit qui l'éloigne le plus des autres. L'optimisation est terminée quand la distance minimale entre les points atteint un seuil donné. L'ensemble de points produit par FPO a de bonnes propriétés, ne contient pas de régularités et est compatible avec l'homogénéité. On remarquera tout de même que l'ensemble de points, bien que très compact, comporte des trous. Cette méthode n'utilise qu'une fonction de distance entre point. Elle peut donc théoriquement fonctionner dans tous les espaces d'échantillonnage. L'article initial décrit la méthode pour l'espace $\mathcal{T}^2$ [Schlömer *et al.*, 2011] et un article récent l'a généralisé à toutes les surfaces 3D [Yan *et al.*, 2014].

4.4.4 Contrôle spectral

Les méthodes dites de contrôle spectral ont été introduites simultanément par Zhou et al. [2012], Heck et al. [2013] et Öztireli et al. [2012]. Des travaux préliminaires, mais dans un domaine de recherche différent, ont aussi été menés par Uche et al. [2006]. Le principe de ces méthodes est de définir une cible sous la forme d'une

fonction de distribution de distances (voir chapitre 5), et d'optimiser itérativement un ensemble de points pour que ses propriétés statistiques se rapprochent le plus possible de cette cible. Le processus d'optimisation varie d'une méthode à l'autre, mais le principe fondamental reste le même. À chaque itération, on compare la fonction de distribution de distances de l'ensemble de points ciblé avec celle de l'état actuel de l'ensemble de points à optimiser. Puis on calcule la différence entre ces deux fonctions que l'on utilise pour décider quels points doivent être éloignés et lesquels doivent être rapprochés. L'algorithme se termine quand la différence entre la cible et l'état de l'ensemble de points optimisé est inférieure à un seuil. On appelle ces méthodes des méthodes de contrôle spectral car la cible est généralement donnée dans le domaine spectral (voir chapitre 6) avant d'être convertie dans le domaine spatial.

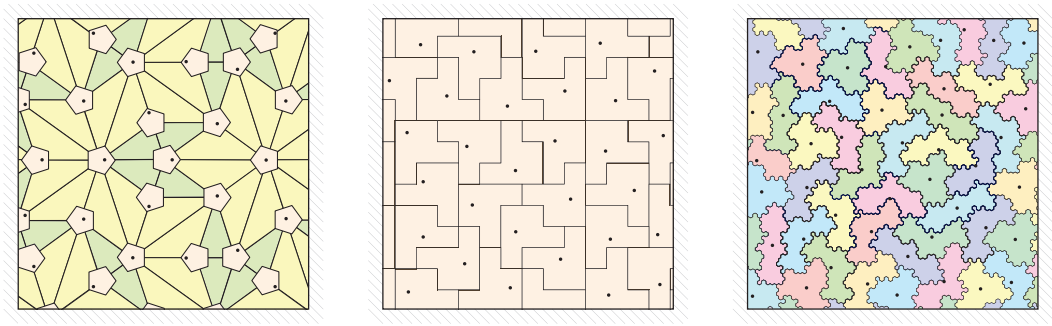
Les méthodes de contrôle spectral ont de nombreux avantages. À partir d'une seule réalisation d'un ensemble de points, elles permettent de générer d'autres réalisations. Les cibles peuvent même être définies théoriquement, sans connaître à l'avance l'aspect de l'ensemble de points. Ce point est particulièrement utile pour explorer de nouveaux types d'ensemble de points. Toutefois, ces avantages sont compensés par le manque de précision du résultat obtenu. Ainsi, même après un grand nombre d'itérations, l'ensemble de points ne correspond jamais totalement à la cible choisie, et parfois même de façon critique. Pour une analyse détaillée de ces défauts, voir le chapitre 8.

4.5 Échantillonnage par pavage hiérarchique

Un système de pavage hiérarchique est une méthode de génération de pavage dont la densité est contrôlable. Le contrôle de densité s'opère à l'aide d'un mécanisme de subdivision. Formellement, un système de pavage hiérarchique est un ensemble de proto-tuiles Σ, couplé à un ensemble de règles de subdivision $\Pi = \{\pi_1, \pi_1, ...\}$. Chaque règle de subdivision π_k permet de partitionner une proto-tuile σ_k en Λ tuiles de taille plus petite à l'aide d'une transformation affine.

L'échantillonnage par pavage hiérarchique consiste à générer un pavage par une système hiérarchique, puis à ajouter des points dans chaque tuile. L'intérêt de ces méthodes est qu'à travers le mécanisme de subdivision, on peut connaître, pour une tuile donnée, quelles sont les proto-tuiles qui sont autour. Ainsi, on a une idée de l'entourage de chaque tuile et l'on peut placer les points à l'intérieur des tuiles en ayant une idée de la position des points qui se trouveront autour. Les méthodes actuelles d'échantillonnage par pavage hiérarchique utilisent toutes un processus d'optimisation durant une phase de pré-calcul afin de créer une table contenant, pour une configuration de proto-tuile (une proto-tuile entourée d'autres), la position des points qui se trouvent à l'intérieur. On appelle cette table une **table de déviations**, où chaque déviation est un vecteur de déplacement exprimé dans un repère locale associé à chaque proto-tuile.

Les propriétés finales de l'ensemble de points dépendent des propriétés du pavage utilisé et du placement des points dans chaque tuile. En particulier, le pavage doit être apériodique et suffisamment riche pour limiter l'impact des répétitions dues au mécanisme de subdivision. Une description détaillée des pavages apériodiques est disponible dans le livre de Grünbaum et Shephard [1986, Chapitre 10].



(a) (b) (c)

FIGURE 4.8 – *Différents pavages et les ensembles de points qui y sont associés.* ***(a) :*** *Un pavage de Penrose.* ***(b) :*** *Un pavage à base de polyominos.* ***(c) :*** *Un pavage à base de trihexes irréguliers.*

Pavage de Penrose Le pavage de Penrose a été découvert dans les années 70 par Roger Penrose [1974]. Quarante ans plus tard, Osrtomoukhov [2004] présente la première méthode d'échantillonnage basée sur ce pavage. Cette méthode d'échantillonnage est compacte, puisqu'elle ne contient que 6 proto-tuiles et une table de déviations de 21 entrées. Dans ce système, seules les tuiles les plus petites contiennent un point, les autres tuiles n'étant utilisées que pour les étapes de subdivision. L'ensemble de points final est uniforme, mais comporte tout de même un certain nombre de régularités. La figure (4.8,a) montre un exemple de pavage de Penrose avec les points d'échantillonnage qui y sont associés.

Pavage à base de polyominos Suite à la méthode utilisant les pavages de Penrose, Ostromoukhov [2007] a présenté une méthode d'échantillonnage basée sur les polyominos. Ce type de pavage possède un plus grand nombre de tuiles et une plus grande diversité dans les arrangements des tuiles. Les régularités y sont donc beaucoup moins prononcées, bien que toujours présentes. La figure (4.8,b) montre un exemple de pavage à base de polyominos avec les points d'échantillonnage qui y sont associés.

Pavage à base de trihexes irréguliers Dans [2014], nous avons proposé ont proposé une méthode d'échantillonnage utilisant un nouveau type de pavage à base de trihexes irréguliers, spécialement créée pour l'échantillonnage. L'avantage de ce pavage est qu'il contient une grande diversité de proto-tuile et qu'il est de nature fractale. Ainsi, les structures régulières y sont presque absentes, permettant un échantillonnage de meilleure qualité. Les auteurs proposent également pour la première fois une approche utilisant le contrôle spectral. Cette méthode de pavage est en effet capable de supporter plusieurs tables de déviations. Dans ce système, chaque table permet d'obtenir des ensembles de points différents, exactement à la manière du contrôle spectral. La figure (4.8,c) montre un exemple de pavage à base de trihexes irréguliers avec les points d'échantillonnage qui y sont associés. Pour plus de détails sur cette méthode, voir la thèse de Wachtel [Wachtel, 2015].

Pavage de Wang Le pavage de Wang est une méthode de pavage apériodique utilisant des carrés avec des marqueurs sur les bords. Dans ce système, deux carrés cote à cote doivent partager le même marqueur dans la bordure qui les relies. L'intérêt de ce système de pavage est qu'il ne nécessite que 16 tuiles différentes [Grunbaum and

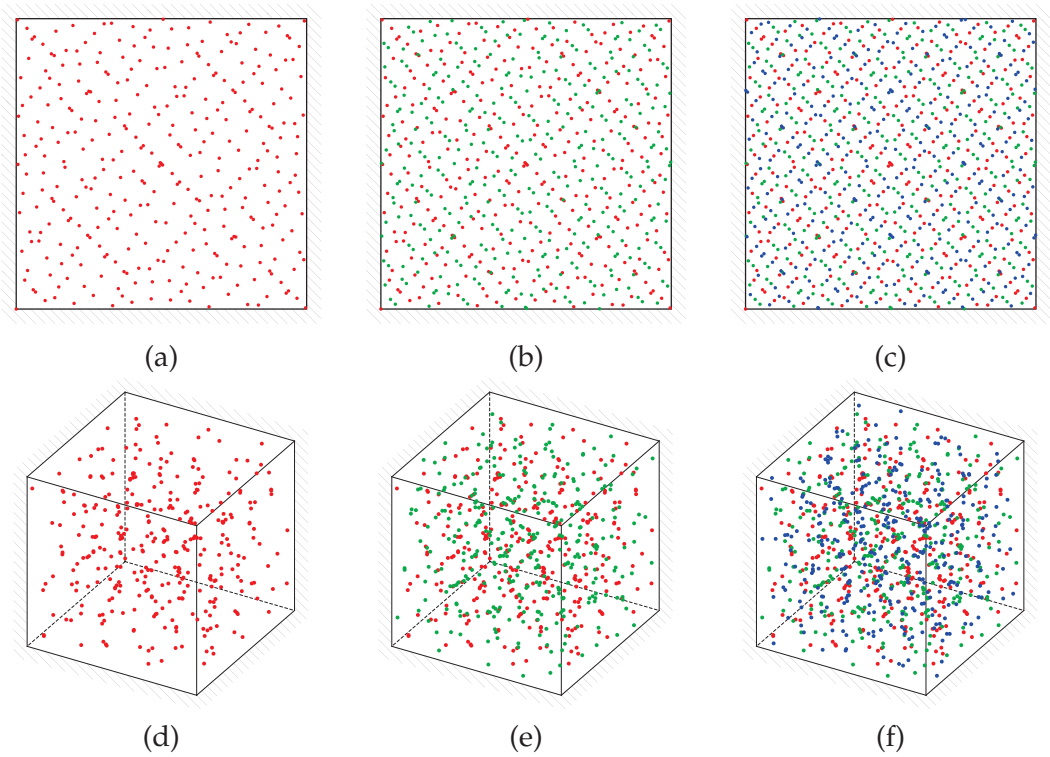


FIGURE 4.9 – *Échantillonnage par séquences à basse discrépance dans différentes dimensions.* ***(a), (b) et (c) :*** *Les* 340, 640 *et* 1024 *premiers points de la séquence de Sobol dans l'espace* $\mathcal{U}^2$. ***(d), (e) et (f) :*** *Les* 340, 640 *et* 1024 *premiers points de la séquence de Halton dans l'espace* $\mathcal{U}^3$. *Les couleurs permettent de voir que les points déjà ajoutés ne sont pas déplacés lorsque le nombre de points augmente.*

Shephard, 1986, Chapitre 11] et que l'on connais à l'avance les tuiles voisines possibles. Hiller et al. [2001] ont proposé une méthode d'échantillonnage utilisant des tuiles de Wang contenant plusieurs points optimisés par relaxation de Lloyd. Cette méthode a, par la suite, été étendue pour générer des ensembles de disques de Poisson [Hiller *et al.*, 2001; Lagae and Dutré, 2006], puis rendu hiérarchique pour supporter les ensembles de points non-uniforme [Kopf *et al.*, 2006].

4.6 Séquences à basse discrépance

Jusqu'à maintenant, nous avons vu des méthodes d'échantillonnage permettant de générer un ensemble de N points. Dans certaines applications, on souhaite augmenter le nombre d'échantillons progressivement afin d'améliorer la qualité de l'estimation de l'intégrale. Dans ce cas, il est préférable de rajouter des points à l'ensemble existant, plutôt que de générer à nouveau un nouvel ensemble de points. On parle alors de séquence de points, où $\mathcal{S}$ est généré en prenant les N premiers points de la séquence. Parmi les séquences existantes, on trouve les séquences à basse discrépance. Nous reviendrons sur la discrépance et son lien avec l'intégration Monte Carlo dans les chapitres 5 et 7. Pour le moment, il est suffisant de savoir que plus la discrépance est faible, plus l'erreur d'intégration l'est aussi. On dit qu'une séquence est à basse discrépance si le comportement asymptotique de la discrépance est en $O_{N\to\infty}\left(log(N)^d N^{-1}\right)$ lorsque N tends vers l'infini [Niederreiter, 1992]. En d'autres termes, quand le nombre de points augmente, la discrépance, et donc l'erreur d'intégration, doit diminuer avec

une certaine vitesse. Les séquences à basses discrépance sont appréciées car elles sont facilement extensible aux espaces de hautes dimensions. Toutefois, ces très bonnes propriétés sont compensées par le fait que les ensembles de points obtenus sont généralement très réguliers et donc propices à l'apparition d' *aliasing* dans le rendu d'image. Parmi les méthodes les plus connus, on trouve les séquences de Halton [Halton, 1960], de Sobol [Sobol, 1967] et de Niederreiter [Niederreiter, 1987]. La figure (4.9) montre un exemple de séquence de Sobol dans l'espace $\mathcal{U}^2$ et un exemple de séquence de Halton dans l'espace $\mathcal{U}^3$.

Conclusion

Ces exemples de méthodes d'échantillonnage montrent les différentes stratégies mises en place pour améliorer l'intégration Monte Carlo. La figure (4.10) montre des ensembles de points générés par chacune de ces méthodes. On remarque une grande diversité dans l'aspect de chaque méthode. La question que l'on se pose est de savoir lequel de ces ensembles de points est le plus performant pour l'intégration Monte Carlo. Comme nous l'avons dit dans l'introduction, les grilles régulières et le bruit blanc représentent deux cas extrêmes. Avec les grilles régulières, les points sont uniformément répartis sur tout le domaine, mais sont très réguliers. Avec le bruit blanc, les points sont indépendants les uns des autres, mais de grandes portions du domaine restent vide. Entre les deux, il existe une myriade de méthodes, dont il est difficile de déterminer la qualité à l'aide des seules notions d'uniformité et de structures régulières. Les méthodes d'optimisation FPO et BNOT semblent par exemple quasi-identique à première vu. Pourtant, de subtiles différences font que l'une d'elles est plus performante que l'autre en terme de réduction de variance. Dans les chapitres suivants, nous détaillons les outils d'analyse permettant de caractériser les différences entre deux ensembles de points, d'une façon beaucoup plus précise que des simples considérations visuelles. Nous avons appliqué ces outils sur certaines méthodes d'échantillonnage que nous venons de présenter. Nos résultats sont compilés dans le chapitre 8.

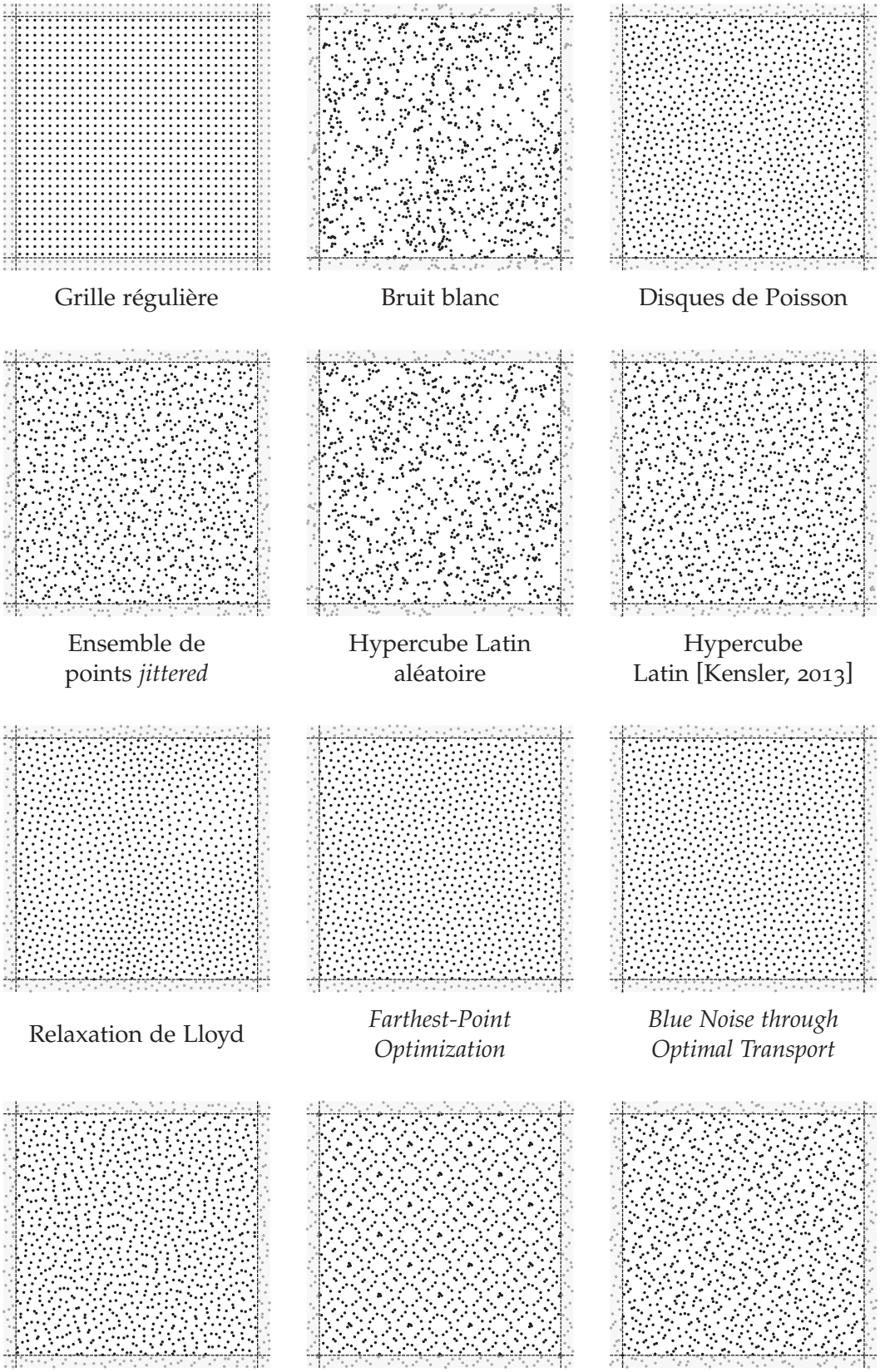


FIGURE 4.10 – *Résultat de plusieurs méthodes d'échantillonnage avec* $N = 1024$ *dans l'espace* $\mathcal{T}^2$.

Analyse spatiale 5

Dans ce chapitre, nous détaillons les principaux outils d'analyse des ensembles de points à travers le domaine spatial. Bien qu'il existe beaucoup de façons d'étudier un ensemble de points, nous avons choisi ici de nous concentrer sur un sous-ensemble de ces outils d'analyses : la densité, la distribution de distances entre points, les diagrammes de Voronoi et la discrépance. Ces outils sont tous connus et utilisés en informatique graphique et d'une manière plus large dans la communauté scientifique. Néanmoins, il nous a semblé utile de les rassembler ici à l'aide d'un formalisme commun et de mettre en avant les relations qui les lient.

Ce chapitre commence par une description détaillée des outils d'analyse de densité. Nous montrons notamment comment s'expriment la densité de points, la densité de paires de points et la densité de distance entre points, et comment ces outils se comportent quand l'ensemble de points est homogène. Nous précisons également quels sont les termes et les différentes normalisations qui existent pour ces outils. Nous aborderons en suite rapidement la question des diagrammes de Voronoi et de leur utilisation dans l'analyse des ensembles de points. Finalement, ce chapitre se termine par une présentation de la discrépance afin de caractériser l'uniformité d'un ensemble de points. Nous exposons deux versions de la discrépance, une version classique pour les ensembles de points déterministes, et une version stochastique pour les ensembles de points stochastiques.

5.1 Densité

La densité $\rho_{\mathcal{S}}$ d'un ensemble de points représente la quantité de points par unité de mesure du domaine. Elle est décrite par la relation

$$\rho_{\mathcal{S}} := \frac{N}{\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)} . \tag{5.1}$$

Bien que la densité ne soit qu'une normalisation du nombre de points, elle permet de représenter plus simplement la quantité de points dans le domaine, notamment dans le cas des ensembles de points uniformément distribués. En particulier, un ensemble de points uniforme dans l'espace euclidien $\mathcal{E}^d$ possède une mesure du domaine d'échantillonnage infinie et un nombre de points infini, conduisant à une densité indeterminée. En réalité, cette densité existe et est définie, mais il n'est pas possible de la connaître avec comme seules informations le nombre de points et la mesure du domaine. On voit donc que si l'on connaît la densité plutôt que le nombre de points, on

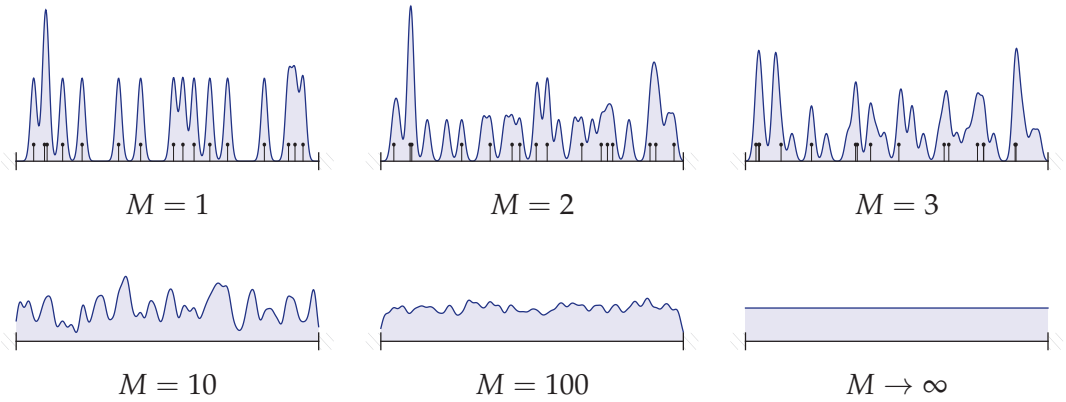


FIGURE 5.1 – *Évolution de l'estimation de la densité locale du bruit blanc dans $\mathcal{U}^1$ pour un nombre de réalisations croissant. La courbe représente un histogramme approximant la densité locale de l'ensemble de points. Dans la première ligne, le dernier ensemble de points utilisé pour mettre à jour la densité locale est affiché.*

gagne une information supplémentaire. Pour cette raison, nous utiliserons ici la densité de l'ensemble de points à la place du nombre de points quand cela est possible.

Définition 5.1 *[Hansen and McDonald, 1990, p. 29] Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$, on appelle* ***densité locale*** *de $\mathcal{S}$, notée $\rho_{\mathcal{S}}^{(1)}$, la fonction $\mathbb{D}_{\mathcal{E}} \to \mathbb{R}^+$ définie par*

$$\rho_{\mathcal{S}}^{(1)}(x) := \sum_{k=1}^{N} \langle \delta(x, \boldsymbol{s}_k) \rangle = \sum_{k=1}^{N} PDF_{\boldsymbol{s}_k}(x) \,. \tag{5.2}$$

La densité locale de $\mathcal{S}$ est construite de la même façon que la fonction de densité de probabilité de $\mathcal{S}$, à la différence qu'elle prend en compte tous les points en même temps. Tout comme une fonction de densité de probabilité, la densité locale prend un sens lorsqu'on l'intègre sur une surface. Ainsi, l'intégrale de la densité locale sur un borélien est égale au nombre de points se trouvant en moyenne à l'intérieur de ce borélien. L'intégrale de la densité locale sur l'ensemble du domaine d'échantillonnage est donc égale à N, ce qui montre bien que malgré les similitudes, la densité locale n'est pas une fonction de densité de probabilité. La figure (5.1) montre un exemple de densité locale pour du bruit blanc en fonction du nombre de réalisations utilisées pour calculer la moyenne de façon empirique. Quand l'ensemble de points est déterministe, l'espérance peut être éliminée, et la densité locale devient simplement égale à la fonction représentative de $\mathcal{S}$. Quand l'ensemble de points est stochastique, on peut remarquer que la densité locale est simplement égale à l'espérance de la fonction représentative de $\mathcal{S}$,

$$\rho_{\mathcal{S}}^{(1)}(x) = \langle S(x) \rangle \,. \tag{5.3}$$

Lemme 5.1 *Soit $\mathcal{S}$ un ensemble dans l'espace $\mathcal{E}$, si $\mathcal{S}$ est homogène, alors la densité locale est constante et égale à la densité globale*

$$\rho_{\mathcal{S}}^{(1)}(x) = \rho_{\mathcal{S}} \,. \tag{5.4}$$

Preuve 5.1 *Ce lemme est la conséquence directe de Def. (5.1) et Lemme (3.2, p. 27).* □

On voit que la densité locale ne porte aucune information dans le cas des ensembles de points homogènes. Il est donc nécessaire d'avoir d'autres outils pour caractériser ces ensembles de points. En particulier, on cherche un outil d'analyse qui se

focalise sur la corrélation entre les points, plutôt que sur les différences de comportement de l'ensemble de points sur le domaine d'échantillonnage. Pour cela, on définit la densité de paires de points, puis la densité différentielle.

5.1.1 Densité de paires

Définition 5.2 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$, on appelle* ***densité de paires****, noté $\rho_{\mathcal{S}}^{(2)}$, la fonction $\mathbb{D}_{\mathcal{E}} \times \mathbb{D}_{\mathcal{E}} \to \mathbb{R}^+$ définie par*

$$\rho_{\mathcal{S}}^{(2)}(x, x') := \sum_{i=1}^{N} \sum_{j=1}^{N} \left\langle \delta(x, \mathbf{s}_i)\delta(x', \mathbf{s}_j) \right\rangle . \tag{5.5}$$

La densité de paires est une extension de la densité locale. Elle représente le nombre de fois qu'un couple de points (x, x') puisse apparaître. Cette définition est inspirée de celle donnée par Hansen [1990, p. 30], à ceci près que la deuxième somme n'exclut pas le cas où $i = j$,

$$\rho_{\mathcal{S}}^{(2*)}(x, x') := \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \left\langle \delta(x, \mathbf{s}_i)\delta(x', \mathbf{s}_j) \right\rangle . \tag{5.6}$$

La différence entre ces deux définitions correspond à

$$\begin{aligned} \rho_{\mathcal{S}}^{(2)}(x, x') - \rho_{\mathcal{S}}^{(2*)}(x, x') &= \sum_{k=1}^{N} \left\langle \delta(x, \mathbf{s}_k)\delta(x', \mathbf{s}_k) \right\rangle \\ &= \delta(x', x) \sum_{k=1}^{N} \left\langle \delta(x, \mathbf{s}_k) \right\rangle \\ &= \delta(x', x)\rho_{\mathcal{S}}^{(1)}(x) . \end{aligned} \tag{5.7}$$

Comme on le voit, les deux définitions diffèrent uniquement dans le cas où l'on souhaite étudier la densité d'ensembles de points avec multiplicité, c'est-à-dire quand $\mathcal{S}$ contient plusieurs fois un même élément. Tout comme la densité locale, la densité de paires est elle aussi reliée à la fonction représentative de $\mathcal{S}$.

Lemme 5.2 *Soit $\mathcal{S}$ un ensemble de points dans l'espace d'échantillonnage $\mathcal{E}$, on a l'égalité suivante*

$$\rho_{\mathcal{S}}^{(2)}(x, x') = Cov(S(x), S(x')) + \rho_{\mathcal{S}}^{(1)}(x)\rho_{\mathcal{S}}^{(1)}(x') . \tag{5.8}$$

Preuve 5.2 *En déplaçant l'espérance à l'extérieur de la somme et en factorisant les deux sommes dans Eq. (5.5), obtient*

$$\rho_{\mathcal{S}}^{(2)}(x, x') = \left\langle \left(\sum_{i=1}^{N} \delta(x, \mathbf{s}_i) \right) \left(\sum_{j=1}^{N} \delta(x', \mathbf{s}_j) \right) \right\rangle , \tag{5.9}$$

ce qui donne, en utilisant Def. (3.2, p. 17),

$$\rho_{\mathcal{S}}^{(2)}(x, x') = \left\langle S(x)S(x') \right\rangle . \tag{5.10}$$

La preuve est terminée en utilisant les définitions de la covariance et de la densité locale. □

En général, on utilise une version normalisée de la densité de paires.

Définition 5.3 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$, on appelle* ***fonction de distribution de paires****, notée $g_{\mathcal{S}}$, la fonction $\mathbb{D}_{\mathcal{E}} \times \mathbb{D}_{\mathcal{E}} \to \mathbb{R}^+$ définie par*

$$g_{\mathcal{S}}(x, x') := \frac{\rho_{\mathcal{S}}^{(2)}(x, x')}{\rho_{\mathcal{S}}^{(1)}(x)\rho_{\mathcal{S}}^{(1)}(x')} . \tag{5.11}$$

Notre définition de la fonction de distribution de paires est la même que celle décrite dans [Hansen and McDonald, 1990], à l'exception de l'expression de la densité de paires qui, comme nous l'avons dit, est différente. Le choix de cette normalisation s'explique par l'expression de la densité de paires dans la forme covariance. Lorsque les points x et x' sont très éloignés l'un de l'autre, la corrélation entre eux tend généralement à être nulle et le terme covariance disparaît[1]. Il ne reste donc plus que la densité locale dans l'expression. En normalisant la densité de paires par la densité locale, on obtient donc une fonction de distribution de paires qui converge vers 1.

5.1.2 Densité de n-uplets

Nous venons de définir la densité de paires, permettant d'étudier des configurations de paires de points. Cette définition peut se généraliser à des configurations de tailles arbitraires.

Définition 5.4 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$ et $n \in \mathbb{N}$, on appelle* ***densité de n-uplets****, notée $\rho_{\mathcal{S}}^{(n)}$, la fonction $\mathbb{D}_{\mathcal{E}}^n \to \mathbb{R}^+$ définie par*

$$\rho_{\mathcal{S}}^{(n)}(x_1, ..., x_n) := \sum_{k_1=1}^{N} ... \sum_{k_n=1}^{N} \left\langle \prod_{i=1}^{n} \delta(x_i, \boldsymbol{s}_{k_i}) \right\rangle . \tag{5.12}$$

Lorsque $n = 1$, on obtient la définition de la densité locale et lorsque $n = 2$, on obtient la définition de la densité de paires. Finalement, pour $n = N$, la densité de n-uplets représente la probabilité de trouver une réalisation particulière dans un échantillonneur stochastique. Elle correspond donc à la densité de probabilité associée à la variable aléatoire $\mathcal{S}$ et liée à la loi de probabilité $\mathbb{P}_{\mathcal{S}}$. En utilisant le même raisonnement que pour la densité de paires, on peut exprimer la densité de n-uplets en fonction de la fonction représentative de $\mathcal{S}$,

$$\rho_{\mathcal{S}}^{(n)}(x_1, ..., x_n) = \langle S(x_1)...S(x_n) \rangle . \tag{5.13}$$

Lemme 5.3 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$ et $1 < n \leq N$, on a l'égalité suivante*

$$\rho_{\mathcal{S}}^{(n-1)}(x_1, ..., x_{n-1}) = \frac{1}{N} \int_{\mathbb{D}_{\mathcal{E}}} \rho_{\mathcal{S}}^{(n)}(x_1, ..., x_n) \mathrm{d}x_n . \tag{5.14}$$

Preuve 5.3 *À partir de Eq. (5.13), on peut écrire que*

$$\frac{1}{N} \int_{\mathbb{D}_{\mathcal{E}}} \rho_{\mathcal{S}}^{(n)}(x_1, ..., x_n) \mathrm{d}x_n = \left\langle \frac{1}{N} \int_{\mathbb{D}_{\mathcal{E}}} S(x_1)...S(x_n) \mathrm{d}x_n \right\rangle . \tag{5.15}$$

1. Cette observation ne s'applique qu'à certains ensembles de points stochastiques n'ayant pas de régularité.

Comme seule $S(x_n)$ est dépendante de la variable locale de l'intégrale, on peut simplifier l'intégrale de la façon suivante, en utilisant le fait que l'intégrale de la fonction représentative de $\mathcal{S}$ est égale à N

$$\begin{aligned}\frac{1}{N}\int_{\mathbb{D}_\mathcal{E}} \rho_\mathcal{S}^{(n)}(x_1,...,x_n)\mathrm{d}x_n &= \left\langle \frac{1}{N} S(x_1)...S(x_{n-1}) \int_{\mathbb{D}_\mathcal{E}} S(x_n)\mathrm{d}x_n \right\rangle \\ &= \langle S(x_1)...S(x_{n-1})\rangle \ . \end{aligned} \tag{5.16}$$

En utilisant Eq. (5.13), on obtient bien Eq. (5.14). □

Sachant que la densité de n-uplets, avec $n = N$, caractérise complètement les propriétés statistiques de l'ensemble de points, on voit que lorsque n diminue, on perd de l'information. Toutefois, il est rarement possible de connaître la densité de N-uplets, ni même de l'approximer, étant donné que c'est une fonction de dimension $N \times d$. Bien souvent, on se limite donc à la densité de paires. Dans la section suivante, nous définissons la densité différentielle, une fonction équivalente à la densité de paires dans le cas homogène, mais dont la dimension est plus faible.

5.2 Densité différentielle

En redéfinissant la paire de points (x, x') comme un couple $(x, x * \tau)$ où τ est une action de $\mathbb{T}_\mathcal{E}$ telle que $x * \tau = x'$, on obtient l'expression de la densité de paires suivante :

$$\rho_\mathcal{S}^{(2)}(x, x*\tau) = \left\langle \sum_{i=1}^{N}\sum_{j=1}^{N} \delta(x, \mathbf{s}_i)\delta(x*\tau, \mathbf{s}_j) \right\rangle . \tag{5.17}$$

Bien que cette relation ne soit qu'une réécriture de la densité de paires et ne change en rien sa valeur, elle permet de mettre en avant la notion de configuration de points. On peut mettre l'emphase sur cette notion de configuration de points en moyennant la densité de paires sur tout le domaine d'échantillonnage. Ainsi, on élimine la variable x pour ne conserver que l'opération de déplacement τ entre les deux points, ce qui caractérise une configuration de deux points.

Définition 5.5 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$, on appelle **densité différentielle**, noté $\check{\rho}_\mathcal{S}^{(2)}$, la fonction $\mathbb{T}_\mathcal{E} \to \mathbb{R}^+$ définie par*

$$\check{\rho}_\mathcal{S}^{(2)}(\tau) := \frac{1}{\mu_\mathcal{E}(\mathbb{D}_\mathcal{E})} \int_{\mathbb{D}_\mathcal{E}} \rho_\mathcal{S}^{(2)}(x, x*\tau)\mathrm{d}x \, . \tag{5.18}$$

Lemme 5.4 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$, on a l'égalité suivante*

$$\check{\rho}_\mathcal{S}^{(2)}(\tau) = \frac{1}{\mu_\mathcal{E}(\mathbb{D}_\mathcal{E})} \sum_{i=1}^{N}\sum_{j=1}^{N} \left\langle \delta(\mathbf{s}_i * \tau, \mathbf{s}_j) \right\rangle . \tag{5.19}$$

Preuve 5.4 *En développant Eq. (5.18) avec Def. (5.2) et en réordonnant l'espérance, l'intégrale et les sommes, on obtient*

$$\begin{aligned}\check{\rho}_\mathcal{S}^{(2)}(\tau) &= \frac{1}{\mu_\mathcal{E}(\mathbb{D}_\mathcal{E})} \int_{\mathbb{D}_\mathcal{E}} \sum_{i=1}^{N}\sum_{j=1}^{N} \left\langle \delta(x, \mathbf{s}_i)\delta(x*\tau, \mathbf{s}_j) \right\rangle \mathrm{d}x \\ &= \frac{1}{\mu_\mathcal{E}(\mathbb{D}_\mathcal{E})} \sum_{i=1}^{N}\sum_{j=1}^{N} \left\langle \int_{\mathbb{D}_\mathcal{E}} \delta(x, \mathbf{s}_i)\delta(x*\tau, \mathbf{s}_j)\mathrm{d}x \right\rangle . \end{aligned} \tag{5.20}$$

L'intégrale peut être simplifiée sachant que $\delta(x, \boldsymbol{s}_i)\delta(x * \tau, \boldsymbol{s}_j)$ est non nul quand $x = \boldsymbol{s}_i$, ce qui nous donne Eq. (5.19). □

Définition 5.6 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$, on appelle **fonction de distribution différentielle**, notée $\breve{g}_{\mathcal{S}}$, la fonction $\mathbb{T}_{\mathcal{E}} \to \mathbb{R}^+$ définie par*

$$\breve{g}_{\mathcal{S}}(\tau) := \frac{\breve{\rho}^{(2)}_{\mathcal{S}}(\tau)}{\rho^2_{\mathcal{S}}} . \tag{5.21}$$

La fonction de distribution différentielle est elle aussi très utilisée, pour les mêmes raisons que la fonction de distribution de paires. La figure (5.2) montre un exemple de fonction de distribution différentielle.

Lemme 5.5 *Soit $\mathcal{S}$ un ensemble de points homogène dans l'espace $\mathcal{E}$, et $x \in \mathbb{D}_{\mathcal{E}}$, on a l'égalité suivante*

$$\breve{\rho}^{(2)}_{\mathcal{S}}(\tau) = \rho^{(2)}_{S}(x, x * \tau) \tag{5.22}$$

Preuve 5.5 *En passant le domaine d'intégration du domaine d'échantillonnage vers le groupe des translations à l'aide de Lemme (3.1, p. 26) sur Eq. (5.18), on obtient*

$$\breve{\rho}^{(2)}_{\mathcal{S}}(\tau) = \frac{1}{\mu_H(\mathbb{T}_{\mathcal{E}})} \int_{\mathbb{T}_{\mathcal{E}}} \rho^{(2)}_{S}(x * \tau', (x * \tau') * \tau) \mathrm{d}\tau . \tag{5.23}$$

ce que l'on peut écrire également comme

$$\breve{\rho}^{(2)}_{\mathcal{S}}(\tau) = \frac{1}{\mu_H(\mathbb{T}_{\mathcal{E}})} \int_{\mathbb{T}_{\mathcal{E}}} \rho^{(2)}_{\tau(S)}(x, x * \tau) \mathrm{d}\tau . \tag{5.24}$$

Finalement, on obtient Eq. (5.22) en utilisant Eq. (3.25, p. 27) du fait que $\mathcal{S}$ soit homogène. □

Comme on le voit, dans le cas des ensembles de points homogènes, il n'est pas nécessaire d'avoir recours à la fonction de distribution de paires. La fonction de distribution différentielle contient les mêmes informations, mais n'a qu'une dimensionnalité de d.

5.2.1 Fonction de distribution de distance

Définition 5.7 *Soit $\mathcal{S}$ un ensemble de points dans $\mathcal{E}$, on appelle **densité de distances**, notée $\breve{\rho}_{\mathcal{S}}$, la fonction $\mathbb{R}^+ \to \mathbb{R}^+$ définie par*

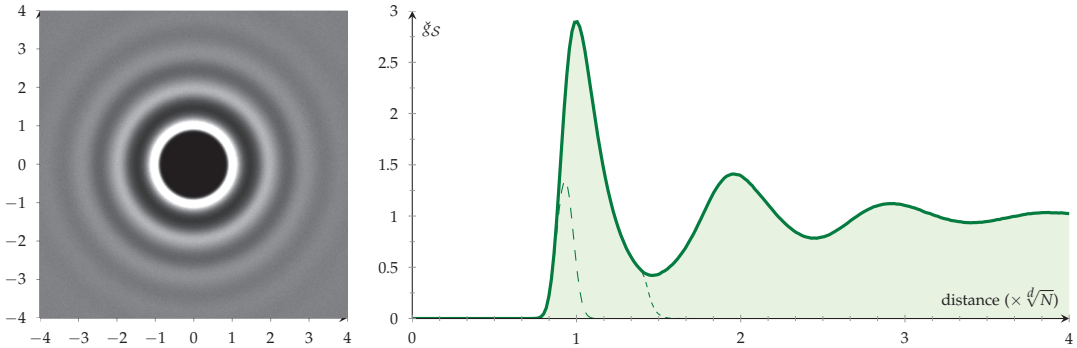
$$\breve{\rho}_{\mathcal{S}}(r) := \frac{1}{\mu_{\mathcal{E}}(\mathbb{D}_{\mathcal{E}})} \int_{\mathbb{D}_{\mathcal{E}}} \left(\frac{1}{\mu_{\mathcal{E}}(\mathbb{D}_{\mathcal{E}}|_{x,r})} \int_{\mathbb{D}_{\mathcal{E}}|_{x,r}} \rho^{(2)}_{\mathcal{S}}(x, x') \mathrm{d}x' \right) \mathrm{d}x , \tag{5.25}$$

où $\mathbb{D}_{\mathcal{E}}|_{x,r}$ est l'ensemble des points de $\mathbb{D}_{\mathcal{E}}$ qui ont une distance de r avec le point x

$$\mathbb{D}_{\mathcal{E}}|_{x,r} := \left\{ x' : x' \in \mathbb{D}_{\mathcal{E}}, dist_{\mathcal{E}}(x, x') = r \right\} . \tag{5.26}$$

Définition 5.8 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$, on appelle **fonction de distribution de distance**, notée $\breve{g}_{\mathcal{S}}$, la fonction $\mathbb{R}^+ \to \mathbb{R}^+$ définie par*

$$\breve{g}_{\mathcal{S}}(\tau) = \frac{\breve{\rho}_{\mathcal{S}}(r)}{\rho^2_{\mathcal{S}}} . \tag{5.27}$$



(a) (b)

FIGURE 5.2 – *Exemple de fonction de distribution différentielle et de distribution de distance d'un ensemble de points optimisé par BNOT dans $\mathcal{T}^2$.* ***(a) :*** *La fonction de distribution différentielle $\breve{\rho}_{\mathcal{S}}^{(2)}$.* ***(b) :*** *La fonction de distribution de distance $\breve{\rho}_{\mathcal{S}}$ (voir 5.2.1). Les fonctions $\breve{g}_{\mathcal{S}}^{del.}$ et $\breve{g}_{\mathcal{S}}^{min}$ sont montrées en pointillés (voir 5.3.3). L'axe des distances est normalisé par $\sqrt{N}$. Dans cet exemple, la fonction de distribution de distance peut être vue comme la moyenne radiale de la fonction de distribution différentielle.*

Concrètement, la densité de distances correspond à l'espérance de la densité de paires sur toutes les paires de points séparés par une distance r. Cette définition peut sembler étonnamment complexe, mais sans plus de connaissance sur l'espace d'échantillonnage, c'est la seule définition, utilisant $\rho_{\mathcal{S}}^{(2)}$, qui reste suffisamment générale.

Lemme 5.6 *Soit $\mathcal{S}$ un ensemble de points dans $\mathcal{E}$ et $x_0 \in \mathbb{D}_{\mathcal{E}}$, on a l'égalité suivante*

$$\breve{\rho}_{\mathcal{S}}(r) = \frac{1}{\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}|_{x_0,r}\right)} \sum_{i=1}^{N}\sum_{j=1}^{N} \left\langle \delta(dist_{\mathcal{E}}(\boldsymbol{s}_i,\boldsymbol{s}_j),r)\right\rangle . \tag{5.28}$$

Preuve 5.6 *$\mathbb{T}_{\mathcal{E}}$ étant transitif sur $\mathbb{D}_{\mathcal{E}}$, on sait qu'il existe $\tau \in \mathbb{T}_{\mathcal{E}}$ tel que $x = x_0 * \tau$, où x_0 est un point arbitraire du domaine d'échantillonnage. On peut donc écrire que*

$$\begin{aligned}\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}|_{x,r}\right) &= \mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}|_{x_0*\tau,r}\right)\\ &= \mu_{\mathcal{E}}\left(\left\{x' : x' \in \mathbb{D}_{\mathcal{E}}, dist_{\mathcal{E}}(x_0*\tau, x') = r\right\}\right).\end{aligned} \tag{5.29}$$

Comme le groupe des translations préserve les distances et la mesure, on peut écrire que

$$\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}|_{x,r}\right) = \mu_{\mathcal{E}}\left(\left\{x' * \tau^{-1} \ : x' \in \mathbb{D}_{\mathcal{E}}, dist_{\mathcal{E}}(x_0, x' * \tau^{-1}) = r\right\}\right). \tag{5.30}$$

En effectuant le changement de variable $x'' = x' * \tau^{-1}$, on obtient

$$\begin{aligned}\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}|_{x,r}\right) &= \mu_{\mathcal{E}}\left(\left\{x'' : x'' \in \mathbb{D}_{\mathcal{E}}, dist_{\mathcal{E}}(x_0, x'') = r\right\}\right)\\ &= \mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}|_{x_0,r}\right),\end{aligned} \tag{5.31}$$

c'est à dire que la valeur de $\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}|_{x,r}\right)$ est la même pour tout x. En reportant ce résultat dans Eq. (5.25), on obtient

$$\breve{\rho}_{\mathcal{S}}(r) = \frac{1}{\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}|_{x_0,r}\right)} \int_{\mathbb{D}_{\mathcal{E}}}\int_{\mathbb{D}_{\mathcal{E}}|_{x,r}} \rho_{\mathcal{S}}^{(2)}(x,x')\mathrm{d}x'\mathrm{d}x . \tag{5.32}$$

On peut changer le domaine d'intégration de la deuxième intégrale de la façon suivante

$$\check{\rho}_{\mathcal{S}}(r) = \frac{1}{\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}|_{x_0,r}\right)} \int_{\mathbb{D}_{\mathcal{E}}} \int_{\mathbb{D}_{\mathcal{E}}|_{x,r}} \rho_{\mathcal{S}}^{(2)}(x,x')\delta(dist_{\mathcal{E}}(x,x'),r)\mathrm{d}x'\mathrm{d}x\,. \tag{5.33}$$

En utilisant Def. (5.2), et en simplifiant les intégrales, on obtient finalement Eq. (5.28). □

Grâce à Lemme (5.6), le calcul de la densité de distances est relativement simple à faire, puisqu'il suffit de calculer un histogramme des distances entres les points. La figure (5.2) montre un exemple de fonction de distribution de distance. On la retrouve souvent sous le nom de *radial distribution function* car elle correspond, pour l'espace euclidien, à la moyenne radiale de la fonction de distribution différentielle.

5.3 Autres fonctions liées à la densité

Comme nous l'avons vu, les fonctions de distribution différentielle et de distribution de distances sont liées à la fonction de densité de paires. Dans cette section, nous discutons d'autres fonctions qui sont liées à la densité de paires. Bien que ces fonctions ne soient pas essentielles dans l'exposé de nos travaux, nous avons jugé utile d'en parler car elles peuvent être intéressantes pour des développements futurs, ou tout simplement car elles permettent de simplifier certaines expressions.

5.3.1 Fonction de corrélation totale

Définition 5.9 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$, on appelle* ***fonction de corrélation totale****, notée $h_{\mathcal{S}}$, la fonction $\mathbb{D}_{\mathcal{E}} \times \mathbb{D}_{\mathcal{E}} \to \mathbb{R}^+$ définie par*

$$h_{\mathcal{S}}(x,x') := \frac{Cov(S(x),S(x'))}{\rho_{\mathcal{S}}^{(1)}(x)\rho_{\mathcal{S}}^{(1)}(x')}\,. \tag{5.34}$$

La fonction de corrélation totale est aussi définie par l'équation de Ornstein–Zernike, qui permet de décomposer $\check{h}_{\mathcal{S}}$ en fonction d'un terme de corrélation directe et indirecte [Hansen and McDonald, 1990, sec. 3.5]. Nous n'aborderons pas cette partie, puisqu'elle repose sur des suppositions propres à la physique. Dans notre contexte, l'intérêt de la fonction de corrélation est son lien avec la fonction de distribution de paires.

Lemme 5.7 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$, on a l'égalité suivante*

$$h_{\mathcal{S}}(x,x') = g_{\mathcal{S}}(x,x') - 1\,. \tag{5.35}$$

Preuve 5.7 *Il suffit d'utiliser Eq. (5.8) dans Eq. (5.34) pour obtenir*

$$\begin{aligned} h_{\mathcal{S}}(x,x') &= \frac{\rho_{\mathcal{S}}^{(2)}(x,x') - \rho_{\mathcal{S}}^{(1)}(x)\rho_{\mathcal{S}}^{(1)}(x')}{\rho_{\mathcal{S}}^{(1)}(x)\rho_{\mathcal{S}}^{(1)}(x')} \\ &= \frac{\rho_{\mathcal{S}}^{(2)}(x,x')}{\rho_{\mathcal{S}}^{(1)}(x)\rho_{\mathcal{S}}^{(1)}(x')} - 1\,. \end{aligned} \tag{5.36}$$

On obtient Eq. (5.35) en remarquant que la fraction correspond à la définition de la fonction de distribution de paires. □

Comme beaucoup d'ensembles de points stochastiques ont une fonction de distribution de paires qui converge vers 1, la fonction de corrélation totale converge généralement vers zéro. Dans certaines applications, elle est donc préférée car son intégrale est finie. Dans ce manuscrit, nous l'utiliserons parfois pour simplifier certaines expressions. Nous utiliserons également la notation $\check{h}_{\mathcal{S}}$ pour désigner la version différentielle de la fonction de corrélation totale, définie par $\check{h}_{\mathcal{S}}(\tau) = \check{g}_{\mathcal{S}}(\tau) - 1$.

5.3.2 Auto-corrélation

Définition 5.10 *Soit f une fonction $\mathbb{D}_{\mathcal{E}} \to \mathbb{C}$, on appelle* ***auto-corrélation*** *de f, noté $(f \oplus f)$, la fonction $\mathbb{T}_{\mathcal{E}} \to \mathbb{C}$ suivante*

$$(f \oplus f)(\tau) := \int_{\mathbb{D}_{\mathcal{E}}} f(x * \tau)\overline{f(x)}\mathrm{d}x\,. \tag{5.37}$$

L'auto-corrélation de la fonction représentative de $\mathcal{S}$ est directement reliée à la fonction de distribution différentielle.

Lemme 5.8 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$, on a l'égalité suivante*

$$\check{g}_{\mathcal{S}}(\tau) = \frac{\langle (S \oplus S)(\tau) \rangle}{N\rho_{\mathcal{S}}}\,. \tag{5.38}$$

Preuve 5.8

$$(S \oplus S)(\tau) = \int_{\mathbb{D}_{\mathcal{E}}} S(x * \tau)S(x)\mathrm{d}x\,. \tag{5.39}$$

En développant cette équation avec Def. (3.2, p. 17), on obtient

$$(S \oplus S)(\tau) = \int_{\mathbb{D}_{\mathcal{E}}} \sum_{i=1}^{N}\sum_{j=1}^{N} \delta(x, \boldsymbol{s}_i)\delta(x * \tau, \boldsymbol{s}_j)\mathrm{d}x\,. \tag{5.40}$$

Si l'on étudie maintenant l'espérance de cette équation, on obtient à un facteur près l'expression de la densité différentielle.

$$\begin{aligned}\langle (S \oplus S)(\tau) \rangle &= \int_{\mathbb{D}_{\mathcal{E}}} \sum_{i=1}^{N}\sum_{j=1}^{N} \left\langle \delta(x, \boldsymbol{s}_i)\delta(x * \tau, \boldsymbol{s}_j) \right\rangle \mathrm{d}x \\ &= \int_{\mathbb{D}_{\mathcal{E}}} \mathrm{d}x \check{\rho}^{(2)}_{\mathcal{S}}(\tau) \\ &= \mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)\check{\rho}^{(2)}_{\mathcal{S}}(\tau)\,.\end{aligned} \tag{5.41}$$

En réordonnant cette équation et en utilisant la définition de la fonction de distribution différentielle, on obtient finalement Eq. (5.38). □

Ce lien montre que, bien que nous ayons défini la fonction de densité de paires à partir de $\mathcal{S}$, il est possible également d'avoir une approche plutôt basée sur le traitement du signal. L'auto-corrélation peut être vue comme l'équivalent de la fonction de distribution de paires pour les fonctions continues. Dans le chapitre 7, nous relions la variance de l'intégration avec la fonction de distribution différentielle et l'auto-corrélation de la fonction à intégrer.

5.3.3 Distribution de voisinage

La fonction de distribution de distance compte le nombre de paires de points appartenant à $\mathcal{S}$ séparés par une distance de r. Soit $\mathcal{N}_{\mathcal{S}}$ l'ensemble des couples de points possibles de $\mathcal{S}$, la fonction de distribution différentielle peut être ré-écrite comme

$$\check{g}_{\mathcal{S}}(r) = \frac{1}{\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}|_{x_0,r}\right)\rho_{\mathcal{S}}^2} \sum_{(\mathbf{s}_i,\mathbf{s}_j)\in\mathcal{N}_{\mathcal{S}}} \left\langle \delta(\mathrm{dist}_{\mathcal{E}}(\mathbf{s}_i,\mathbf{s}_j), r)\right\rangle . \tag{5.42}$$

Il arrive que l'on souhaite restreindre la fonction de distribution de distance à un sous-ensemble de $\mathcal{N}_{\mathcal{S}}$. Une solution consiste à ne prendre en compte que le plus proche voisin de chaque point. L'ensemble de couples de points devient donc

$$\mathcal{N}_{\mathcal{S}}^{\min} := \left\{(\mathbf{s}_i,\mathbf{s}_j) : \mathbf{s}_i,\mathbf{s}_j \in \mathcal{S}, \forall \mathbf{s}_k \in \mathcal{S}, \mathrm{dist}_{\mathcal{E}}(\mathbf{s}_i,\mathbf{s}_j) \leq \mathrm{dist}_{\mathcal{E}}(\mathbf{s}_i,\mathbf{s}_k)\right\} . \tag{5.43}$$

Cette relation de voisinage est par exemple utilisée dans la méthode d'optimisation FPO [Schlömer *et al.*, 2011]. Soit $\check{g}_{\mathcal{S}}^{\min}$ la fonction de distribution de distance restreinte à $\mathcal{N}_{\mathcal{S}}^{\min}$, on a la propriété suivante

$$\check{g}_{\mathcal{S}}^{\min}(r) \leq \check{g}_{\mathcal{S}}(r) . \tag{5.44}$$

On appelle **distance minimale moyenne** de $\mathcal{S}$, noté $\delta_{\mathcal{S}}^{\min}$, l'espérance de $\check{g}_{\mathcal{S}}^{\min}(r)$ par rapport à r

$$\delta_{\mathcal{S}}^{\min} = \int_0^{\infty} r\, \check{g}_{\mathcal{S}}^{\min}(r)\mathrm{d}r . \tag{5.45}$$

Une autre solution de restriction de voisinage consiste à prendre les paires de points correspondant à la triangulation de Delaunay. Ce type de voisinage, noté $\mathcal{N}_{\mathcal{S}}^{\mathrm{del.}}$, correspond au premier anneau de voisin autour d'un point, incluant donc le plus proche voisin de chaque point. Soit $\check{g}_{\mathcal{S}}^{\mathrm{del.}}$ la fonction de distribution de distance restreinte à $\mathcal{N}_{\mathcal{S}}^{\mathrm{del.}}$, on a donc la propriété suivante

$$\check{g}_{\mathcal{S}}^{\min}(r) \leq \check{g}_{\mathcal{S}}^{\mathrm{del.}}(r) \leq \check{g}_{\mathcal{S}}(r) . \tag{5.46}$$

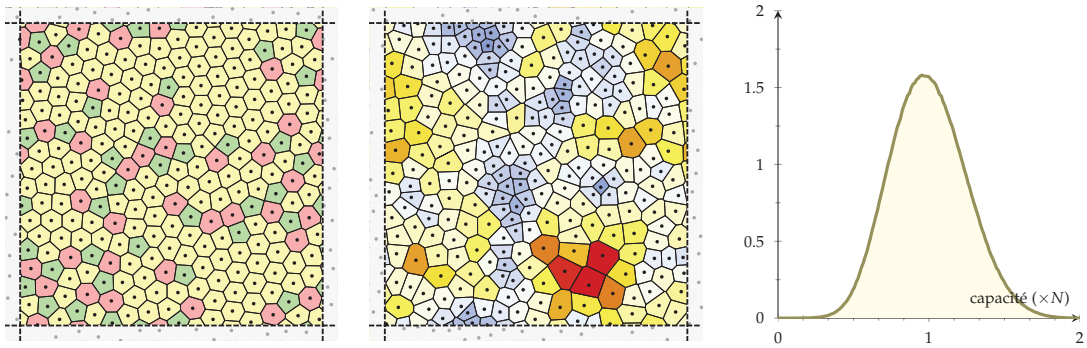
La figure (5.2) montre un exemple de fonctions de distribution de distances restreintes à $\mathcal{N}_{\mathcal{S}}^{\mathrm{del.}}$ et $\mathcal{N}_{\mathcal{S}}^{\min}$. Ces fonctions de distribution de distances restreintes ont une certaine importance dans les méthodes d'optimisation. Plutôt que d'optimiser un ensemble de points pour que sa fonction de distribution de distance corresponde à une cible donnée, on travaille à la place sur ces versions restreintes, qui sont beaucoup plus simples à traiter.

Dans la section suivante, nous verrons brièvement comment utiliser les cellules de Voronoi pour obtenir quelques caractéristiques sur un ensemble de points.

5.4 Diagramme de Voronoi

Les diagrammes de Voronoi permettent de partitionner un domaine à partir d'un ensemble de points. À chaque point $\mathbf{s}_k$ de l'ensemble correspond une cellule de Voronoi ν_k. Cette cellule représente tous les éléments du domaine dont le point le plus proche est le point $\mathbf{s}_k$

$$\nu_k := \left\{x : x \in \mathbb{D}_{\mathcal{E}}, \forall \mathbf{s}_i \in \mathcal{S}, \mathrm{dist}_{\mathcal{E}}(x,\mathbf{s}_k) \leq \mathrm{dist}_{\mathcal{E}}(x,\mathbf{s}_i), \quad x \in \mathbb{D}_{\mathcal{E}}\right\} . \tag{5.47}$$



(a) (b) (c)

FIGURE 5.3 – *Utilisation des diagrames de Voronoi dans l'analyse spatiale des ensembles de points.* ***(a) :*** *un diagramme de Voronoi d'un ensemble de points optimisé par relaxation de Lloyd. Les cellules de couleur jaune sont des hexagones, celles de couleur verte des pentagones et celles de couleur rouge de heptagones. Comme l'ensemble de points est régulier, on retrouve des amas d'hexagones de grandes tailles.* ***(b) :*** *un diagramme de Voronoi d'un ensemble de points non-uniforme. Les cellules bleues ont une faible capacité, et les cellules rouges ont une grande capacité. Les amas de cellules rouges et bleues montrent que la densité n'est pas la même partout.* ***(c) :*** *la fonction de distribution de capacité d'un ensemble de points* jittered.

Définition 5.11 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$, on appelle* ***fonction de distribution de capacité****, notée $CDF_{\mathcal{S}}$, la fonction $\mathbb{R}^+ \to \mathbb{R}^+$ définie par*

$$CDF_{\mathcal{S}}(c) := \frac{1}{N} \sum_{k=1}^{N} \langle \delta(c, \mu_{\mathcal{E}}(v_k)) \rangle \ , \tag{5.48}$$

où la fonction δ de Dirac est ici une fonction $\mathbb{R}^+ \to \mathbb{R}$.

La capacité d'un point représente l'aire de sa cellule de Voronoi. La fonction de distribution de capacité correspond donc, à un facteur près, à la probabilité qu'un point ait une certaine capacité c. La figure (5.3) montre un exemple de distribution de capacité dans un ensemble de points non-uniforme.

On utilise également les cellules de Voronoi pour mesurer les régularités d'un ensemble de points. En général, on considère qu'un ensemble de points est régulier quand ses cellules de Voronoi sont des hexagones réguliers. On néglige le cas où les cellules forment des structures périodiques basées sur des triangles ou des carrés car ces configurations sont rarement rencontrées dans les méthodes d'optimisation. Pour mesurer le niveau de régularité d'un ensemble de points, on mesure donc la distance entre la géométrie de ses cellules de Voronoi et un hexagone régulier. Une façon naïve de le faire est de vérifier que le nombre de faces formant la cellule de Voronoi est égale à 6. La figure (5.3) montre un exemple de visualisation de régularité à l'aide d'un diagramme de Voronoi.

5.5 Uniformité

Dans l'introduction, nous avons dit que l'uniformité d'un ensemble de points est un des critères les plus importants pour l'intégration Monte Carlo. Nous détaillons dans cette dernière section les outils d'analyse permettant de mesurer le degré d'uniformité d'un ensemble de points. Puisque l'on a *a priori* aucune information sur la

fonction à échantillonner, on cherche en général à donner la même importance à chaque partie du domaine d'échantillonnage. On parle alors d'uniformité pour décrire à quel point un ensemble de points couvre bien le domaine. En particulier, on veut que pour toute sous-partie A du domaine d'échantillonnage, le nombre de points se trouvant à l'intérieur de A soient égale à la mesure de Lebesgue de A,

$$\psi_A(\mathcal{S}) \approx \rho_{\mathcal{S}} \mu_{\mathcal{E}}(A), \tag{5.49}$$

où $\psi_A(\mathcal{S})$ est le nombre de points se trouvant dans A, défini par

$$\psi_A(\mathcal{S}) := \sum_{k=1}^{N} \chi_A(\mathbf{s}_k). \tag{5.50}$$

Ici, la fonction $\chi_A \mathcal{S}$ est la fonction caractéristique de A, défini dans Eq. (2.7, p. 13). Dans cette section, nous allons décrire les outils et les relations permettant de quantifier le dégrée d'uniformité d'un ensemble de points.

5.5.1 Test de Kolmogorov-Smirnov

Une des façons de quantifier l'uniformité d'un ensemble de points est d'utiliser le **test de Kolmogorov-Smirnov**. Le test de Kolmogorov-Smirnov permet de mesurer la distance entre deux distributions de probabilité en mesurant l'écart maximum entre leur fonction de répartition. Soient f et g deux distributions de probabilité définies sur $\mathbb{R}$ et $\mathcal{KS}(f,g)$ le résultat du test de Kolmogorov-Smirnov entre f et g, est donné par

$$\mathcal{KS}(f,g) := \sup_{x \in \mathbb{R}} \left| \int_{-\infty}^{x} f(y) - g(y) \mathrm{d}y \right|. \tag{5.51}$$

Pour tester l'uniformité d'un ensemble de points dans $\mathcal{U}^1$, on cherche à comparer la fonction S définie dans Def. (3.2, p. 17) avec la fonction de distribution uniforme $U(x) = 1$ sur l'intervalle $[0,1[$. Le teste de Kolmogorov-Smirnov devient donc

$$\mathcal{KS}(U,S) = \sup_{x \in [0,1[} \left| \frac{1}{\rho_{\mathcal{S}}} \int_{0}^{x} S(y) \mathrm{d}y - x \right|, \tag{5.52}$$

où la densité $\rho_{\mathcal{S}}$ permet de normaliser S de telle sorte que son intégrale soit égale à 1. Comme la partie intégrale représente le nombre de points entre 0 et x, on peut réécrire l'équation de la façon suivante

$$\mathcal{KS}(U,S) = \sup_{x \in [0,1[} \left| \frac{\psi_{B_{0,x}}(\mathcal{S})}{\rho_{\mathcal{S}}} - \mu_{\mathcal{E}}(B_{0,x}) \right|, \tag{5.53}$$

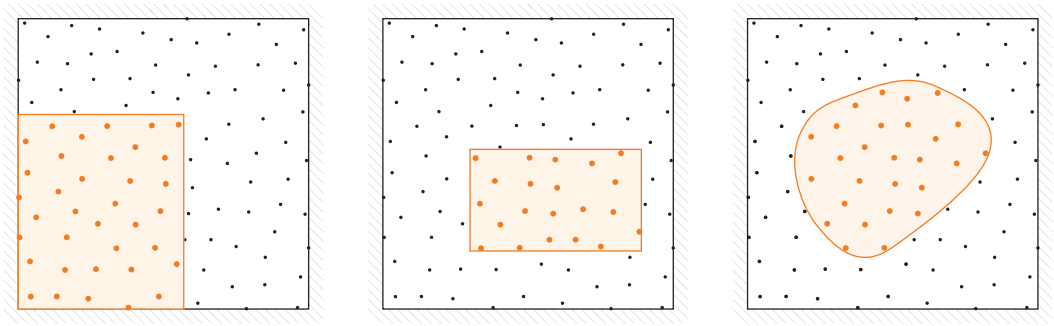
où $B_{0,x}$ représente le sous-domaine $[0,x]$ et $\psi_{B_{0,x}}(\mathcal{S})$ le nombre de points de $\mathcal{S}$ se trouvant dans ce sous-domaine.

5.5.2 Discrépance

Définition 5.12 *[2009, p. 139] Soit $\mathcal{S}$ un ensemble de points dans $\mathcal{U}^d$, on appelle **discrépance à l'origine**, notée $D^*(S)$, la valeur*

$$D^*(S) := \sup_{x \in \mathbb{D}_{\mathcal{U}^d}} \left| \frac{\psi_{B_x}(\mathcal{S})}{\rho_{\mathcal{S}}} - \mu_{\mathcal{E}}(B_{0,x}) \right|, \tag{5.54}$$

où $B_{0,x}$ est un hypercube défini par le vecteur origine et x.



$\mu_{\mathcal{E}}(A) = 0.3819$
$\psi_A(\mathcal{S}) = 36$
(a)

$\mu_{\mathcal{E}}(A) = 0.2065$
$\psi_A(\mathcal{S}) = 21$
(b)

$\mu_{\mathcal{E}}(A) = 0.2585$
$\psi_A(\mathcal{S}) = 27$
(c)

FIGURE 5.4 – *Trois méthodes de calcul de la discrépance, sur une réalisation de la méthode des disques de Poisson dans $\mathcal{U}^2$ avec $N = 100$.* ***(a)*** *La discrépance à l'origine,* ***(b)*** *la discrépance extrême et* ***(c)*** *la discrépance isotrope.*

La discrépance à l'origine, où *star discrepancy*, est une généralisation du teste de Kolmogorov-Smirnov aux espaces unitaires hyper-cubiques $\mathcal{U}^d$. La discrépance à l'origine représente donc la plus grande différence possible entre l'aire d'une boite et le nombre de points qui s'y trouvent, pour toutes les boites contenant l'origine du domaine (voir figure (5.4,a)). La **discrépance extrême**, notée $D(S)$, est une variante où l'on prend en compte toutes les boites, y compris celles qui ne contiennent pas l'origine [Niederreiter, 1992, chap. 2] (voir figure (5.4,b))

$$D(S) = \sup_{x,x' \in \mathbb{D}_{\mathcal{U}^d}} \left| \frac{\psi_{B_{x,x'}}(\mathcal{S})}{\rho_{\mathcal{S}}} - \mu_{\mathcal{E}}\left(B_{x,x'}\right) \right| . \tag{5.55}$$

Cependant, cette valeur n'est pas souvent utilisée étant donné qu'elle est encore plus difficile à calculer que la discrépance à l'origine. De plus, il a été prouvé que la discrépance extrême peut être bornée par la discrépance à l'origine,

$$D^*(S) \leq D(S) \leq 2^d D^*(S) . \tag{5.56}$$

Les détails de cette propriété sont expliqués dans le livre de Kuipers et Niederreiter [1974, p. 93]. Une autre version de la discrépance consiste à prendre en compte d'autres types de sous-domaine de $\mathbb{D}_{\mathcal{E}}$ que les hypercubes. Soit $\mathcal{P}$ l'ensemble de tous les sous-ensembles convexes du domaine $\mathbb{D}_{\mathcal{E}}$, la **discrépance isotrope** (voir figure (5.4,c)) est définie par

$$J(S) = \sup_{A \in \mathcal{P}} \left| \frac{\psi_A(\mathcal{S})}{\rho_{\mathcal{S}}} - \mu_{\mathcal{E}}\left(A\right) \right| . \tag{5.57}$$

La discrépance isotrope est elle aussi bornée par la discrépance à l'origine [Kuipers and Niederreiter, 1974, p. 93]

$$D^*(S) \leq J(S) \leq (4d\sqrt{d} + 1)\sqrt[d]{D^*(S)} . \tag{5.58}$$

Toutes ces mesures d'uniformité ont en commun qu'elles ne cherchent à mesurer l'uniformité que d'une seule réalisation de $\mathcal{S}$ à la fois. Dans le cas des ensembles de points

stochastiques, il est pourtant nécessaire de prendre en compte toutes les réalisations. Il est bien sûr possible d'étudier l'espérance de la discrépance $\langle D(\mathcal{S})\rangle$ sur l'ensemble des réalisations de $\mathcal{S}$. Cependant, cette solution n'est pas satisfaisante étant donné que la discrépance cherche à quantifier le pire cas d'uniformité. Or une valeur basée sur «le pire cas moyen d'uniformité » ne permettrait, ni de connaître le comportement moyen de l'échantillonnage, ni de calculer des bornes supérieurs d'erreur d'intégration. Considérer la discrépance la plus forte sur toutes les réalisations de $\mathcal{S}$ n'est pas non plus une solution satisfaisante. Le bruit blanc est un bon contre-exemple : Puisque toutes les configurations sont possibles avec le bruit blanc, la discrépance maximal doit prendre en compte le cas où tous les points de l'ensemble tombent au même endroit. Dans ce cas là, la mesure de discrépance maximale devient alors très forte, alors que la réalisation responsable de cette forte valeur est statistiquement insignifiante.

Une approche alternative consiste à étudier la variance et le nombre de points se trouvant à l'intérieur de toute forme géométrique mesurable. On intègre donc dans la discrépance une dimension stochastique, tout en supprimant l'opération de supérieur. Le résultat de l'analyse n'est donc pas un scalaire, comme pour la discrépance, mais une fonction donnant l'espérance et la variance pour une forme géométrique donnée.

5.5.3 Discrépance stochastique

Soit A un sous-ensemble mesurable de $\mathbb{D}_{\mathcal{E}}$, on appelle **discrépance stochastique** l'étude de l'espérance et la variance de $\psi_A(\mathcal{S})$. On retrouve cet outil d'analyse, sous le nom de *local density fluctuation* dans [Torquato and Stillinger, 2003]. Comme nous l'avons déjà dit, le nombre de points $\psi_A(\mathcal{S})$ se trouvant dans le sous-domaine A s'exprime à l'aide de la fonction caractéristique χ_A de A. Cette quantité peut aussi s'exprimer par la fonction représentative S de l'ensemble de points $\mathcal{S}$:

$$\psi_A(\mathcal{S}) = \int_A S(x)\mathrm{d}x\,. \tag{5.59}$$

Un point intéressant consiste à étudier comment évolue l'espérance et la variance de $\psi_A(\mathcal{S})$ quand la densité de l'ensemble de points augmente. Pour étudier cette évolution, il est préférable d'avoir d'abord une expression de la discrépance stochastique en fonction de la densité.

Lemme 5.9 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$, on a l'égalité suivante*

$$\langle\psi_A(\mathcal{S})\rangle = \int_{\mathbb{D}_{\mathcal{E}}} \rho_{\mathcal{S}}^{(1)}(x)\chi_A(x)\mathrm{d}x\,. \tag{5.60}$$

Preuve 5.9 *À partir de Eq. (5.59) et de Def. (3.2, p. 17), on peut écrire que*

$$\langle\psi_A(\mathcal{S})\rangle = \int_A \left\langle \sum_{k=1}^{N} \delta(x, \mathbf{s}_k) \right\rangle \mathrm{d}x\,. \tag{5.61}$$

On remarquant que la partie à l'intérieur de l'intégrale correspond à la définition de la densité locale, et en changeant le domaine d'intégration de l'intégrale à l'aide de la fonction caractéristique de A, on obtient Eq. (5.60). □

Lemme 5.10 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$, on a l'égalité suivante*

$$Var(\psi_A(\mathcal{S})) = \int_{\mathbb{D}_{\mathcal{E}}}\int_{\mathbb{D}_{\mathcal{E}}} \rho_{\mathcal{S}}^{(1)}(x)\rho_{\mathcal{S}}^{(1)}(x')\check{h}_{\mathcal{S}}(x,x')\chi_A(x)\chi_A(x')\mathrm{d}x'\mathrm{d}x\,. \tag{5.62}$$

Preuve 5.10 *À partir de l'expression de la variance donnée dans Eq. (3.7, p. 19) et de Eq. (5.59), on peut écrire que*

$$\begin{aligned} Var(\psi_A(\mathcal{S})) &= \langle \psi_A(\mathcal{S})^2 \rangle - \langle \psi_A(\mathcal{S}) \rangle^2 \\ &= \int_{\mathbb{D}_\mathcal{E}} \int_{\mathbb{D}_\mathcal{E}} \left(\langle S(x)S(x') \rangle - \langle S(x) \rangle \langle S(x') \rangle \right) \chi_A(x)\chi_A(x') \mathrm{d}x' \mathrm{d}x \,. \end{aligned} \tag{5.63}$$

Le terme $\langle S(x)S(x') \rangle$ correspond à la définition de la densité de paires, et les termes $\langle S(x) \rangle$ et $\langle S(x') \rangle$ à la définition de la densité locale. On obtient donc

$$\begin{aligned} Var(\psi_A(\mathcal{S})) &= \int_{\mathbb{D}_\mathcal{E}} \int_{\mathbb{D}_\mathcal{E}} \left(\rho_\mathcal{S}^{(2)}(x,x') - \rho_\mathcal{S}^{(1)}(x)\rho_\mathcal{S}^{(1)}(x') \right) \chi_A(x)\chi_A(x') \mathrm{d}x' \mathrm{d}x \\ &= \int_{\mathbb{D}_\mathcal{E}} \int_{\mathbb{D}_\mathcal{E}} \rho_\mathcal{S}^{(1)}(x)\rho_\mathcal{S}^{(1)}(x') \left(\frac{\rho_\mathcal{S}^{(2)}(x,x')}{\rho_\mathcal{S}^{(1)}(x)\rho_\mathcal{S}^{(1)}(x')} - 1 \right) \chi_A(x)\chi_A(x') \mathrm{d}x' \mathrm{d}x \,. \end{aligned} \tag{5.64}$$

On obtient Eq. (5.62) en remarquant que le terme entre parenthèses correspond à la définition de fonction de corrélation totale. □

Comme le montrent ces deux lemmes, la densité locale et la densité de paires jouent un rôle crucial dans la discrépance stochastique. Dans le cas des ensembles de points homogènes, les expressions de l'espérance et de la variance de $\psi_A(\mathcal{S})$ se simplifient grandement.

Lemme 5.11 *Soit $\mathcal{S}$ un ensemble de points homogène dans un espace $\mathcal{E}$, on a l'égalité suivante*

$$\langle \psi_A(\mathcal{S}) \rangle = \rho_\mathcal{S} \mu_\mathcal{E}(A) \,. \tag{5.65}$$

Preuve 5.11 *Ce lemme est une conséquence directe de Lemme (5.9) et Lemme (5.1)* □

Lemme 5.12 *Soit $\mathcal{S}$ un ensemble de points homogène dans un espace $\mathcal{E}$, on a l'égalité suivante*

$$Var(\psi_A(\mathcal{S})) = \rho_\mathcal{S}^2 \int_{\mathbb{T}_\mathcal{E}} h_\mathcal{S}(\tau)(\chi_A \oplus \chi_A)(\tau) \mathrm{d}\tau \,. \tag{5.66}$$

Preuve 5.12 *À partir de Eq. (5.62), et à l'aide du changement de variable $x' = x * \tau$, on obtient*

$$Var(\psi_A(\mathcal{S})) = \int_{\mathbb{D}_\mathcal{E}} \int_{\mathbb{T}_\mathcal{E}} \rho_\mathcal{S}^{(1)}(x)\rho_\mathcal{S}^{(1)}(x*\tau)\check{h}_\mathcal{S}(x, x*\tau)\chi_A(x)\chi_A(x*\tau) \mathrm{d}\tau \mathrm{d}x \,. \tag{5.67}$$

En utilisant les propriétés des ensembles de points homogènes, on peut écrire que

$$Var(\psi_A(\mathcal{S})) = \rho_\mathcal{S}^2 \int_{\mathbb{T}_\mathcal{E}} \check{h}_\mathcal{S}(\tau) \int_{\mathbb{D}_\mathcal{E}} \chi_A(x)\chi_A(x*\tau) \mathrm{d}x \mathrm{d}\tau \,. \tag{5.68}$$

En remarquant que l'intégrale sur le domaine d'échantillonnage correspond à la définition de la corrélation de la fonction χ_A, on retrouve Eq. (5.66). □

Dans le cas homogène, on remarque que l'expression de l'espérance de $\psi_A(\mathcal{S})$ ne fait pas intervenir la position des points de $\mathcal{S}$. Pour étudier l'impacte de la position des points sur la discrépance stochastique, il faut donc se tourner vers la variance de $\psi_A(\mathcal{S})$. L'expression de la variance de $\psi_A(\mathcal{S})$ est très proche de celle de la variance

de l'intégration Monte Carlo donnée dans le chapitre 7. En effet, on cherche ici à ce que le nombre de points dans A soit le plus proche possible de sa mesure, ce qui revient à calculer l'intégrale de la fonction caractéristique de A à l'aide du motif d'échantillonnage S. Il faut donc voir Lemme (5.12) comme un cas particulier d'une expression plus générale de la variance de l'intégration, qui sera détaillée dans le chapitre 7.

Conclusion

Dans ce chapitre, nous avons rassemblé les méthodes d'analyse des ensembles de points dans le domaine spatial. Ces méthodes ne sont pas nouvelles, mais nous avons attaché une importance à les traduire au sein du formalisme présenté dans le chapitre 3. Ce faisant, les outils d'analyse présentés ici deviennent indépendants de l'espace d'échantillonnage utilisé et plus orientés vers le rendu d'image.

Parmi tous les outils que nous avons présentés, deux en particulier apparaissent comme fondamentaux : la densité locale et la densité de paires. Le chapitre 7 détaille comment ces outils peuvent être reliés à l'intégration Monte Carlo. Dans le chapitre 8, nous montrons également les fonctions de distribution différentielle et de distribution de distances pour un ensemble de méthodes d'échantillonnage. Toutefois, avant de commencer à traiter les applications de l'analyse spatiale, nous présentons dans le chapitre suivant comment analyser un ensemble de points dans le domaine spectral.

Analyse spectrale 6

Dans ce chapitre, nous abordons l'analyse des ensembles de points dans le domaine spectral. Nous commençons par définir ce qu'est la transformée de Fourier et le domaine spectral dans l'espace euclidien, puis nous décrivons les outils théoriques permettant de généraliser cette notion aux espaces d'échantillonnage définis dans le chapitre 3. Nous précisons aussi comment, concrètement, construire le domaine spectral dans les espaces que nous avons décrit dans le premier chapitre. Une fois ces bases établies, nous montrons comment étudier les ensembles de points à travers l'analyse spectrale et quelles sont les propriétés utiles qui leur correspondent.

6.1 Transformée de Fourier

Définition 6.1 *Soit f une fonction $\mathbb{R}^d \to \mathbb{C}$, on appelle* ***transformée de Fourier*** *la fonction $\mathcal{F}_f : \mathbb{R}^d \to \mathbb{C}$, définie par*

$$\mathcal{F}_f(\omega) := \int_{\mathbb{R}^d} f(x) e^{-i2\pi(x\cdot\omega)} \mathrm{d}x \,, \tag{6.1}$$

où $x \cdot \omega$ désigne ici le produit scalaire entre x et ω.

Cette définition utilise la normalisation «System 1 » décrite dans [Bracewell, 1999, chap. 2]. On utilisera parfois la notation $\mathcal{F}\left[f(x)\right](\omega)$ à la place de $\mathcal{F}_f(\omega)$ pour des raisons de lisibilité. La transformée de Fourier permet de décomposer une fonction en somme de sinusoïdes pondérées par une valeur complexe. On appelle $\mathcal{F}_f(\omega)$ le ω-ième coefficient de la transformée de Fourier, lequel représente le poids associé à une sinusoïde de fréquence ω. Ce poids est une valeur complexe, telle que la norme de $\mathcal{F}_f(\omega)$ est associée à l'amplitude de la sinusoïde et l'argument à la phase de la sinusoïde. Il existe aussi une transformée de Fourier inverse [Bracewell, 1999, chap. 2], définie par

$$\mathcal{F}_f^{-1}(x) := \int_{\mathbb{R}^d} f(\omega) e^{i2\pi(\omega\cdot x)} \mathrm{d}\omega \,, \tag{6.2}$$

et vérifiant la propriété suivante

$$f(x) = \mathcal{F}\left[\mathcal{F}_f(\omega)\right]^{-1}(x) \,. \tag{6.3}$$

Le fait qu'il soit possible de revenir à la fonction originale après une transformée de Fourier et une transformée de Fourier inverse, permet de dire que f et $\mathcal{F}_f$ sont des fonctions duales. On appelle **domaine spatial** la représentation classique de f

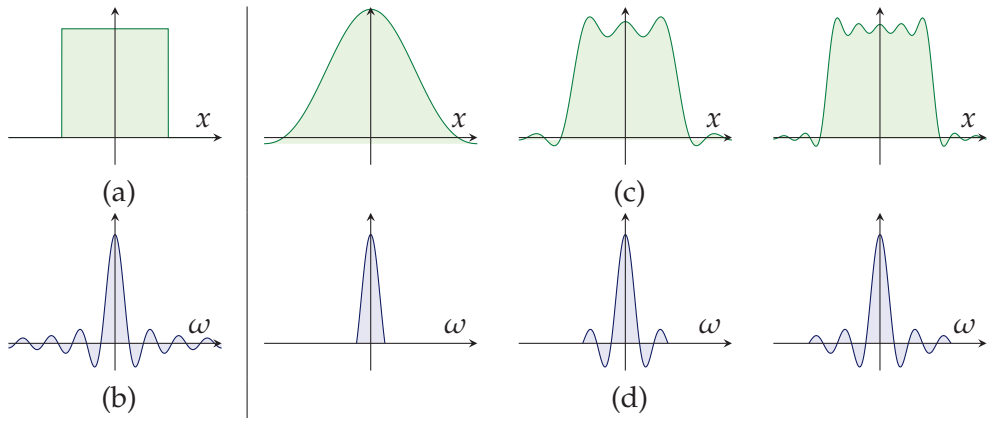


FIGURE 6.1 – *Une même fonction dans le domaine spatial et spectral.* ***(a) :*** *La fonction* Π *dans le domaine spatial.* ***(b) :*** *La fonction* Π *dans le domaine spectral.* ***(c) et (d) :*** *La fonction* Π *partiellement reconstruite par des fonctions sinusoïdales, respectivement dans le domaine spatial et spectral. La reconstruction de la fonction* Π *est de plus en plus fidèle à mesure que l'on ajoute des coefficients de Fourier.*

et **domaine spectrale** la représentation de f par fréquences. Les coefficients de Fourier ayant des valeurs complexes, on peut les décomposer en partie réelle et imaginaire, ou en amplitude et phase. On définit donc $\mathcal{R}_f : \mathbb{R}^d \to \mathbb{R}$ et $\mathcal{I}_f : \mathbb{R}^d \to \mathbb{R}$ tel que

$$\mathcal{F}_f(\omega) = \mathcal{R}_f(\omega) + i\,\mathcal{I}_f(\omega)\,, \tag{6.4}$$

ainsi que $\mathcal{A}_f : \mathbb{R}^d \to \mathbb{R}^+$ et $\Phi_f : \mathbb{R}^d \to [0, 2\pi[$ tel que

$$\mathcal{F}_f(\omega) = \mathcal{A}_f(\omega) e^{i\Phi_f(\omega)}\,. \tag{6.5}$$

Propriété 6.1 *Soit* f *et* f' *deux fonctions* $\mathbb{R}^d \to \mathbb{C}$ *et* $\alpha \in \mathbb{C}$*, on a les propriétés suivantes :*

$$\begin{aligned}
\mathcal{F}\left[\alpha\, f(x) + f'(x)\right](\omega) &= \alpha\, \mathcal{F}_f(\omega) + \mathcal{F}_{f'}(\omega) \\
\mathcal{F}\left[f(\alpha\, x)\right] &= \frac{1}{\alpha^d} \mathcal{F}_f\left(\frac{\omega}{\alpha}\right) \\
\mathcal{F}\left[f(x) f'(x)\right](\omega) &= (\mathcal{F}_f \otimes \mathcal{F}_{f'})(\omega) \\
\mathcal{F}\left[(f \otimes f')(x)\right](\omega) &= \mathcal{F}_f(\omega) \mathcal{F}_{f'}(\omega)\,.
\end{aligned} \tag{6.6}$$

Ici, $\otimes$ représente l'opération de convolution, à ne pas confondre avec l'autocorrélation, noté $\oplus$, bien que ces deux concepts soient proches. La combinaison linéaire de deux fonctions dans le domaine spatial correspond aussi à une combinaison linéaire dans le domaine spectral. Mais toutes les opérations dans le domaine spatial ne sont pas les mêmes dans le domaine spectral. Ainsi, une dilatation par un facteur α dans le domaine spatial correspond à une contraction par ce même facteur dans le domaine spectral. Cette dualité entre opérateurs est particulièrement utile dans certains cas. Comme le montre la troisième et la quatrième propriété, une convolution dans un domaine devient une multiplication dans l'autre. La convolution, qui est une opération généralement coûteuse en terme de calcul, devient trivial dans le domaine spectral. Pour calculer une convolution, il suffit donc de calculer la transformée des deux fonctions, de les multiplier fréquence par fréquence, et de finalement calculer la transformée inverse du résultat.

6.1.1 Transformée de Fourier et intégrales

Propriété 6.2 *Soit f une fonction $\mathbb{R}^d \to \mathbb{C}$, on a la propriété suivante*

$$\mathcal{F}_f(0) = \int_{\mathbb{R}^d} f(x)\mathrm{d}x\,. \tag{6.7}$$

Ainsi, l'opération d'intégration correspond simplement à la valeur du coefficient zéro de la transformée de Fourier. Toujours à propos des intégrales, le théorème de Parseval, aussi appelé théorème de Plancherel, permet de relier l'intégrale du produit de deux fonctions dans le domaine spatial et spectral :

$$\int_{\mathbb{R}^d} f(x)\overline{f'(x)}\mathrm{d}x = \int_{\mathbb{R}^d} \mathcal{F}_f(\omega)\overline{\mathcal{F}_{f'}(\omega)}\mathrm{d}\omega\,, \tag{6.8}$$

avec $f, f' : \mathbb{R}^d \to \mathbb{C}$.

6.1.2 Spectre de puissance

Définition 6.2 *Soit f une fonction $\mathbb{R}^d \to \mathbb{C}$, on appelle* ***spectre de puissance*** *de f, noté $\mathcal{P}_f$, la norme au carré de la transformée de Fourier de f*

$$\mathcal{P}_f(\omega) := |\mathcal{F}_f(\omega)|^2\,. \tag{6.9}$$

Le spectre de puissance représente l'énergie que contient une fonction pour une fréquence donnée. Puisque le spectre de puissance à valeurs dans $\mathbb{R}$, l'information de phase est perdue. En conséquence, le spectre de puissance est invariant par translation, c'est à dire que pour toute fonction f et pour tout $\tau \in \mathbb{T}_{\mathcal{E}^d}$,

$$\mathcal{P}\left[f(x * \tau)\right] = \mathcal{P}\left[f(x)\right]\,. \tag{6.10}$$

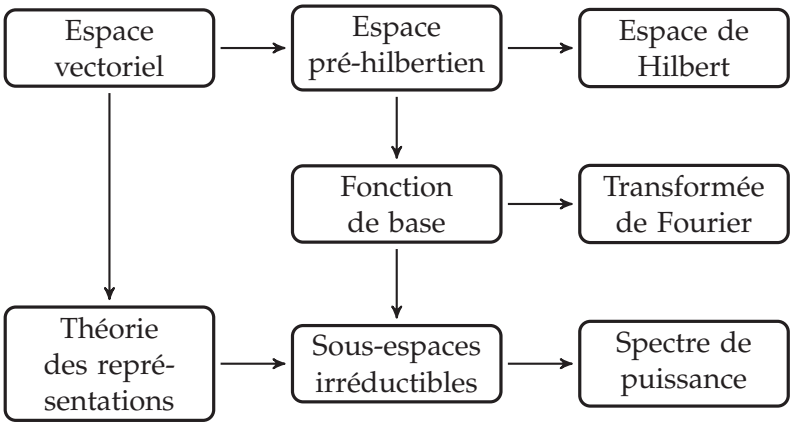
Cela fait du spectre de puissance un outil de choix pour l'étude des ensembles de points homogènes.

Le théorème de Parseval montre que l'intégrale d'une fonction f au carré est égale à l'intégrale du spectre de puissance de f

$$\int_{\mathbb{R}^d} |f(x)|^2\mathrm{d}x = \int_{\mathbb{R}^d} \mathcal{P}_f(\omega)\mathrm{d}\omega\,. \tag{6.11}$$

6.2 Généralisation de la transformée de Fourier

La transformée de Fourier, telle que vue dans la section précédente, est définie pour les fonctions évoluant dans l'espace euclidien. Or, nous souhaitons disposer d'un outil d'analyse spectral permettant de traiter les espaces d'échantillonnage décrits dans le chapitre 3. Nous détaillons dans cette section les éléments théoriques permettant de définir cette généralisation. À l'aide des espaces de Hilbert, il est possible de définir un espace de fonctions décomposables à l'aide d'une somme pondérée de fonctions élémentaires. Les coefficients associés à cette décomposition forme la généralisation de la transformée de Fourier. Pour le spectre de puissance, nous faisons appel à la théorie des représentations afin de trouver une forme invariante, par le groupe $\mathbb{T}_{\mathcal{E}}$ des actions, de la transformée de Fourier.



6.2.1 Espace de Hilbert

Définition 6.3 *Soit K un corps commutatif, généralement $\mathbb{R}$ ou $\mathbb{C}$, dont la multiplication est un produit implicite tel que $\forall k, k' \in K, \quad kk' \in K$, on dit que $(H, K, +, \cdot)$ est un K-**espace vectoriel** si les propriétés suivantes sont respectées*

- *$(H, +)$ est un groupe où l'operateur $+$ est commutatif,*
- *l'opérateur $\cdot$ est une loi de composition externe à gauche : $K \times H \to H$,*
- *$\forall f \in H, \quad k, k' \in K, \quad k \cdot (k' \cdot f) = (kk') \cdot f$,*
- *$\exists k_0 \in K$ tel que $\forall f \in H, \quad k_0 \cdot f = f$,*
- *$\forall k \in K, \quad f, f' \in H, \quad k \cdot (f + f') = k \cdot f + k \cdot f'$ et*
- *$\forall k, k' \in K, \quad f \in H, \quad (k + k') \cdot f = k \cdot f + k' \cdot f$.*

On dira par abus de langage que H est un espace K-vectoriel, ou tout simplement un espace vectoriel. On note 0, l'élément qui vérifie $\alpha \cdot 0 = 0$ et $0 + f = f$ pour tout $\alpha \in K$ et pour tout $f \in H$. L'exemple d'espace vectoriel le plus simple est l'espace euclidien. Si l'on considère $H = \mathbb{R}^d$ et $K = \mathbb{R}$, $+$ comme l'addition et $\cdot$ comme le produit, on obtient un espace vectoriel de dimension d. Un autre exemple, qui sera celui pour lequel nous utiliserons les espaces vectoriels, est l'espace des fonctions complexes. En effet, soit $F(X, \mathbb{C})$ l'ensemble des fonctions de X vers $\mathbb{C}$, il est par exemple toujours possible d'écrire des égalités du type $k(f(x) + f'(x)) = kf(x) + kf'(x)$ où $k \in \mathbb{C}$ et $f, f' \in F(X, \mathbb{C})$.

Définition 6.4 *Soit H un K-espace vectoriel, on dit que H est un **espace pré-hilbertien** si il est muni d'un opérateur $\langle \cdot, \cdot \rangle : H \times H \to H$, appelé produit scalaire.*

En anglais, on trouve le terme *inner product space* et le terme *pre-Hilbert space* pour désigner les espaces pré-hilbertiens. L'espace euclidien muni du produit scalaire classique forme un espace pré-hilbertien. L'espace des fonctions $F(X, \mathbb{C})$ peut aussi être un espace pré-hilbertien si on lui associe le produit scalaire suivant

$$\langle f, f' \rangle = \int_X f(x) \overline{f'(x)} \mathrm{d}x \,, \tag{6.12}$$

où f et f' sont deux fonctions de X vers $\mathbb{C}$. Dans un espace pré-hilbertien, on définit la norme, noté $||f||$, comme la racine de $\langle f, f \rangle$.

Définition 6.5 *Soit H un espace pré-hilbertien, on dit que H est un **espace de Hilbert** s'il est un espace métrique complet, c'est à dire si pour toute suite de Cauchy dans H, la limite de cette suite appartient à H.*

Ici encore, l'espace euclidien est un espace de Hilbert. En revanche, l'espace des fonctions à valeurs réelles ou complexes n'est pas un espace de Hilbert. Pour obtenir un espace de fonctions qui soit un espace de Hilbert, il faut se restreindre aux fonctions appartenant à $L^2(X, \mathbb{C})$, c'est à dire les fonctions $f : X \to \mathbb{C}$ telles que

$$\int_X |f(x)|^2 \mathrm{d}x < \infty . \tag{6.13}$$

Définition 6.6 *Soit H un espace de Hilbert et $u : H \to H$ une fonction surjective de H vers H, on dit que u est une opération unitaire si elle préserve les produits scalaires*

$$\langle u(f), u(f') \rangle = \langle f, f' \rangle , \tag{6.14}$$

où $f, f' \in H$.

L'intérêt des espaces vectoriels (pré-hilbertiens ou de Hilbert) est qu'il est possible de les décomposer à l'aide d'une base. Nous allons voir comme se définit cette base, y compris pour des espace vectoriels particulier comme les espaces de fonctions.

6.2.2 Décomposition d'un espace pré-hilbertien

Définition 6.7 *On dit que $M \subset H$ est un **sous-espace** de H si M est lui aussi un espace vectoriel.*

En pratique, il est suffisant de montrer que pour tout $f, f' \in M$, $a \in K$, on a $a \cdot f + f' \in M$. On dit que deux vecteurs $f, f' \in H$ sont orthogonaux si $\langle f, f' \rangle = 0$. De la même façon, on dit que deux sous-ensembles M et M' de H sont orthogonaux si $\forall f \in M, f' \in M', \langle f, f' \rangle = 0$. On écrira alors que $f \perp f'$ et que $M \perp M'$. On note $M^\perp$ l'ensemble des éléments de H qui sont orthogonaux à tous les éléments de M :

$$M^\perp := \{ f : f \in H, \forall f' \in M, f \perp f' \} . \tag{6.15}$$

Définition 6.8 *Soit H un espace pré-hilbertien et $\mathcal{B} = \{ b^\omega : \omega \in \mathbb{F} \}$ un sous-ensemble de H indexé par $\mathbb{F}$, on dit que l'ensemble $\mathcal{B}$ forme une **base orthogonale** de H si les propriétés suivantes sont respectées :*
- $\forall \omega, \omega' \in \mathbb{F}, \quad \omega \neq \omega', \quad b^\omega \perp b^{\omega'}$,
- $\forall f \in H \setminus \{0\}, \exists \omega \in \mathbb{F}$ *tel que* $\langle f, b^\omega \rangle \neq 0$.

La deuxième condition signifie qu'il n'est pas possible de trouver un élément de H qui soit orthogonal à tous les éléments de la base orthogonale, à l'exception de l'élément zéro. On appelle généralement les éléments de la base $\mathcal{B}$ des vecteurs. Le dernier point de la définition signifie que tout élément f de H peut s'exprimer par une somme pondérée des vecteurs de $\mathcal{B}$. On dit alors que $Span(\mathcal{B}) = H$, où $Span(\mathcal{B})$ est l'ensemble de toutes les combinaisons linéaires des vecteurs de la base $\mathcal{B}$. La relation entre $f \in H$ et la base $\mathcal{B}$ est donnée par

$$f = \sum_{\omega \in \mathbb{F}} \mathcal{F}_f(\omega) b^\omega , \tag{6.16}$$

où le coefficient $\mathcal{F}_f(\omega)$ associé à chaque vecteur de la base est donné par la relation

$$\mathcal{F}_f(\omega) = \langle f, b^\omega \rangle . \tag{6.17}$$

Ici, la notation utilisée n'est pas due au hasard puisque le coefficient $\mathcal{F}_f(\omega)$ est l'équivalent de la transformée de Fourier que nous avons définie précédemment. On dit qu'une base est orthonormale si la norme de chacun des vecteurs de la base est égale à un. Tout espace de Hilbert admet une base orthogonale. En revanche, il existe des espaces pré-hilbertiens qui n'ont pas de base orthogonales . Pour cette raison, beaucoup d'ouvrages font l'hypothèse que l'espace utilisé est un espace de Hilbert. Un dernier point concerne la cardinalité de $\mathbb{F}$. La notation que nous avons utilisé dans Eq. (6.16) implique que $\mathbb{F}$ est dénombrable. Or ce n'est pas toujours le cas, même pour les espaces de Hilbert. On dit qu'un espace pré-hilbertien est **séparable** si les éléments de sa base orthogonale sont dénombrables.

On note $H_\omega := Span(\{b^\omega\})$ le sous-espace des éléments de H qui sont générés par b^ω. À partir de ces sous-espaces, on peut reconstruire l'espace H à l'aide de la somme directe, notée

$$H = H_{\omega_0} \oplus H_{\omega_1} \oplus ... = \bigoplus_{\omega \in \mathbb{F}} H_\omega \,. \tag{6.18}$$

Cette notation indique que les éléments de H peuvent être retrouvés en sommant ensemble des éléments pris dans chaque sous-espace H_ω. La somme directe de deux sous-espaces ne doit donc pas être confondue avec l'union de deux sous-espaces.

Comme nous n'utilisons les espaces pré-hilbertiens que dans le cas des espaces de fonctions, nous allons à partir de maintenant restreindre les définitions que nous donnerons. Lorsque nous parlerons d'espaces pré-hilbertiens, nous utiliserons l'espace de fonctions $F(\mathbb{D}_\mathcal{E}, \mathbb{C})$ muni du produit scalaire défini par Eq. (6.12), et lorsque nous parlerons d'espaces de Hilbert, nous utiliserons l'espace de fonctions $L^2(\mathbb{D}_\mathcal{E}, \mathbb{C})$ muni du même produit scalaire. Nous faisons de plus l'hypothèse que l'espace de fonction admet une base orthogonale. On notera $\mathcal{B}$ la base orthogonale associée à l'espace de fonction et $b^\omega, \omega \in \mathbb{F}$ les vecteurs de cette base.

6.2.3 Similitude avec la transformée de Fourier

Avec la décomposition des espaces pré-hilbertiens, nous avons une forme de généralisation de la transformée de Fourier. Pour compléter cette généralisation, il est cependant nécessaire de faire le parallèle entre les propriétés de la transformée de Fourier et la décomposition des espaces pré-hilbertiens.

Transformée inverse L'équation Eq. (6.17) est l'équivalent de la transformée de Fourier. En utilisant la définition du produit scalaire, on obtient effectivement une définition proche de la transformée de Fourier

$$\mathcal{F}_f(\omega) = \int_{\mathbb{D}_\mathcal{E}} f(x)\overline{b^\omega(x)}\mathrm{d}x \,. \tag{6.19}$$

La seule différence concerne l'ensemble d'intégration qui devient $\mathbb{D}_\mathcal{E}$ et la fonction exponentielle qui devient une fonction b^ω. L'équation Eq. (6.16) est quant à elle l'équivalent de la transformée inverse. Une différence majeure est que l'intégrale a été remplacée par une somme discrète. Comme nous l'avons dit dans l'introduction, section 2.2.4, il est toujours possible de réécrire une somme discrète en intégrale en utilisant une mesure de comptage. On obtient alors une expression de la transformée inverse suivante

$$f(x) = \int_{\mathbb{F}} \mathcal{F}_f(\omega) b^\omega(x)\mathrm{d}\omega \,. \tag{6.20}$$

Dans le cas de la transformée de Fourier classique, le domaine spectral $\mathbb{F}$ est constitué par l'espace euclidien $\mathbb{R}^d$ muni d'une mesure continue. Dans le cas des espaces de Hilbert, le domaine spectral $\mathbb{F}$ est constitué d'un ensemble dénombrable muni d'une mesure de comptage. L'intégrale nous permet donc ici de garder le même formalisme pour ces deux outils.

Théorème de Parseval Le théorème de Parseval est un autre exemple de similitude entre la transformée de Fourier et notre généralisation. D'après [Rudin, 1987, chap. 4], si un espace de Hilbert possède une base orthonormale, alors le théorème de Parseval existe. Dans notre cas, en se restreignant à l'espace de Hilbert $L^2(\mathbb{D}_\mathcal{E}, \mathbb{C})$, on obtient pour $f, f' \in L^2(\mathbb{D}_\mathcal{E}, \mathbb{C})$ l'égalité suivante

$$\int_{\mathbb{D}_\mathcal{E}} f(x)\overline{f'(x)}\mathrm{d}x = \sum_{\omega\in\mathbb{F}} \mathcal{F}_f(\omega)\overline{\mathcal{F}_{f'}(\omega)}\,. \tag{6.21}$$

Ici encore, on peut retrouver une forme intégrale en utilisant la mesure sur l'ensemble $\mathbb{F}$

$$\int_{\mathbb{D}_\mathcal{E}} f(x)\overline{f'(x)}\mathrm{d}x = \int_{\mathbb{F}} \mathcal{F}_f(\omega)\overline{\mathcal{F}_{f'}(\omega)}\mathrm{d}\omega\,. \tag{6.22}$$

Transformée et intégrale Dans la transformée de Fourier, la fréquence $\omega_0 = 0$ représente l'intégrale de la fonction. On souhaite un comportement similaire avec la décomposition d'un espace de fonctions. Il faut d'abord remarquer que l'intégrale d'une fonction peut s'écrire comme

$$\int_{\mathbb{D}_\mathcal{E}} f(x)\mathrm{d}x = \int_{\mathbb{D}_\mathcal{E}} \sum_{\omega\in\mathbb{F}} \mathcal{F}_f(\omega) b^\omega(x)\mathrm{d}x \tag{6.23}$$

$$= \sum_{\omega\in\mathbb{F}} \mathcal{F}_f(\omega) \int_{\mathbb{D}_\mathcal{E}} b^\omega(x)\mathrm{d}x\,. \tag{6.24}$$

Une façon simple si l'on veut qu'une fréquence ω_0 représente l'intégrale d'une fonction est de choisir une base telle que l'intégrale de b^{ω_0} soit non-nulle et que l'intégrale de tout les autres vecteurs $b^\omega, \omega \neq \omega_0$ soit égales à zéro. Dans le lemme suivant, nous montrons que si un des vecteurs de la base est une fonction constante, alors l'intégrale de tous les autres vecteurs est nulle. Cela permet de s'assurer qu'un seul coefficient suffit pour déterminer l'intégrale d'une fonction.

Lemme 6.1 *Soit $\{b^\omega : \omega \in \mathbb{F}\}$ une base orthogonale d'un espace de fonctions pré-hilbertien et $\omega_0 \in \mathbb{F}$ une fréquence telle que $b^{\omega_0}(x) = \alpha$, $\alpha \in \mathbb{C} \setminus \{0\}$, l'intégrale de $b^\omega(x), \omega \neq \omega_0$ est égale à zéro.*

Preuve 6.1 *Supposons qu'il existe $\omega_0 \in \mathbb{F}$ tel que*

$$b^{\omega_0}(x) = \alpha\,, \tag{6.25}$$

alors pour tout $\omega \in \mathbb{F}$ tel que $\omega \neq \omega_0$, on a

$$\begin{aligned}\langle b^{\omega_0}, b^\omega\rangle &= \int_{\mathbb{D}_\mathcal{E}} b^{\omega_0}(x)\overline{b^\omega(x)}\mathrm{d}x \\ &= \alpha \int_{\mathbb{D}_\mathcal{E}} \overline{b^\omega(x)}\mathrm{d}x\,. \end{aligned}\tag{6.26}$$

Or, comme b^{ω_0} et b^ω appartiennent à $\mathcal{B}$, leur produit scalaire est égal à zéro et donc l'intégrale de b^ω est égale à zéro. □

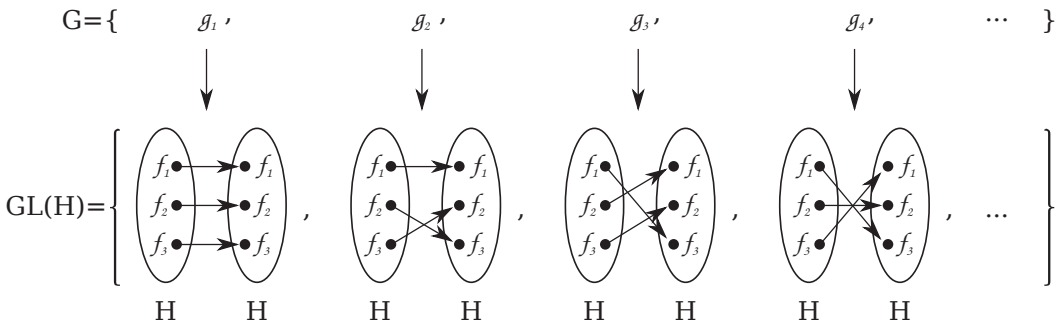


FIGURE 6.2 – *Illustration d'une representation de G vers $GL(H)$. G est un groupe composé des éléments $g_1, g_2, g_3, g_4, ...$ H est un espace vectoriel composé des éléments f_1, f_2 et f_3. $GL(H)$ est l'ensemble de toutes les bijections de H vers H. Il est important de noter que toutes les bijections de H ne sont pas forcément reliées à un élément du groupe G.*

Une telle fréquence ω_0 existe dans toutes les bases orthogonales des espaces d'échantillonnage présentés dans le chapitre 3. Par convention, nous assumons que $\alpha = 1$ et donc que

$$\int_{\mathbb{D}_{\mathcal{E}}} f(x)\mathrm{d}x = \mathcal{F}_f(\omega_0)\,. \tag{6.27}$$

Nous venons de définir une généralisation de la transformée de Fourier. Il reste cependant un point important que nous n'avons pas abordé : le spectre de puissance. Dans le début de ce chapitre, nous avons défini le spectre de puissance comme le carré de l'amplitude de la transformée de Fourier. Cette définition pourrait toujours s'appliquer à notre généralisation de la transformée de Fourier. Mais la propriété d'invariance par translation, ou plutôt par action dans le cadre général, serait alors perdue. Nous avons donc choisi de définir différemment le spectre de puissance de façon à conserver cette propriété. Notre définition du spectre de puissance fait intervenir la théorie des représentations, que nous allons maintenant introduire.

6.2.4 Théorie des représentations

Définition 6.9 *[Barut and Rączka, 1986] Soit $(G, \star)$ un groupe et H un espace vectoriel, on dit que la fonction $\phi : G \to GL(H)$ est une* ***représentation*** *de G si*

$$\phi(g \star g') = \phi(g) \circ \phi(g')\,, \tag{6.28}$$

où $g, g' \in G$ et $\circ$ est l'opération de composition de fonction.

L'ensemble $GL(H)$ est **le groupe général linéaire** de l'espace vectoriel H, c'est à dire l'ensemble des bijections de H vers H muni de l'opération de composition de fonction. Une représentation permet donc, à partir d'un élément du groupe G, de définir une bijection dans l'espace vectoriel. La figure (6.2) illustre les relations entre G, H et $GL(H)$. Les espaces pré-hilbertiens étant des espaces vectoriels, ils sont tout à fait compatibles avec la théorie des représentations. Dans le cas qui nous intéresse, il faut voir le groupe G comme l'ensemble des actions $\mathbb{T}_{\mathcal{E}}$ et l'espace vectoriel comme un espace de fonctions. La représentation ϕ transforme donc une fonction f en une autre fonction du même espace de fonctions. On utilisera la représentation suivante

$$[\phi(\tau)(f)]\,(x) = f(x * \tau^{-1})\,, \tag{6.29}$$

où $\tau \in \mathbb{T}_{\mathcal{E}}$, f une fonction de l'espace de fonctions et $x \in \mathbb{D}_{\mathcal{E}}$. Par soucis de concision, et par cohérence avec les notations précédemment introduites, nous écrirons simplement $\tau(f)$ à la place de $\phi(\tau)(f)$, tout en continuant à nommer la représentation ϕ. Cette transformation est compatible avec la théorie des représentations puisque

$$\begin{aligned}\left[(\tau \star \tau')(f)\right](x) &= f(x * (\tau \star \tau')^{-1}) \\ &= f(x * (\tau'^{-1} \star \tau^{-1})) \\ &= f((x * \tau'^{-1}) * \tau^{-1}) \\ &= \left[\tau(\tau'(f))\right](x)\,, \end{aligned} \tag{6.30}$$

où le passage de la ligne 1 à 2 se fait à l'aide des propriétés des groupes et le passage de la ligne 2 à 3 se fait à l'aide des propriétés des actions de groupes, présentées dans la section 3.2. On remarque que la représentation $\phi(\tau)$ est unitaire par rapport à l'espace de fonctions puisque les actions de $\mathbb{T}_{\mathcal{E}}$ préservent les distances. La représentation ϕ est aussi une application linéaire, puisque l'on a

$$\begin{aligned}\tau(\alpha f + f')(x) &= \alpha f(x * \tau^{-1}) + f'(x * \tau^{-1}) \\ &= \alpha\tau(f)(x) + \tau(f')(x)\,, \end{aligned} \tag{6.31}$$

où $f, f' \in F(\mathbb{D}_{\mathcal{E}}, \mathbb{C})$ et $\alpha \in \mathbb{C}$.

Définition 6.10 *[Barut and Rączka, 1986] Soit G un groupe, H un espace de Hilbert, $\phi : G \to GL(H)$ une représentation de G et $M \subset H$ un sous-espace de H, on dit que M est* ***stable*** *par rapport à ϕ si pour tout $g \in G, f \in M$, $\phi(g)(f) \in H$.*

Toute représentation admet au minimum deux sous-espaces stables : l'ensemble H et l'ensemble vide. On dit qu'une représentation est **irréductible** si elle ne possède que ces deux sous-espaces stables. Soit M un sous-espace stable de H, si pour tout $\tau \in \mathbb{T}_{\mathcal{E}}$, $\phi(\tau)$ est unitaire par rapport à H, alors $M^{\perp}$ est lui aussi stable [Barut and Rączka, 1986]. L'espace H peut alors s'écrire comme la somme directe de deux sous-espaces stables : $H = M \oplus M^{\perp}$. Une façon de décomposer l'espace H lorsque la représentation est unitaire, consiste à trouver un sous-espace stable irréductible H^{r_1} de H, puis de prendre le sous-espace orthogonal $H^{r_1\perp}$ et de chercher à l'intérieur un nouveau sous-espace stable irréductible H^{r_2}. Le processus se répète jusqu'à ce que le sous-espace orthogonal du dernier sous-espace stable irréductible trouvé soit lui aussi un sous-espace stable irréductible. On écrit alors

$$H = \bigoplus_{r \in \mathbb{P}} H^r\,, \tag{6.32}$$

où $\mathbb{P}$ est un ensemble indexant les sous-espaces stables irréductibles trouvés. Si l'espace possède une base orthogonale $\mathcal{B}$, alors on peut construire un ensemble de sous-espaces stables et irréductibles tel que la base orthogonale de chacun de ces sous-espaces est un sous-ensemble de $\mathcal{B}$. Puisque chaque sous-espace est orthogonal aux autres, un vecteur de la base ne peut appartenir qu'à un et un seul sous-espace et le nombre de sous-espaces ne peut pas être supérieur au nombre de vecteurs de la base. Par contre, le nombre de sous-espaces peut être inférieur au nombre de vecteurs de la base si l'action est capable de transformer un vecteur de la base en un autre. On note $\mathcal{B}^r$ la base orthogonale du sous-espace H^r, définie comme l'ensemble $\{b^{\omega} : \omega \in \mathbb{F}^r\}$, où $\mathbb{F}^r \subset \mathbb{F}$ est l'ensemble des indices de la base de H^r. On note $dim(H^r)$ le nombre de vecteurs de la base orthogonale $\mathcal{B}^r$.

Lemme 6.2 *Soit $\phi : G \to GL(H)$ une représentation telle que $\phi(g)$ soit une application linéaire et unitaire, et H^r un sous-espace stable et irréductible de H ayant une base orthogonale $\mathcal{B}^r := \{b^\omega : \omega \in \mathbb{F}^r\}$. On a, pour tout $g \in G, f \in H$, l'égalité suivante*

$$\sum_{\omega \in \mathbb{F}^r} \mathcal{F}_{\phi(g)(f)}(\omega)\overline{\mathcal{F}_{\phi(g)(f)}(\omega)} = \sum_{\omega \in \mathbb{F}^r} \mathcal{F}_f(\omega)\overline{\mathcal{F}_f(\omega)}\,. \tag{6.33}$$

Preuve 6.2 *On considère la fonction f^r telle que*

$$f^r = \sum_{\omega \in \mathbb{F}^r} \mathcal{F}_{\phi(g)(f)}(\omega) b^\omega\,, \tag{6.34}$$

représentant la fonction $\phi(g)(f)$ projetée dans le sous-espace H^r. Comme cette fonction est une projection de f, on a pour tout $\omega \in \mathbb{F}^r$ l'égalité $\mathcal{F}_{f^r}(\omega) = \mathcal{F}_{\phi(g)(f)}(\omega)$ et pour tout $\omega \notin \mathbb{F}^r$, on a $\mathcal{F}_{f^r}(\omega) = 0$. D'après le théorème de Parseval, on peut écrire que

$$\sum_{\omega \in \mathbb{F}^r} |\mathcal{F}_{f^r}(\omega)|^2 = \|f^r\|^2\,, \tag{6.35}$$

ce qui donne, en fonction de $\phi(g)(f)$, la relation suivante

$$\sum_{\omega \in \mathbb{F}^r} |\mathcal{F}_{\phi(g)(f)}(\omega)|^2 = \|\sum_{\omega \in \mathbb{F}^r} \mathcal{F}_{\phi(g)(f)}(\omega) b^\omega\|^2\,. \tag{6.36}$$

De façon équivalente, on peut appliquer $\phi(g)$ sur la base, plutôt que sur f. On obtient alors

$$\sum_{\omega \in \mathbb{F}^r} |\mathcal{F}_{\phi(g)(f)}(\omega)|^2 = \|\sum_{\omega \in \mathbb{F}^r} \mathcal{F}_f(\omega)\phi(g^{-1})(b^\omega)\|^2\,. \tag{6.37}$$

Comme $\phi(g^{-1})$ est une application linéaire, on peut écrire que

$$\sum_{\omega \in \mathbb{F}^r} |\mathcal{F}_{\phi(g)(f)}(\omega)|^2 = \|\phi(g^{-1})\left(\sum_{\omega \in \mathbb{F}^r} \mathcal{F}_f(\omega) b^\omega\right)\|^2\,. \tag{6.38}$$

Finalement, comme $\phi(g^{-1})$ est unitaire, on peut le supprimer et obtenir

$$\sum_{\omega \in \mathbb{F}^r} |\mathcal{F}_{\phi(g)(f)}(\omega)|^2 = \|\sum_{\omega \in \mathbb{F}^r} \mathcal{F}_f(\omega) b^\omega\|^2\,, \tag{6.39}$$

ce qui, en appliquant à nouveau le théorème de Parseval sur la deuxième partie, permet de retrouver Eq. (6.33). □

Donc, l'énergie portée par les coefficients de Fourier d'un sous-espace stable irréductible reste la même après l'application d'une action. Ce constat sera la base de notre définition du spectre de puissance. Le spectre de puissance a pour but de quantifier l'énergie contenue dans chaque sous-espace stable irréductible de l'espace de fonctions. Une bonne façon de le faire est de regarder la moyenne des coefficients de Fourier dans le sous-espace H^r. Ainsi, la définition du spectre de puissance généralisé devient

$$\begin{aligned}\mathcal{P}_f(r) &= \frac{1}{dim(H^r)} \sum_{\omega \in \mathbb{F}^r} \mathcal{F}_f(\omega)\overline{\mathcal{F}_f(\omega)} \\ &= \frac{1}{dim(H^r)} \sum_{\omega \in \mathbb{F}^r} ||\mathcal{F}_f(\omega)||^2\,.\end{aligned} \tag{6.40}$$

Une fois encore, il est possible d'utiliser une intégrale à la place de la somme. On note μ_F la mesure de comptage associée à $\mathbb{F}^r$. Puisque $dim(H^r)$ correspond au de nombre de fréquences dans $\mathbb{F}^r$, on peut écrire que $dim(H^r) = \mu_F(\mathbb{F}^r)$. On obtient alors la définition du spectre de puissance suivante,

$$\mathcal{P}_f(r) = \frac{1}{\mu_F(\mathbb{F}^r)} \int_{\mathbb{F}^r} ||\mathcal{F}_f(\omega)||^2 \mathrm{d}\omega\,, \tag{6.41}$$

où la somme a été remplacée par une intégrale par application de Lemme (2.1, p. 14). Ici, la principale différence avec la première définition du spectre de puissance donnée par Def. (6.2) est que le domaine de définition n'est plus $\mathbb{F} \to \mathbb{R}$ mais $\mathbb{P} \to \mathbb{R}$. Cette subtilité vient du fait que le nombre de sous-espaces stables et irréductibles peut-être inférieur au nombre de vecteurs de la base $\mathcal{B}$. Dans l'espace euclidien, une translation ne permet pas de changer la fréquence ni l'orientation d'une fonction cosinus. Il existe donc autant de sous-espaces stable et irréductibles que de vecteurs dans la base et donc, $dim(H^r) = 1$. Si l'on considère maintenant le groupe des rotations à la place du groupe des actions dans l'espace euclidien, alors il devient possible de changer l'orientation d'une fonction cosinus par l'opération τ et donc de passer d'un vecteur de la base à un autre. Dans ce cas-là, $dim(H^r)$ est égale au nombre de vecteurs de la base qui ont la même fréquence mais pas forcement la même orientation. Le domaine $\mathbb{P}$ devient mono-dimensionnel et le spectre de puissance ne contient ni l'information de phase, ni l'information d'orientation. Le spectre de puissance est donc, par la théorie des représentations, directement lié au groupe des actions $\mathbb{T}_{\mathcal{E}}$, là où la transformée de Fourier ne dépend que du domaine d'échantillonnage $\mathbb{D}_{\mathcal{E}}$.

Lemme 6.3 *[Kazhdan* et al.*, 2015] Soit G un groupe et $\phi : G \to GL(H^r)$ une représentation telle que $\phi(g)$ soit une application linéaire, unitaire et irréductible, si $f_1, f_2, f_3, f_4 \in H^r$, alors*

$$\frac{1}{\mu_H(G)} \int_G \langle \phi(g)(f_1), f_2 \rangle \langle \phi(g)(f_3), f_4 \rangle \,\mathrm{d}g = \frac{1}{dim(H^r)} \langle f_1, f_3 \rangle \langle f_2, f_4 \rangle\,. \tag{6.42}$$

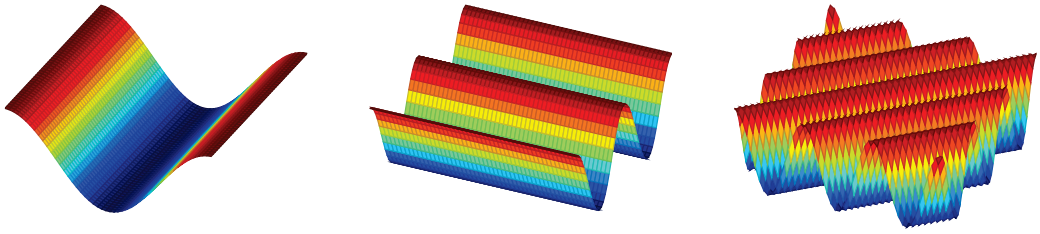
Lemme 6.4 *[Kazhdan* et al.*, 2015] Soit G un groupe, $\phi : G \to GL(H^r)$ et $\phi' : G \to GL(H^{r'})$ deux représentations linéaires, unitaires et irréductibles, si $f_1, f_2 \in H^r, f_3, f_4 \in H^{r'}$ et que ϕ et ϕ' sont isomorphes, alors*

$$\int_G \langle \phi(g)(f_1), f_2 \rangle \langle \phi'(g)(f_3), f_4 \rangle \,\mathrm{d}g = 0\,. \tag{6.43}$$

Ces lemmes nous permettrons de montrer, en prenant $f_1 = f_3 = S$, que dans le cas des ensembles de points homogènes, la moyenne de la transformée de Fourier est égale au spectre de puissance. Mais avant cela, nous allons parler plus concrètement de ce que nous entendons par transformée de Fourier de l'espace d'échantillonnage, et de la façon dont les ensembles de points peuvent être analysés à travers ces outils spectraux.

6.3 Transformée de Fourier de l'espace d'échantillonnage

Dans les chapitres suivants, nous utiliserons à de nombreuses reprises la transformée de Fourier. Pour pouvoir l'utiliser indépendamment de l'espace d'échantillonnage, nous nous restreindrons à n'utiliser que les propriétés qu'ont en commun les



$b^{(1,0)}$ $b^{(0,2)}$ $b^{(3,3)}$

FIGURE 6.3 – *Exemple de fonctions de base dans l'espace $\mathcal{T}^2$ pour différentes fréquences. Les fonctions sont représentées comme des déformations de la surface d'un carré. Les valeurs fortes (en rouge) sont élevées et les valeurs faibles (en bleu) sont abaissées. On remarque que plus la fréquence augmente, plus le nombre d'oscillations augmente.*

transformées associées à chaque espace d'échantillonnage. Par transformée de Fourier, nous entendrons donc une opération de décomposition d'un espace de fonctions $\mathcal{L}_{\mathcal{E}}$ par une base $\mathcal{B} = \{b^{\omega} : \omega \in \mathbb{F}\}$, où b^{ω_0} est une fonction constante. Nous assumons que les vecteurs de la base sont normalisés de façon à ce que leur norme soit égale à 1, à l'exception du vecteur b^{ω_0}. Cette décomposition doit permettre une transformée inverse et doit être compatible avec le théorème de Parseval. Nous assumerons également qu'il existe une partition de $\mathbb{F}$ en sous-ensembles $\mathbb{F}^r, r \in \mathbb{P}$ définissant le spectre de puissance.

6.3.1 Espace euclidien $\mathcal{E}^d$

Comme nous l'avons déjà vu, la transformée de Fourier classique dans l'espace euclidien respect ces propriétés. Pour cela, il faut utiliser l'espace de fonction $\mathcal{L}_{\mathcal{E}^d} = L^2(\mathbb{D}_{\mathcal{E}}, \mathbb{C})$ et la base $\mathcal{B} = \{b^{\omega} : \omega \in \mathbb{F}\}$ telle que $\mathbb{F} = \mathbb{R}^d$ et $b^{\omega}(x) = e^{i2\pi x \cdot \omega}$. La partition de $\mathbb{F}$ se fait de façon triviale en considérant que chaque $\mathbb{F}^r$ ne contient qu'un seul élément.

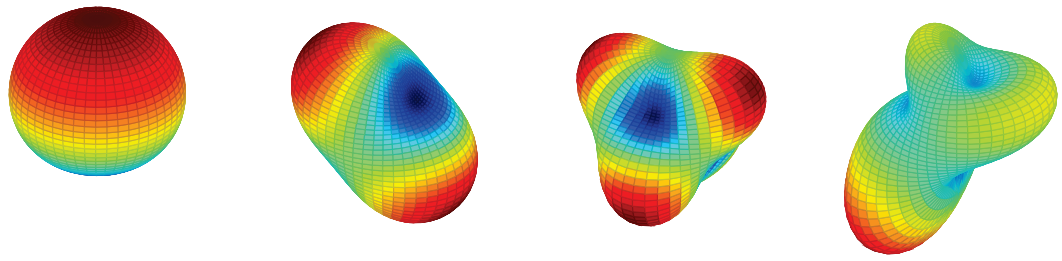
Dans cet espace, puisque $\mathbb{F}$ est lui aussi un espace euclidien, on peut définir une version radiale du spectre de puissance, notée $\breve{\mathcal{P}}_f$, et définie par

$$\breve{\mathcal{P}}_f(\nu) = \frac{1}{\mu\left(\mathbb{S}^{d-1}\right)} \int_{\mathbb{S}^{d-1}} \mathcal{P}_f(\nu \cdot n) \mathrm{d}n \,, \tag{6.44}$$

où $\mathbb{S}^{d-1}$ est une hyper-sphère formée par l'ensemble des vecteurs n de $\mathbb{R}^d$ dont la norme est égale à 1, et μ une mesure au sens de Lebesgue. Dans cette définition, n et ν représentent des coordonnées hyper-sphériques telles que n est une orientation et ν une distance. Cette définition correspond à la moyenne radiale du spectre de puissance. Le spectre de puissance radial permet donc de ne pas tenir compte de l'orientation de la fonction.

6.3.2 Espace hyper-cubique periodique $\mathcal{T}^d$

La transformée de Fourier de l'espace hyper-cubique periodique est étroitement liée à celle de l'espace euclidien, puisqu'elle correspond à la transformée de Fourier d'un signal periodique dans $\mathbb{R}^d$. Il est connu que la transformée de Fourier dans $\mathcal{T}^d$ est



$b^{(1,0)}$ $b^{(2,1)}$ $b^{(3,2)}$ $b^{(4,1)}$

FIGURE 6.4 – *Exemple d'harmoniques sphériques pour différentes fréquences. Les fonctions sont représentées comme des déformations de la surface d'une sphère. Les valeurs fortes (en rouge) sont éloignées du centre de la sphère et les valeurs faibles (en bleu) sont rapprochées du centre. On remarque que plus la fréquence l augmente, plus le nombre de «lobes » augmente.*

une transformée de Fourier restreinte aux fréquences entières $\mathbb{F} = \mathbb{Z}^d$. L'expression des vecteurs de la base $\mathcal{B}$ est quant à elle la même que celle de l'espace euclidien $b^{\omega}(x) = e^{i2\pi x\cdot\omega}$. Pour la fréquence $\omega_0 = 0$, on a bien $b^{\omega_0}(x) = 1$. Cette base étant un espace de Hilbert, l'ensemble des propriétés nécessaires est donc vérifié. La norme des vecteurs de la base est égale à 1. La figure (6.3) montre plusieurs exemples de fonctions appartenant à la base orthonormale de l'espace $\mathcal{T}^2$.

6.3.3 Espace sphérique $\mathcal{S}^2$

Les harmoniques sphériques forment une base orthogonale des fonctions $L^2(\mathbb{D}_{\mathcal{S}^2}, \mathbb{C})$ [Groemer, 1996; Wieczorek and Simons, 2005]. Soit $\mathbb{F} = \{(l,m) : l \in \mathbb{N}, m \in \mathbb{Z}, -l \leq m \leq l\}$ le domaine spectral de l'espace sphérique, les harmoniques sphériques sont définies par

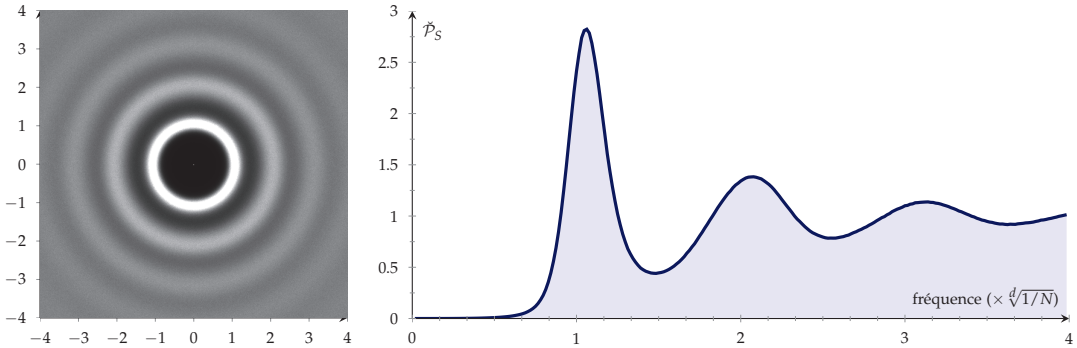
$$b^{(l,m)}(\phi,\theta) = (-1)^m \sqrt{\frac{(2l+1)}{4\pi}\frac{(l-m)!}{(l+m)!}} P_l^m(cos(\phi))e^{im\theta}\,, \tag{6.45}$$

où P_l^m sont les polynômes associés de Legendre. Cette définition est normalisée de telle sorte que la norme de chacun des vecteurs de la base est égale à 1. La figure (6.4) montre des exemples d'harmoniques sphériques pour différents valeurs de l et m.

6.4 Transformée de Fourier et ensemble de points

La transformée de Fourier peut être utilisée pour étudier un ensemble de points dans $\mathbb{D}_{\mathcal{E}}$ en utilisant la fonction représentative de $\mathcal{S}$. Son expression peut être donnée directement par

$$\begin{aligned}\mathcal{F}_S(\omega) &= \int_{\mathbb{D}_{\mathcal{E}}} S(x)\overline{b^{\omega}(x)}\mathrm{d}x \\ &= \sum_{k=1}^{N}\int_{\mathbb{D}_{\mathcal{E}}} \delta(x,\mathbf{s}_k)\overline{b^{\omega}(x)}\mathrm{d}x \\ &= \sum_{k=1}^{N}\overline{b^{\omega}(\mathbf{s}_k)}\,. \end{aligned} \tag{6.46}$$



(a) (b)

FIGURE 6.5 – *Exemple de spectre de puissance d'un ensemble de points optimisé par BNOT dans $\mathcal{T}^2$.* ***(a) :*** *Le spectre de puissance $\mathcal{P}_S$.* ***(b) :*** *La moyenne radiale $\breve{\mathcal{P}}_S$ du spectre de puissance $\mathcal{P}_S$. L'axe des fréquences est normalisé par $1/\sqrt{N}$.*

À ce stade, la question de savoir si la fonction delta de Dirac fait partie de l'ensemble de fonctions $\mathcal{L}_\mathcal{E}$ est légitime. Il est difficile de répondre à cette question car la fonction delta de Dirac n'est pas complètement définie. On ne sait pas, par exemple, quelle est sa norme L^2 car on ne sait pas comment calculer le carré d'une fonction delta de Dirac. Nous remarquerons simplement que, comme nous venons de le montrer, la fonction delta de Dirac est décomposable à l'aide de la base $\mathcal{B}$, et qu'elle est compatible avec le théorème de Parseval sous certaines conditions. Une condition suffisante est que f soit bornée, c'est à dire que pour tout x, on a $|f(x)| < \infty$. Dans ce cas on a, d'après les propriétés de la fonction delta de Dirac, l'égalité

$$\begin{aligned}\int_{\mathbb{D}_\mathcal{E}} f(x)\overline{\delta(x_0, x)}\mathrm{d}x &= f(x_0) \\ &= \sum_{k=1}^{N} \mathcal{F}_f(\omega)\overline{b^\omega(x_0)} \\ &= \sum_{k=1}^{N} \mathcal{F}_f(\omega)\overline{\mathcal{F}_\delta(\omega)}\,, \end{aligned} \tag{6.47}$$

qui valide donc le théorème de Parseval.

Puisque $\mathcal{S}$ est une variable aléatoire, sa transformée de Fourier l'est aussi et chaque coefficient $\mathcal{F}_S(\omega)$ est une variable aléatoire à valeur dans $\mathbb{C}$. On peut donc utiliser la densité de probabilité de $\mathcal{F}_S(\omega)$, notée $PDF_{\mathcal{F}_S(\omega)}$.

La figure (6.5) montre un exemple de spectre de puissance d'un ensemble de points. Dans cette figure, et celles du même type que nous montrerons dans le chapitre 8, chaque pixel correspond à un vecteur de la base. Comme le nombre de vecteur de la base $\mathcal{B}$ est infini, cette image ne montre qu'un sous-ensemble des coefficients de Fourier. Au centre, on retrouve la fréquence ω_0, aussi appelée DC. Les fréquences proches du centre sont qualifiées de «basses fréquences » et les fréquences en périphérie du centre sont qualifiées de hautes fréquences.

6.4.1 Ensembles de points homogènes

Lemme 6.5 *Soit $\mathcal{S}$ un ensemble de points homogène dans l'espace $\mathcal{E}$, on a*

$$\langle \mathcal{F}_S(\omega) \rangle = \begin{cases} N & si\ \omega = \omega_0 \\ 0 & sinon\,. \end{cases} \tag{6.48}$$

Preuve 6.3 *Comme l'ensemble de points est homogène, on peut utiliser Eq. (3.25, p. 27) et écrire que*

$$\langle \mathcal{F}_S(\omega) \rangle = \left\langle \frac{1}{\mu_H(\mathbb{T}_\mathcal{E})} \int_{\mathbb{T}_\mathcal{E}} \mathcal{F}_{\tau(S)}(\omega) \mathrm{d}\tau \right\rangle . \tag{6.49}$$

Oublions la moyenne et regardons l'intégrale. D'après la définition d'un coefficient de Fourier, donnée par Eq. (6.17), et sachant que la représentation est unitaire, on peut écrire que

$$\frac{1}{\mu_H(\mathbb{T}_\mathcal{E})} \int_{\mathbb{T}_\mathcal{E}} \mathcal{F}_{\tau(S)}(\omega) \mathrm{d}\tau = \frac{1}{\mu_H(\mathbb{T}_\mathcal{E})} \int_{\mathbb{T}_\mathcal{E}} \left\langle S, \tau^{-1}(b^\omega) \right\rangle \mathrm{d}\tau\,. \tag{6.50}$$

Comme nous avons défini dans Eq. (6.12) le produit scalaire avec une intégrale, on a

$$\begin{aligned} \frac{1}{\mu_H(\mathbb{T}_\mathcal{E})} \int_{\mathbb{T}_\mathcal{E}} \mathcal{F}_{\tau(S)}(\omega) \mathrm{d}\tau &= \frac{1}{\mu_H(\mathbb{T}_\mathcal{E})} \int_{\mathbb{T}_\mathcal{E}} \int_{\mathbb{D}_\mathcal{E}} S(x) \overline{b^\omega(x * \tau)} \mathrm{d}x \mathrm{d}\tau \\ &= \frac{1}{\mu_H(\mathbb{T}_\mathcal{E})} \int_{\mathbb{D}_\mathcal{E}} S(x) \int_{\mathbb{T}_\mathcal{E}} \overline{b^\omega(x * \tau)} \mathrm{d}\tau \mathrm{d}x\,. \end{aligned} \tag{6.51}$$

En appliquant Lemme (3.1, p. 26), on obtient

$$\frac{1}{\mu_H(\mathbb{T}_\mathcal{E})} \int_{\mathbb{T}_\mathcal{E}} \mathcal{F}_{\tau(S)}(\omega) \mathrm{d}\tau = \frac{1}{\mu_H(\mathbb{T}_\mathcal{E})} \int_{\mathbb{D}_\mathcal{E}} S(x) \int_{\mathbb{D}_\mathcal{E}} \overline{b^\omega(x')} \mathrm{d}x' \mathrm{d}x\,. \tag{6.52}$$

La preuve est terminée en sachant que l'intégrale de b^ω est égale à zéro quand $\omega \neq \omega_0$ et 1 quand $\omega = \omega_0$, et que l'intégrale de la fonction représentative de $\mathcal{S}$ est égale à N. La variable aléatoire $\mathcal{S}$ étant supprimée de l'équation, la moyenne peut être supprimée. □

Ce lemme montre que l'espérance de la transformée de Fourier d'un ensemble de points homogène est égale à $N\delta(\omega, \omega_0)$. Le fait que le spectre de puissance soit invariant par action en fait un candidat idéal pour l'étude des ensembles de points homogènes. En supprimant l'information propre au groupe des actions, le spectre de puissance met ainsi en lumière les caractéristiques déterminantes des ensembles de points homogènes. Le lemme suivant illustre le fait que dans le cas des ensembles de points homogènes, étudier l'espérance de la transformée de Fourier revient à étudier le spectre de puissance.

Lemme 6.6 *Soit $\mathcal{S}$ un ensemble de points homogène dans l'espace $\mathcal{E}$, on a*

$$Cov(\mathcal{F}_S(\omega), \overline{\mathcal{F}_S(\omega')}) = \begin{cases} \mathcal{P}_S(r) & si\ \omega = \omega'\ et\ \omega \neq \omega_0 \\ 0 & sinon\,, \end{cases} \tag{6.53}$$

avec r choisi tel que H^r soit le sous-espace stable irréductible contenant b^ω, c'est à dire tel que $\omega \in \mathbb{F}^r$.

Preuve 6.4 *En utilisant Lemme (6.5) et la définition de la covariance donnée par Eq. (3.6, p. 18), on obtient*

$$Cov(\mathcal{F}_S(\omega), \overline{\mathcal{F}_S(\omega')}) = \left\langle \mathcal{F}_S(\omega)\overline{\mathcal{F}_S(\omega')} \right\rangle - N\delta(\omega, \omega_0)\,. \tag{6.54}$$

Comme l'ensemble de points est homogène, on peut utiliser Eq. (3.25, p. 27) sur l'espérance et écrire que

$$\left\langle \mathcal{F}_S(\omega)\overline{\mathcal{F}_S(\omega')} \right\rangle = \left\langle \frac{1}{\mu_H(\mathbb{T}_\mathcal{E})} \int_{\mathbb{T}_\mathcal{E}} \mathcal{F}_{\tau(S)}(\omega)\overline{\mathcal{F}_{\tau(S)}(\omega')}\mathrm{d}\tau \right\rangle\,. \tag{6.55}$$

D'après la définition d'un coefficient de Fourier, donnée par Eq. (6.17), on peut réécrire cette équation sous la forme

$$\left\langle \mathcal{F}_S(\omega)\overline{\mathcal{F}_S(\omega')} \right\rangle = \left\langle \frac{1}{\mu_H(\mathbb{T}_\mathcal{E})} \int_{\mathbb{T}_\mathcal{E}} \langle \tau(S), b^\omega \rangle \overline{\langle \tau(S), b^{\omega'} \rangle}\mathrm{d}\tau \right\rangle\,. \tag{6.56}$$

On pose $r, r' \in \mathbb{P}$ tel que $\omega \in \mathbb{F}^r$ et $\omega' \in \mathbb{F}^{r'}$. Dans les deux produits scalaires, seules les projections de S dans les sous-espaces H^r et $H^{r'}$ contribuent au résultat. On écrit alors

$$\left\langle \mathcal{F}_S(\omega)\overline{\mathcal{F}_S(\omega')} \right\rangle = \left\langle \frac{1}{\mu_H(\mathbb{T}_\mathcal{E})} \int_{\mathbb{T}_\mathcal{E}} \langle \tau(S^r), b^\omega \rangle \overline{\langle \tau(S^{r'}), b^{\omega'} \rangle}\mathrm{d}\tau \right\rangle\,, \tag{6.57}$$

où S^r est la fonction S projetée dans le sous-espace H^r et $S^{r'}$ est la fonction S projetée dans le sous-espace $H^{r'}$. Lorsque $r \neq r'$, on peut appliquer Lemme (6.4) et trouver que $\left\langle \mathcal{F}_S(\omega)\overline{\mathcal{F}_S(\omega')} \right\rangle = 0$. Lorsque $r = r'$, on peut appliquer Lemme (6.3) et trouver que

$$\left\langle \mathcal{F}_S(\omega)\overline{\mathcal{F}_S(\omega')} \right\rangle = \frac{1}{dim(H^r)} \langle S^r, S^r \rangle \overline{\langle b^\omega, b^{\omega'} \rangle}\,. \tag{6.58}$$

Si $\omega \neq \omega'$, le produit scalaire $\left\langle b^\omega, b^{\omega'} \right\rangle$ est nul. À l'inverse, quand $\omega = \omega'$, ce produit scalaire vaut $\|b^\omega\|^2 = 1$ et on peut écrire que

$$\left\langle \mathcal{F}_S(\omega)\overline{\mathcal{F}_S(\omega)} \right\rangle = \frac{1}{dim(H^r)} \langle S^r, S^r \rangle\,. \tag{6.59}$$

En exprimant le produit scalaire $\langle S^r, S^r \rangle$ sous forme d'intégrale et en utilisant le théorème de Parseval, on obtient

$$\begin{aligned} \left\langle \mathcal{F}_S(\omega)\overline{\mathcal{F}_S(\omega)} \right\rangle &= \frac{1}{dim(H^r)} \int_{\mathbb{D}_\mathcal{E}} S^r(x)S^r(x)\mathrm{d}x \\ &= \frac{1}{dim(H^r)} \int_{\mathbb{F}^r} \mathcal{F}_S(\omega')\overline{\mathcal{F}_S(\omega')}\mathrm{d}\omega'\,. \end{aligned} \tag{6.60}$$

En utilisant la définition du spectre de puissance, on obtient

$$\left\langle \mathcal{F}_S(\omega)\overline{\mathcal{F}_S(\omega)} \right\rangle = \mathcal{P}_S(r)\,. \tag{6.61}$$

La preuve est terminée en remarquant que pour la fréquence ω_0, les deux termes de la covariance s'annulent. □

Cette preuve servira de base, dans le chapitre 7, pour donner une expression de la variance de l'intégration Monte Carlo en fonction du spectre de puissance d'un ensemble de points homogène.

6.5 Lien avec l'analyse spatiale

Pour terminer ce chapitre, nous parlons ici du lien entre le domaine spatial et le domaine spectral. Comme nous l'avons vu, à une fonction f définie dans le domaine spatial, est associée une fonction $\mathcal{F}_f$ définie dans le domaine spectral. Cette dualité existe également entre la densité locale et l'espérance de la transformée de Fourier.

Lemme 6.7 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$, la densité locale est liée à la transformée de Fourier par la relation suivante*

$$\langle \mathcal{F}_S(\omega) \rangle = \mathcal{F}_{\rho_{\mathcal{S}}^{(1)}}(\omega)\,. \tag{6.62}$$

Preuve 6.5 *Il suffit de remarquer que $\rho_{\mathcal{S}}^{(1)}(x) = \langle S(x) \rangle$.* □

Le spectre de puissance et la fonction de distribution différentielle peuvent aussi être reliés. Pour les espaces d'échantillonnage utilisant la transformée de Fourier classique, comme l'espace euclidien ou l'espace hyper-cubique periodique, une expression simple existe pour passer d'un domaine à l'autre. Pour cela, on utilise le lien entre l'auto-corrélation et le spectre de puissance [Bracewell, 1999, chap. 6] :

$$\mathcal{P}_f(\omega) = \mathcal{F}\left[(f \oplus f)(x)\right](\omega)\,. \tag{6.63}$$

En utilisant Lemme (5.8, p. 53), on obtient alors le lien suivant

$$\begin{aligned} \langle \mathcal{P}_S(\omega) \rangle &= \mathcal{F}\left[\langle (S \oplus S)(x) \rangle\right](\omega) \\ &= \mathcal{F}\left[N\rho_{\mathcal{S}}\check{g}_{\mathcal{S}}(x) + N\delta(x,0)\right](\omega)\,. \end{aligned} \tag{6.64}$$

En normalisant le spectre de puissance par N, et en utilisant le fait que la transformée de Fourier d'une constante est un Dirac, on obtient finalement

$$\begin{aligned} \frac{\langle \mathcal{P}_S(\omega) \rangle}{N} - \rho_{\mathcal{S}}\delta(\tau,0) &= 1 + \rho_{\mathcal{S}}\mathcal{F}\left[\check{g}_{\mathcal{S}}(\tau) - 1\right](\omega) \\ &= 1 + \rho_{\mathcal{S}}\mathcal{F}_{\check{h}_{\mathcal{S}}}(\omega)\,. \end{aligned} \tag{6.65}$$

Cette relation est au cœur des méthodes d'optimisation de contrôle spectral [Zhou *et al.*, 2012; Heck *et al.*, 2013; Öztireli and Gross, 2012]. En effet, les algorithmes d'optimisation utilisent des propriétés spatiales pour déplacer les points, comme par exemple une descente de gradient d'une certaine énergie. Il est donc nécessaire de traduire le spectre de puissance cible dans le domaine spatial pour pouvoir optimiser l'ensemble de points.

Conclusion

Dans ce chapitre, nous avons introduit l'analyse spectrale des ensembles de points. Nous l'avons aussi généralisée dans le cadre théorique définie dans le chapitre 3, à l'aide des espaces de Hilbert et de la théorie des représentations. Ces outils seront appliqués sur les méthodes d'échantillonnage de l'état de l'art dans le chapitre 8. Dans le chapitre suivant, nous abordons les liens entre l'intégration Monte Carlo et l'analyse spatiale et spectrale. La partie spectrale y joue un rôle clé, puisqu'elle se révèle plus efficace pour caractériser l'erreur d'intégration. Nous verrons en particulier comment les basses fréquences de l'ensemble de points et les hautes fréquences de l'intégrande affectent la vitesse de convergence de l'intégration Monte Carlo.

Intégration Monte Carlo 7

Les chapitres précédents nous ont permis de définir des outils d'analyse pour les ensembles de points, que ce soit dans le domaine spatial ou spectral. Dans ce chapitre, nous décrivons comment ces outils peuvent être utilisés pour quantifier l'erreur dans l'intégration Monte Carlo. S'il est déjà connu que la discrépance permet de borner l'erreur d'intégration grâce à l'inégalité de Koksma-Hlawka, nous montrons qu'il est possible de donner une expression exacte de l'erreur d'intégration à l'aide, soit de la fonction de distribution de paires $g_{\mathcal{S}}$ définie par Def. (5.3, p. 48), soit de la transformée de Fourier $\mathcal{F}_S$ définie dans le chapitre 6. Lorsque l'ensemble de points est homogène, nous verrons que la fonction de distribution différentielle $\check{g}_{\mathcal{S}}$, définie par Def. (5.6, p. 50), et le spectre de puissance $\mathcal{P}_S$, défini dans le chapitre 6, sont suffisants pour donner une expression exacte de l'erreur d'intégration. Toujours dans le cas homogène, nous montrons comment la forme du spectre de puissance de l'ensemble de points dans les basses fréquences affecte la vitesse de convergence de l'intégration Monte Carlo.

7.1 Définition formelle

Nous rappelons ici que par intégration Monte Carlo, nous entendons tout type d'intégration par échantillonnage, qu'il soit stochastique ou non. Dans la littérature, on retrouve généralement le terme intégration Monte Carlo pour désigner l'approximation d'une intégrale à l'aide d'un ensemble de points bruit blanc. Le terme intégration Quasi Monte Carlo est quant à lui réservé pour les méthodes utilisant des processus déterministes comme les séquences à basse discrépance. Il nous a semblé naturel de reprendre le terme d'intégration Monte Carlo pour tous les ensembles de points stochastiques. Cependant, comme notre formalisme, à base de variable aléatoire, est compatible avec les ensembles de points déterministes, nous appliquons donc le terme intégration Monte Carlo quelque soit le type de l'ensemble de points utilisé.

L'intégration Monte Carlo permet d'approximer l'intégrale I_f d'une fonction f sur le domaine d'échantillonnage. Elle est définie généralement, comme par exemple dans [Niederreiter, 1992, chap. 2], par l'expression suivante,

$$\frac{\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)}{N} \sum_{k=1}^{N} f\left(\mathbf{s}_k\right). \tag{7.1}$$

Pour des raisons pratiques, nous utiliserons une définition légèrement différente. À la place d'évaluer l'intégrale de f, nous évaluerons l'intégrale du conjugué de f. Cette

astuce permet de simplifier les dérivations et les résultats présentés dans ce chapitre. Dans tous les cas, étant donné qu'en informatique graphique, on cherche généralement à intégrer des fonctions réelles, les résultats que nous présentons peuvent être utilisés sans tenir compte de la partie imaginaire. Nous avons également utilisé la densité $\rho_{\mathcal{S}}$ de l'ensemble de points, définie par Eq. (5.1, p. 45), car comme nous l'avons expliqué dans le chapitre 5, la densité porte plus d'information que le nombre de points. Notre expression de l'intégration Monte Carlo se résume donc par la définition suivante.

Définition 7.1 *Soit f une fonction $\mathbb{D}_{\mathcal{E}} \to \mathbb{C}$, on appelle* ***intégration Monte-Carlo****, notée $I_{\mathcal{S},f}$, l'approximation du conjugué de l'intégrale de f donnée par*

$$I_{\mathcal{S},f} := \frac{1}{\rho_{\mathcal{S}}} \sum_{k=1}^{N} \overline{f(\mathbf{s}_k)} \,. \tag{7.2}$$

L'intégration Monte Carlo, définie de cette façon, cherche à approximer l'intégrale suivante

$$I_f := \int_{\mathbb{D}_{\mathcal{E}}} \overline{f(x)} \mathrm{d}x \,. \tag{7.3}$$

En utilisant les propriétés de la fonction delta de Dirac, on peut obtenir une expression de l'intégration Monte Carlo en forme intégrale [Durand, 2011; Subr and Kautz, 2013]

$$\begin{aligned} I_{\mathcal{S},f} &= \frac{1}{\rho_{\mathcal{S}}} \sum_{k=1}^{N} \int \delta(\mathbf{s}_k, x) \overline{f(x)} \mathrm{d}x \\ &= \frac{1}{\rho_{\mathcal{S}}} \int_{\mathbb{D}_{\mathcal{E}}} S(x) \overline{f(x)} \mathrm{d}x \,. \end{aligned} \tag{7.4}$$

Il est important de se rappeler ici que l'ensemble de points $\mathcal{S}$ est une variable aléatoire. En conséquence, l'intégration Monte Carlo est différente chaque fois qu'elle est effectuée. On considère donc $I_{\mathcal{S},f}$ comme une variable aléatoire à valeurs dans $\mathbb{C}$.

7.2 Erreur d'intégration

L'intégration Monte-Carlo étant une estimation, on souhaite quantifier l'erreur d'intégration afin de mieux comprendre comment la réduire. Pour une seule réalisation de l'ensemble de points, l'erreur de l'intégration Monte Carlo est donnée par

$$Err(I_{\mathcal{S},f}) := |I_{\mathcal{S},f} - I_f| \,. \tag{7.5}$$

Pour les ensembles de points déterministes, observer et quantifier l'erreur d'intégration est donc suffisant puisque toutes les réalisations sont les mêmes. Pour les ensembles de points stochastiques, il faut prendre en compte l'ensemble des réalisations de $\mathcal{S}$, puisque le résultat de l'intégration est une variable aléatoire dépendante de $\mathcal{S}$. On utilise généralement pour cela l'erreur quadratique moyenne, ou *Mean Square Error* (MSE), définie par

$$MSE(I_{\mathcal{S},f}) := \left\langle Err(I_{\mathcal{S},f})^2 \right\rangle \,. \tag{7.6}$$

La MSE peut aussi s'écrire en fonction de la variance et du biais de l'estimateur

$$MSE(I_{\mathcal{S},f}) = Var(I_{\mathcal{S},f}) + Bias(I_{\mathcal{S},f})^2 \,, \tag{7.7}$$

où la variance est définie par

$$Var(I_{\mathcal{S},f}) = \left\langle I_{\mathcal{S},f}\overline{I_{\mathcal{S},f}}\right\rangle - \left\langle I_{\mathcal{S},f}\right\rangle \left\langle \overline{I_{\mathcal{S},f}}\right\rangle . \tag{7.8}$$

et le biais par

$$Bias(I_{\mathcal{S},f}) = \left\langle I_{\mathcal{S},f}\right\rangle - I_f . \tag{7.9}$$

Pour qu'une méthode d'échantillonnage dédiée à l'intégration Monte Carlo soit acceptable, elle doit au moins vérifier la propriété de convergence suivante [Niederreiter, 1992, chap. 2] :

$$\lim_{\rho_{\mathcal{S}}\to\infty} I_{\mathcal{S},f} = I_f . \tag{7.10}$$

Cette propriété implique notamment que l'erreur converge vers zéro quand le nombre de points, c'est à dire la densité, tend vers l'infini

$$\lim_{\rho_{\mathcal{S}}\to\infty} MSE(I_{\mathcal{S},f}) = 0 . \tag{7.11}$$

On dira qu'un ensemble de points $\mathcal{S}_1$ est meilleur qu'un autre ensemble de points $\mathcal{S}_2$ si l'erreur de l'intégration Monte Carlo converge plus vite en utilisant $\mathcal{S}_1$ que $\mathcal{S}_2$ pour une classe de fonctions donnée.

7.3 Ensemble de points homogène et biais

Nous venons de voir comment est définie l'erreur d'intégration dans le cas déterministe et dans le cas stochastique. Lorsque l'ensemble de points est stochastique, l'erreur se manifeste sous forme de biais et de variance. Regardons maintenant ce qui se passe lorsque l'ensemble de points est homogène.

Lemme 7.1 *Si $\mathcal{S}$ est un ensemble de points homogène, alors le biais de l'intégration Monte Carlo est nul.*

Preuve 7.1 *On veux montrer que $Bias(I_{\mathcal{S},f}) = 0$ quand $\mathcal{S} \in \mathfrak{M}_{Homo.}$. Pour cela, il faut exprimer le biais en fonction de la position des points de l'ensemble $\mathcal{S}$ à l'aide de Eq. (7.2) et Eq. (7.9)*

$$Bias(I_{\mathcal{S},f}) = \frac{1}{\rho_{\mathcal{S}}}\sum_{k=1}^{N}\left\langle \overline{f(\mathbf{s}_k)}\right\rangle - I_f . \tag{7.12}$$

On peut ensuite réécrire l'espérance sous une forme intégrale en utilisant la densité de probabilité de $\mathbf{s}_k$

$$Bias(I_{\mathcal{S},f}) = \frac{1}{\rho_{\mathcal{S}}}\sum_{k=1}^{N}\int_{\mathbb{D}_{\mathcal{E}}} \overline{f(x)}PDF_{\mathbf{s}_k}(x)\mathrm{d}x - I_f . \tag{7.13}$$

Or on sait d'après Lemme (3.2, p. 27) que la fonction de densité de probabilité est uniforme dans le cas des ensembles de points homogènes, ce qui donne donc

$$\begin{aligned} Bias(I_{\mathcal{S},f}) &= \frac{1}{\rho_{\mathcal{S}}}\sum_{k=1}^{N}\int_{\mathbb{D}_{\mathcal{E}}} \frac{1}{\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)}\overline{f(x)}\mathrm{d}x - I_f \\ &= \frac{1}{\rho_{\mathcal{S}}}\frac{N}{\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)}\int_{\mathbb{D}_{\mathcal{E}}} \overline{f(x)}\mathrm{d}x - I_f \\ &= 0 . \end{aligned} \tag{7.14}$$

Ici, le passage de la deuxième ligne à la troisième ligne s'est fait à l'aide de la définition de la densité Eq. (5.1, p. 45) et en remarquant que l'intégrale restante est égale à I_f. □

Comme on le voit, les ensembles de points homogènes ne produisent pas de biais dans l'intégration Monte Carlo. L'erreur d'intégration se manifeste donc essentiellement sous forme de variance. Cette observation pourrait n'être qu'anecdotique si les ensembles de points homogènes ne jouaient pas un rôle essentiel dans l'intégration Monte Carlo. Comme nous l'avons dit dans le chapitre 4, de nombreuses méthodes stochastiques sont homogènes. Ceci s'explique par le fait qu'en générale, l'intégrande est inconnue et que l'on ne peut donc pas supposer qu'une partie du domaine d'intégration soit plus importante qu'une autre. L'homogénéité intervient donc pour garantir que tout le domaine d'intégration est couvert de la même façon par l'ensemble de points. D'une autre façon, on peut dire que pour un ensemble de points fixé, on veut que l'erreur de l'intégration Monte Carlo soit la même quelque soit la «position » de l'intégrande dans le domaine. Dans le formalisme que nous avons introduit dans le chapitre 3, où $\mathbb{T}_{\mathcal{E}}$ est un ensemble d'action (translation, rotation, ...), ce revient à écrire que pour tout $\tau \in \mathbb{T}_{\mathcal{E}}$,

$$MSE(I_{\mathcal{S},f}) = MSE(I_{\mathcal{S},\tau^{-1}(f)}) = MSE(I_{\tau(\mathcal{S}),f}) \,. \tag{7.15}$$

Comme on le voit, en imposant que la position de l'intégrande n'affecte pas le résultat de l'intégration Monte Carlo, on impose de la même façon que l'ensemble de points soit homogène. Cela ne veux pas dire que les ensembles de points qui ne sont pas homogènes ne sont pas intéressants. Mais si l'on assume que la position de l'intégrande ne doit pas affecter l'erreur d'intégration, alors il faut considérer que l'ensemble de points est homogène, soit par nature, soit par une opération d'homogénéisation, comme décrite dans Lemme (3.3, p. 28).

Dans les sections suivantes, nous montrons comment sont reliés la discrépance, la fonction de distribution de paires et le spectre de puissance avec l'erreur d'intégration, le biais et la variance. Puis nous nous restreindrons aux ensembles de points homogènes pour analyser la vitesse de convergence de la variance dans l'intégration Monte Carlo.

7.4 Lien avec la discrépance à l'origine

L'inégalité de Koksma-Hlawka [Lemieux, 2009, p. 184] permet de borner l'erreur générée par un ensemble de points déterministe dans l'intégration Monte-Carlo. Pour cela, on restreint les intégrandes à une classe de fonction particulière.

Définition 7.2 *[Brandolini* et al.*, 2013] Soit $[a,b] \subset \mathbb{R}$ un intervalle et f une fonction $[a,b] \to \mathbb{C}$, on appelle **variation totale** de f, notée $\mathcal{V}(f)$, l'opérateur à valeur dans $\mathbb{R}^+$ défini par*

$$\mathcal{V}(f) := \sup \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \,, \tag{7.16}$$

où les x_k forment une partition de $[a,b]$ tels que $a = x_0 < x_1 < .... < x_n = b$ et le supremum est choisi sur l'ensemble des partitions possibles de $[a,b]$.

On dit qu'une fonction f est à variation bornée si $\mathcal{V}(f)$ est finie, et on nomme l'ensemble de ces fonctions, la classe des fonctions à variation bornée. Lorsque l'on se restreint à cette classe de fonctions et que l'on se place dans l'espace d'échantillonnage restreint $\mathcal{U}^1$, on peut relier l'erreur de l'intégration Monte Carlo avec la discrépance,

définie par Def. (5.12, p. 56), et la variation totale de f à l'aide de l'inégalité de Koksma-Hlawka [Niederreiter, 1992, theo. 2.9],

$$Err(I_{\mathcal{S},f}) \leq D^*(\mathcal{S})\mathcal{V}(f)\,. \tag{7.17}$$

Pour généraliser cette relation aux dimensions supérieurs, il faut étendre la notion de variation totale aux dimensions supérieures, tout en préservant l'inégalité de Koksma-Hlawka. On utilise pour cela la variation de Hardy-Krause, basée sur la variation de Vitali.

7.4.1 Variation de Hardy-Krause

Cette définition de la variance de Hardy-Krause est inspirée de [Niederreiter, 1992], [Lemieux, 2009] et [Liflyand and Stadtmüller, 2013]. Soit f une fonction $[0,1]^d \to \mathbb{C}$ et $\{e_1, ..., e_d\}$ les vecteurs de la base canonique de $\mathbb{R}^d$, on définit la fonction Δ_{h_k} par

$$\Delta_{h_k}(f,x) := f(x + h_k \cdot e_k) - f(x)\,, \tag{7.18}$$

et la fonction $\Delta_{h_1,...,h_k}$ par

$$\Delta_{h_1,...,h_k}(f,x) := \Delta_{h_k}(\Delta_{h_{k-1}}, x)\,, \tag{7.19}$$

tel que $x \in [0,1]^d$ et $x + e_k \cdot h_k \in [0,1]^d$. On considère maintenant l'ensemble Π des partitions de $[0,1]^d$ ayant la forme d'une grille alignée aux axes. Chaque partition $\pi \in \Pi$ est formée en subdivisant chaque dimension à l'aide d'un vecteur $\pi_k = (x_k^{(1)}, ..., x_k^{(n_k)})$ ayant la propriété suivante

$$0 = x_k^{(1)} < ... < x_k^{(n_k)} = 1\,. \tag{7.20}$$

Une partition π est donc formée de $(n_1 - 1) \cdot ... \cdot (n_k - 1)$ pavés. La variation de Vitali est définie par

$$\mathcal{V}_{\text{Vitali}}(f) := \sup_{\pi \in \Pi} \sum_{J \in \pi} |\Delta_{h_1,...,h_k}(f,())|\,, \tag{7.21}$$

où J est un des pavés de la partition π et h_k est la taille du pavé J pour la dimension k. La variation de Hardy-Krause, notée $V_{HK}(f)$, est définie comme la somme des variations de Vitali pour toutes les projections de f sur un sous-ensemble de dimensions. Chaque projection f^S est définie sur un sous-ensemble $S = \{i_1, ..., i_s\}$ des dimensions $1, ..., d$ et vérifie $f^S(x_{i_1}, ..., x_{i_s}) = f(\widetilde{x_1}, ..., \widetilde{x_d})$ tel que

$$\widetilde{x_k} = \begin{cases} x_k, & \text{si } k \in S \\ 1, & \text{sinon}\,. \end{cases} \tag{7.22}$$

La classe des fonctions à variation bornée selon Hardy-Krause est l'ensemble des fonctions f telles que $V_{HK}(f) < \infty$. Certaines fonctions non-continues peuvent avoir une variation de Hardy-Krause bornée, à condition que les discontinuités soient alignées aux axes. À l'inverse, certaines fonctions dont les discontinuités ne sont pas alignées aux axes n'ont pas une variation bornée [Lemieux, 2009, p. 185]. Ce point est important car il signifie que les fonctions caractéristiques de formes simples comme des cercles ou des triangles ne font pas partie de la classe des fonctions à variation bornée. Or ces types de fonctions sont très répandus dans le rendu graphique, par exemple dans les tests de visibilité.

7.4.2 Inégalité de Koksma-Hlawka multi-dimensionnelle

À l'aide de la variation de Hardy-Krause, il est donc possible d'étendre l'inégalité de Koksma-Hlawka aux fonctions de dimensions supérieurs.

Lemme 7.2 *[Niederreiter, 1992, theo. 2.11] Soit f une fonction dont la variation de Hardy-Krause $V_{HK}(f)$ est finie et $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{U}^d$, la borne supérieur du biais de $I_{\mathcal{S},f}$ est donnée par*

$$Err(I_{\mathcal{S},f}) \leq D^*(\mathcal{S})V_{HK}(f)\,. \tag{7.23}$$

La leçon principale à tirer de l'inégalité de Koksma-Hlawka est que plus la discrépance est faible, plus l'erreur d'intégration l'est également. Ainsi, la réduction de la discrépance est un des enjeux des algorithmes d'échantillonnage. En générale, on associe la densité de l'ensemble de points dans la réduction de la discrépance. Concrètement, on cherche à ce que lorsque le nombre de points utilisés pour échantillonner l'intégrande augmente, l'erreur diminue le plus rapidement possible. On parle alors de vitesse de convergence d'une méthode d'échantillonnage. Du point de vue de l'inégalité de Koksma-Hlawka, l'objectif d'un ensemble de points est donc de diminuer la discrépance à l'origine le plus rapidement possible quand la densité augmente. C'est sur cet objectif que se sont bâties les séquences à basse discrépance, dont nous avons parlé dans le chapitre 4. Ainsi, les séquences à basse discrépance ont une discrépance se comportant en $O_{N\to\infty}\left(log(N)^d N^{-1}\right)$, ce qui implique donc que l'erreur de l'intégration Monte Carlo, $Err(I_{\mathcal{S},f})$, converge elle aussi en $O_{N\to\infty}\left(log(N)^d N^{-1}\right)$. La vitesse de convergence de $Err(I_{\mathcal{S},f})$ dans le cas du bruit blanc étant en $O_{N\to\infty}\left(N^{-1/2}\right)$, on voit que les séquences à basse discrépance, lorsque le nombre d'échantillons N est suffisamment grand, surpasse le bruit blanc.

L'inégalité de Koksma-Hlawka nous montre ainsi comment, à l'aide d'une expression de l'erreur d'intégration, fixer un objectif clair pour l'amélioration des ensembles de points. Ici, l'objectif est de créer un ensemble de points dont la discrépance diminue le plus rapidement possible en fonction de la densité. Cependant, il ne faut pas oublier que dans le cas multi-dimensionnel, l'inégalité de Koksma-Hlawka est restreinte aux fonctions dont la variation de Hardy-Krause est bornée, c'est-à-dire un ensemble de fonctions qui ne correspondent pas actuellement aux besoins de la synthèse d'images. Ainsi, la discrépance ne nous dit pas quels doivent être les efforts à faire sur les ensembles de points pour intégrer correctement une fonction simple telle que la fonction caractéristique d'un triangle ou d'un disque. Il est donc nécessaire de dépasser l'inégalité de Koksma-Hlawka en utilisant d'autres outils pour donner une expression de l'erreur d'intégration pour une classe de fonction plus grande que les fonctions à variation bornée.

Nous proposons deux approches pour exprimer l'erreur d'intégration d'un ensemble de points stochastique, l'une basée sur la fonction de distribution de paires, et l'autre sur la transformée de Fourier. Puis en nous restreignant aux ensembles de points homogènes, nous montrons qu'à la manière de l'inégalité de Koksma-Hlawka, il est possible de définir des critères simples pour améliorer la vitesse de convergence d'une méthode d'échantillonnage.

7.5 Lien avec l'analyse spatiale

Lemme 7.3 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$ et f une fonction $\mathbb{D}_\mathcal{E} \to \mathbb{C}$, le biais de l'intégration Monte Carlo de f est donné par*

$$Bias(I_{\mathcal{S},f}) = \frac{1}{\rho_\mathcal{S}} \int_{\mathbb{D}_\mathcal{E}} \left(\rho_\mathcal{S}^{(1)}(x) - \rho_\mathcal{S}\right) \overline{f(x)} \mathrm{d}x \,, \tag{7.24}$$

où $\rho_\mathcal{S}^{(1)}$ est la densité locale de l'ensemble de points définie dans Def. (5.1, p. 46).

Preuve 7.2 *En développant Eq. (7.9), on obtient*

$$Bias(I_{\mathcal{S},f}) = \frac{1}{\rho_\mathcal{S}} \int_{\mathbb{D}_\mathcal{E}} \langle S(x) \rangle \overline{f(x)} \mathrm{d}x - \frac{1}{\rho_\mathcal{S}} \int_{\mathbb{D}_\mathcal{E}} \overline{f(x)} \mathrm{d}x \,. \tag{7.25}$$

Sachant que l'espérance de $S(x)$ correspond à $\rho_\mathcal{S}^{(1)}(x)$, voir Eq. (5.3, p. 46), et en regroupant les intégrales, on obtient bien Eq. (7.24). □

On remarque tout de suite dans cette forme pourquoi les ensembles de points homogènes ne produisent pas de biais, puisque l'on a vu que dans ce cas, la densité locale est égale à la densité globale (voir chapitre 5). Cependant, cette propriété est plus générale, puisque des ensembles de points non-homogènes peuvent vérifier l'égalité $\rho_\mathcal{S}^{(1)}(x) = \rho_\mathcal{S}$. C'est par exemple le cas des ensembles de points *jittered* (voir chapitre 8).

Lemme 7.4 *Soit $\mathcal{S}$ un ensemble de points dans l'espace $\mathcal{E}$ et f une fonction $\mathbb{D}_\mathcal{E} \to \mathbb{C}$, la variance de l'intégration Monte Carlo de f est donnée par*

$$Var(I_{\mathcal{S},f}) = \frac{1}{\rho_\mathcal{S}^2} \int_{\mathbb{D}_\mathcal{E}} \int_{\mathbb{D}_\mathcal{E}} \rho_\mathcal{S}^{(1)}(x) \rho_\mathcal{S}^{(1)}(x') h_\mathcal{S}(x,x') \overline{f(x)} f(x') \mathrm{d}x' \mathrm{d}x \,, \tag{7.26}$$

où $h_\mathcal{S}$ est la fonction de corrélation totale définie dans Def. (5.9, p. 52).

Preuve 7.3 *À partir de Eq. (7.8) et de Eq. (7.4), on obtient le développement suivant*

$$\begin{aligned} Var(I_{\mathcal{S},f}) &= \frac{1}{\rho_\mathcal{S}^2} \int_{\mathbb{D}_\mathcal{E}} \int_{\mathbb{D}_\mathcal{E}} \left\langle S(x)\overline{S(x')} \right\rangle \overline{f(x)} f(x') \mathrm{d}x' \mathrm{d}x \\ &- \frac{1}{\rho_\mathcal{S}^2} \int_{\mathbb{D}_\mathcal{E}} \int_{\mathbb{D}_\mathcal{E}} \langle S(x) \rangle \overline{\langle S(x') \rangle} \overline{f(x)} f(x') \mathrm{d}x' \mathrm{d}x \\ &= \frac{1}{\rho_\mathcal{S}^2} \int_{\mathbb{D}_\mathcal{E}} \int_{\mathbb{D}_\mathcal{E}} \left(\left\langle S(x)\overline{S(x')} \right\rangle - \langle S(x) \rangle \overline{\langle S(x') \rangle} \right) \overline{f(x)} f(x') \mathrm{d}x' \mathrm{d}x \,. \end{aligned} \tag{7.27}$$

À l'aide de la définition de la densité locale et de la densité de paires, on obtient l'égalité suivante

$$\left\langle S(x)\overline{S(x')} \right\rangle - \langle S(x) \rangle \overline{\langle S(x') \rangle} = \rho_\mathcal{S}^{(2)}(x,x') - \rho_\mathcal{S}^{(1)}(x) \rho_\mathcal{S}^{(1)}(x') \,. \tag{7.28}$$

La forme finale de Eq. (7.26) est obtenue en factorisant par la densité locale et en utilisant la définition de la fonction de corrélation totale. □

La fonction de corrélation totale est une simple normalisation de la densité de paires. Notre expression de la variance est donc dépendante de la densité locale et de la densité de paires de l'ensemble de points. Lorsque l'ensemble de points est homogène, nous avons vu dans le chapitre 5 que la densité locale et la densité de paires peuvent se simplifier. De la même façon, l'expression de la variance de l'intégration Monte Carlo dans le domaine spatial se simplifie dans le cas homogène.

Lemme 7.5 *Soit $\mathcal{S}$ un ensemble de points homogène dans l'espace $\mathcal{E}$ et f une fonction $\mathbb{D}_\mathcal{E} \to \mathbb{C}$, la variance de l'intégration Monte Carlo de f est donnée par*

$$Var(I_{\mathcal{S},f}) = \frac{\mu_\mathcal{E}(\mathbb{D}_\mathcal{E})}{\mu_H(\mathbb{T}_\mathcal{E})} \int_{\mathbb{T}_\mathcal{E}} h_\mathcal{S}(\tau)(f \oplus f)(\tau)\mathrm{d}\tau\,, \tag{7.29}$$

où $h_\mathcal{S}$ est la version différentielle de la fonction de corrélation totale définie page 52 du chapitre 5, et $(f \oplus f)$ l'auto-corrélation de la fonction f, définie par Def. (5.10, p. 53).

Preuve 7.4 *À l'aide de Eq. (3.18), on applique le changement de variable $x' = x * \tau$ sur Eq. (7.26). On obtient alors*

$$Var(I_{\mathcal{S},f}) = \frac{\mu_\mathcal{E}(\mathbb{D}_\mathcal{E})}{\mu_H(\mathbb{T}_\mathcal{E})} \frac{1}{\rho_\mathcal{S}^2} \int_{\mathbb{D}_\mathcal{E}} \int_{\mathbb{T}_\mathcal{E}} \rho_\mathcal{S}^{(1)}(x)\rho_\mathcal{S}^{(1)}(x*\tau)h_\mathcal{S}(x, x*\tau)\overline{f(x)}f(x*\tau)\mathrm{d}\tau\mathrm{d}x\,. \tag{7.30}$$

Comme l'ensemble de points est homogène, on sait que la densité locale est égale à la densité globale, que la densité de paires est égale à la densité différentielle et donc que $h_\mathcal{S}(x, x*\tau) = \check{h}_\mathcal{S}(\tau)$. L'équation se simplifie donc de la façon suivante

$$\begin{aligned} Var(I_{\mathcal{S},f}) &= \frac{\mu_\mathcal{E}(\mathbb{D}_\mathcal{E})}{\mu_H(\mathbb{T}_\mathcal{E})} \frac{1}{\rho_\mathcal{S}^2} \int_{\mathbb{D}_\mathcal{E}} \int_{\mathbb{T}_\mathcal{E}} \rho_\mathcal{S}^2 \check{h}_\mathcal{S}(\tau)\overline{f(x)}f(x*\tau)\mathrm{d}\tau\mathrm{d}x \\ &= \frac{\mu_\mathcal{E}(\mathbb{D}_\mathcal{E})}{\mu_H(\mathbb{T}_\mathcal{E})} \int_{\mathbb{T}_\mathcal{E}} \check{h}_\mathcal{S}(\tau) \int_{\mathbb{D}_\mathcal{E}} \overline{f(x)}f(x*\tau)\mathrm{d}x\mathrm{d}\tau\,. \end{aligned} \tag{7.31}$$

La deuxième intégrale correspond à la définition de l'auto-corrélation de f, ce qui donne bien Eq. (7.29). □

Cette expression de la variance dans le cas homogène est plus lisible que la version générale. On peut par exemple tirer comme conclusion que là où l'auto-corrélation de l'intégrande est élevée, la fonction de corrélation totale de l'ensemble de points doit être la plus faible possible. Cependant, comme l'auto-corrélation de l'intégrande n'est généralement pas connue, cette expression de la variance est difficilement utilisable pour déterminer les propriétés essentielles que doit avoir un ensemble de points. Dans la section suivante, nous verrons que la transformée de Fourier, et plus précisément le spectre de puissance, sont des outils plus à même de trouver ces propriétés.

7.6 Lien avec l'analyse spectrale

En utilisant le théorème de Parseval et la définition de l'intégration Monte-Carlo donnée dans Eq. (7.4), on obtient l'expression de l'intégration Monte-Carlo dans le domaine spectral

$$I_{\mathcal{S},f} = \frac{1}{\rho_\mathcal{S}} \int_{\mathbb{F}} \mathcal{F}_\mathcal{S}(\omega)\overline{\mathcal{F}_f(\omega)}\mathrm{d}\omega\,. \tag{7.32}$$

On retrouve cette forme notamment dans [Durand, 2011] et [Subr and Kautz, 2013] pour l'espace d'échantillonnage restreint $\mathcal{U}^d$. Ici, le formalisme que nous avons introduit dans les chapitres 3 et 6 nous permet de généraliser cette expression à d'autres espaces d'échantillonnage.

Lemme 7.6 *Soit $\mathcal{S}$ un ensemble de points dans un espace d'échantillonnage $\mathcal{E}$ et f une fonction $\mathbb{D}_\mathcal{E} \to \mathbb{C}$, le biais de l'intégration Monte Carlo de f s'exprime dans le domaine*

spectral par

$$Bias(I_{\mathcal{S},f}) = \frac{1}{\rho_{\mathcal{S}}} \int_{\mathbb{F}\setminus\{\omega_0\}} \langle \mathcal{F}_S(\omega) \rangle \overline{\mathcal{F}_f(\omega)} \mathrm{d}\omega \,, \tag{7.33}$$

où ω_0, définie p. 67, est la fréquence représentant l'intégrale du signal.

Preuve 7.5 *D'après les propriétés de la transformée de Fourier, on sait qu'il existe une fréquence ω_0 telle que $I_f = \mathcal{F}_f(\omega_0)$. Le biais s'exprimant dans le domaine spectral comme*

$$Bias(I_{\mathcal{S},f}) = \frac{1}{\rho_{\mathcal{S}}} \int_{\mathbb{F}} \langle \mathcal{F}_S(\omega) \rangle \overline{\mathcal{F}_f(\omega)} \mathrm{d}\omega - I_f \,, \tag{7.34}$$

il suffit de changer le domaine d'intégration par $\mathbb{F} \setminus \{\omega_0\}$ pour trouver l'équation Eq. (7.33). □

Cette expression du biais dans l'intégration Monte Carlo a été formulée par Subr et al. [2013] dans le cas de l'espace d'échantillonnage restreint $\mathcal{U}^d$. Notre version, donnée dans Lemme (7.6), est une version généralisée à tous les espaces d'échantillonnage. Regardons maintenant la variance de l'intégration.

Lemme 7.7 *Soit $\mathcal{S}$ un ensemble de points dans un espace d'échantillonnage $\mathcal{E}$ et f une fonction $\mathbb{D}_{\mathcal{E}} \to \mathbb{C}$, la variance de l'intégration Monte Carlo de f s'exprime dans le domaine spectral par*

$$Var(I_{\mathcal{S},f}) = \frac{1}{\rho_{\mathcal{S}}^2} \int_{\mathbb{F}} \int_{\mathbb{F}} Cov(\mathcal{F}_S(\omega), \overline{\mathcal{F}_S(\omega')}) \overline{\mathcal{F}_f(\omega)} \mathcal{F}_f(\omega') \mathrm{d}\omega' \mathrm{d}\omega \,. \tag{7.35}$$

Preuve 7.6 *À partir de Eq. (7.32) et de Eq. (7.8), on obtient l'expression de la variance suivante*

$$\begin{aligned} Var(I_{\mathcal{S},f}) = & \frac{1}{\rho_{\mathcal{S}}^2} \left\langle \left(\int_{\mathbb{F}} \mathcal{F}_S(\omega) \overline{\mathcal{F}_f(\omega)} \mathrm{d}\omega \right) \left(\int_{\mathbb{F}} \overline{\mathcal{F}_S(\omega')} \mathcal{F}_f(\omega') \mathrm{d}\omega' \right) \right\rangle \\ & - \frac{1}{\rho_{\mathcal{S}}^2} \left\langle \int_{\mathbb{F}} \mathcal{F}_S(\omega) \overline{\mathcal{F}_f(\omega)} \mathrm{d}\omega \right\rangle \left\langle \int_{\mathbb{F}} \overline{\mathcal{F}_S(\omega')} \mathcal{F}_f(\omega') \mathrm{d}\omega' \right\rangle \,. \end{aligned} \tag{7.36}$$

En développant cette expression, on obtient

$$\begin{aligned} Var(I_{\mathcal{S},f}) = & \frac{1}{\rho_{\mathcal{S}}^2} \left(\int_{\mathbb{F}} \int_{\mathbb{F}} \left\langle \mathcal{F}_S(\omega) \overline{\mathcal{F}_S(\omega')} \right\rangle \overline{\mathcal{F}_f(\omega)} \mathcal{F}_f(\omega') \mathrm{d}\omega' \mathrm{d}\omega \right) \\ & - \frac{1}{\rho_{\mathcal{S}}^2} \int_{\mathbb{F}} \int_{\mathbb{F}} \langle \mathcal{F}_S(\omega) \rangle \left\langle \overline{\mathcal{F}_S(\omega')} \right\rangle \overline{\mathcal{F}_f(\omega)} \mathcal{F}_f(\omega') \mathrm{d}\omega' \mathrm{d}\omega \\ = & \frac{1}{\rho_{\mathcal{S}}^2} \int_{\mathbb{F}} \int_{\mathbb{F}} \left(\left\langle \mathcal{F}_S(\omega) \overline{\mathcal{F}_S(\omega')} \right\rangle - \langle \mathcal{F}_S(\omega) \rangle \left\langle \overline{\mathcal{F}_S(\omega')} \right\rangle \right) \overline{\mathcal{F}_f(\omega)} \mathcal{F}_f(\omega') \mathrm{d}\omega' \mathrm{d}\omega \,. \end{aligned} \tag{7.37}$$

La preuve est terminée en utilisant la définition de la covariance donnée dans Eq. (3.6). □

Cette expression de la variance dépend de la corrélation entre les coefficients de Fourier de l'ensemble de points. À l'inverse de l'expression de l'intégration Monte Carlo et de l'expression du biais, notre expression de la variance est différente de celle donnée par Subr et al. [2013] pour l'espace d'échantillonnage restreint $\mathcal{U}^d$. Selon

leur article, la variance de l'intégration Monte Carlo peut-être exprimée de la façon suivante

$$Var(I_{\mathcal{S},f})_{\text{Subr et al.}} = \frac{1}{\rho_{\mathcal{S}}} \int_{\mathbb{D}_{\mathcal{E}}} Var(\mathcal{F}_f(\omega))|\mathcal{F}_f(\omega)|^2 \mathrm{d}\omega\,, \tag{7.38}$$

où la variance d'un nombre complexe est calculée selon son amplitude. Cette expression montre que la corrélation entre fréquences ne joue aucun rôle dans la variance. Comme nous l'avons vu dans le chapitre 6, la covariance entre deux fréquences différentes est effectivement nulle lorsque l'ensemble de points est homogène. Mais cette propriété n'est pas généralisable à tout type d'ensemble de points. Plus important, les auteurs montrent par cette équation que la phase de l'intégrande n'intervient pas dans l'expression de la variance [1]. Dans notre formalisme, cela voudrait dire que la variance de l'intégration Monte Carlo est invariante par le groupe des actions. Or, de nombreux contre-exemples peuvent être trouvés. Il suffit par exemple de prendre un ensemble de points constitué d'un seul point, lequel est aléatoirement déplacé dans un faible rayon. Prenons maintenant une intégrande possédant une région dont les variations sont faibles, et une région dont les variations sont élevées. Dans ce cas-là, selon l'emplacement de l'intégrande par rapport à l'ensemble de points, la variance sera soit faible, soit élevée. En l'absence de plus d'éclairage sur ces points, nous utiliserons donc notre version de l'expression de la variance de l'intégration Monte Carlo.

Lorsque l'ensemble de points est homogène, il est possible de donner une expression plus compacte de la variance de l'intégration Monte Carlo, en utilisant cette fois le spectre de puissance.

Lemme 7.8 *Soit $\mathcal{S}$ un ensemble de points homogène dans un espace d'échantillonnage $\mathcal{E}$ et f une fonction $\mathbb{D}_{\mathcal{E}} \to \mathbb{C}$, la variance de l'intégration Monte Carlo de f s'exprime dans le domaine spectral par*

$$Var(I_{\mathcal{S},f}) = \frac{1}{\rho_{\mathcal{S}}^2} \int_{\mathbb{P}\setminus\{r_0\}} \mu_F\left(\mathbb{F}^r\right) \langle \mathcal{P}_S(r)\rangle\, \mathcal{P}_f(r) \mathrm{d}r\,. \tag{7.39}$$

Preuve 7.7 *En éliminant les termes égaux à zéro à l'aide de Lemme (6.6, p. 75) sur Eq. (7.35), on obtient*

$$Var(I_{\mathcal{S},f}) = \frac{1}{\rho_{\mathcal{S}}^2} \int_{\mathbb{F}\setminus\{\omega_0\}} Cov(\mathcal{F}_S(\omega), \overline{\mathcal{F}_S(\omega)})\overline{\mathcal{F}_f(\omega)}\mathcal{F}_f(\omega) \mathrm{d}\omega\,. \tag{7.40}$$

Puis on décompose le domaine d'intégration $\mathbb{F}$ en sous-domaines $\mathbb{F}^{r_0}, \mathbb{F}^{r_1}, \ldots$ de la façon suivante,

$$Var(I_{\mathcal{S},f}) = \frac{1}{\rho_{\mathcal{S}}^2} \int_{\mathbb{P}\{r_0\}} \int_{\mathbb{F}^r} Cov(\mathcal{F}_S(\omega), \overline{\mathcal{F}_S(\omega)})\overline{\mathcal{F}_f(\omega)}\mathcal{F}_f(\omega) \mathrm{d}\omega \mathrm{d}r\,. \tag{7.41}$$

On peut maintenant utiliser le lien entre la covariance et le spectre de puissance établi par Lemme (6.6),

$$Var(I_{\mathcal{S},f}) = \frac{1}{\rho_{\mathcal{S}}^2} \int_{\mathbb{P}\{r_0\}} \langle \mathcal{P}_S(r)\rangle \int_{\mathbb{F}^r} \overline{\mathcal{F}_f(\omega)}\mathcal{F}_f(\omega) \mathrm{d}\omega \mathrm{d}r\,. \tag{7.42}$$

La deuxième intégrale correspond à la définition du spectre de puissance donnée dans Eq. (6.41), ce qui donne bien Eq. (7.39). □

1. Dans [Subr and Kautz, 2013], la dernière ligne de la section 2.3 souligne très clairement ce point

Cette expression de la variance en fonction du spectre de puissance est, à notre connaissance, nouvelle dans notre domaine. Elle repose sur notre définition de l'espace d'échantillonnage, ainsi que sur la généralisation de la transformée de Fourier décrite dans le chapitre 6. Dans l'espace d'échantillonnage euclidien $\mathcal{E}^d$, cette expression devient

$$\boxed{Var(I_{\mathcal{S},f}) = \frac{1}{\rho_{\mathcal{S}}^2} \int_{\mathbb{P}} \langle \mathcal{P}_{\mathcal{S}}(r) \rangle \, \mathcal{P}_f(r) \mathrm{d}r} \,, \tag{7.43}$$

avec $\mathbb{P} = \mathbb{R}^d$. Dans l'espace d'échantillonnage hyper-cubique périodique $\mathcal{T}^d$, cette expression devient

$$\boxed{Var(I_{\mathcal{S},f}) = \frac{1}{N^2} \sum_{r \in \mathbb{P}} \langle \mathcal{P}_{\mathcal{S}}(r) \rangle \, \mathcal{P}_f(r)} \,, \tag{7.44}$$

avec $\mathbb{P} = \mathbb{Z}^d$. La mesure du domaine d'intégration étant égale à 1, la densité de l'ensemble de points est égale au nombre de points N. Finalement, pour l'espace d'échantillonnage sphérique $\mathcal{S}^2$, le lemme 7.8 devient

$$\boxed{Var(I_{\mathcal{S},f}) = \frac{16\pi^2}{N^2} \sum_{r \in \mathbb{P}} (2r+1) \, \langle \mathcal{P}_{\mathcal{S}}(r) \rangle \, \mathcal{P}_f(r)} \,, \tag{7.45}$$

où $\mathbb{P} = \mathbb{N}$. Pour chacune de ces déclinaisons, nous avons utilisé les définitions de la transformée de Fourier et du spectre de puissance données dans la section 6.3, page 71. Le lemme 7.8, et ses déclinaisons par espace d'échantillonnage, sont une des contributions principales de ce manuscrit. Bien qu'ils ne soient que la version spectrale de Lemme (7.5), leur porté est toute autre. Du point de vue pratique, le spectre de puissance est l'outil d'analyse le plus utilisé pour les ensembles de points stochastiques. Une expression de la variance en fonction du spectre de puissance a donc nécessairement un attrait pratique plus fort que d'autres expressions basées sur des outils plus difficiles d'accès. Du point de vue théorique, nous allons voir dans la section suivante que la forme du spectre de puissance, de l'ensemble de points comme de l'intégrande, joue un rôle clé dans la vitesse de convergence de l'intégration Monte Carlo.

7.7 Vitesse de convergence

Nous souhaitons étudier la vitesse de convergence d'une méthode d'échantillonnage homogène en fonction de la densité de l'ensemble de points. On suppose donc que la méthode d'échantillonnage est capable de générer des ensembles de points de densité arbitraire, et que les ensembles de points sont tous homogènes. Ici, nous parlons de densité plutôt que de nombre de points pour prendre en compte les espaces d'échantillonnage infinis, comme l'espace euclidien. Dans le cas des espaces d'échantillonnage à mesure finie, tels que les espaces sphériques $\mathcal{S}^2$ ou hyper-cubiques $\mathcal{T}^d$, la densité peut être remplacée par le nombre de points sans perte de généralité. On note $\mathcal{S}_\rho$ l'ensemble de points généré par un échantillonneur tel que sa densité soit égale à ρ, et S_ρ sa fonction représentative telle que définie dans Def. (3.2, p. 17). Le lecteur notera ici le renversement de notation par rapport aux précédants chapitres. Jusqu'à maintenant, la notation $\rho_{\mathcal{S}}$ représentait la densité d'un ensemble de points $\mathcal{S}$, soulignant que celle-ci était dépendante de l'ensemble de points. En utilisant la notation $\mathcal{S}_\rho$, nous

exprimons à l'inverse que l'ensemble de points est dépendant d'une densité fixée de façon externe. Ce choix est justifié par le fait que nous voulons étudier l'impacte la densité sur l'ensemble de points et sur l'intégration Monte Carlo.

La variance de l'intégration Monte Carlo utilisant l'ensemble de points $\mathcal{S}_\rho$ est donnée, à l'aide de Lemme (7.5), par

$$Var(I_{\mathcal{S}_\rho,f}) = \frac{1}{\rho^2}\int_{\mathbb{P}} \mu_F\left(\mathbb{F}^r\right)\left\langle \mathcal{P}_{S_\rho}(r)\right\rangle \mathcal{P}_f(r)\mathrm{d}r\,. \tag{7.46}$$

Comme on le voit, la densité de l'ensemble de points intervient en deux endroits : dans la fraction en début d'équation, et dans le spectre de puissance de l'ensemble de points. Puisque l'on veut étudier le comportement de la variance lorsque la densité augmente, il nous reste donc à comprendre comment la densité affecte le spectre de puissance de l'ensemble de points. Malheureusement, il ne peut y avoir de réponses claires sur ce point car cela dépend entièrement de l'échantillonneur. Il est par exemple possible d'imaginer un échantillonneur produisant des ensembles de points bruit blanc lorsque la densité est faible, puis des grilles régulières lorsque la densité est élevée. Cela n'aurait pas de sens d'un point de vue pratique, mais cette situation reste théoriquement possible. Il faut donc ajouter certaines conditions pour que la vitesse de convergence de la variance puisse être étudiée du point de vue spectral. Le plus simple est d'imposer que l'ensemble de points $\mathcal{S}_\rho$ ait les mêmes propriétés, à un facteur d'échelle près, quelque soit la densité choisie.

7.7.1 Spectre de puissance et densité dans l'espace euclidien

Dans l'espace euclidien, il y a deux façons d'augmenter la densité. Soit on ajoute des points dans un ensemble de points existant, soit on effectue une mise à l'échelle de l'ensemble de points. La deuxième solution est possible car l'ensemble de points est homogène et possède un nombre de points infini [2]. Comme on veut que les propriétés, à un facteur d'échelle près, soit les mêmes quand on augmente la densité, la deuxième solution semble la plus pratique à modéliser. On peut donc écrire l'ensemble de points $\mathcal{S}_\rho$ de la façon suivante,

$$\mathcal{S}_\rho = \left\{ \frac{1}{\sqrt[d]{\rho}} \cdot s : s \in \mathcal{S}_1 \right\}, \tag{7.47}$$

où $\mathcal{S}_1$ est l'ensemble de points à la densité $\rho = 1$, s représente un point de $\mathcal{S}_1$ sous forme vectoriel et $1/\sqrt[d]{\rho} \cdot s$ est une opération de multiplication entre un scalaire et un vecteur de $\mathbb{R}^d$. Le facteur $\sqrt[d]{\rho}$ s'explique par l'exemple suivant. En dimension deux, augmenter la densité de quatre fois veut dire que les points seront deux fois plus proches les uns des autres. En dimension plus élevée, si le nombre de points augmente de ρ fois, les points deviennent $\sqrt[d]{\rho}$ fois plus proches les uns des autres.

Regardons maintenant la fonction représentative de $\mathcal{S}_\rho$. Comme l'ensemble de points est mis à l'échelle lorsque la densité augmente, la position des fonctions delta de Dirac dans la fonction représentative de $\mathcal{S}_\rho$ se déplacent. On peut alors écrire que

$$S_\rho(x) = \rho S_1(\sqrt[d]{\rho} \cdot x)\,. \tag{7.48}$$

Le facteur ρ permet de s'assurer que l'intégrale de chaque fonction delta de Dirac est toujours égale à 1. Avec cette expression de la fonction représentative de $\mathcal{S}_\rho$, nous

2. La mesure du domaine d'échantillonnage est infini, le densité est fini, donc le nombre de point est infini

sommes maintenant prêts pour donner une expression du spectre de puissance en fonction de ρ. Comme nous l'avons dit dans Eq. (6.6, p. 62), une mise à l'échelle par un facteur α dans le domaine spatial implique une mise à l'échelle par un facteur $1/\alpha$ dans le domaine spectral. On obtient donc la relation suivante pour l'espace euclidien,

$$\begin{aligned}\mathcal{P}_{S_\rho}(\omega) &= \mathcal{F}_{S_\rho}(\omega)\overline{\mathcal{F}_{S_\rho}(\omega)} \\ &= \mathcal{F}_{S_1}\left(\frac{1}{\sqrt[d]{\rho}}\cdot\omega\right)\overline{\mathcal{F}_{S_1}\left(\frac{1}{\sqrt[d]{\rho}}\cdot\omega\right)} \\ &= \mathcal{P}_{S_1}\left(\frac{1}{\sqrt[d]{\rho}}\cdot\omega\right). \end{aligned} \tag{7.49}$$

Ainsi, lorsque la densité augmente, le spectre de puissance est «dilaté ». En reportant cette expression du spectre de puissance en fonction de la densité dans Eq. (7.43), on obtient alors

$$Var(I_{S_\rho,f}) = \frac{1}{\rho^2}\int_{\mathbb{P}}\left\langle \mathcal{P}_{S_1}\left(\frac{1}{\sqrt[d]{\rho}}\cdot\omega\right)\right\rangle \mathcal{P}_f(r)\mathrm{d}r\,. \tag{7.50}$$

Cette équation n'est valide que dans l'espace euclidien. Pour en obtenir une similaire dans les autres espaces, il est nécessaire d'effectuer quelques modifications.

7.7.2 Spectre de puissance et densité dans les espaces à mesure finie

Dans les autres espaces que nous avons défini, il n'est pas possible d'utiliser une mise à l'échelle pour augmenter la densité du nombre de points. Si le nombre de points est fini, la mise à l'échelle ne change pas la densité. L'équation (7.49) ne peut pas être directement utilisée, mais elle nous offre une définition acceptable de ce que veut dire «garder les mêmes propriétés quelque soit la densité ». Comme nous l'avons dit, il ne peut pas exister de relation générale liant le spectre de puissance d'un ensemble de points et la densité car cette relation dépend du choix d'implémentation fait dans l'échantillonneur. Nous assumerons donc que le choix fait est celui décrit par Eq. (7.49), qui devient dans le cas des espaces d'échantillonnage fermés,

$$\left\langle \mathcal{P}_{S_\rho}(r)\right\rangle = \rho\left\langle \mathcal{P}_{S_1}\left(\frac{1}{\sqrt[d]{\rho}}\cdot r\right)\right\rangle. \tag{7.51}$$

Le facteur ρ permet ici de modéliser le fait que le nombre de points augmente quand la densité augmente et que par conséquence, l'énergie contenue dans le spectre de puissance augmente également. La dimension d correspond à la dimension topologique de l'espace d'échantillonnage. Nos observations nous ont montré que cette modélisation s'applique à de nombreuses méthodes d'échantillonnage stochastiques.

Plusieurs points restent tout de même à éclaircir dans cette équation. Tout d'abord, lorsqu'un ensemble de points homogène ne contient qu'un seul point, il est nécessairement équivalent au bruit blanc. L'ensemble de points $\mathcal{S}_1$ n'est donc pas l'ensemble de points qui reflète le mieux les propriétés de l'échantillonneur. À la place, nous définirons donc un profil spectral, noté $P_{\mathcal{S}}$, vers lequel doit tendre le spectre de puissance de l'ensemble de points quand la densité augmente,

$$\lim_{\rho\to\infty}\left\langle \mathcal{P}_{S_\rho}(r)\right\rangle = \rho P_{\mathcal{S}}\left(\frac{1}{\sqrt[d]{\rho}}\cdot r\right). \tag{7.52}$$

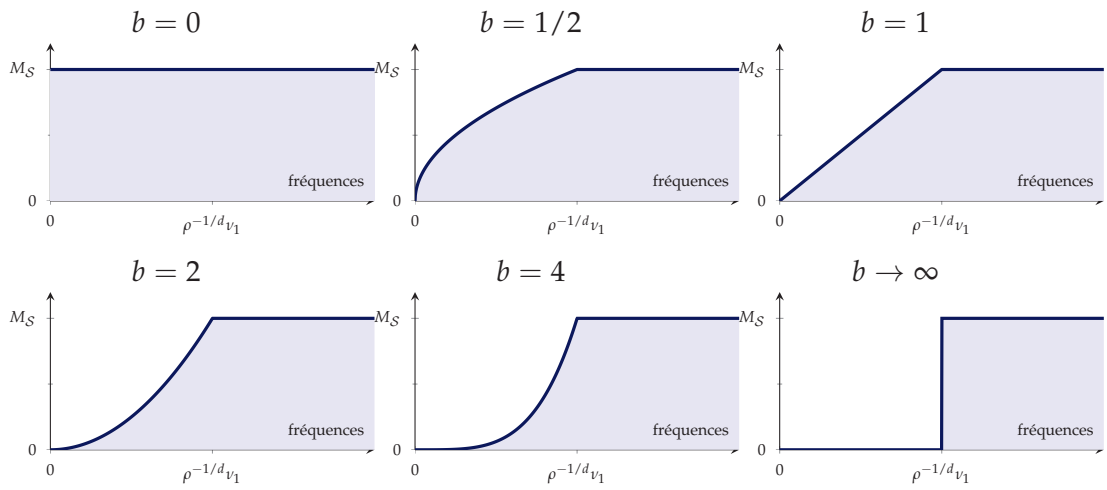


FIGURE 7.1 – *Exemple de profils spectraux de l'ensemble de points pour différentes valeurs de b. Lorsque $b = 0$, le profil est dit constant, et lorsque $b \to \infty$, le profil est dit discontinu.*

Ensuite, dans les espaces d'échantillonnage à mesure finie, le domaine spectral $\mathbb{P}$ est discret ce qui empêche en pratique d'utiliser des opérations de mise à l'échelle. Il faut donc que le profil spectral soit défini sur un domaine continu contenant le domaine $\mathbb{P}$. Pour l'espace hyper-cubique périodique, nous prendrons le domaine $\mathbb{R}^d$ et pour l'espace sphérique, le domaine $\mathbb{R}^+$. Avec la définition du profil spectral, on obtient finalement une expression de la variance en fonction de la densité dans le cas des espaces à mesure finie,

$$Var(I_{\mathcal{S}_\rho,f}) = \frac{1}{\rho}\int_{\mathbb{P}} \mu_F\left(\mathbb{F}^r\right) P_{\mathcal{S}}\left(\frac{1}{\sqrt[d]{\rho}}\cdot r\right)\mathcal{P}_f(r)\mathrm{d}r\,. \tag{7.53}$$

7.7.3 Borne supérieur de la variance

Nous venons de voir que quand la densité augmente, le spectre de puissance de l'ensemble de points est «dilaté ». Cela signifie que relativement au spectre de puissance de l'intégrande, les basses fréquences de l'ensembles de points prennent de plus en plus d'importance. Si la limite du profil spectral de l'ensemble de points vaut zéro en r_0, alors quand la densité tend vers l'infini, $\mathcal{P}_S$ devient presque égale à zéro partout. Comme la variance de l'intégration Monte Carlo est égale à l'intégral du produit entre le spectre de puissance de l'intégrande et de l'ensemble de points, on peut en déduire que la variance tend vers zéro quand la densité tend vers l'infini, si la limite du profil spectral $P_{\mathcal{S}}$ tend vers zéro en r_0. Pour mieux comprendre ce phénomène, nous proposons d'étudier le comportement de la variance en fonction du comportement du profil spectral dans les basses fréquences.

On pose $L(r)$ une fonction $\mathbb{P} \to \mathbb{R}^+$ tel que $P_{\mathcal{S}} \in O_{r\to r_0}(L(r))$. En utilisant la notation en grand O pour les basses fréquences, nous nous assurons que le profil spectral de l'ensemble de points est toujours inférieur à la fonction L, à une constante $M_{\mathcal{S}}$ près et pour r suffisamment proche de r_0. On note ν_1 la distance à partir de laquelle pour tout r tel que $|r| \leq \nu$, on a $M_{\mathcal{S}}L(r) \geq P_{\mathcal{S}}(r)$. La notation $|r|$ désigne ici la distance entre r et r_0 dans l'espace $\mathbb{P}$. Si la limite de L en r_0 est zéro, alors la limite de $P_{\mathcal{S}}$ est elle aussi zéro en r_0. On suppose également que le profil spectral de l'ensemble de points

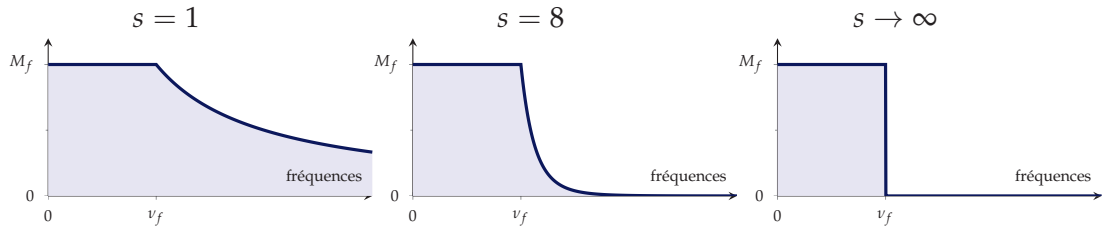


FIGURE 7.2 – *Exemples de profils spectraux de l'intégrande pour différentes valeurs de s. Lorsque $s \to \infty$, le profil a un support fini.*

est toujours borné, c'est-à-dire que pour tout $r \in \mathbb{P}$, il existe $M'_{\mathcal{S}}$ tel que $P_{\mathcal{S}}(r) < M'_{\mathcal{S}}$. À l'aide de cette borne supérieur et de la notation en grand O, on peut écrire que

$$P_{\mathcal{S}}(r) \leq \begin{cases} M_{\mathcal{S}} L(r) & \text{si } |r| \leq \nu_1 \\ M'_{\mathcal{S}} & \text{sinon}\,. \end{cases} \tag{7.54}$$

En reportant cette inéquation dans Eq. (7.53), on obtient l'inégalité suivante

$$\begin{aligned} Var(I_{\mathcal{S}_\rho,f}) \leq & \left(\frac{M_{\mathcal{S}}}{\rho} \int_{R(\sqrt[d]{\rho}\nu_1)} \mu_F\left(\mathbb{F}^r\right) L\left(\frac{1}{\sqrt[d]{\rho}} \cdot r\right) \mathcal{P}_f(r) \mathrm{d}r \right) \\ & + \left(\frac{M'_{\mathcal{S}}}{\rho} \int_{\mathbb{P}\setminus R(\sqrt[d]{\rho}\nu_1)} \mu_F\left(\mathbb{F}^r\right) \mathcal{P}_f(r) \mathrm{d}r \right), \end{aligned} \tag{7.55}$$

où $R(\sqrt[d]{\rho}\nu_1)$ est le sous-ensemble de $\mathbb{P}$ des fréquences telles que $|r| < \sqrt[d]{\rho} \cdot \nu_1$. Comme on le voit, le deuxième terme dépend uniquement des propriétés de l'intégrande. Si l'énergie de f est concentrée dans les hautes fréquences, la variance diminuera lentement. À l'inverse, si l'énergie est concentrée dans les basses fréquences, ce terme convergera rapidement vers zéro. Il est donc nécessaire d'introduire des classes, dont la répartition de l'énergie entre basses et hautes fréquences est connue.

7.7.4 Spectre de puissance de l'intégrande

Il existe de nombreuses relations entre le spectre de puissance et les propriétés d'intégrabilité et de continuité d'un signal. En dimension 1, le théorème de Riemann-Lebesgue nous dit par exemple que si une fonction appartient à $L^1(\mathbb{R}^d)$, c'est-à-dire si elle est intégrable, alors sa transformée de Fourier tend vers zéro quand la fréquence tend vers l'infini [Strichartz, 2003, sec. 7.1]. On a donc une classe de fonctions, $L^1(\mathbb{R}^d)$, qui possède une propriété particulière dans le domaine spectral. En restreignant la classe de fonctions, on peut trouver d'autres propriétés. Brandolini et al. [2001] ont déterminé la vitesse de décroissance du spectre de puissance pour une classe de fonctions particulière. Ainsi, soit $f : \mathbb{R}^d \to \mathbb{C}$ une fonction infiniment dérivable et fenêtrée par un borélien dont les frontières sont elles aussi infiniment dérivables, alors le spectre de puissance radial $\breve{\mathcal{P}}_f(\nu)$, défini par Eq. (6.44, p. 72), est en $O_{\nu\to\infty}\left(\nu^{-(d+1)}\right)$. Les auteurs ont également prouvé ce résultat pour l'espace sphérique. Quelques années plus tard, Brandolini et al. [2003] ont montré que cette propriété est valide également pour tout borélien convexe. Ces classes de fonctions sont très intéressantes pour l'informatique graphique, car les fonctions de test de visibilité en font partie.

En s'inspirant de ces travaux, nous avons défini une classe de fonctions compatible avec les besoins en synthèse d'image et notre expression de la variance.

Définition 7.3 *Soit f une fonction $\mathbb{D}_{\mathcal{E}} \to \mathbb{C}$, l'espace de fonctions que nous considérons est noté FuncClassDecays et, on dit que $f \in \Phi^s$ si*

$$\frac{\mathrm{d}E_f(\nu)}{\mathrm{d}\nu} \in O_{\nu\to\infty}\left(\frac{1}{\nu^{s+1}}\right), \tag{7.56}$$

où E_f est définie par

$$E_f(\nu) = \int_{R(\nu)} \mu_F\left(\mathbb{F}^r\right)\mathcal{P}_f(r)\mathrm{d}r\,, \tag{7.57}$$

et $R(\nu)$ est l'ensemble des fréquences r telles que $|r| < \nu$.

Dans cette définition, la dérivée permet de faire abstraction du spectre de puissance radial. Cela nous permet de traiter les espaces hyper-cubiques périodiques sans perte de généralité. Cependant, comme nous allons le voir dans le lemme suivant, la dérivée de E_f peut toujours être reliée au spectre de puissance radial.

Lemme 7.9 *Soit f une fonction $\mathbb{R}^d \to \mathbb{C}$ tel que $\check{\mathcal{P}}_f(\nu) \in O_{\nu\to\infty}\left(\nu^{-q}\right)$, alors $f \in \Phi^{d-q}$.*

Preuve 7.8 *Dans l'espace euclidien, on peut écrire l'intégrale du spectre de puissance en coordonnées polaires en utilisant le spectre de puissance radial,*

$$\int_{\mathbb{P}} \mathcal{P}_f(r)\mathrm{d}r = \int_{\mathbb{R}^+} \mu\left(\mathbb{D}_{\mathcal{S}^{d-1}}\right)\nu^{d-1}\check{\mathcal{P}}_f(\nu)\mathrm{d}\nu\,. \tag{7.58}$$

Le domaine d'intégration $R(\nu)$ correspond à toutes les fréquences qui ont une norme inférieure à ν. On peut alors écrire que

$$E_f(\nu) = \int_{R(\nu)} \mathcal{P}_f(r)\mathrm{d}r = \int_{[0,\nu[} \mu\left(\mathbb{D}_{\mathcal{S}^{d-1}}\right)\psi^{d-1}\check{\mathcal{P}}_f(\psi)\mathrm{d}\psi\,. \tag{7.59}$$

D'un autre coté, on peut aussi écrire que

$$E_f(\nu) = \int_{[0,\nu[} \frac{\mathrm{d}E_f(\psi)}{\mathrm{d}\psi}\mathrm{d}\psi\,. \tag{7.60}$$

En comparant ces deux équations, on en déduit que

$$\frac{\mathrm{d}E_f(\nu)}{\mathrm{d}\nu} = \mu\left(\mathbb{D}_{\mathcal{S}^{d-1}}\right)\nu^{d-1}\check{\mathcal{P}}_f(\nu)\,. \tag{7.61}$$

Comme $\check{\mathcal{P}}_f(\nu) \in O_{\nu\to\infty}\left(\nu^{-q}\right)$, on trouve que

$$\frac{\mathrm{d}E_f(\nu)}{\mathrm{d}\nu} \in O_{\nu\to\infty}\left(\frac{1}{\nu^{d-1-q}}\right), \tag{7.62}$$

ce qui correspond au critère pour qu'une fonction soit dans Φ^{d-q}. □

Ainsi, les classes de fonctions définies par Brandolini et al. [2001; 2003], où $q = d+1$, sont des sous-ensembles de la classe Φ^1. La figure (7.2) montre une illustration du spectre de puissance radial de l'intégrande en fonction du paramètre s.

Lemme 7.10 *Soit f une fonction $\mathbb{R}^d \to \mathbb{C}$ appartenant à Φ^s, alors*

$$\widetilde{E}_f(\nu) \in O_{\nu\to\infty}\left(\frac{1}{\nu^s}\right), \tag{7.63}$$

avec

$$\widetilde{E}_f(\nu) = \int_{\mathbb{P}\setminus R(\nu)} \mu_F\left(\mathbb{F}^r\right)\mathcal{P}_f(r)\mathrm{d}r\,. \tag{7.64}$$

Preuve 7.9 *D'après la définition de E_f, et en remarquant que la limite à l'infini de $R(\nu)$ vaut $\mathbb{P}$, on peut écrire que*

$$\int_{\mathbb{P}} \mu_F\left(\mathbb{F}^r\right)\mathcal{P}_f(r)\mathrm{d}r = \lim_{\nu\to\infty} E_f(\nu) = \int_{\mathbb{R}^+} \frac{\mathrm{d}E_f(\psi)}{\mathrm{d}\psi}\mathrm{d}\psi\,. \tag{7.65}$$

$\widetilde{E}_f$ peut alors s'écrire comme

$$\widetilde{E}_f(\nu) = \lim_{\nu\to\infty} E_f(\nu) - E_f(\nu) = \int_{[\nu,\infty[} \frac{\mathrm{d}E_f(\psi)}{\mathrm{d}\psi}\mathrm{d}\psi \tag{7.66}$$

Comme f appartient à Φ^s, on sait que la dérivée de E_f est en $O_{\nu\to\infty}\left(\nu^{-s-1}\right)$. On obtient alors

$$\begin{aligned}\widetilde{E}_f(\nu) &\in \int_{[\nu,\infty[} O_{\psi\to\infty}\left(\frac{1}{\psi^{s+1}}\right)\mathrm{d}\psi \\ &\in O_{\nu\to\infty}\left(\int_{[\nu,\infty[} \frac{1}{\psi^{s+1}}\mathrm{d}\psi\right) \\ &\in O_{\nu\to\infty}\left(\frac{1}{\nu^s}\right)\,.\end{aligned} \tag{7.67}$$

□

Comme on le voit, ces classes de fonctions vont nous permettre de déterminer le comportement asymptotique du deuxième terme de l'équation Eq. (7.55).

7.7.5 Vitesse de convergence par classe de fonctions

Maintenant que nous avons défini un ensemble de classes de fonctions et énoncé leurs propriétés dans le domaine spectral, revenons à l'équation Eq. (7.55). Pour simplifier cette équation, nous proposons de modéliser la fonction L par un polynôme de degré b tel que $L(r) = |r|^b$. Dans le chapitre suivant, nous montrerons que cette modélisation est valide pour une grande partie des méthodes d'échantillonnage stochastiques. La figure (7.1) montre une illustration du profile spectral de l'ensemble de points en fonction du paramètre b.

Lemme 7.11 *Soit f une fonction $\mathbb{D}_{\mathcal{E}} \to \mathbb{C}$ tel que $f \in \Phi^s$, et $\mathcal{S}_\rho$ un ensemble de points homogène tel que le profil spectral de $\mathcal{S}_\rho$ est en $O_{r\to r_0}\left(|r|^b\right)$, alors*

$$Var(I_{\mathcal{S}_\rho,f}) \in O_{\rho\to\infty}\left(\rho^{-\frac{d+b}{d}} + \rho^{-\frac{d+s}{d}}\right)\,. \tag{7.68}$$

Preuve 7.10 *À partir de Eq. (7.55) et sachant que nous avons fixé $L(r) = |r|^b$, on peut écrire que*

$$Var(I_{\mathcal{S}_\rho,f}) \leq \left(\frac{M_{\mathcal{S}}}{\rho\sqrt[d]{\rho}^b}\int_{R(\sqrt[d]{\rho}\nu_1)} \mu_F\left(\mathbb{F}^r\right)|r|^b\mathcal{P}_f(r)\mathrm{d}r\right) + \left(\frac{M'_{\mathcal{S}}}{\rho}\int_{\mathbb{P}\setminus R(\sqrt[d]{\rho}\nu_1)} \mu_F\left(\mathbb{F}^r\right)\mathcal{P}_f(r)\mathrm{d}r\right)\,. \tag{7.69}$$

D'après Lemme (7.10), on sait que la deuxième intégrale est en $O_{\rho\to\infty}\left(\rho^{-(d+s)/d}\right)$. Il nous reste donc à trouver le comportement asymptotique de la première intégrale. Nous savons déjà, d'après Def. (7.3), que la dérivée de E_f est en $O_{\rho\to\infty}\left(\rho^{-1-s/d}\right)$. Comme c'est une notation en

grand O, on pose ν_f la constante à partir de laquelle la notation en grand O est valide. On peut alors couper la première intégrale en deux pour obtenir

$$\frac{M_{\mathcal{S}}}{\rho\sqrt[d]{\rho}^b}\int_{R(\sqrt[d]{\rho}\nu_1)}\mu_F\left(\mathbb{F}^r\right)|r|^b\mathcal{P}_f(r)\mathrm{d}r = \left(\frac{M_{\mathcal{S}}}{\rho\sqrt[d]{\rho}^b}\int_{R(\nu_f)}\mu_F\left(\mathbb{F}^r\right)|r|^b\mathcal{P}_f(r)\mathrm{d}r\right) + \left(\frac{M_{\mathcal{S}}}{\rho\sqrt[d]{\rho}^b}\int_{R(\sqrt[d]{\rho}\nu_1)\setminus R(\nu_f)}\mu_F\left(\mathbb{F}^r\right)|r|^b\mathcal{P}_f(r)\mathrm{d}r\right). \tag{7.70}$$

La première partie est en $O_{\rho\to\infty}\left(\rho^{-(d+b)/d}\right)$ car l'intégrale est indépendante de ρ. Pour la partie restante, sachant que $|r|^b$ ne dépend que de la norme de r, on peut utiliser la dérivée de E_f pour réécrire l'intégrale de la façon suivante

$$\frac{M_{\mathcal{S}}}{\rho\sqrt[d]{\rho}^b}\int_{R(\sqrt[d]{\rho}r_1)\setminus R(\nu_f)}\mu_F\left(\mathbb{F}^r\right)|r|^b\mathcal{P}_f(r)\mathrm{d}r = \frac{M_{\mathcal{S}}}{\rho\sqrt[d]{\rho}^b}\int_{[r_f,\sqrt[d]{\rho}\nu_1[}\nu^b\frac{\mathrm{d}E_f(\nu)}{\mathrm{d}\nu}\mathrm{d}\nu\,. \tag{7.71}$$

On appelle M_f la constante associée à la notation en grand O de la dérivée de E_f. On obtient alors l'inégalité suivante,

$$\frac{M_{\mathcal{S}}}{\rho\sqrt[d]{\rho}^b}\int_{R(\sqrt[d]{\rho}r_1)\setminus R(\nu_f)}\mu_F\left(\mathbb{F}^r\right)|r|^b\mathcal{P}_f(r)\mathrm{d}r \leq \frac{M_{\mathcal{S}}M_f}{\rho\sqrt[d]{\rho}^b}\int_{[\nu_f,\sqrt[d]{\rho}\nu_1[}\frac{\nu^b}{\nu^{s+1}}\mathrm{d}\nu\,. \tag{7.72}$$

Cette intégrale est simple à résoudre et on obtient finalement

$$\frac{M_{\mathcal{S}}}{\rho\sqrt[d]{\rho}^b}\int_{R(\sqrt[d]{\rho}r_1)\setminus R(\nu_f)}\mu_F\left(\mathbb{F}^r\right)|r|^b\mathcal{P}_f(r)\mathrm{d}r \leq \frac{M_{\mathcal{S}}M_f\nu_1^b}{\rho\sqrt[d]{\rho}^s\nu_1^s} - \frac{M_{\mathcal{S}}M_f\nu_f^b}{\rho\sqrt[d]{\rho}^b\nu_f^s}\,, \tag{7.73}$$

On remarquant que cette expression est en $O_{\rho\to\infty}\left(\rho^{-(d+b)/d}+\rho^{-(d+s)/d}\right)$, et en réutilisant les notations en grand O que nous avons trouvées pour les autres parties de Eq. (7.69), on trouve que

$$\begin{aligned} Var(I_{\mathcal{S}_\rho,f}) &\in O_{\rho\to\infty}\left(\rho^{-(d+b)/d}\right) + O_{\rho\to\infty}\left(\rho^{-(d+b)/d}+\rho^{-(d+s)/d}\right) + O_{\rho\to\infty}\left(\rho^{-(d+s)/d}\right) \\ &\in O_{\rho\to\infty}\left(\rho^{-(d+b)/d}+\rho^{-(d+s)/d}\right)\,. \end{aligned} \tag{7.74}$$

□

Comme on le voit, la vitesse de convergence dépend de la classe de fonctions étudiée et de la forme du spectre de puissance du profil spectral de l'ensemble de points. Lorsque $b > s$, la variance est en $O_{\rho\to\infty}\left(\rho^{-\frac{d+s}{d}}\right)$. À l'inverse, lorsque $b < s$, la variance est en $O_{\rho\to\infty}\left(\rho^{-b/d}\right)$. La figure (7.3) illustre comment la variance diminue en fonction de la densité et du spectre de puissance de l'ensemble de points et de l'intégrande. Le tableau (7.1) montre la vitesse de convergence en fonction des paramètre s et b.

Ce tableau se lit de deux façons. Il montre que si $f \in \Phi^s$, alors le profil spectral de l'ensemble de points doit avoir au minimum une vitesse de convergence en $O_{r\to r_0}\left(|r|^s\right)$. De la même façon, si le spectre de puissance radial de l'ensemble de points a une vitesse de convergence en $O_{r\to r_0}\left(|r|^b\right)$, alors il ne sera pas capable d'intégrer de façon optimale les intégrandes de la classe $\Phi^s, s > b$. Le tableau (7.2) montre

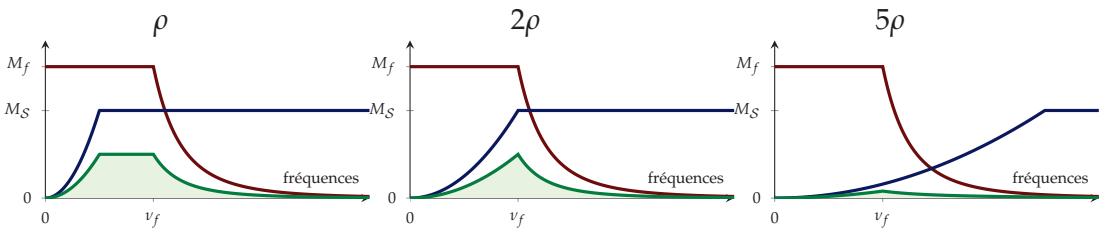


FIGURE 7.3 – *Évolution du produit de deux spectres de puissance en fonction de la densité de l'ensemble de points.* ***Bleu :*** *Le spectre de puissance de l'ensemble de points.* ***Rouge :*** *Le spectre de puissance de l'intégrande.* ***Vert :*** *Le produit des deux spectres. L'augmentation de la densité entraîne une dilatation du spectre de puissance de l'ensemble de points. On remarque que l'intégrale du produit diminue à mesure que la densité augmente, ce qui explique la diminution de la variance quand la densité augmente.*

	$b=0$	$b=1/2$	$b=1$	$b=2$	$b\to\infty$
$s\to 0$	$O\left(\rho^{-1}\right)$	$O\left(\rho^{-1}\right)$	$O\left(\rho^{-1}\right)$	$O\left(\rho^{-1}\right)$	$O\left(\rho^{-1}\right)$
$s=1/2$	$O\left(\rho^{-1}\right)$	$O\left(\rho^{-\frac{d+1/2}{d}}\right)$	$O\left(\rho^{-\frac{d+1/2}{d}}\right)$	$O\left(\rho^{-\frac{d+1/2}{d}}\right)$	$O\left(\rho^{-\frac{d+1/2}{d}}\right)$
$s=1$	$O\left(\rho^{-1}\right)$	$O\left(\rho^{-\frac{d+1/2}{d}}\right)$	$O\left(\rho^{-\frac{d+1}{d}}\right)$	$O\left(\rho^{-\frac{d+1}{d}}\right)$	$O\left(\rho^{-\frac{d+1}{d}}\right)$
$s=2$	$O\left(\rho^{-1}\right)$	$O\left(\rho^{-\frac{d+1/2}{d}}\right)$	$O\left(\rho^{-\frac{d+1}{d}}\right)$	$O\left(\rho^{-\frac{d+2}{d}}\right)$	$O\left(\rho^{-\frac{d+2}{d}}\right)$
$s\to\infty$	$O\left(\rho^{-1}\right)$	$O\left(\rho^{-\frac{d+1/2}{d}}\right)$	$O\left(\rho^{-\frac{d+1}{d}}\right)$	$O\left(\rho^{-\frac{d+2}{d}}\right)$	$O\left(0\right)$

TABLE 7.1 – *Évolution de la vitesse de convergence de la variance de l'intégration Monte Carlo en fonction des paramètre s et b. La notation en grand O correspond ici au cas où $\rho\to\infty$. La couleur rouge représente les vitesses de convergence lentes et la couleur verte les vitesses de convergence rapides.*

	$b=0$	$b=1/2$	$b=1$	$b=2$	$b\to\infty$
$s\to 0$	$O\left(N^{-1}\right)$	$O\left(N^{-1}\right)$	$O\left(N^{-1}\right)$	$O\left(N^{-1}\right)$	$O\left(N^{-1}\right)$
$s=1/2$	$O\left(N^{-1}\right)$	$O\left(N^{-\frac{d+1/2}{d}}\right)$	$O\left(N^{-\frac{d+1/2}{d}}\right)$	$O\left(N^{-\frac{d+1/2}{d}}\right)$	$O\left(N^{-\frac{d+1/2}{d}}\right)$
$s=1$	$O\left(N^{-1}\right)$	$O\left(N^{-\frac{d+1/2}{d}}\right)$	$O\left(N^{-\frac{d+1}{d}}\right)$	$O\left(N^{-\frac{d+1}{d}}\right)$	$O\left(N^{-\frac{d+1}{d}}\right)$
$s=2$	$O\left(N^{-1}\right)$	$O\left(N^{-\frac{d+1/2}{d}}\right)$	$O\left(N^{-\frac{d+1}{d}}\right)$	$O\left(N^{-\frac{d+2}{d}}\right)$	$O\left(N^{-\frac{d+2}{d}}\right)$
$s\to\infty$	$O\left(N^{-1}\right)$	$O\left(N^{-\frac{d+1/2}{d}}\right)$	$O\left(N^{-\frac{d+1}{d}}\right)$	$O\left(N^{-\frac{d+2}{d}}\right)$	$O\left(0\right)$

TABLE 7.2 – *Évolution de la vitesse de convergence de la variance de l'intégration Monte Carlo en fonction des paramètre s et b. La notation en grand O correspond ici au cas où $N\to\infty$. La couleur rouge représente les vitesses de convergence lentes et la couleur verte les vitesses de convergence rapides.*

le même comportement, mais cette fois en fonction du nombre de points N. On voit donc que les ensembles des points ayant un coefficient b élevé seront plus efficaces que les ensembles de points avec un b faible. Par exemple, le profile constant ($b = 0$), n'a une vitesse de convergence pour la variance qu'en $O_{\rho\to\infty}\left(\rho^{-1}\right)$, alors qu'à l'autre extrême, le profile discontinu ($b \to \infty$) peut réussir à reconstruire parfaitement les intégrandes dont le spectre de puissance a un support fini ($s \to \infty$).

Conclusion

Dans ce chapitre, nous avons donné une expression du biais et de la variance de l'intégration Monte Carlo dans le domaine spatial et le domaine spectral. Lorsque l'ensemble de points est homogène, l'expression de la variance prend une forme simple à l'aide de la fonction de distribution différentielle et du spectre de puissance. Ces expressions permettent, lorsque l'intégrande et la méthode d'échantillonnage sont connus, de quantifier exactement la variance. Lorsque l'intégrande n'est pas connue, mais que l'on connaît certaines informations sur son spectre de puissance, il est possible de borner la variance. Nous avons montré que le comportement du spectre de puissance de l'ensemble de points dans les basses fréquences suffit à déterminer la vitesse de convergence de la méthode d'échantillonnage. Ce travail est une généralisation de ce qui a été présenté dans [Pilleboue *et al.*, 2015]. Dans le chapitre suivant, nous montrerons des exemples concrets où des différences subtiles dans les basses fréquences se traduisent de façon claire en terme de réduction de variance.

Analyse de l'état de l'art 8

Nous abordons ici le dernier chapitre de ce manuscrit. Dans le chapitre 4, nous avons parlé des méthodes d'échantillonnage développées jusqu'à maintenant pour diminuer l'erreur d'intégration. Puis nous avons introduit des méthodes d'analyse spatiale et spectrale pour finalement les relier à la variance de l'intégration Monte Carlo. Il est maintenant temps de confronter l'état de l'art à toutes ces méthodes d'analyse et de valider les déductions théoriques du chapitre 7 de façon expérimentale.

Nous avons sélectionné les méthodes d'échantillonnage stochastiques : bruit blanc (voir page 29), disques de Poisson (voir page 30), *jittered sampling* (voir page 34), ainsi que les méthodes d'optimisation *Blue Noise through Optimal Transport* (BNOT, voir page 38), *Farthest-Point Optimization* (FPO, voir page 39) et une méthode d'optimisation par contrôle spectral (voir page 39). Nous avons ajouté également à cette étude la séquence à basse discrépance de Halton (voir page 42).

Pour chacune de ces méthodes, nous avons calculé la fonction de distribution de capacité CDF_S (voir page 55), la fonction de distribution différentielle $\check{g}_S$ (voir page 50), la fonction de distribution de distance $\check{g}_S$ (voir page 50), le spectre de puissance $\mathcal{P}_S$ (voir chapitre 6) et sa version radiale $\check{\mathcal{P}}_S$.

8.1 Méthodologie expérimentale

Notre analyse de l'état de l'art porte sur plusieurs méthodes d'échantillonnage stochastiques homogénéisés. Pour l'homogénéisation, nous avons utilisé la méthode décrite dans Lemme (3.3, p. 28). Nous avons choisi comme espace d'échantillonnage, l'espace carré périodique $\mathcal{T}^2$. De cette façon, les ensembles de points sont simples à visualiser, tout comme la fonction de distribution différentielle et le spectre de puissance, qui restent ainsi des fonctions de dimension deux. Pour une analyse similaire à celle présentée dans ce chapitre, mais dans le domaine sphérique $\mathcal{S}^2$, voir [Singh, 2015].

8.1.1 Catalogue d'ensembles de points

Les ensembles de points que nous avons étudiés sont majoritairement stochastiques. Comme nous l'avons dit dans le chapitre 3, les ensembles de points stochastiques sont des variables aléatoires. À chaque fois que l'échantillonneur est exécuté, une nouvelle réalisation de l'ensemble de points est générée. Pour approximer l'espérance d'une fonction de l'ensemble de points, il est nécessaire d'utiliser une approche empirique en générant un grand nombre de réalisations. Nous avons observé qu'une

moyenne empirique utilisant 1000 réalisations produisait généralement des résultats satisfaisants. Pour chaque méthode, nous avons donc généré un catalogue d'un millier d'ensembles de points, contenant chacun 4096 points. Le choix de $\rho = 4096$ repose sur un compromis de plusieurs critères. Premièrement, le nombre de points doit être le carré d'un nombre entier, pour générer facilement des ensembles de points *jittered*. Ensuite, il doit être assez grand pour que les propriétés statistiques soient suffisamment proches du cas des ensembles de points uniforme dans l'espace euclidien. Pour finir, il doit être d'une taille acceptable afin d'éviter des dépassements de mémoire et des temps de calcul prohibitifs lors de la génération des ensembles de points.

8.1.2 Fonction de distribution différentielle

Pour l'analyse spatiale, nous nous sommes concentrés sur la fonction de distribution différentielle et la fonction de distribution de distance. Ces outils étant invariants par action, ils permettent d'extraire facilement les caractéristiques importantes des ensembles de points homogènes. La fonction de distribution différentielle est calculée en utilisant un histogramme, où chaque classe de l'histogramme représente un sous-ensemble du groupe d'actions $\mathbb{T}_{\mathcal{E}}$. De la même façon, la fonction de distribution de distance est calculée en utilisant un histogramme où chaque classe représente un intervalle de distance.

8.1.3 Spectre de puissance

La méthodologie utilisée est la même que celle présentée dans [Schlömer and Deussen, 2010]. L'ensemble de points est discrétisé sur une grille de taille $(32\sqrt{N}) \times (32\sqrt{N})$, la valeur de chaque case est égale au nombre de points qui s'y trouvent. La transformée de Fourier est calculée en utilisant la bibliothèque FFTW et le spectre de puissance en utilisant Eq. (6.9, p. 63). L'espérance du spectre de puissance sur l'ensemble des réalisations est ensuite normalisée de la façon suivante :

— l'axe des fréquences est divisé par $\sqrt[d]{N}$ de façon à ce que la fréquence 1 corresponde à la fréquence de la grille régulière,
— l'axe des puissances est divisé par N de façon à ce que le spectre de puissance converge vers 1 dans le cas des ensembles de points stochastiques.

Il est important de noter que la normalisation de l'axe des fréquences permet de simuler une densité $\rho_{\mathcal{S}}$ (voir page 45) égale à 1. Nous avons retenu l'intervalle de fréquences normalisées $[-4, 4]^2$ comme objet d'analyse, les basses fréquences étant l'élément qui nous a semblé le plus important. Toutefois, la taille de cette fenêtre de visualisation est suffisamment grande pour permettre de voir un certain nombre d'oscillations.

8.1.4 Intégrandes

Afin de tester expérimentalement la vitesse de convergence de l'intégration Monte Carlo, nous avons sélectionné deux types d'intégrandes reflétant différents cas d'intégration. Ces intégrandes sont toutes définies dans l'espace carré périodique.

Fonction gaussienne Nous avons utilisé une intégrande de type gaussien définie par

$$f_G(x) = 8e^{-32\pi\|x\|^2}. \tag{8.1}$$

Le spectre de puissance radial de cette fonction est donné par

$$\check{\mathcal{P}}_{f_G}(\nu) = \frac{1}{16} e^{-\frac{\pi\nu^2}{16}} . \tag{8.2}$$

Cette fonction a l'avantage d'avoir un spectre de puissance négligeable à partir d'une certaine fréquence. Son spectre de puissance radial est en $O_{\nu\to\infty}\left(e^{-\nu^2}\right)$. On peut donc dire que cette intégrande est simple à intégrer.

Fonction caractéristique d'un disque Le deuxième type d'intégrande que nous avons testée est un disque défini par

$$f_D(x) = \begin{cases} \frac{4}{\sqrt{\pi}} & \text{si } \|x\| < \frac{1}{4} \\ 0 & \text{sinon} . \end{cases} \tag{8.3}$$

Son spectre de puissance radial est donné par

$$\check{\mathcal{P}}_{f_D}(\nu) = \frac{J_1(\frac{\pi}{2}\nu)^2}{\pi\nu^2} , \tag{8.4}$$

où J_1 est la fonction de Bessel de première espèce et d'ordre 1 [Chow, 2000, eq. 7.76]. Puisque la fonction $J_1(\nu)$ est en $O_{\nu\to\infty}\left(\nu^{-1/2}\right)$ [Olenko, 2006], le spectre de puissance radial de f_D est en $O_{\nu\to\infty}\left(\nu^{-3}\right)$. Elle appartient donc à la classe de fonctions Φ^1, définie dans Def. (7.3, p. 94). Cette fonction est intéressante à étudier car elle contient des discontinuités, phénomène que l'on rencontre dans presque toutes les scènes en rendu d'image. Une autre choix d'intégrande aurait été de prendre un carré ou un triangle au lieu d'un disque. Mais ces deux fonctions ont un spectre de puissance radial qui converge plus rapidement que celui du disque et sont donc plus simple à intégrer. Or, nous souhaitons tester un cas d'intégration difficile, par opposition à la fonction gaussienne.

Maintenant que nous avons décrit l'environnement expérimental, nous pressentons nos résultats expérimentaux pour chaque méthode d'échantillonnage. Ces résultats sont compilés à la fin de ce chapitre. Les figures (8.1), (8.3), (8.5), (8.7), (8.9), (8.11) et (8.13) présentent les résultats de l'analyse des ensembles de points, et les figures (8.2), (8.4), (8.6), (8.8), (8.10) et (8.12) présentent les résultats de l'intégration Monte Carlo des intégrandes que nous avons décrites. En plus de ces résultats, nous présentons, quand cela est possible, une expression formelle de la fonction de distribution différentielle et du spectre de puissance.

8.2 Bruit blanc

La première méthode d'échantillonnage que nous avons étudiée est le bruit blanc. Nous avons pu déterminer une expression exacte de la fonction de distribution différentielle et du spectre de puissance du bruit blanc. Ces résultats sont déjà connus dans la communauté scientifique, mais nous les re-démontrons ici à l'aide de notre formalisme pour tous les espaces d'échantillonnage.

8.2.1 Fonction de distribution différentielle

Lemme 8.1 *Soit $\mathcal{S} \in \mathfrak{M}_{WN}$ un ensemble de points bruit blanc dans l'espace $\mathcal{E}$, la fonction de distribution de paires est donnée par*

$$g_{\mathcal{S}}(x, x') = \frac{N-1}{N} + \frac{\delta(x, x')}{\rho_{\mathcal{S}}} . \tag{8.5}$$

Corollaire 8.1 *Soit $\mathcal{S} \in \mathfrak{M}_{WN}$ un ensemble de points bruit blanc dans l'espace $\mathcal{E}$, les fonctions de distribution différentielle et distribution de distance sont données par*

$$\begin{aligned} \check{g}_{\mathcal{S}}(\tau) &= \frac{N-1}{N} + \frac{\delta(\tau, \tau_0)}{\rho_{\mathcal{S}}} \\ \check{g}_{\mathcal{S}}(r) &= \frac{N-1}{N} + \frac{\delta(r, 0)}{\rho_{\mathcal{S}}} . \end{aligned} \tag{8.6}$$

Preuve 8.1 *A partir de la définition de la densité de paires, Def. (5.2, p. 47), on sépare le cas où $i = j$ du reste.*

$$\rho_{\mathcal{S}}^{(2)}(x, x') = \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \left\langle \delta(x, \boldsymbol{s}_i) \delta(x', \boldsymbol{s}_j) \right\rangle + \sum_{i=1}^{N} \left\langle \delta(x, \boldsymbol{s}_i) \delta(x', \boldsymbol{s}_i) \right\rangle . \tag{8.7}$$

La première partie de l'équation se simplifie en remarquant que lorsque $i \neq j$, les variables aléatoires $\boldsymbol{s}_i$ et $\boldsymbol{s}_j$ sont indépendantes. La deuxième partie se simplifie à l'aide de Eq. (5.7, p. 47). On obtient alors

$$\rho_{\mathcal{S}}^{(2)}(x, x') = \sum_{i=1}^{N} \langle \delta(x, \boldsymbol{s}_i) \rangle \sum_{\substack{j=1 \\ j \neq i}}^{N} \left\langle \delta(x', \boldsymbol{s}_j) \right\rangle + \delta(x', x) \rho_{\mathcal{S}}^{(1)}(x) , \tag{8.8}$$

où $\rho_{\mathcal{S}}^{(1)}$ est la densité locale de $\mathcal{S}$, définie par Eq. (5.2, p. 46). Les deux espérances sont égales à la définition de la fonction de densité de probabilité des points d'échantillonnage, Eq. (3.5, p. 18). Or, comme l'ensemble de points est homogène, on sait d'après Eq. (3.26, p. 27) que la fonction de densité de probabilité est égale à l'inverse de la mesure du domaine d'échantillonnage. On sait également pour les ensembles de points homogènes que la densité locale est égale à la densité globale. En conséquence, on obtient

$$\begin{aligned} \rho_{\mathcal{S}}^{(2)}(x, x') &= \sum_{i=1}^{N} \frac{1}{\mu_{\mathcal{E}}(\mathbb{D}_{\mathcal{E}})} \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{1}{\mu_{\mathcal{E}}(\mathbb{D}_{\mathcal{E}})} + \delta(x', x) \rho_{\mathcal{S}}^{(1)}(x) \\ &= \frac{N(N-1)}{\mu_{\mathcal{E}}(\mathbb{D}_{\mathcal{E}})^2} + \delta(x', x) \rho_{\mathcal{S}} . \end{aligned} \tag{8.9}$$

□

Nos résultats expérimentaux sont en accord avec ces résultats théoriques, puisque l'on constate, dans la figure (8.1,f) que la fonction de distribution différentielle est constante sur tout le domaine d'étude. Il est intéressant de noter également que lorsque le nombre de points tend vers l'infini, la fonction de distribution de paires, et donc la fonction de distribution différentielle, tend vers 1.

8.2.2 Spectre de puissance

Lemme 8.2 *Soit $\mathcal{S} \in \mathfrak{M}_{WN}$ un ensemble de points bruit blanc dans l'espace $\mathcal{E}$, le spectre de puissance moyen est donné par*

$$\langle \mathcal{P}_S(r) \rangle = \delta(r, r_0) N(N-1) + \frac{\rho_{\mathcal{S}}}{\mu_F(\mathbb{F}^r)} \int_{\mathbb{F}^r} \|b^\omega\|^2 \mathrm{d}\omega \,. \tag{8.10}$$

Preuve 8.2 *D'après Eq. (6.46, p. 73), on a l'égalité suivante*

$$\left\langle \mathcal{F}_S(\omega)\overline{\mathcal{F}_S(\omega)} \right\rangle = \sum_{i=1}^{N}\sum_{j=1}^{N} \left\langle \overline{b^\omega(\boldsymbol{s}_i)} b^\omega(\boldsymbol{s}_j) \right\rangle \,. \tag{8.11}$$

On peut séparer le cas où $i \neq j$ du reste et simplifier l'espérance dans ce cas là puisque les variables aléatoires deviennent indépendantes

$$\left\langle \mathcal{F}_S(\omega)\overline{\mathcal{F}_S(\omega)} \right\rangle = \sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N} \left\langle \overline{b^\omega(\boldsymbol{s}_i)} \right\rangle \left\langle b^\omega(\boldsymbol{s}_j) \right\rangle + \sum_{i=1}^{N} \left\langle \overline{b^\omega(\boldsymbol{s}_i)} b^\omega(\boldsymbol{s}_i) \right\rangle \,. \tag{8.12}$$

Comme $\boldsymbol{s}_i$ et $\boldsymbol{s}_j$ sont des variables aléatoires uniformément réparties sur le domaine d'échantillonnage, l'espérance devient une intégrale sur le domaine d'échantillonnage.

$$\begin{aligned}\left\langle \mathcal{F}_S(\omega)\overline{\mathcal{F}_S(\omega)} \right\rangle =& \frac{N(N-1)}{\mu_{\mathcal{E}}(\mathbb{D}_{\mathcal{E}})^2} \left(\int_{\mathbb{D}_{\mathcal{E}}} \overline{b^\omega(x)} \mathrm{d}x \right) \left(\int_{\mathbb{D}_{\mathcal{E}}} b^\omega(x) \mathrm{d}x \right) \\ &+ \frac{N}{\mu_{\mathcal{E}}(\mathbb{D}_{\mathcal{E}})} \int_{\mathbb{D}_{\mathcal{E}}} \overline{b^\omega(x)} b^\omega(x) \mathrm{d}x \,. \end{aligned} \tag{8.13}$$

On sait que l'intégrale des vecteurs de la base est égale à zéro, sauf pour la fréquence ω_0 où elle est égale à $\mu_{\mathcal{E}}(\mathbb{D}_{\mathcal{E}})$. En remarquant que la deuxième partie correspond à la norme au carré de b^ω, on obtient

$$\left\langle \mathcal{F}_S(\omega)\overline{\mathcal{F}_S(\omega)} \right\rangle = \delta(\omega, \omega_0) N(N-1) + \rho_{\mathcal{S}} \|b^\omega\|^2 \,. \tag{8.14}$$

La preuve est terminée en utilisant la définition du spectre de puissance et en remarquant que $\mu_F(\mathbb{F}^r) = 1$. □

Dans le cas de l'espace hyper-cubique périodique, où la norme des vecteurs de la base $\mathcal{B}$ est toujours égale à 1, le spectre de puissance est tout simplement constant et égal à 1, à l'exception de la fréquence r_0. C'est d'ailleurs de cette propriété que vient le nom de "bruit blanc" par analogie avec les spectres lumineux où un tel spectre est visible sous la forme de lumière blanche. La figure (8.1,h) montre le spectre de puissance obtenu expérimentalement pour un ensemble de points bruit blanc.

8.2.3 Variance de l'intégration Monté Carlo

Le spectre de puissance du bruit blanc étant constant, on en déduit que

$$\mathcal{P}_S(r) \in O_{r\to 0}(1) \,, \tag{8.15}$$

et que l'on peut utiliser Eq. (7.11, p. 95) avec $b = 0$. On obtient alors une vitesse de convergence pour la variance en $O_{\rho\to\infty}\left(\rho^{-1} + \rho^{-\frac{d+s}{d}}\right)$. Dans le cas de la fonction disque, où l'on a $s = 1$, la vitesse de convergence est en $O_{\rho\to\infty}\left(\rho^{-1}\right)$, tout comme dans

le cas de la fonction gaussienne. Ce résultat est connu, en particulier pour l'espace carré $\mathcal{U}^2$ [Lemieux, 2009, chap. 1], sous la forme

$$\sqrt{Var(I_{\mathcal{S},f})} \in O_{N\to\infty}\left(\frac{1}{\sqrt{N}}\right) . \tag{8.16}$$

8.3 Disques de Poisson

La deuxième méthode d'échantillonnage que nous avons analysé est la méthode des disques de Poisson. Contrairement au bruit blanc, nous n'avons pas pu obtenir une expression exacte de la fonction de distribution différentielle et du spectre de puissance. Cependant, nous avons pu déterminer, sur la base d'observation dans le domaine spectral, la vitesse de convergence de la variance dans l'intégration Monte Carlo.

8.3.1 Fonction de distribution différentielle

Les ensembles de disques de Poisson ont été largement étudiés par la communauté. La recherche de la fonction de distribution différentielle a fait l'objet d'une attention particulière. Dans l'espace euclidien mono-dimensionnel, une expression formelle existe [Bonnier *et al.*, 1994]. Mais pour les dimensions supérieures, il est nécessaire d'avoir recours à des résultats expérimentaux. La figure (8.3,f et g) montre la fonction de distribution différentielle et la fonction de distribution de distance dans l'espace euclidien $\mathcal{E}^2$ pour un ensemble de 4096 points. On remarque que la distance minimale entre deux points est d'environ 0.816, ce qui correspond bien à deux fois la taille des disques utilisés.

8.3.2 Spectre de puissance

Le spectre de puissance des ensembles de disques de Poisson pour des dimensions supérieures à 1 est lui aussi inconnu. Des études expérimentales ont été menées pour en calculer une approximation [Torquato *et al.*, 2006]. En particulier, le spectre de puissance des ensembles de disques de Poisson est connu pour être borné en bas par une constante d'environ 1/20 [Bonnier *et al.*, 1994; Torquato *et al.*, 2006]. La figure (8.3,i) montre des résultats similaires effectués dans notre environnement de test. On observe que le spectre converge vers cette constante quand la fréquence tend vers zéro. La plus grande valeur du spectre de puissance correspond à la fréquence $r = 1.125$ et à une valeur de 1.744. Le spectre est donc borné en haut et en bas par des fonctions constantes, ce qui permet de dire que

$$\mathcal{P}_S(r) \in O_{r\to 0}(1) . \tag{8.17}$$

8.3.3 Variance de l'intégration Monte Carlo

Comme le spectre de puissance est en $O_{r\to 0}(1)$, la vitesse de convergence de la variance est la même que celle du bruit blanc : $O_{\rho\to\infty}\left(\rho^{-1} + \rho^{-\frac{d+s}{d}}\right)$. Ce résultat est intéressant car il montre que même si la méthode d'échantillonnage est plus sophistiquée, la vitesse de converge ne s'en trouve pas nécessairement améliorée. Le fait d'imposer que les points ne soient pas proches les uns des autres ne permet pas d'agir sur la

vitesse de convergence. En comparant les figures (8.2,e et f) et (8.4,e et f), on observe tout de même une différence entre le bruit blanc et les disques de Poisson. L'ordre de grandeur entre ces deux méthodes n'est pas le même. Mais pour un grand nombre d'échantillons, il vaudra mieux privilégier d'autres méthodes d'échantillonnage.

8.4 Échantillonnage *jittered*

Pour les ensembles de points *jittered*, nous montrons qu'il est possible de trouver une expression formelle de la densité locale, de la fonction de distribution différentielle et du spectre de puissance. En conséquence, nous avons pu déterminer exactement la vitesse de convergence de cette méthode, dans sa version homogénéisée.

8.4.1 Densité locale

Lemme 8.3 *Soit $\mathcal{S} \in \mathfrak{M}_{Jittered}$ un ensemble de points* jittered *dans l'espace $\mathcal{E}$, la densité locale est égale à la densité globale*

$$\rho_{\mathcal{S}}^{(1)}(x) = \rho_{\mathcal{S}} \,. \tag{8.18}$$

Preuve 8.3 *Comme l'espace est partitionné à l'aide de tuiles, chaque point du domaine appartient à une et une seule tuile. On appelle $\hat{\sigma}_k$ la tuile contenant le point x et $\mathbf{s}_k$ le point de $\mathcal{S}$ attaché à cette tuile. La densité locale est donc donnée par*

$$\rho_{\mathcal{S}}^{(1)}(x) = PDF_{\mathbf{s}_k}(x) \,, \tag{8.19}$$

les autres points ne pouvant pas se trouver dans la tuile $\hat{\sigma}_k$. Comme chaque point est distribué uniformement dans chaque tuile, et que la mesure de chaque tuile est égale à $\mu_{\mathcal{E}}(\mathbb{D}_{\mathcal{E}})/N$, on a

$$\rho_{\mathcal{S}}^{(1)}(x) = \frac{1}{\mu_{\mathcal{E}}(\hat{\sigma}_k)} = \frac{N}{\mu_{\mathcal{E}}(\mathbb{D}_{\mathcal{E}})} = \rho_{\mathcal{S}} \,. \tag{8.20}$$

□

Ainsi, les ensembles de points *jittered* sont des exemples d'ensembles de points non-homogènes, mais qui ont une densité locale égale à leur densité globale.

8.4.2 Fonction de distribution différentielle

Lemme 8.4 *Soit $\mathcal{S} \in \mathfrak{M}_{Jittered}$ un ensemble de points* jittered *dans l'espace $\mathcal{E}$, la fonction de distribution différentielle est donnée par la relation suivante*

$$\check{g}_{\mathcal{S}}(\tau) = 1 - \frac{1}{N}\sum_{i=1}^{N}(\chi_{\hat{\sigma}_i} \oplus \chi_{\hat{\sigma}_i})(\tau) + \frac{1}{\rho_{\mathcal{S}}}\delta(\tau, \tau_0) \,, \tag{8.21}$$

où τ_0 est l'action neutre du groupe des actions $\mathbb{T}_{\mathcal{E}}$, définie dans le chapitre 3.

Preuve 8.4 *Puisque chaque point dans les ensembles de points* jittered *est indépendant des autres, ont peut reprendre l'équation Eq. (8.8) et remplacer l'espérance par la fonction de densité*

de probabilité de chaque point donnée en Eq. (4.9, p. 34). On obtient alors

$$\rho_{\mathcal{S}}^{(2)}(x,x') = \sum_{i=1}^{N}\left(\frac{\chi_{\hat{\sigma}_i}(x)}{\mu_{\mathcal{E}}\left(\hat{\sigma}_i\right)}\sum_{\substack{j=1\\ j\neq i}}^{N}\frac{\chi_{\hat{\sigma}_j}(x')}{\mu_{\mathcal{E}}\left(\hat{\sigma}_j\right)}\right) + \delta(x',x)\rho_{\mathcal{S}}$$
$$= \rho_{\mathcal{S}}^{2}\sum_{i=1}^{N}\left(\chi_{\hat{\sigma}_i}(x)\sum_{\substack{j=1\\ j\neq i}}^{N}\chi_{\hat{\sigma}_j}(x')\right) + \delta(x',x)\rho_{\mathcal{S}}\,. \tag{8.22}$$

La deuxième somme vérifie que le point x' est dans une tuile différente de i. Comme l'union des tuiles est égale au domaine, la deuxième somme peut être remplacée par $(1-\chi_{\hat{\sigma}_i}(x'))$. On obtient alors l'expression suivante

$$\rho_{\mathcal{S}}^{(2)}(x,x') = \rho_{\mathcal{S}}^{2}\sum_{i=1}^{N}\chi_{\hat{\sigma}_i}(x)(1-\chi_{\hat{\sigma}_i}(x')) + \rho_{\mathcal{S}}\delta(x',x)\,. \tag{8.23}$$

Cette expression peut être incorporée dans la définition de la densité différentielle Eq. (5.18, p. 49) :

$$\check{\rho}_{\mathcal{S}}^{(2)}(\tau) = \frac{\rho_{\mathcal{S}}}{\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)}\left(\sum_{i=1}^{N}\rho_{\mathcal{S}}\int_{\mathbb{D}_{\mathcal{E}}}\chi_{\hat{\sigma}_i}(x)(1-\chi_{\hat{\sigma}_i}(x+\tau))\mathrm{d}x + \int_{\mathbb{D}_{\mathcal{E}}}\delta(x+\tau,x)\mathrm{d}x\right)\,. \tag{8.24}$$

Dans la première partie, l'intégrale correspond à la définition de l'auto-corrélation. Dans la deuxième partie, on remarque que l'intégrale est égale à $\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)$ quand $x = x+\tau$, c'est-à-dire quand $\tau = \tau_0$. On obtient alors

$$\check{\rho}_{\mathcal{S}}^{(2)}(\tau) = \frac{\rho_{\mathcal{S}}}{\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)}\left(\sum_{i=1}^{N}\left(\rho_{\mathcal{S}}\int_{\mathbb{D}_{\mathcal{E}}}\chi_{\hat{\sigma}_i}(x)\mathrm{d}x - (\chi_{\hat{\sigma}_i}\oplus\chi_{\hat{\sigma}_i})(\tau)\right) + \mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)\delta(\tau,\tau_0)\right)$$
$$= \frac{\rho_{\mathcal{S}}}{\mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)}\left(\sum_{i=1}^{N}(1-(\chi_{\hat{\sigma}_i}\oplus\chi_{\hat{\sigma}_i})(\tau)) + \mu_{\mathcal{E}}\left(\mathbb{D}_{\mathcal{E}}\right)\delta(\tau,\tau_0)\right)\,, \tag{8.25}$$

qui est égale à Eq. (8.21). □

8.4.3 Spectre de puissance

Nous avons obtenu une expression analytique du spectre de puissance moyen des ensembles de points *jittered* en utilisant le lien entre l'auto-corrélation d'une fonction et le spectre de puissance. Ce lien n'étant valide que dans le cas des espaces euclidiens, nous avons restreint ce résultat aux espaces d'échantillonnage $\mathcal{E}^d$ et $\mathcal{T}^d$.

Lemme 8.5 *Soit $\mathcal{S} \in \mathfrak{M}_{Jittered}$ un ensemble de points* jittered *dans l'espace $\mathcal{E}^d$ ou $\mathcal{T}^d$, le spectre de puissance moyen de l'ensemble de points est donné par*

$$\frac{\langle\mathcal{P}_{\mathcal{S}}(\omega)\rangle}{N} = 1 - \frac{\rho_{\mathcal{S}}}{N}\sum_{i=1}^{N}\mathcal{P}_{\chi_{\hat{\sigma}_i}}(\omega) + \rho_{\mathcal{S}}\delta(\omega,\omega_0)\,. \tag{8.26}$$

Preuve 8.5 *On utilisant Eq. (8.21), et la transformée de Fourier de la fonction δ de Dirac, on obtient*

$$\begin{aligned}\frac{\langle \mathcal{P}_S(\omega)\rangle}{N} &= \rho_S \mathcal{F}\left[\frac{1}{\rho_S}\delta(\tau,\tau_0)\right](\omega) - \frac{1}{N}\sum_{i=1}^{N}(\chi_{\hat{\sigma}_i} \oplus \chi_{\hat{\sigma}_i})(\tau) + 1 \\ &= 1 - \frac{\rho_S}{N}\sum_{i=1}^{N}\mathcal{F}\left[(\chi_{\hat{\sigma}_i} \oplus \chi_{\hat{\sigma}_i})(\tau)\right](\omega) + \rho_S\delta(\omega,\omega_0)\,. \end{aligned} \tag{8.27}$$

On obtient Eq. (8.26) en remarquant que la transformée de Fourier de l'auto-corrélation correspond au spectre de puissance. □

Pour nos expérimentations, nous avons utilisé un pavage a base de tuiles hyper-cubiques dont chaque coté à une longueur de $\sqrt[d]{N}$. Le pavage étant régulier, chaque tuile est une version décalée des autres. On appelle $\chi_{\hat{\sigma}}$ la fonction caractéristique de la tuile hyper-cubique centrée à l'origine. On peut donc écrire $\chi_{\hat{\sigma}}$ comme le résultat de la convolution entre des fonctions Π dans chaque dimension :

$$\chi_{\hat{\sigma}}(x) = (\Pi(\sqrt[d]{N}x|_1) \otimes ... \otimes \Pi(x|_d)(\sqrt[d]{N}x)\,, \tag{8.28}$$

où $x|_k$ désigne ici la k-ième coordonnée de x, avec x exprimé sous la forme d'un vecteur de dimension d. D'après les liens entre la convolution et la transformée de Fourier, et connaissant la transformée de Fourier de la fonction Π, on trouve l'expression de la transformée de Fourier suivante :

$$\mathcal{F}_{\chi_{\hat{\sigma}}}(\omega) = \sqrt[d]{N}\prod_{k=1}^{N} sinc(\sqrt[d]{N}\omega|_k)\,, \tag{8.29}$$

où ici aussi, la notation $\omega|_k$ désigne la k-ième coordonnée de ω. Le spectre de puissance de $\chi_{\hat{\sigma}}$ est donné par le carré de $\mathcal{F}_{\chi_{\hat{\sigma}}}$. Comme le spectre de puissance est invariant par translation, on en déduit que

$$\begin{aligned}\mathcal{P}_{\chi_{\hat{\sigma}_i}}(\omega) &= \mathcal{P}_{\chi_{\hat{\sigma}}}(\omega) \\ &= \sqrt[d]{N}^2 \prod_{k=1}^{N} sinc(\sqrt[d]{N}\omega|_k)^2\,. \end{aligned} \tag{8.30}$$

L'expression analytique du spectre de puissance moyen est finalement obtenue en utilisant cette équation avec Eq. (8.26) :

$$\frac{\langle \mathcal{P}_S(\omega)\rangle}{N} = 1 - \rho_S \sqrt[d]{N}^2 \prod_{k=1}^{N} sinc(\sqrt[d]{N}\omega|_k)^2 + \rho_S\delta(\omega,\omega_0)\,. \tag{8.31}$$

8.4.4 Variance de l'intégration Monte Carlo

Nous venons de donner l'expression analytique du spectre de puissance des ensembles de points *jittered*. Pour pouvoir appliquer Eq. (7.11, p. 95), nous devons maintenant considérer la version homogénéisée de ces ensembles de points, par la méthode décrite dans Lemme (3.3, p. 28). Le spectre de puissance moyen des ensembles de points *jittered* se comporte de façon quadratique quand $\omega \to \omega_0$. C'est aussi vrai pour la version homogénéisée puisque le spectre de puissance est invariant par le groupe

des actions. On peut donc le borner avec un profil spectral où le paramètre b est fixé à 2. On peut alors écrire que

$$\mathcal{P}_S \in O_{r\to 0}\left(|r|^2\right) . \tag{8.32}$$

En conséquence, la variance de l'intégration Monte Carlo est en $O_{\rho\to\infty}\left(\rho^{-\frac{d+2}{d}} + \rho^{-\frac{d+s}{d}}\right)$. Lorsque $s < 2$, la variance est en $O_{\rho\to\infty}\left(\rho^{-\frac{d+s}{d}}\right)$. Dans l'autre cas, elle est en $O_{\rho\to\infty}\left(\rho^{-\frac{d+2}{d}}\right)$. Ce résultat est un des plus surprenant de notre étude. Les ensembles de points *jittered* sont asymptotiquement plus efficaces que les disques de Poisson. Ainsi, dans la figure (8.15), on remarque que la variance générée par les ensembles de points *jittered* est plus faible à partir d'environ $\rho = 200$ échantillons. C'est particulièrement surprenant car visuellement, les points ont l'air mieux répartis dans le cas des ensembles de disques de Poisson.

8.5 *Farthest-Point Optimization*

Pour générer les ensembles de points optimisés par FPO, nous avons utilisé l'implémentation donnée par Schlömer [2011]. À partir d'une distribution bruit blanc, nous avons appliqué la méthode globale de l'optimisation jusqu'à ce que la distance minimale entre les points atteigne $\delta_X = 0.925$, comme conseillé par les auteurs. Ici, δ_X est la distance minimale entre deux points, normalisée par la distance minimale entre deux points d'une grille hexagonale. La distance minimale, non-normalisée, est donc égale à $0.925 \times \sqrt{2/\sqrt{3}N} \approx 0.994/\sqrt{N}$.

La figure (8.7) montre la fonction de distribution différentielle et la fonction de distribution de distance pour les ensembles de points optimisés par FPO. On remarque qu'en accord avec l'algorithme d'optimisation, la distance minimale entre les points est de $0.994/\sqrt{N}$. En comparaison avec les méthodes que nous avons évaluées jusqu'à maintenant, les ensembles de points optimisés par FPO ont une distribution de capacité relativement compacte. En revanche, l'uniformité de cette méthode d'optimisation n'est pas optimale, comme en témoigne la répartition des capacités dans le domaine.

Ce défaut d'uniformité se traduit dans le domaine spectral par une énergie résiduelle dans les basses fréquences. Dans la figure (8.8,d), on remarque que le spectre de puissance ne converge pas vers zéro quand la fréquence tend vers zéro. À la place, on voit que le spectre de puissance est toujours supérieur à une constante égale à 6.9×10^{-3}. Il remonte même légèrement pour les fréquences au plus près de zéro. La figure (8.7,h et i) montre le spectre de puissance moyen dans sa représentation classique en $2D$ et en moyenne radiale. On observe que $\mathcal{P}_S$ est bien isotrope et que de fortes oscillations apparaissent dans les hautes fréquences.

8.5.1 Variance de l'intégration Monte Carlo

Comme le spectre de puissance est borné en bas par une constante, on en déduit que

$$\mathcal{P}_S(r) \in O_{r\to 0}(1) , \tag{8.33}$$

et donc que la vitesse de convergence de cette méthode est en $O_{\rho\to\infty}\left(\rho^{-1} + \rho^{-\frac{d+s}{d}}\right)$. Les résultats d'intégration pour le cas du disque et de la gaussienne, figure (8.8,e et f), montrent que notre estimation de la vitesse de convergence de la variance est cohérente. Cette méthode d'échantillonnage est une illustration parfaite de la difficulté

d'obtenir un bon motif d'échantillonnage. Lorsque l'on compare visuellement un ensemble de points optimisés par FPO avec un ensemble de points *jittered*, il semble à première vue évident que la première surpasse la deuxième. Les points semblent mieux distribués et aucune paire de points proches n'est visible. Mais des défauts visuellement imperceptibles d'uniformité suffisent à briser tout cela. En terme de vitesse de convergence, cette méthode est donc similaire aux disques de Poisson et au bruit blanc. Toutefois, comme pour les disques de Poisson, l'ordre de grandeur de la variance est beaucoup plus faible que dans le cas du bruit blanc. Dans la figure (8.15), on remarque qu'il faut attendre $N > 2000$ pour que l'ensemble de points *jittered* soit plus efficace.

8.6 *Blue Noise through Optimal Transport*

Pour la méthode d'échantillonnage BNOT, nous avons choisi l'implémentation fournie dans [de Goes *et al.*, 2012]. L'ensemble de points a été obtenu en optimisant un ensemble de points bruit blanc en utilisant les paramètres d'optimisation par défaut. La fonction de distribution de capacité est ici proche de celle de FPO. Toutefois, comme le montre la figure (8.9,d), la répartition au sein du domaine d'échantillonnage est plus uniforme. Cette amélioration de l'uniformité est directement visible dans le domaine spectral. En regardant le spectre de puissance moyen en échelle logarithmique, voir figure (8.10,d), on observe qu'il converge vers zéro quand la fréquence tend vers zéro. La vitesse de convergence est quadratique et on est donc tenté d'utiliser un profil spectral polynomial avec un paramètre $b = 2$ pour encadrer le spectre de puissance. Cependant, il faut se rappeler que pour que le profil spectral soit valide, le spectre de puissance doit tendre vers le profil spectral quand $\rho \to \infty$. Or dans le cas de BNOT, ce n'est pas exactement le cas si l'on choisit un profil spectral avec $b = 2$. Dans les basses fréquences, si la courbe ressemble bien à une rampe quadratique, on constate qu'elle devient une courbe de dégrée 3 puis 4 immédiatement après. Quand la densité augmente, la partie quadratique devient de plus en plus petite et la partie de degré 4 de plus en plus importante. Bien qu'imprécis, il semble qu'utiliser un profil spectrale avec $b = 4$ soit plus proche de la réalité que le profil quadratique dans le cas de BNOT.

8.6.1 Variance de l'intégration Monte Carlo

Le choix du profile spectral de degré 4 est confirmé par nos résultats expérimentaux, montrés dans la figure (8.10,e et f). Ainsi, la vitesse de convergence dans le cas de la fonction gaussienne est en $O_{\rho\to\infty}\left(\rho^{-3}\right)$ et la vitesse de convergence dans le cas du disque est en $O_{\rho\to\infty}\left(\rho^{-3/2}\right)$, ce qui correspond bien à la vitesse de convergence d'un profil spectral polynomial avec $b = 4$. On en déduit donc que la vitesse de convergence de la variance est en $O_{\rho\to\infty}\left(\rho^{-\frac{d+4}{d}} + \rho^{-\frac{d+s}{d}}\right)$. En comparaison avec les ensembles de points *jittered*, cette méthode d'optimisation a une meilleur vitesse de convergence pour les intégrandes appartenant à Φ^s, $s > 2$.

8.7 *Step Blue Noise*

Pour analyser les méthodes de contrôle spectral, nous avons choisi la méthode décrite dans [Heck *et al.*, 2013], en utilisant la même implémentation que celle des auteurs. Le choix de la cible spectrale est ici important. Notre choix s'est arrêté sur ce que les auteurs appellent *Step Blue Noise*, c'est-à-dire un spectre de puissance discontinu donnée par

$$\frac{\langle \mathcal{P}_S(r) \rangle}{N} = \begin{cases} 0 & \text{si } |r| < \sqrt{N/\pi} \\ 1 & \text{sinon}\,. \end{cases} \tag{8.34}$$

Le spectre de puissance moyen que nous avons obtenu, voir (8.11,h et i) et (8.12,c et d), est proche de la cible spectrale, mais présente deux défauts. Le premier défaut concerne les oscillations du spectre. Ces oscillations proviennent du fait que l'optimisation est effectuée sur un sous-domaine de la fonction de distribution différentielle. Le deuxième défaut porte quant à lui sur les basses fréquences. La cible spectrale choisie ne comporte aucune basse fréquence. Mais le spectre de puissance moyen en échelle logarithmique montre clairement le résidu d'énergie dans les basses fréquences. À cause des oscillations, il est difficile de mesurer la vitesse de convergence en zéro du spectre moyen. Une solution acceptable consiste à prendre un profil spectral linéaire ($b = 1$).

8.7.1 Variance de l'intégration Monte Carlo

En utilisant comme profil spectral une rampe linéaire, on obtient une vitesse de convergence dans le cas gaussien et le cas du disque en $O_{\rho\to\infty}\left(\rho^{-3/2}\right)$. La figure (8.12,e et f) montre ces résultats. De manière générale, la vitesse de convergence est en $O_{\rho\to\infty}\left(\rho^{-\frac{d+1}{d}} + \rho^{-\frac{d+s}{d}}\right)$. Dans le cas du disque, la vitesse de convergence est semblable à celle d'un profil spectral des méthodes avec un coefficient b plus élevé, comme la méthode d'optimisation BNOT. En revanche, dans le cas de la gaussienne, la vitesse de convergence de cette méthode est beaucoup plus lente.

8.8 Séquence de Halton

Pour terminer cette étude de l'état de l'art, nous avons appliqué notre méthodologie à une séquence à basse discrépance. Nous avons choisi la séquence de Halton, avec la base 2 pour la dimension 1 et la base 3 pour la dimension 2. Ce que montrent les résultats présentés dans la figure (8.13), c'est que les outils que nous avons proposés s'appliquent mal aux séquences à basses discrépance. Cela n'est pas surprenant car ces ensembles de points ne sont pas stochastiques et possèdent de nombreuse régularités. La fonction de distribution différentielle et le spectre de puissance se manifestent donc comme un ensemble de pics. Ces fonctions ne sont pas continues par morceaux ce qui les rend difficiles à utiliser. Il n'est par exemple pas possible d'associer un profil spectral au spectre de puissance et donc, de trouver une expression du spectre de puissance en fonction de la densité de l'ensemble de points. On observera simplement que la distribution de la capacité (8.13,d et e) semble du même ordre que celle des ensembles de points *jittered*, et que la vitesse de convergence pour les deux intégrandes, voir figure (8.14,e et f), est semblable à celle des ensembles de points *jittered*.

8.9 Intégration d'une texture

Afin d'illustrer un cas concret d'intégration, nous avons choisi l'intégration d'une image *High-Dynamic-Range* (HDR), aussi appelée en français image à grande gamme dynamique. En prenant une image représentant une scène naturelle, on peut ainsi se rapprocher d'un cas concret d'intégration en rendu d'image. Il faut toutefois faire attention à ce que le nombre de pixels composant l'image soit bien supérieur au nombre d'échantillons. S'il y a plus d'un échantillon par pixel, alors l'erreur d'intégration sera liée principalement à la méthode d'interpolation utilisée entre les pixels, au lieu du contenu de l'image à intégrer. Nous avons choisi une résolution d'image de 1600×1600 pixels, permettant de tester des ensembles contenant plus de deux millions de points. L'image sélectionnée provient de la bibliothèque d'image HDR sIBL Archive [1]. Nous avons extrait la moitié gauche de l'image et converti les couleurs en niveau de gris en moyennant ensemble chaque canal de couleur.

Les résultats sont présentés dans la figure (8.15). La vitesse de convergence de la séquence de Halton semble meuilleure que celle des autres méthodes. Ceci peut s'expliquer par le fait que l'intégrande fait partie de l'ensemble de fonctions compatibles avec l'équation de Koksma-Hlawka, puisque les discontinuités produites par les pixels sont alignés aux axes X et Y de l'image. Viennent ensuite les méthodes uniformes dont le paramètre b du profil spectral est supérieur ou égale à 1. Ainsi, les ensembles de points *jittered*, BNOT et *Step Blue Noise* ont une vitesse de convergence similaire. Finalement, le bruit blanc, les disques de Poisson et FPO sont les méthodes avec la vitesse de convergence la plus lente. Il faut cependant rappeler que la vitesse de convergence reflète le comportement de la méthode d'échantillonnage quand la densité est très grande. Quand la densité est faible, certaines méthodes, dont la vitesse de convergence est lente, sont quand même meilleures que d'autres. Ainsi, les disques de Poisson sont meilleurs que les ensembles de points *jittered* jusqu'à environ $\rho = 100$, FPO est meilleur que les ensembles de points *jittered* jusqu'à environ $\rho = 1500$, et BNOT est meilleur que la séquence de Halton jusqu'à environ $\rho = 1500$.

8.10 Test *zoneplate*

Le test *zoneplate* est un test classique de rendu d'image artificielle. Il vise à montrer la sensibilité d'une méthode d'échantillonnage au contenu fréquentiel de l'intégrande. La scène est composée d'une fonction sinus radiale dont la fréquence augmente en fonction de la distance au centre,

$$f_{\text{zoneplate}}(x) = sin(||x||^2)\,. \tag{8.35}$$

Le niveau de gris de chaque pixel est calculé en utilisant une moyenne pondérée des échantillons proches du pixel. Dans notre implémentation, nous avons choisi d'utiliser un filtre de Mitchell-Netravali pour pondérer chaque échantillon. L'image est échantillonnée en répétant périodiquement un ensemble de points généré dans l'espace $\mathcal{T}^2$. Contrairement au test habituel, nous avons calibré le test de façon à ce que chaque pixel ne contienne environ qu'une seule fréquence. Ainsi, la fréquence moyenne dans chaque pixel du test *zoneplate* correspond à une fréquence dans le spectre de puissance. Les figures (8.2), (8.4), (8.6), (8.8), (8.10), (8.12) et (8.14) montrent le spectre de puissance et le resultat du test *zoneplate* cote à cote.

1. http ://www.hdrlabs.com/sibl/archive.html

Les résultats sont similaires pour chacune des méthodes d'échantillonnage. Le niveau de bruit dans chaque image, c'est-à-dire la variance, est proportionnel au spectre de puissance. Ce test est ainsi une illustration du lien entre variance et spectre de puissance du motif d'échantillonnage. On notera toutefois que lorsque le spectre de puissance contient des pics (des valeurs très grandes, concentrées sur une fréquence particulière), des artefacts apparaissent dans le test. La figure (8.14) montre ces artefacts dans le cas de la séquence de Halton. C'est en partie ce phénomène qui justifie que nous ayons focalisé notre étude sur les ensembles de points stochastiques dépourvus de pics.

Ensemble de points idéal

Pour conclure ce chapitre, nous dressons ici les conditions et propriétés qui nous semblent nécessaires pour qu'un ensemble de points soit efficace pour l'intégration Monte Carlo. Nous avons déjà justifié qu'à notre sens, en l'absence d'information sur l'intégrande, l'ensemble de points doit être homogène. Ainsi, la méthode d'échantillonnage produira en moyenne la même quantité d'erreur quelque soit la position de l'intégrande. L'autre avantage est que l'erreur d'intégration ne se manifeste qu'en terme de variance, permettant de se concentrer sur la réduction d'un seule mesure d'erreur. Comme nous avons relié la variance de l'intégration Monte Carlo au spectre de puissance, nous définirons dans le domaine spectral, les bonnes propriétés qu'un ensemble de points doit avoir. Sur ce point, notre approche est la même que celle d'Ulichney avec sa définition du bruit bleu [Ulichney, 1987].

Basses fréquences Puisque le spectre de puissance de l'intégrande est décroissant, l'énergie de son spectre de puissance est principalement localisé dans les basses fréquences. C'est d'autant plus vrai que la fonction est «simple à intégrer », c'est-à-dire qu'elle appartient à la classe de fonctions Φ^s, où s est grand. Dans Lemme (7.8, p. 88), nous avons vu que la variance de l'intégration est le résultat de l'intégrale du produit du spectre de puissance de l'intégrande et de l'ensemble de points. Si l'essentiel de l'énergie de l'intégrande est dans les basses fréquences, il faut donc à l'inverse, le moins d'énergie possible dans cette partie pour l'ensemble de points. Cette observation suit celle d'Ulichhney et n'est pas nouvelle en soi.

Moyennes et hautes fréquences Si les basses fréquences sont primordiales, les moyennes et hautes fréquences peuvent aussi avoir des effets négatifs sur la qualité de l'intégration Monte Carlo. Ainsi, les grilles régulières n'ont pas d'énergie dans les basses fréquences, mais sont très sensibles à certaines fréquences car leur spectre de puissance contient des pics. En réalité, toujours en conséquence de Lemme (7.8, p. 88), une forte valeur dans le spectre de puissance de l'ensemble de points peut engendrer une forte variance selon le contenu fréquentiel de l'intégrande. Il vaux donc mieux que l'énergie dans le spectre de puissance soit répartie sur les moyennes et hautes fréquences, plutôt que concentrée dans des pics. Sur ce point, FPO et la séquence de Halton ne sont pas optimaux. On remarque par exemple dans le test *zoneplate*, figures (8.8,b) et (8.14,b), que la variance est très élevée là où le spectre de puissance l'est aussi.

Comportement dans les très basses fréquences Comme nous l'avons montré dans le chapitre précédant, et comme nous l'avons observé avec les différentes méthodes d'échantillonnage dans ce chapitre, le comportement dans les très basses fréquences est très important pour la vitesse de convergence. Le spectre de puissance doit converger vers zéro à la fréquence zéro, sans quoi, comme le bruit blanc, les disques de Poisson et FPO, la vitesse de convergence sera en $O_{\rho\to\infty}\left(\rho^{-1}\right)$. Cependant, pour avoir la meilleure vitesse de convergence possible, il faut que les basses fréquences, de la fréquence zéro jusqu'à un seuil donné, ne possèdent absolument aucune énergie. C'est actuellement la condition, à l'exception des grilles régulières, qui n'est respectée par aucune méthode.

D'autres conditions ont été énoncée dans la littérature. Ulichney a par exemple proposé qu'une grande partie de l'énergie du spectre de puissance soit concentrée vers une fréquence proche de l'inverse de la distance entre les points. On constate ce phénomène avec les disques de Poisson, FPO et BNOT. De notre expérience, ce point n'est pas directement relié à l'intégration Monte Carlo. Cela résulte plutôt du fait que l'énergie contenue dans les basses fréquences doit être distribuée ailleurs, ce qui se fait généralement en remplissant les fréquences qui suivent directement. Une autre condition, formulée par Heck et al. [2013] est que l'amplitude des oscillations dans spectre de puissance doit être limitée. La figure (8.7,i) montre un ensemble de points possédant des oscillations fortes. Cependant, comme nous l'avons montré, ce sont les basses fréquences, et non les hautes fréquences, qui jouent un rôle déterminant. Ce point ne nous semble donc pas primordial.

Reste la question de savoir si cette description de l'ensemble de points idéal est réalisable. À l'heure actuelle, la seule condition connue pour qu'un ensemble de points soit réalisable est que le spectre de puissance et la fonction de distribution différentielle soit entièrement positives [Torquato and Stillinger, 2003]. Ces conditions sont nécessaires, mais pas forcement suffisantes, bien que nous n'ayons pas à ce jour de contre-exemples. La fonction *Step Blue Noise*, dont nous avons présenté une approximation par contrôle spectral, semble par exemple atteignable. Mais comme nous n'avons à ce jour aucune implémentation parfaite, la question de son existence est toujours posée.

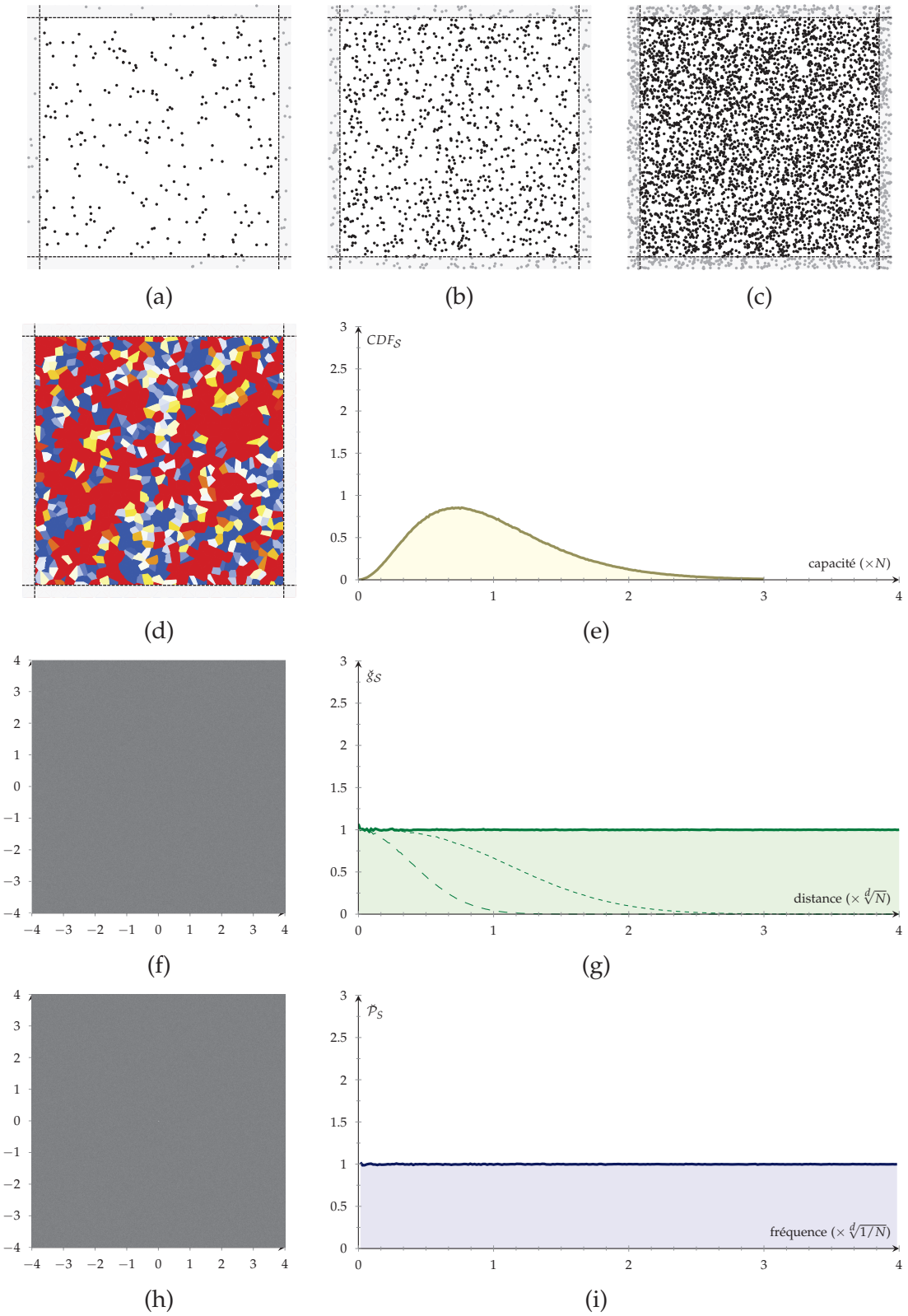


FIGURE 8.1 – *Analyse spatiale et spectrale des ensembles de points bruit blanc.* ***(a), (b) et (c) :*** *Des réalisations de l'ensemble de points contenant respectivement* 256, 1024 *et* 4096 *points.* ***(d) :*** *La distribution spatiale des capacités de chaque point dans une ensemble de* 1024 *points.* ***(e) :*** *La fonction distribution de capacité.* ***(f) :*** *La fonction de distribution différentielle.* ***(f) :*** *La fonction de distribution de distance.* ***(g) :*** *Le spectre de puissance moyen.* ***(h) :*** *Le spectre de puissance moyen radial.*

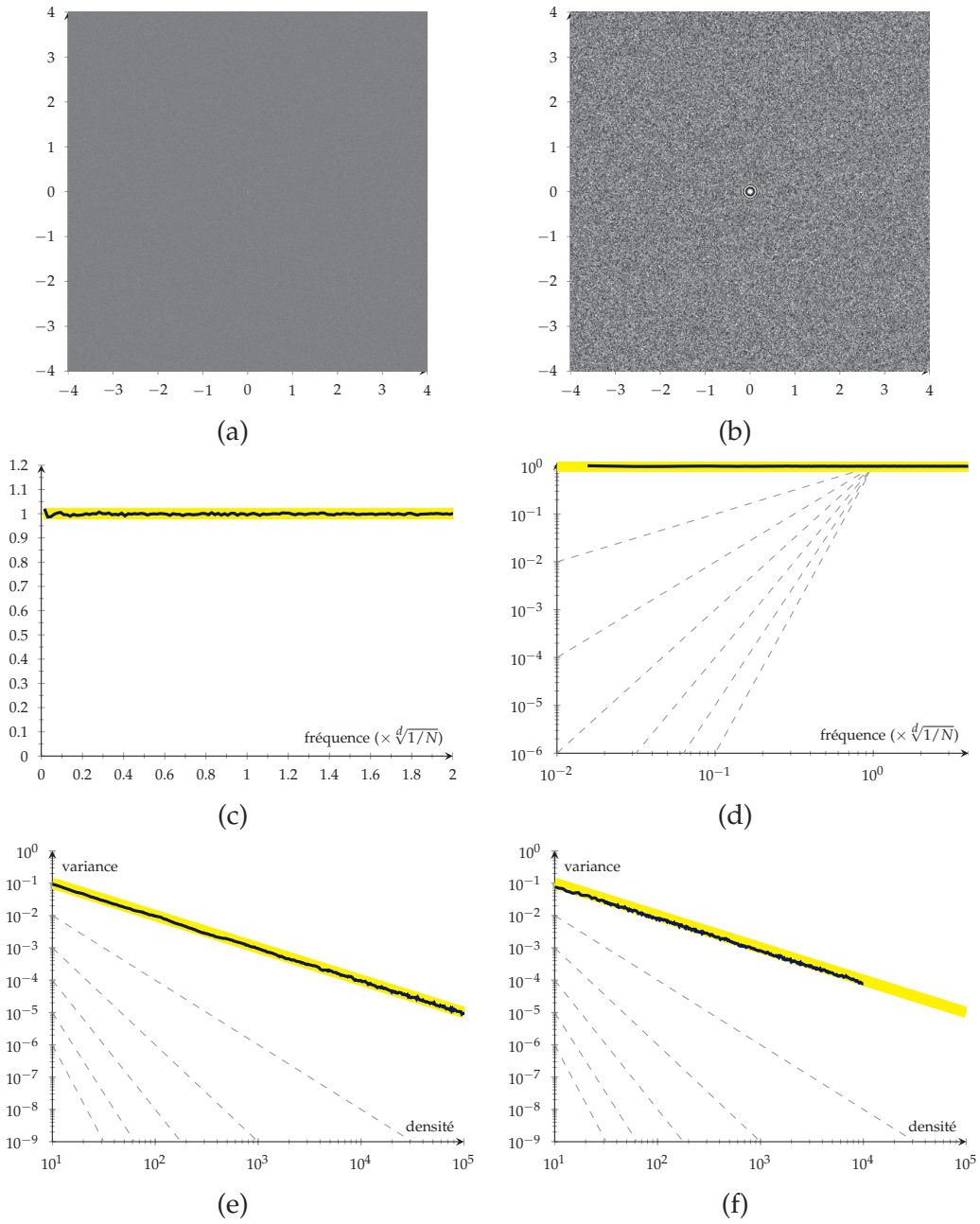


FIGURE 8.2 – *Test d'intégration avec les ensembles de points bruit blanc.* ***(a) :*** *Le spectre de puissance moyen.* ***(b) :*** *Le test* zoneplate. ***(c) :*** *Le spectre de puissance moyen radial avec en bleu, celui de l'ensemble de points avec* $N = 4096$*, en jaune, le spectre de puissance théorique, et en gris, des repères correspondant des profiles spectraux polynomiaux de différents degrés.* ***(d) et (e) :*** *La variance de l'intégration d'une gaussienne et d'un disque avec en bleu, la variance générée par l'ensemble de points avec différentes densités, en jaune, la variance prédite théoriquement, et en gris, des repères correspondant à différentes vitesses de convergence.*

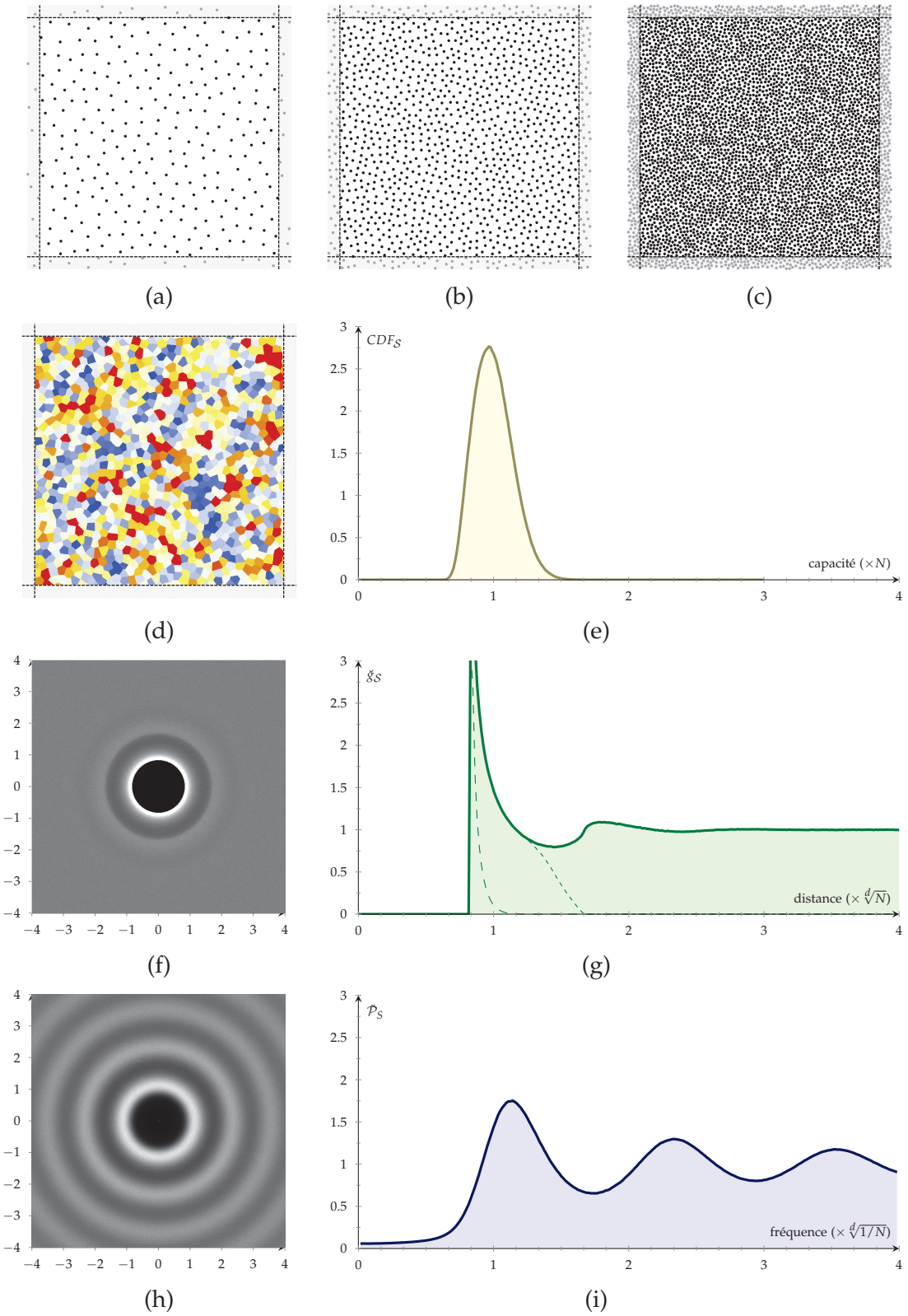


FIGURE 8.3 – *Analyse spatiale et spectrale des ensembles de disques de Poisson.* ***(a), (b) et (c) :*** *Des réalisations de l'ensemble de points contenant respectivement* 256, 1024 *et* 4096 *points.* ***(d) :*** *La distribution spatiale des capacités de chaque point dans une ensemble de* 1024 *points.* ***(e) :*** *La fonction distribution de capacité.* ***(f) :*** *La fonction de distribution différentielle.* ***(f) :*** *La fonction de distribution de distance.* ***(g) :*** *Le spectre de puissance moyen.* ***(h) :*** *Le spectre de puissance moyen radial.*

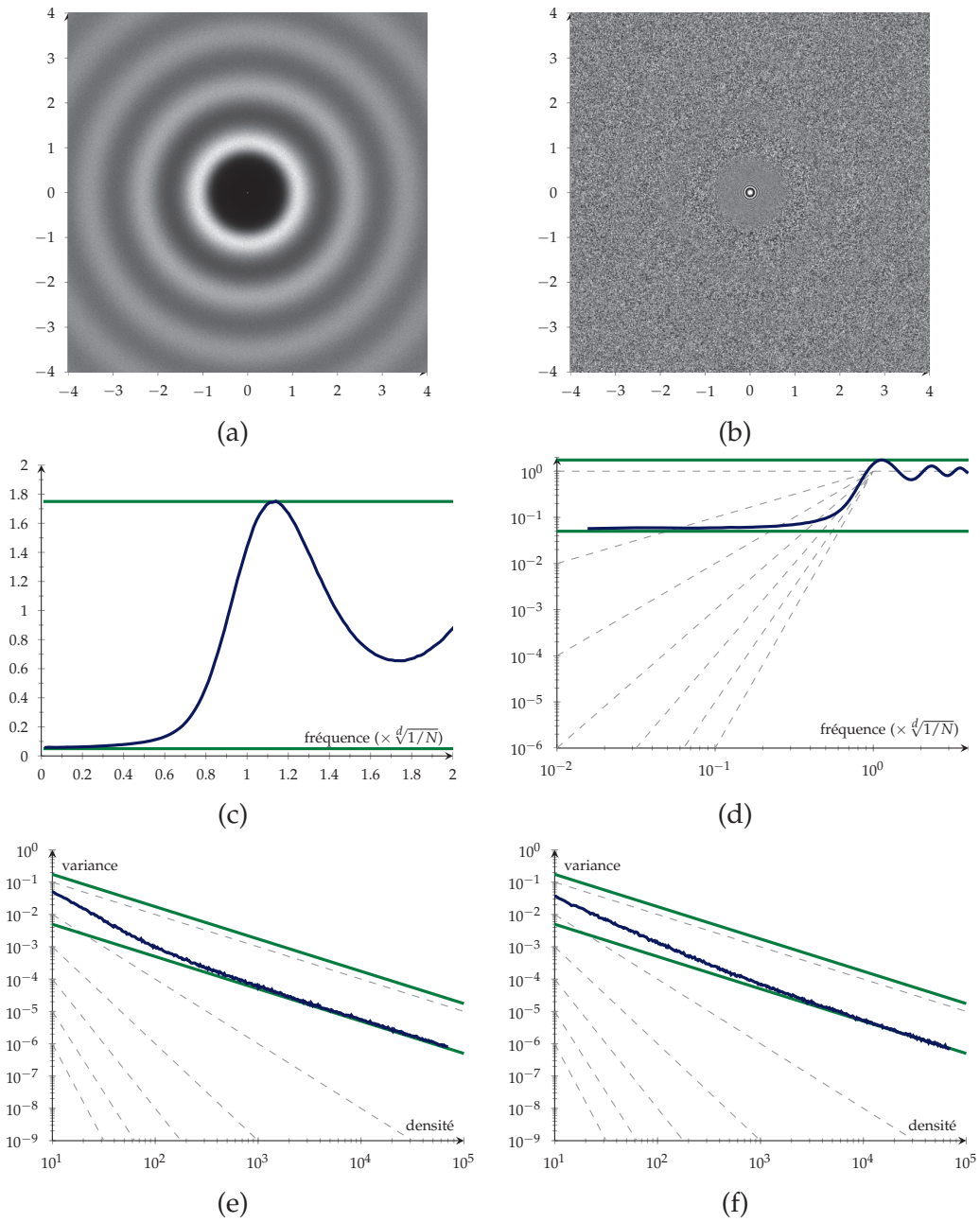


FIGURE 8.4 – *Test d'intégration avec les ensembles de disques de Poisson.* ***(a) :*** *Le spectre de puissance moyen.* ***(b) :*** *Le test* zoneplate. ***(c) et (d) :*** *Le spectre de puissance moyen radial, en échelle normale et logarithmique avec en bleu, celui de l'ensemble de points avec* $N = 4096$*, en vert, les bornes supérieur et inférieur du spectre de puissance, et en gris, des repères correspondant des profiles spectraux polynomiaux de différents degrés.* ***(e) et (f) :*** *La variance de l'intégration d'une gaussienne et d'un disque avec en bleu, la variance générée par l'ensemble de points avec différentes densités, en vert, la variance encadrée, et en gris, des repères correspondant à différentes vitesses de convergence.*

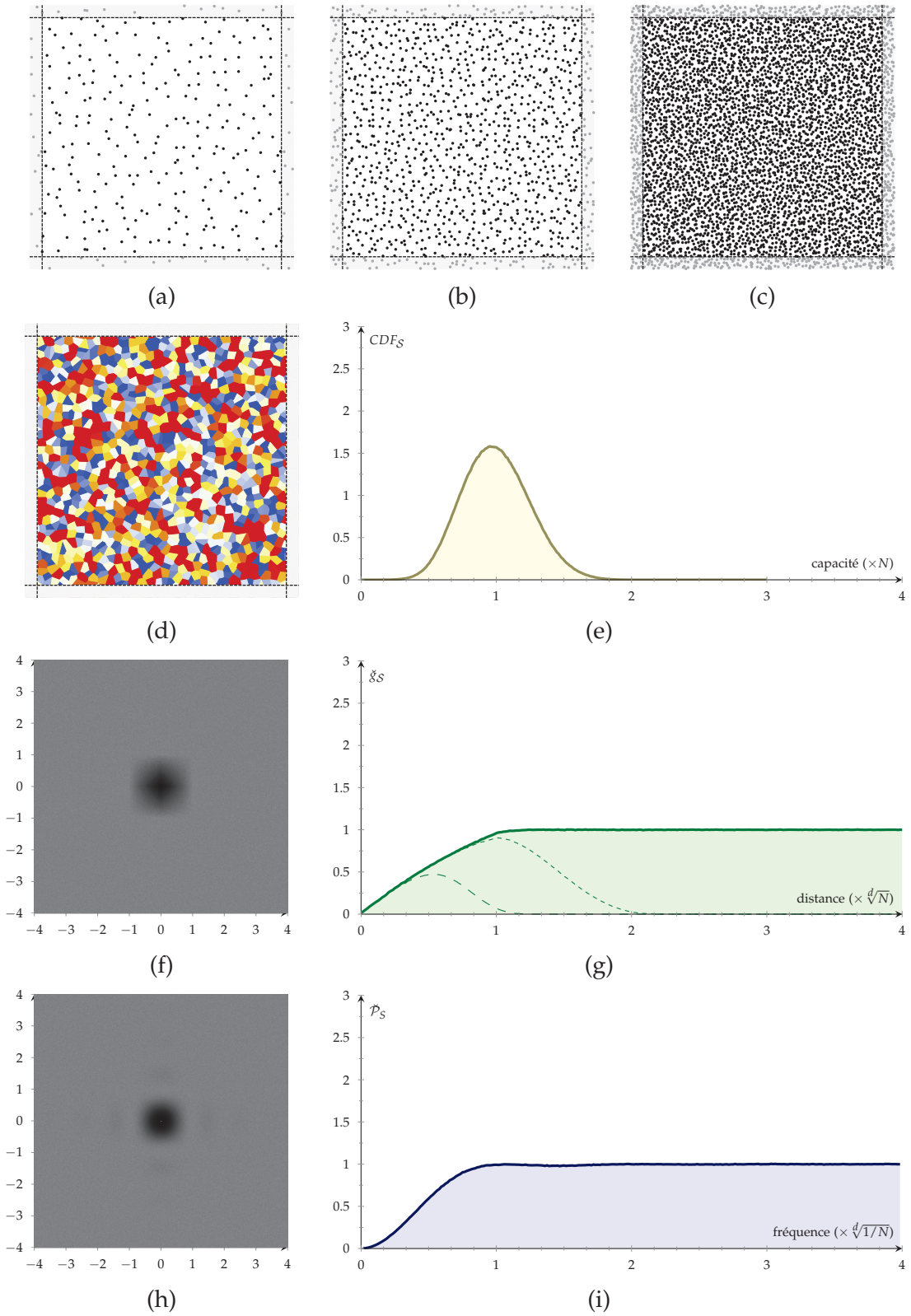


FIGURE 8.5 – *Analyse spatiale et spectrale des ensembles de points* jittered. ***(a), (b) et (c) :*** *Des réalisations de l'ensemble de points contenant respectivement* 256, 1024 *et* 4096 *points.* ***(d) :*** *La distribution spatiale des capacités de chaque point dans une ensemble de* 1024 *points.* ***(e) :*** *La fonction distribution de capacité.* ***(f) :*** *La fonction de distribution différentielle.* ***(f) :*** *La fonction de distribution de distance.* ***(g) :*** *Le spectre de puissance moyen.* ***(h) :*** *Le spectre de puissance moyen radial.*

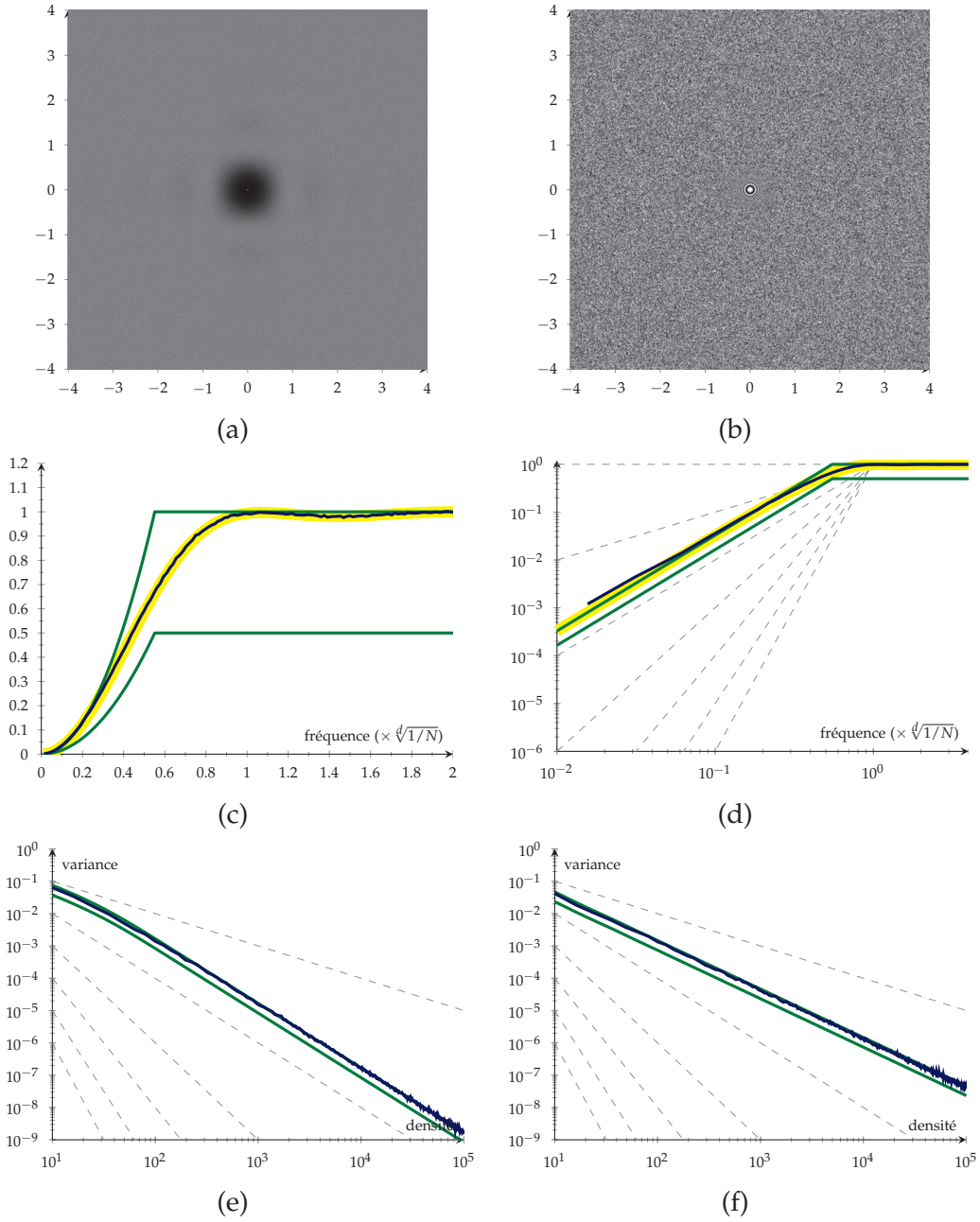


FIGURE 8.6 – *Test d'intégration avec les ensembles de points* jittered. ***(a) :*** *Le spectre de puissance moyen.* ***(b) :*** *Le test* zoneplate. ***(c) et (d) :*** *Le spectre de puissance moyen radial, en échelle normale et logarithmique avec en bleu, celui de l'ensemble de points avec* $N = 4096$*, en jaune, le spectre de puissance théorique, en vert, les bornes supérieur et inférieur du spectre de puissance, et en gris, des repères correspondant des profiles spectraux polynomiaux de différents degrés.* ***(e) et (f) :*** *La variance de l'intégration d'une gaussienne et d'un disque avec en bleu, la variance générée par des ensembles de points bruit blanc de différentes densités, en jaune, la variance prédite théoriquement, en vert, la variance encadrée, et en gris, des repères correspondant à différentes vitesses de convergence.*

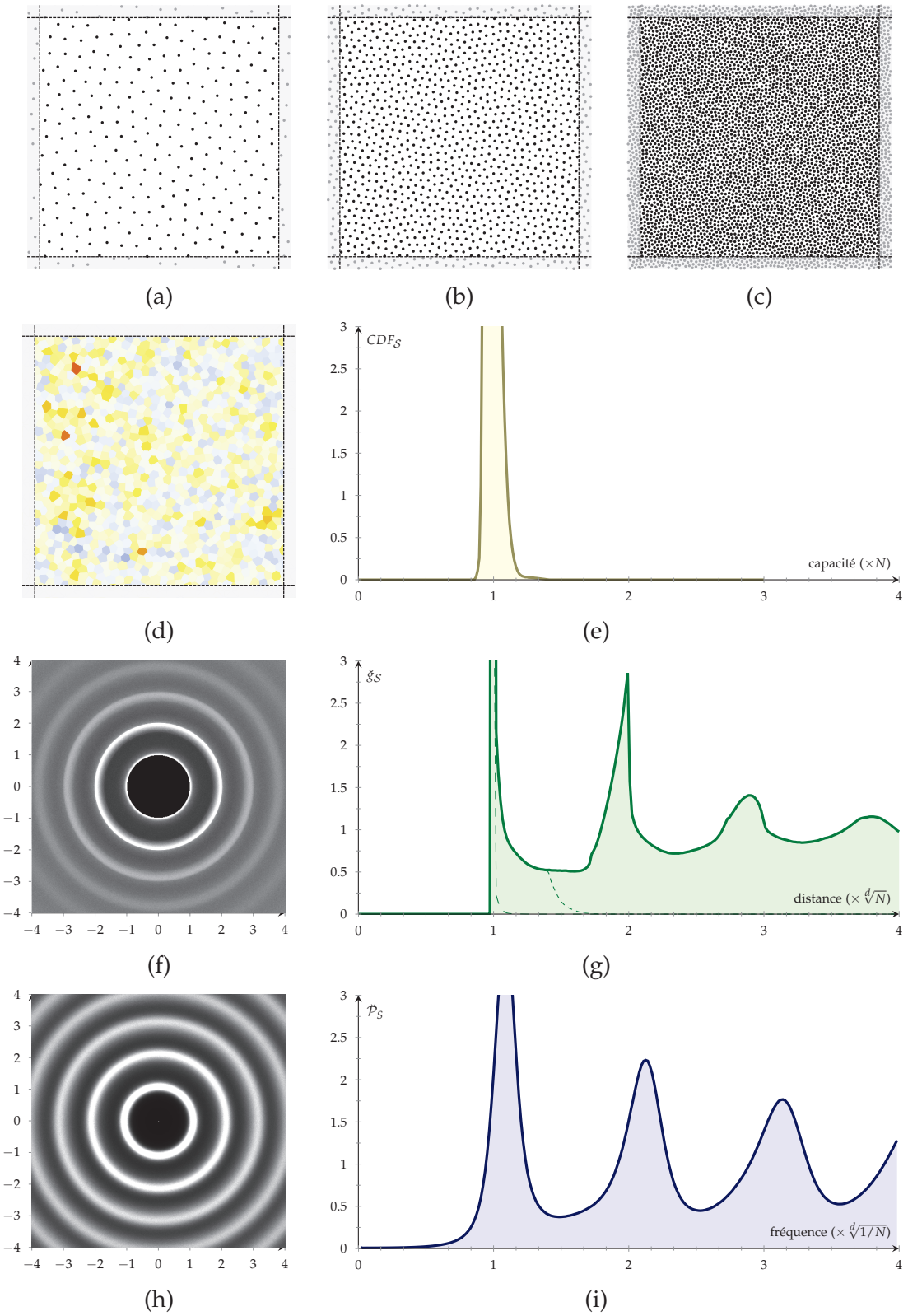


FIGURE 8.7 – *Analyse spatiale et spectrale des ensembles de points optimisés par FPO.* ***(a), (b) et (c) :*** *Des réalisations de l'ensemble de points contenant respectivement* 256, 1024 *et* 4096 *points.* ***(d) :*** *La distribution spatiale des capacités de chaque point dans une ensemble de* 1024 *points.* ***(e) :*** *La fonction distribution de capacité.* ***(f) :*** *La fonction de distribution différentielle.* ***(f) :*** *La fonction de distribution de distance.* ***(g) :*** *Le spectre de puissance moyen.* ***(h) :*** *Le spectre de puissance moyen radial.*

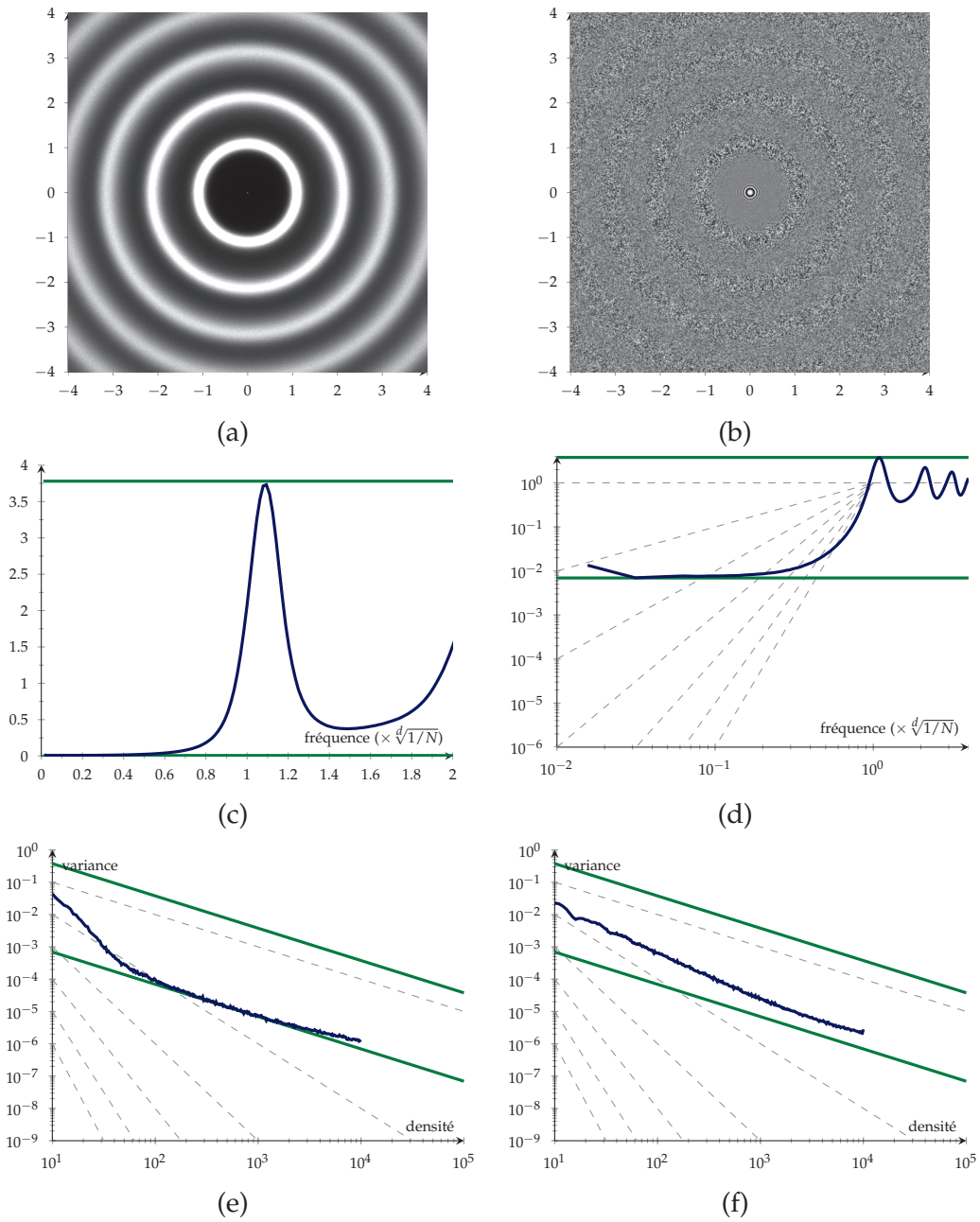


FIGURE 8.8 – *Test d'intégration avec les ensembles de points optimisés par FPO.* ***(a) :*** *Le spectre de puissance moyen.* ***(b) :*** *Le test* zoneplate. ***(c) et (d) :*** *Le spectre de puissance moyen radial, en échelle normale et logarithmique avec en bleu, celui de l'ensemble de points avec* $N = 4096$*, en vert, les bornes supérieur et inférieur du spectre de puissance, et en gris, des repères correspondant des profiles spectraux polynomiaux de différents degrés.* ***(e) et (f) :*** *La variance de l'intégration d'une gaussienne et d'un disque avec en bleu, la variance générée par l'ensemble de points avec différentes densités, en vert, la variance encadrée, et en gris, des repères correspondant à différentes vitesses de convergence.*

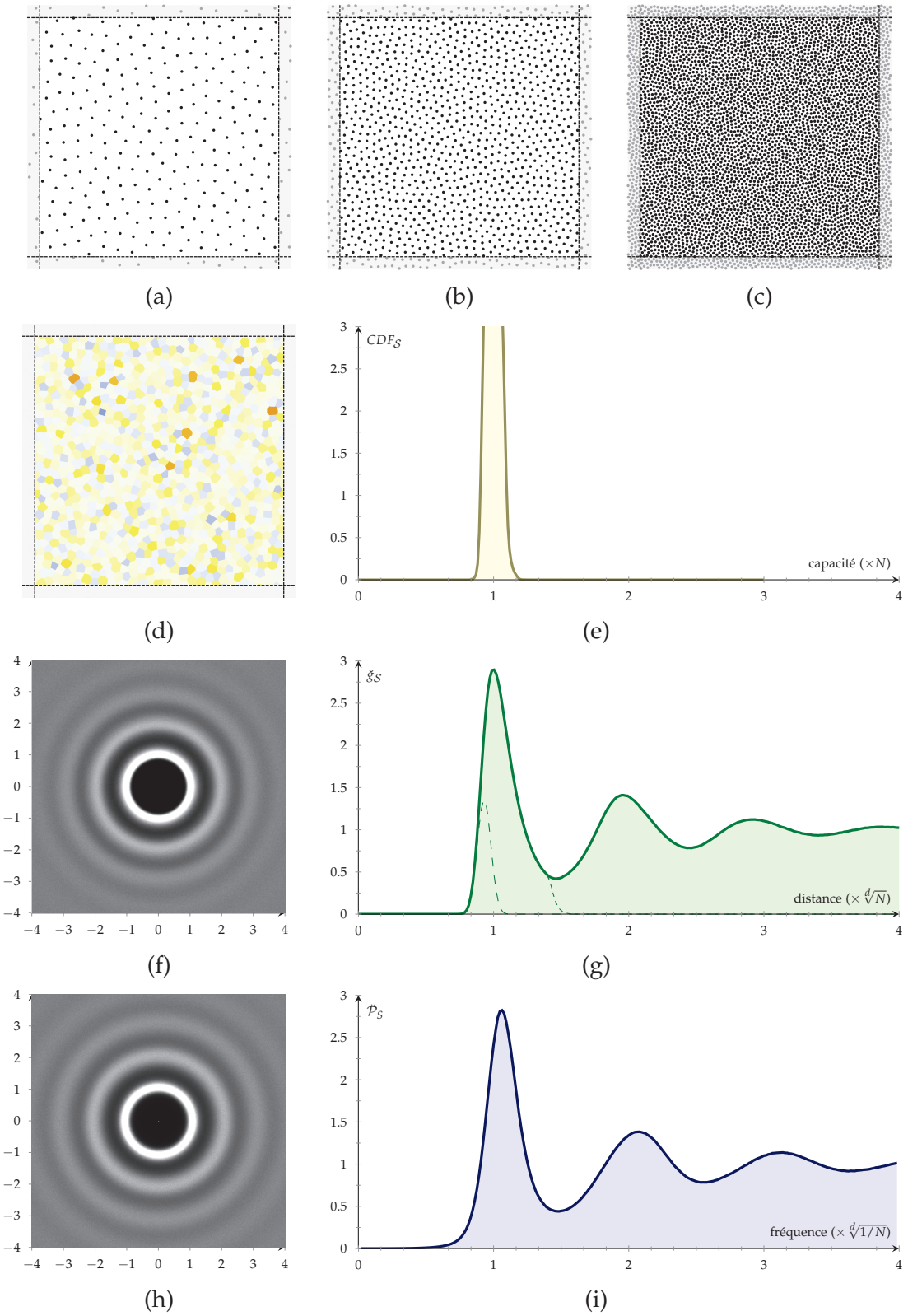


FIGURE 8.9 – *Analyse spatiale et spectrale des ensembles de points optimisés par BNOT.* ***(a), (b) et (c) :*** *Des réalisations de l'ensemble de points contenant respectivement* 256, 1024 *et* 4096 *points.* ***(d) :*** *La distribution spatiale des capacités de chaque point dans une ensemble de* 1024 *points.* ***(e) :*** *La fonction distribution de capacité.* ***(f) :*** *La fonction de distribution différentielle.* ***(f) :*** *La fonction de distribution de distance.* ***(g) :*** *Le spectre de puissance moyen.* ***(h) :*** *Le spectre de puissance moyen radial.*

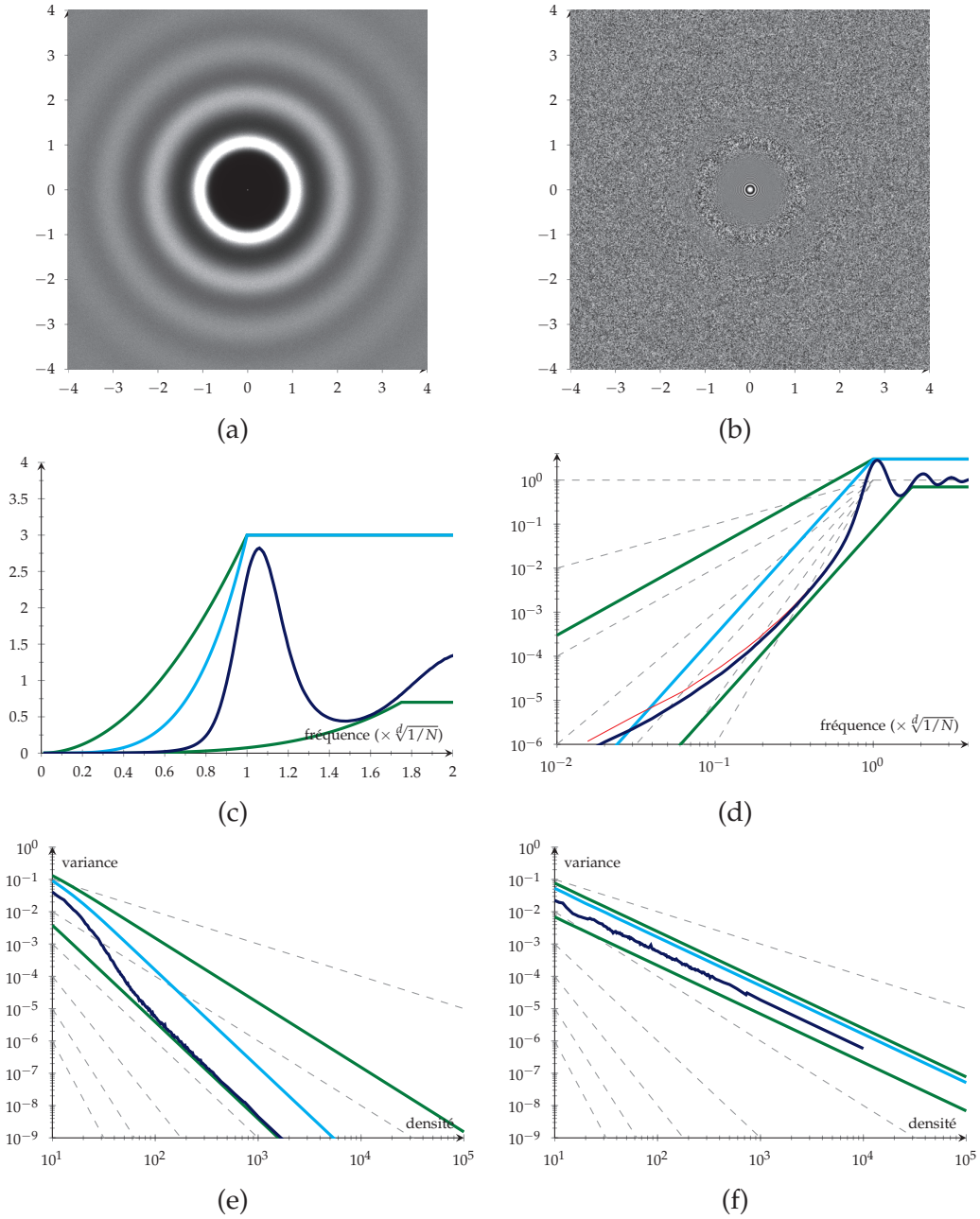


FIGURE 8.10 – *Test d'intégration avec les ensembles de points optimisés par BNOT.* ***(a) :*** *Le spectre de puissance moyen.* ***(b) :*** *Le test* zoneplate. ***(c) et (d) :*** *Le spectre de puissance moyen radial, en échelle normale et logarithmique avec en bleu, celui de l'ensemble de points avec $N = 10000$, en rouge, celui de l'ensemble de points avec $N = 4096$, en vert, les bornes supérieur et inférieur du spectre de puissance, en cyan, une borne supérieur alternative, et en gris, des repères correspondant des profiles spectraux polynomiaux de différents degrés.* ***(e) et (f) :*** *La variance de l'intégration d'une gaussienne et d'un disque avec en bleu, la variance générée par l'ensemble de points avec différentes densités, en vert, la variance encadrée, en cyan, la variance en utilisant la borne supérieur alternative, et en gris, des repères correspondant à différentes vitesses de convergence.*

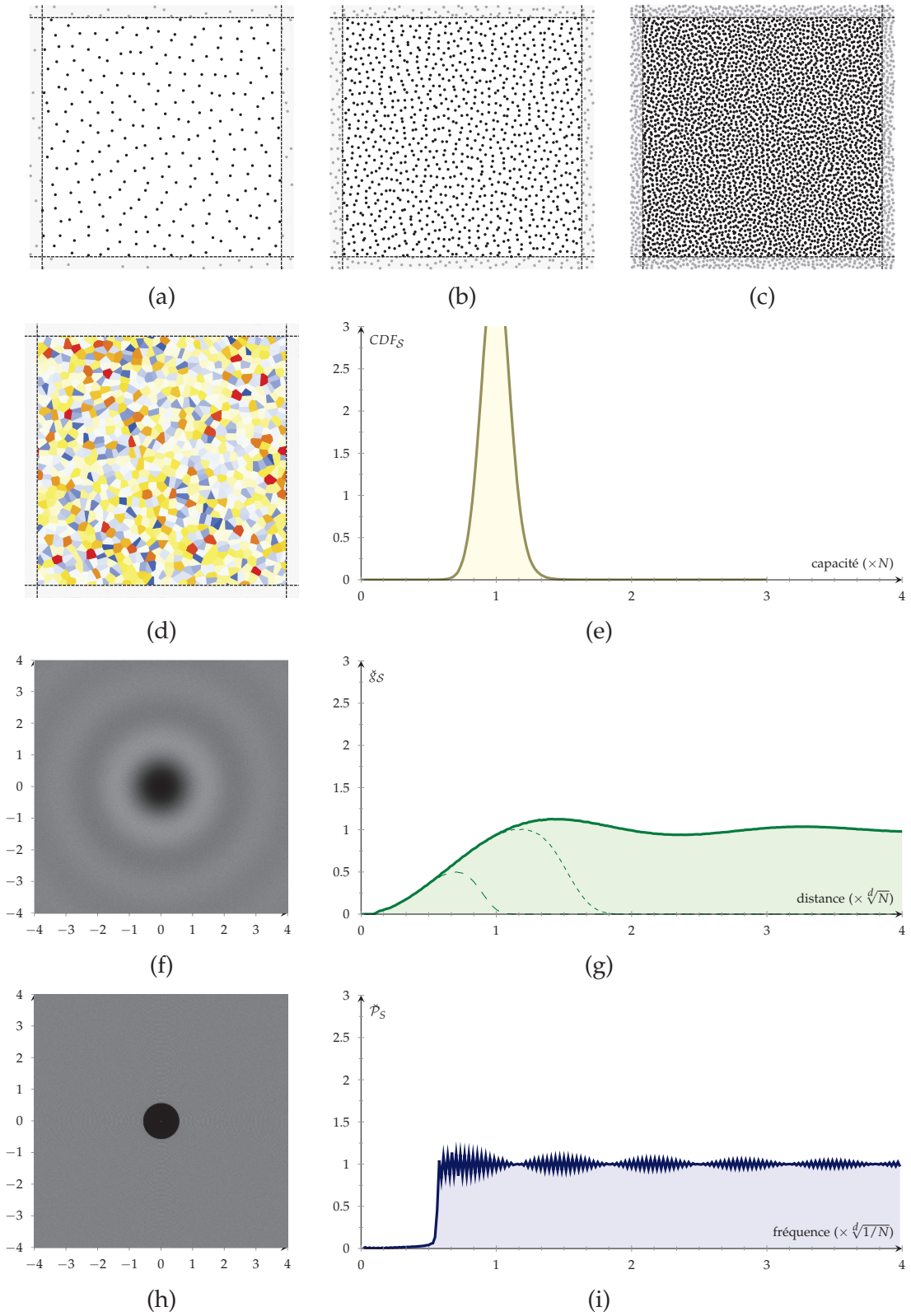


FIGURE 8.11 – *Analyse spatiale et spectrale des ensembles de points* step blue noise. ***(a), (b) et (c) :*** *Des réalisations de l'ensemble de points contenant respectivement* 256, 1024 *et* 4096 *points.* ***(d) :*** *La distribution spatiale des capacités de chaque point dans une ensemble de* 1024 *points.* ***(e) :*** *La fonction distribution de capacité.* ***(f) :*** *La fonction de distribution différentielle.* ***(f) :*** *La fonction de distribution de distance.* ***(g) :*** *Le spectre de puissance moyen.* ***(h) :*** *Le spectre de puissance moyen radial.*

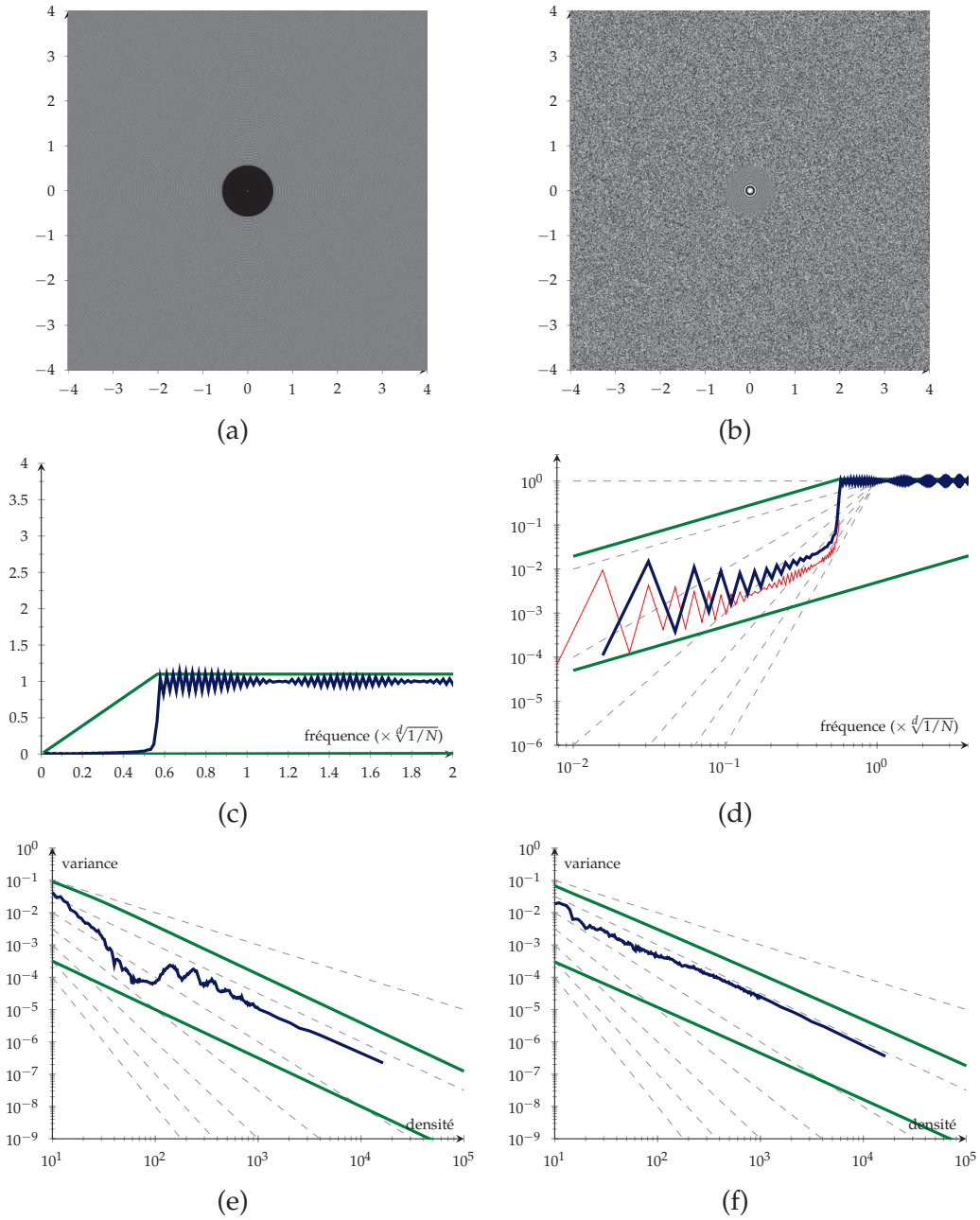


FIGURE 8.12 – *Test d'intégration avec les ensembles de points* step blue noise. ***(a) :*** *Le spectre de puissance moyen.* ***(b) :*** *Le test* zoneplate. ***(c) et (d) :*** *Le spectre de puissance moyen radial, en échelle normale et logarithmique avec en bleu, celui de l'ensemble de points avec $N = 16384$, en rouge, celui de l'ensemble de points avec $N = 4096$, en vert, les bornes supérieur et inférieur du spectre de puissance, et en gris, des repères correspondant des profiles spectraux polynomiaux de différents degrés.* ***(e) et (f) :*** *La variance de l'intégration d'une gaussienne et d'un disque avec en bleu, la variance générée par l'ensemble de points avec différentes densités, en vert, la variance encadrée, et en gris, des repères correspondant à différentes vitesses de convergence.*

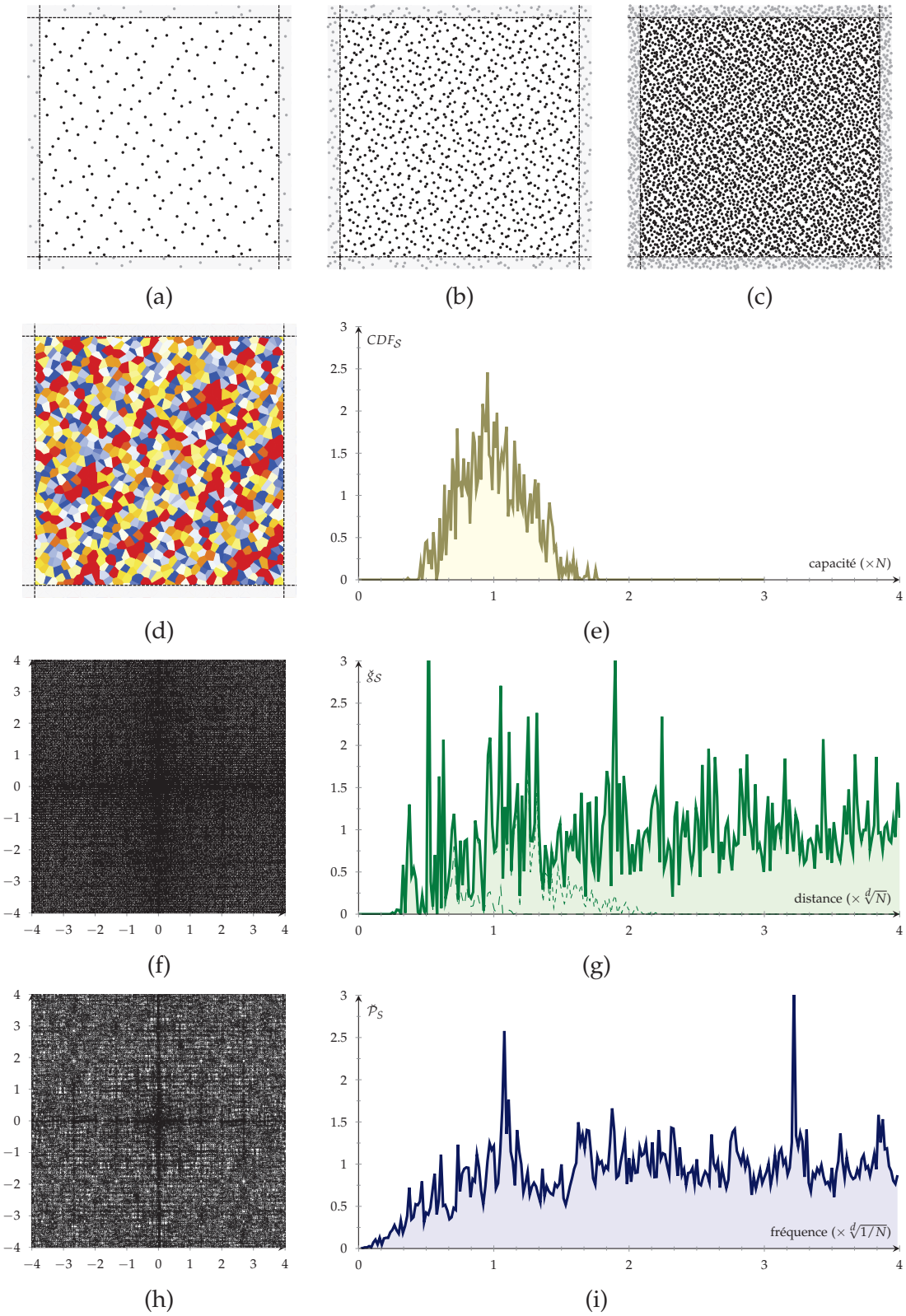


FIGURE 8.13 – *Analyse spatiale et spectrale de la séquence de points de Halton.* ***(a), (b) et (c) :*** *Des réalisations de l'ensemble de points contenant respectivement* 256, 1024 *et* 4096 *points.* ***(d) :*** *La distribution spatiale des capacités de chaque point dans une ensemble de* 1024 *points.* ***(e) :*** *La fonction distribution de capacité.* ***(f) :*** *La fonction de distribution différentielle.* ***(f) :*** *La fonction de distribution de distance.* ***(g) :*** *Le spectre de puissance moyen.* ***(h) :*** *Le spectre de puissance moyen radial.*

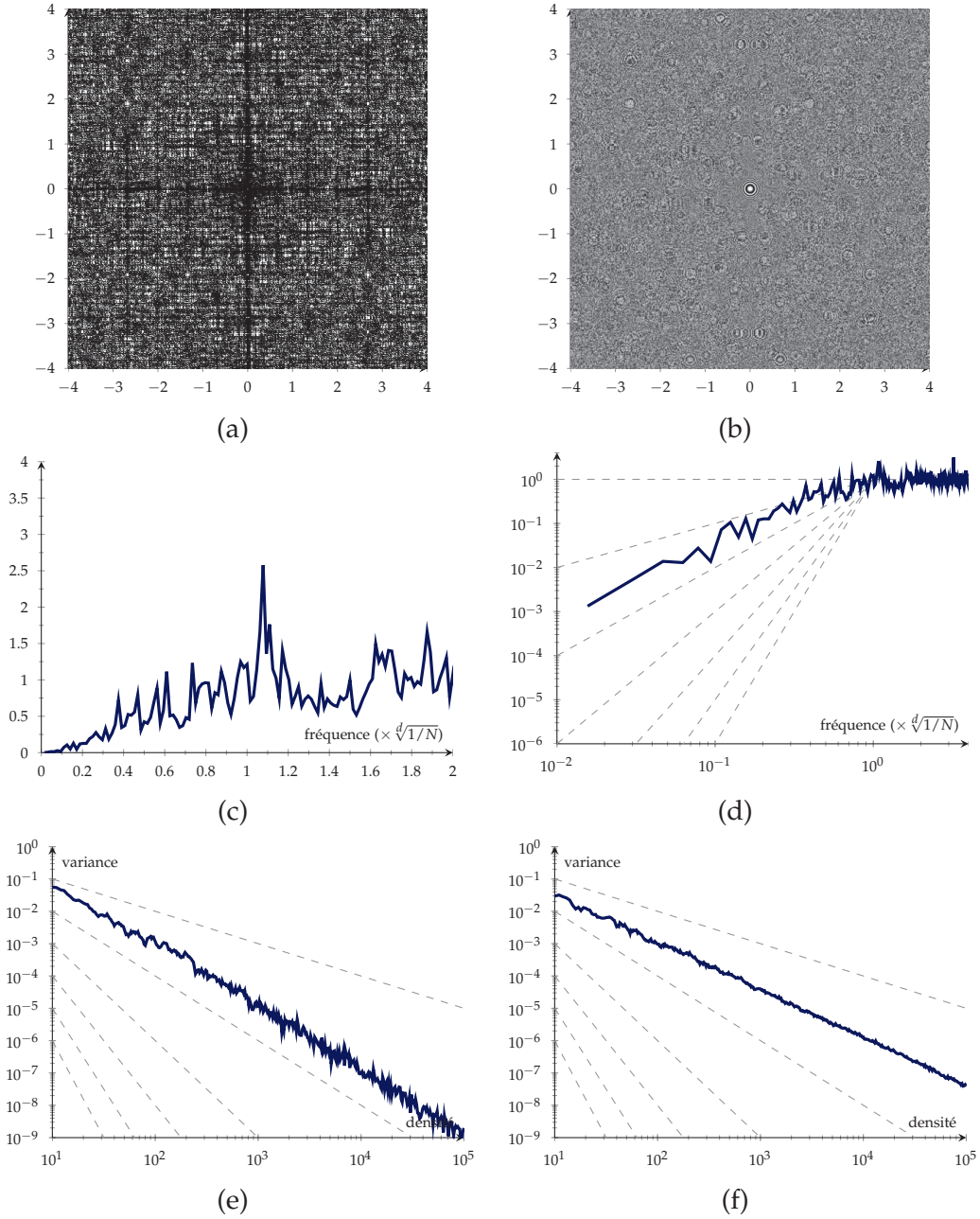


FIGURE 8.14 – *Test d'intégration avec la séquence de Halton.* ***(a) :*** *Le spectre de puissance.* ***(b) :*** *Le test* zoneplate. ***(c) et (d) :*** *Le spectre de puissance radial, en échelle normale et logarithmique avec en bleu, celui de l'ensemble de points avec* $N = 4096$ *et en gris, des repères correspondant des profiles spectraux polynomiaux de différents degrés.* ***(e) et (f) :*** *La variance de l'intégration d'une gaussienne et d'un disque avec en bleu, la variance générée par l'ensemble de points avec différentes densités et en gris, des repères correspondant à différentes vitesses de convergence.*

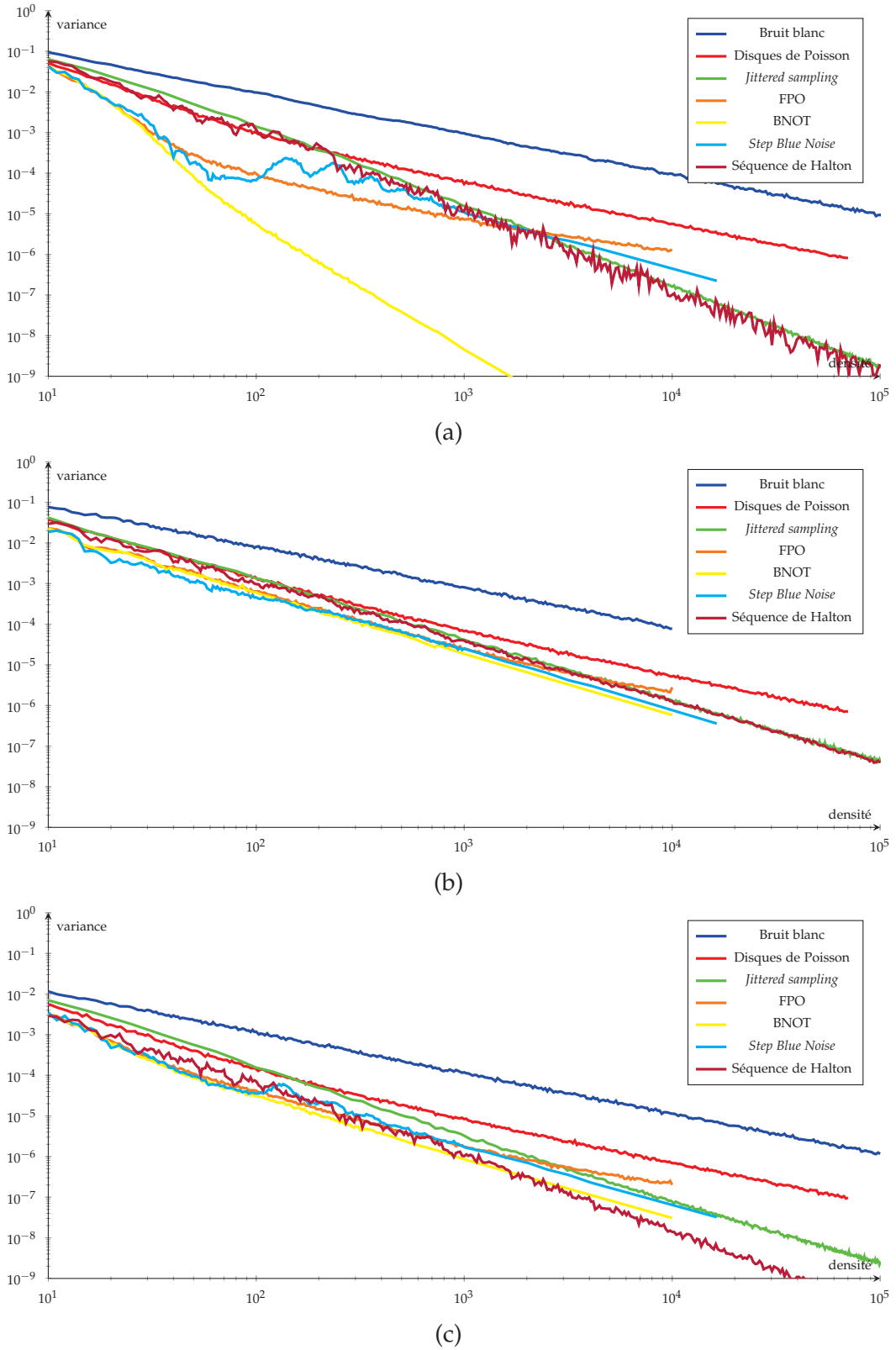


FIGURE 8.15 – *Variance de l'intégration pour différentes méthodes et différentes intégrandes.* ***(a) :*** *Intégration d'une gausienne.* ***(b) :*** *Intégration d'un disque.* ***(c) :*** *Intégration d'une image HDR.*

Conclusion

Dans ce manuscrit, nous avons présenté une approche originale pour formaliser les ensembles de points dans le but de les relier à l'intégration Monte Carlo. En modélisant de façon abstraite l'espace d'échantillonnage, nous avons pu démontrer un certain nombre de propriétés sur les ensembles de points et les méthodes d'échantillonnage. Nos contributions complètent les connaissances actuelles dans le domaine, soit en généralisant des résultats existants à d'autres domaines d'échantillonnage, soit en mettant en lumière de nouvelles relations entre les ensembles de points et leurs outils d'analyse.

Notre première contribution, détaillée dans le chapitre 3, consiste en une nouvelle formalisation de l'espace d'échantillonnage. Nous avons proposé ce formalisme dans le but de mieux comprendre l'intégration Monte Carlo. Ainsi, l'espace d'échantillonnage incorpore tous les éléments nécessaires pour l'étude des ensembles de points homogènes. Il comprend aussi les pré-requis nécessaires pour utiliser les outils d'analyse que nous avons définis. Ces outils d'analyse, détaillés dans les chapitres 5 et 6, forment la deuxième contribution présentée dans ce manuscrit. Ils permettent d'étudier les ensembles de points à travers des points de vue différents et de révéler leurs propriétés fondamentales. Rassemblés ensembles dans un formalisme commun, ces outils peuvent être utilisés par des méthodes d'optimisation, à la manière des méthodes d'échantillonnage par contrôle spectral. L'expression de la variance de l'intégration en fonction du spectre de puissance de l'ensemble de points et de l'intégrande, rendue possible par notre formalisme et les outils d'analyse que nous avons exposés, constitue notre troisième contribution. Cette expression joue un rôle particulier dans ce manuscrit. C'est sur elle que repose l'analyse de la vitesse de convergence de l'intégration Monte Carlo que nous avons menée. C'est elle aussi qui nous a permis de définir des critères explicites pour qu'un ensemble de points soit optimal pour l'intégration. Finalement, notre dernière contribution repose sur une analyse méthodique de l'état de l'art, à la fois dans le domaine spatial et le domaine spectral. Nous avons également, à la lumière de nos précédentes contributions, étudié la vitesse de convergence de l'intégration Monte Carlo pour ces différentes méthodes.

Si notre approche peut sembler parfois trop formelle, elle a, à notre sens, fait ses preuves en permettant d'exprimer de façon indépendante du domaine d'échantillonnage la variance de l'intégration. Nous avons ainsi pu donner une expression de la variance dans l'espace euclidien, l'espace hyper-cubique périodique et l'espace sphérique. Notre travail a aussi des conséquences pratiques puisqu'il montre, par une étude théorique et une validation expérimentale, que des méthodes d'échantillonnage comme les disques de Poisson et les ensembles de points *jittered* ont des comportements bien différents. Il apparaît clairement que si le nombre d'échantillons est faible, la première méthode est satisfaisante. Mais lorsque le nombre d'échantillons est suffisamment élevé, c'est les ensembles de points *jittered* qui sont les plus performants.

8.11 Limitations et perspectives

Relation avec l'*Aliasing* L'*Aliasing* n'est pas abordé dans ce manuscrit. Ce type d'erreur d'intégration ne se manifeste que lorsque l'on compare le résultat de plusieurs intégrations. Ainsi, l'*aliasing* apparaît si le motif d'échantillonnage utilisé pour intégrer un pixel est corrélé à ceux utilisés pour les pixels voisins. Il en résulte que l'erreur du pixel est corrélée à l'erreur des autres pixels, faisant apparaître des artefacts. Sous cette forme, l'*aliasing* peut donc être traité par notre formalisme. Toutefois, pour rester proche des réalités propres au rendu d'images, il serait nécessaire d'introduire également les notions de fenêtrage et de reconstruction. Or la diversité des méthodes utilisées rend cette tache difficile. La recherche d'une expression du même type celle de la variance en fonction du spectre de puissance, mais portant cette fois sur l'erreur de reconstruction, nous semble un objectif important dans ce domaine.

Prise en compte de l'intégrande Notre formalisme suppose que l'intégrande est totalement inconnue. Or, dans de nombreux cas, il est possible d'en connaître certaines informations, comme la position des discontinuités ou bien encore les zones où les hautes fréquences du signale sont concentrées. Les ensembles de points doivent alors s'adapter à ces informations. Deux approches sont possibles. Soit les outils d'analyse que nous avons introduits doivent être modifiés pour intégrer ces données, en quantifiant à quel point les ensemble de points sont proches des propriétés souhaitées. Soit l'espace d'échantillonnage est modifié de façon à ce qu'il contienne déjà en lui les propriétés de l'intégrande. Cette dernière méthode est souvent utilisée, par exemple à l'aide des opérateurs laplaciens. Elle présente l'avantage de garder le même formalisme et de faire abstraction de ces informations dans la génération des ensembles de points. La bonne solution se trouve probablement entre les deux. Une étude méthodique du type d'information que l'on peut avoir sur l'intégrande, et la façon dont elle peut être traduite théoriquement, pourrait permettre de définir un formalisme plus général que celui présenté ici.

Classe de fonctions pour le rendu Les classes de fonctions que nous avons définies sont très générales. Une étude approfondie des caractéristiques spectrales des intégrandes doit être faite dans le cadre du rendu graphique. La dimensionnalité de ces intégrandes est élevée car il faut prendre en compte les réflexions mutuelles entre objets et donc simuler plusieurs rebonds pour chaque photon. Lorsque l'on simule deux rebonds, il faut intégrer le pixel, ainsi que deux espaces hémisphériques. La dimension devient ainsi égale à 6. En prenant en compte d'autres effets comme le flou cinétique et les effets de lentilles, on arrive facilement à plusieurs dizaines de dimensions. L'analyse spectrale de cette intégrande multidimensionnelle doit permettre d'identifier les fréquences qui sont importantes pour pouvoir définir les priorités du motif d'échantillonnage.

Réalisation d'un motif d'échantillonnage optimal Comme nous l'avons dit, les conditions nécessaires et suffisantes pour qu'un ensemble de points soit réalisable ne sont actuellement pas connues. Il faut donc essayer de comprendre les limites de réalisabilité des ensembles de points. Une façon d'approcher le problème est d'essayer de réaliser de façon exacte des ensembles de points comme *Step Blue Noise*. Même si l'algorithme générant cet ensemble de points se trouve être inefficace, le simple fait

de pouvoir étudier expérimentalement ce type d'ensemble de points serait déjà une contribution importante.

Relation entre les cellules de Voronoi et l'ensemble de points Les cellules de Voronoi sont un outil d'analyse très utilisé car elles sont simples à calculer. Mais les liens entre les propriétés de ces cellules (aires, formes, répartition) et les autres outils d'analyse sont encore flous. On ne sait pas, par exemple, calculer à partir d'un spectre de puissance, la distribution des capacités. Ce type de lien serait pourtant très utile, car il donnerait un moyen algorithmique puissant pour le contrôle spectral. En effet, les méthodes actuelles se basent sur la fonction de distribution différentielle, qui n'est pas la fonction la plus pratique pour l'optimisation. Une approche basée sur les cellules de Voronoi permettrait d'améliorer la qualité et la rapidité de ces méthodes.

Index terminologique

Step Blue Noise . 110
σ-algèbre . 12

Action de groupe . 20
Action transitive . 20
Aliasing . 6, 43
Auto-corrélation . 53

Base orthogonale . 65
Base orthonormale . 66
Blue Noise through Optimal Transport **(BNOT)** . 39, 109
Borélien . 12
Boule ouverte . 19
Bruit blanc . 29, 58, 101

Capacity-Constrained Voronoi Tesselation **(CCVT)** . 38
Capacité . 38
Constante de Rényi . 32
Contrôle spectral . 39, 41, 77, 110
Covariance . 18

Densité . 45
Densité de n-uplets . 48
Densité de distances . 50
Densité de paires . 47
Densité différentielle . 49
Densité locale . 46, 77
Diagramme de Voronoi . 36, 38, 54
Discrépance . 42, 56, 82
Discrépance extrême . 57
Discrépance isotrope . 57
Discrépance stochastique . 58
Discrépance à l'origine . 56
Disques de Poisson . 30, 42, 104

Distance minimale moyenne . 54
Distribution de Poisson . 29
Domaine spatial . 61
Domaine spectral . 62

Échantillonneur . 18
Échantillonneur déterministe . 18
Échantillonneur stochastique . 18, 58
Ensemble compact . 12
Ensemble de points . 17
Ensemble de points homogène . 26, 46, 50, 59, 75, 81, 85, 88
Ensemble de points hypercube Latin . 35
Ensemble de points *jittered* . 34, 105
Ensemble métrisable . 19
Espace d'échantillonnage . 19, 20
Espace d'échantillonnage discret régulier . 24
Espace d'échantillonnage euclidien . 22, 72
Espace d'échantillonnage hexagonal périodique . 23
Espace d'échantillonnage hyper-cubique périodique . 22, 72
Espace d'échantillonnage hyper-cubique unitaire . 24
Espace d'échantillonnage hémisphérique . 25
Espace d'échantillonnage restreint . 24
Espace d'échantillonnage sphérique . 23, 73
Espace de Hausdorff . 12
Espace de Hilbert . 64, 66
Espace homogène . 20
Espace mesurable . 13
Espace pré-hilbertien . 64
Espace topologique localement compact . 12
Espace vectoriel . 64
Espérance . 18

***Farthest-Point Optimization* (FPO)** . 39, 54, 108
Fonction de corrélation totale . 52
Fonction de densité de probabilité . 18, 46
Fonction de distribution de capacité . 55
Fonction de distribution de distance . 40, 50, 54
Fonction de distribution de paires . 48
Fonction de distribution différentielle . 50, 53, 77, 100

Fonction représentative . 17
Fonction étagée . 13

Grille régulière . 33
Groupe . 20
Groupe topologique . 20

Harmonique sphérique . 73
HEALPix . 34

Intégrale . 13
Intégration Monte Carlo . 5, 42, 79
Intégration Quasi Monte Carlo . 5

L'inégalité de Koksma-Hlawka . 82

Mesure . 13
Mesure de Haar . 25
Métrique . 19

Opération unitaire . 65

Parking de Rényi . 32
Pavage . 33
Pavage barycentrique de Voronoi . 36
Pavage de Penrose . 41
Pavage de Wang . 41
Pavage hiérarchique . 40
Pavage régulier . 33
Problème de Thomson . 34
Proto-tuile . 33

Relaxation de Lloyd . 36
Représentation . 68

Sequence de Halton . 43, 110
Sequence de Niederreiter . 43
Sequence de Sobol . 43
Somme directe . 66
Sous-espace . 65
Sous-espace irréductible . 69
Sous-espace stable . 69
Spectre de puissance . 63, 71, 77, 90, 100

Séquence à basse discrépance . 42, 84

Table de déviations . 40
Test de Kolmogorov-Smirnov . 56
Test *zoneplate* . 111
Théorème de Parseval . 63, 67
Théorème de Plancherel . 63
Topologie . 12
Transformée de Fourier . 61, 65, 71, 73, 74
Transport optimal . 39
Triangulation de Delaunay . 54
Tuile . 33

Variance . 19, 58
Variation de Hardy-Krause . 83
Variation de Vitali . 83
Variation totale . 82

Widesense stationary process . 27

Table des symboles

$H^1 \oplus H^2$ Somme directe entre le espaces H^1 et H^2. 66

$M^\perp$ Ensemble orthogonal de M. 65

$P_\mathcal{S}$ Profil spectral de l'ensemble de points. 91

$Bias(I_{\mathcal{S},f})$ Biais de l'intégration Monte Carlo $I_{\mathcal{S},f}$. 81

$f(x) \in O_{x\to a}(g(x))$ Notation indiquant que f se comporte comme g quand x tend vers a. 16

$CDF_\mathcal{S}$ Fonction de distribution de capacité de l'ensemble de points $\mathcal{S}$. 55

i Unité imaginaire, tel que $i^2 = -1$. 11

$\mathbb{C}$ Ensemble des nombres complexes. 11

$Cov(...,...)$ Covariance par rapport aux réalisations de l'ensemble de points. 18

$\check{g}_\mathcal{S}$ Fonction de distribution différentielle de l'ensemble de points $\mathcal{S}$. 50

$\check{g}_\mathcal{S}$ Fonction de distribution de distances de l'ensemble de points $\mathcal{S}$. 50

$g_\mathcal{S}$ Fonction de distribution de paires de l'ensemble de points $\mathcal{S}$. 48

$:=$ Définition. 11

$\rho_\mathcal{S}$ Densité de l'ensemble de points $\mathcal{S}$. 45

$\check{\rho}^{(2)}_\mathcal{S}$ Densité différentielle de l'ensemble de points $\mathcal{S}$. 49

$\check{\rho}_\mathcal{S}$ Densité de distances de l'ensemble de points $\mathcal{S}$. 50

$\rho^{(1)}_\mathcal{S}$ Densité locale de l'ensemble de points $\mathcal{S}$. 46

$\rho^{(2)}_\mathcal{S}$ Densité de paires de l'ensemble de points $\mathcal{S}$. 47

$\rho^{(n)}_\mathcal{S}$ Densité de n-uplets de l'ensemble de points $\mathcal{S}$. 48

$\mathfrak{M}_{\textbf{Determ.}}$ Classe des ensembles de points déterministes. 18

τ_0 Action neutre de $\mathbb{T}_\mathcal{E}$. 20

$\mathcal{L}^d$ Espace d'échantillonnage discret régulier. 24

$Err(I_{\mathcal{S},f})$ Erreur de l'intégration Monte Carlo $I_{\mathcal{S},f}$. 80

$\mathcal{E}^d$ Espace d'échantillonnage euclidien de dimension d. 22

$D(\mathcal{S})$ Discrépance extrême de l'ensemble de points $\mathcal{S}$. 57

$\mathcal{F}_f$ Transformée de Fourier d'une fonction f. 61, 65

$\mathcal{A}_f$ Amplitude de la transformée de Fourier de la fonction f. 62

$\mathbb{F}$ Ensemble des fréquences de l'espace spectral. 65

$\mathcal{I}_f$ Partie imaginaire de la transformée de Fourier de la fonction f. 62

Φ_f Phase de la transformée de Fourier de la fonction f. 62

$\mathcal{R}_f$ Partie réelle de la transformée de Fourier de la fonction f. 62

ω Variable fréquentielle. 61

ω_0 Fréquence zéro, qui représente l'intégrale du signal. 67

Φ^s Classe de fonctions dont le spectre de puissance décroit selon un facteur s. 93

$\mathcal{HS}^2$ Espace d'échantillonnage hémisphérique. 25

$\mathcal{H}^2$ Espace d'échantillonnage hexagonal. 23

$\mathcal{B}$ Base orthogonale d'une espace pré-hilbertien. 65, 66

b^ω Vecteur de la base orthogonale $\mathcal{B}$. 65, 66

$\mathfrak{M}_{\mathbf{Homo.}}$ Classe des ensembles de points homogènes. 27

$\mathbb{Z}$ Ensemble des nombres entiers relatifs. 11

f Intégrande. 79

I_f Intégrale de l'intégrande f. 79

$I_{\mathcal{S},f}$ Intégration Monte Carlo de l'intégrande f avec l'ensemble de points $\mathcal{S}$. 80

$J(\mathcal{S})$ Discrépance isotrope de l'ensemble de points $\mathcal{S}$. 57

$\mathfrak{M}_{\mathbf{LHC}}$ Classe des ensembles de points hypercube Latin. 35

$MSE(I_{\mathcal{S},f})$ Erreur quadratique moyenne de l'intégration Monte Carlo $I_{\mathcal{S},f}$. 80

$\langle ... \rangle$ Espérance par rapport aux réalisations de l'ensemble de points. 18

$\delta_{\mathcal{S}}^{\mathbf{min}}$ Distance minimale moyenne. 54

$\mathbb{N}$ Ensemble des nombres entiers naturels. 11

N Nombre de points de l'ensemble $\mathcal{S}$. 17

$\psi_A(\mathcal{S})$ Nombre de points de l'ensemble de points $\mathcal{S}$ appartenant au sous-domaine A. 56

$PDF_{...}(...)$ Fonction de densité de probabilité d'une variable aléatoire, par rapport aux réalisations de l'ensemble de points. 18

Π Fonction valant 1 sur l'interval $[-1/2, 1/2]$, sinon 0. 15

$\mathfrak{M}_{\mathbf{P.Disk}}$ Classe des ensembles de disques de Poisson. 32

$\mathbb{P}$ Domaine de définition du spectre de puissance, ensemble indexant les sous espace stables et irréductibles d'un espace pré-hilbertien. 69

$\mathcal{P}_f$ Spectre de puissance de la fonction f. 63

σ_k Proto-tuile de la k-ième tuile d'un pavage. 33

$\mathbf{s}_k$ Variable aléatoire représentant le k-ième point de l'ensemble $\mathcal{S}$. 17

$\mathcal{S}$ Variable aléatoire représentant l'ensemble de points. 17

$\mathcal{S}[m]$ m-ième réalisation de l'ensemble de points $\mathcal{S}$. 18

$\mathbb{R}$ Ensemble des nombres réels. 11

$\mathfrak{M}_{\mathbf{Regular}}$ Classe des ensembles de points générés par pavage régulier. 34

R Constante de Rényi. 32

S Fonction représentative de l'ensemble $\mathcal{S}$. 17

$\mathbb{P}_{\mathcal{S}}$ Loi de probabilité associée à la variable aléatoire $\mathcal{S}$. 18

$\mathcal{E}$ Espace dans lequel évoluent les points d'un ensemble de points. 19

$\mathbf{dist}_{\mathcal{E}}$ Fonction de distance de l'espace d'échantillonnage $\mathcal{E}$. 20

$\mathbb{D}_{\mathcal{E}}$ Domaine de l'espace $\mathcal{E}$. 17, 20

$Span(\mathcal{B})$ Ensemble des vecteurs engendrés par la combinaison linéaire des vecteurs appartenant à $\mathcal{B}$. 65

x Variable spatiale appartenant à $\mathbb{D}_{\mathcal{E}}$. 17

$\mathcal{S}^2$ Espace d'échantillonnage sphérique. 23

$D^*(\mathcal{S})$ Discrépance à l'origine de l'ensemble de points $\mathcal{S}$. 56

$\mathfrak{M}_{\mathbf{Stocha.}}$ Classe des ensembles de points stochastiques. 19

$\mathfrak{M}_{\mathbf{Jittered}}$ Classe des ensembles de points *jittered*. 34

$h_{\mathcal{S}}$ Fonction de corrélation totale de l'ensemble de points $\mathcal{S}$. 52

$\hat{\sigma}_k$ k-ième tuile d'un pavage. 33

$\mathfrak{M}_{\mathbf{Pavage}}$ Classe des ensembles de points générés par pavage. 33

$\mathcal{T}^d$ Espace d'échantillonnage hyper-cubique périodique de dimension d. 22

$\mathcal{V}(f)$ Variation totale de la fonction f. 82

$\mathbb{T}_{\mathcal{E}}$ Groupe des actions de l'espace d'échantillonnage $\mathcal{E}$. 20

$\mathcal{U}^d$ Espace d'échantillonnage hyper-cubique unitaire. 24

$Var(...)$... Variance par rapport aux réalisations de l'ensemble de points. 19

$V_{HK}(f)$ Variation de Hardy-Krause de la fonction f. 83

v_k Cellule de Voronoi du k-ième point de l'ensemble de points. 54

$\mathfrak{M}_{\mathbf{WN}}$ Classe des ensembles de points bruit blanc. 29

χ_A Fonction caractéristique de A. 13, 56

δ Fonction δ de Dirac. 15

$\emptyset$ Ensemble vide. 11

$\mathcal{V}_{\mathbf{Vitali}}(f)$ Variation de Vitali de la fonction f. 83

μ_H Mesure de Haar. 25

ν_F Mesure de l'espace de Fourier. 71

ϕ Représentation d'un groupe. 68

b Degré du profil spectral de $\mathcal{S}$ dans les basses fréquences. 95

$f \otimes g$ Convolution entre les fonctions f et g. 62

$f \oplus g$ Corrélation entre les fonctions f et g. 53

$f \perp g$ Condition d'orthogonalité entre f et g. 65

Bibliographie

[Balzer *et al.*, 2009] Michael Balzer, Thomas Schlömer, and Oliver Deussen. Capacity-constrained point distributions : A variant of Lloyd's method. *ACM Trans. on Graphics*, 28(3) :86 :1–8, 2009. (Cité page 38.)

[Barut and Rączka, 1986] A.O. Barut and R. Rączka. *Theory of Group Representations and Applications*. World Scientific, 1986. (Cité pages 12, 25, 68 et 69.)

[Bonnier *et al.*, 1994] B Bonnier, D Boyer, and P Viot. Pair correlation function in random sequential adsorption processes. *Journal of Physics A : Mathematical and General*, 27(11) :3671, 1994. (Cité page 104.)

[Bracewell, 1999] Ronald Bracewell. *The Fourier Transform & Its Applications*. McGraw-Hill Science, 3 edition, 1999. (Cité pages 15, 61 et 77.)

[Brandolini *et al.*, 2001] Luca Brandolini, Leonardo Colzani, Andrea Torlaschi, et al. Mean square decay of fourier transforms in euclidean and non euclidean spaces. *Tohoku Mathematical Journal*, 53(3) :467–478, 2001. (Cité pages 93 et 94.)

[Brandolini *et al.*, 2003] L Brandolini, S Hofmann, and A Iosevich. Sharp rate of average decay of the fourier transform of a bounded set. *Geometric & Functional Analysis GAFA*, 13(4) :671–680, 2003. (Cité pages 93 et 94.)

[Brandolini *et al.*, 2013] Luca Brandolini, Leonardo Colzani, Giacomo Gigante, and Giancarlo Travaglini. On the koksma–hlawka inequality. *Journal of Complexity*, 29(2) :158–172, 2013. (Cité page 82.)

[Bredon, 1972] Glen E Bredon. *Introduction to compact transformation groups*. Elsevier, 1972. (Cité page 20.)

[Chen *et al.*, 2012] Zhonggui Chen, Zhan Yuan, Yi-King Choi, Ligang Liu, and Wenping Wang. Variational blue noise sampling. *Visualization and Computer Graphics, IEEE Transactions on*, 18(10) :1784–1796, 2012. (Cité page 39.)

[Chow, 2000] Tai L Chow. *Mathematical Methods for Physicists : A concise introduction*. Cambridge University Press, 2000. (Cité page 101.)

[Cline *et al.*, 2009] D. Cline, S. Jeschke, K. White, A. Razdan, and P. Wonka. Dart throwing on surfaces. In *Proc. EGSR '09*, pages 1217–1226. Eurographics Association, 2009. (Cité page 32.)

[Cook, 1986] Robert L. Cook. Stochastic sampling in computer graphics. *ACM Trans. Graph.*, 5(1) :51–72, 1986. (Cité pages 30 et 34.)

[Cornwell, 1997] John F Cornwell. *Group theory in physics : An introduction*, volume 1. Academic press, 1997. (Cité page 20.)

[de Goes *et al.*, 2012] F. de Goes, K. Breeden, V. Ostromoukhov, and M. Desbrun. Blue noise through optimal transport. *Proc. SIGGRAPH Asia '12*, 31 :171 :1–171 :10, 2012. (Cité pages 39 et 109.)

[Dippé and Wold, 1985] Mark A. Z. Dippé and Erling Henry Wold. Antialiasing through stochastic sampling. In *ACM SIGGRAPH*, pages 69–78, July 1985. (Cité pages 27 et 34.)

[Du *et al.*, 1999] Qiang Du, Vance Faber, and Max Gunzburger. Centroidal voronoi tessellations : applications and algorithms. *SIAM review*, 41(4) :637–676, 1999. (Cité page 38.)

[Dunbar and Humphreys, 2006] Daniel Dunbar and Greg Humphreys. A spatial data structure for fast poisson-disk sample generation. In *ACM Transactions on Graphics (TOG)*, volume 25, pages 503–508. ACM, 2006. (Cité page 32.)

[Durand, 2011] Fredo Durand. A frequency analysis of Monte-Carlo and other numerical integration schemes. *MIT CSAIL Technical report TR-2011-052*, 2011. (Cité pages 7, 8, 80 et 86.)

[Dvoretzky and Robbins, 1985] A. Dvoretzky and H. Robbins. On the „parking“ problem. In T.L. Lai and D. Siegmund, editors, *Herbert Robbins Selected Papers*, pages 477–493. Springer New York, 1985. (Cité page 32.)

[Gamito and Maddock, 2009] Manuel N. Gamito and Steve C. Maddock. Accurate multidimensional poisson-disk sampling. *ACM Trans. Graph.*, 29(1) :8 :1–8 :19, December 2009. (Cité page 32.)

[Gorski *et al.*, 2005] Krzysztof M Gorski, Eric Hivon, AJ Banday, Benjamin D Wandelt, Frode K Hansen, Mstvos Reinecke, and Matthia Bartelmann. Healpix : a framework for high-resolution discretization and fast analysis of data distributed on the sphere. *The Astrophysical Journal*, 622(2) :759, 2005. (Cité page 34.)

[Groemer, 1996] Helmut Groemer. *Geometric Applications of Fourier Series and Spherical Harmonics*. Cambridge University Press, 1996. Cambridge Books Online. (Cité page 73.)

[Grunbaum and Shephard, 1986] Branko Grunbaum and Geoffrey C Shephard. *Tilings and patterns*. W.H. Freeman & Company, 1986. (Cité pages 23, 33, 40 et 42.)

[Halton, 1960] John H Halton. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik*, 2(1) :84–90, 1960. (Cité page 43.)

[Hansen and McDonald, 1990] Jean-Pierre Hansen and Ian R McDonald. *Theory of simple liquids*. Elsevier, 1990. (Cité pages 27, 46, 47, 48 et 52.)

[Heck *et al.*, 2013] Daniel Heck, Thomas Schlömer, and Oliver Deussen. Blue noise sampling with controlled aliasing. *ACM Trans. on Graphics*, 32(3) :25 :1–25 :12, 2013. (Cité pages 39, 77, 110 et 113.)

[Hiller *et al.*, 2001] Stefan Hiller, Oliver Deussen, and Alexander Keller. Tiled blue noise samples. 2001. (Cité page 42.)

[Kazhdan *et al.*, 2015] M. Kazhdan, G. Singh, A. Pilleboue, D. Coeurjolly, and V. Ostromoukhov. Variance Analysis for Monte Carlo Integration : A Representation-Theoretic Perspective. *ArXiv e-prints*, May 2015. (Cité pages 8 et 71.)

[Kensler, 2013] Andrew Kensler. Correlated multi-jittered sampling. *Technical Report, TM-13-01, Pixar Animation Studio*, 2013. (Cité pages 35 et 44.)

[Knuth, 1968] D Knuth. *The Art of Computer Programming 1 : Fundamental Algorithms 2 : Seminumerical Algorithms 3 : Sorting and Searching*. MA : Addison-Wesley, 1968. (Cité page 30.)

[Kopf *et al.*, 2006] Johannes Kopf, Daniel Cohen-Or, Oliver Deussen, and Dani Lischinski. Recursive wang tiles for real-time blue noise. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2006)*, 25(3) :509–518, 2006. (Cité page 42.)

[Kuipers and Niederreiter, 1974] Lauwerens Kuipers and H Niederreiter. *Uniform distribution of sequences*. 1974. (Cité page 57.)

[Lagae and Dutré, 2006] Ares Lagae and Philip Dutré. An alternative for wang tiles : colored edges versus colored corners. *ACM Transactions on Graphics (TOG)*, 25(4) :1442–1459, 2006. (Cité page 42.)

[Lee, 2000] John Marshall Lee. *Introduction to topological manifolds*. Graduate texts in mathematics. Springer, New York, Berlin, Heidelberg, 2000. (Cité page 12.)

[Lemieux, 2009] Christiane Lemieux. *Monte Carlo and Quasi Monte Carlo Sampling*. Springer, 2009. (Cité pages 56, 82, 83 et 104.)

[Liflyand and Stadtmüller, 2013] E Liflyand and U Stadtmüller. A multidimensional euler-maclaurin formula and an application. *Complex Analysis and Dynamical Systems V*, 591 :183, 2013. (Cité page 83.)

[Liu *et al.*, 2009] Yang Liu, Wenping Wang, Bruno Lévy, Feng Sun, Dong-Ming Yan, Lin Lu, and Chenglei Yang. On centroidal voronoi tessellation—energy smoothness and fast computation. *ACM Transactions on Graphics (ToG)*, 28(4) :101, 2009. (Cité page 38.)

[Lloyd, 1983] SA Lloyd. An optimization approach to relaxation labelling algorithms. *Image and Vision Computing*, 1(2) :85–91, 1983. (Cité pages 36 et 38.)

[Marsaglia and others, 1972] George Marsaglia et al. Choosing a point from the surface of a sphere. *The Annals of Mathematical Statistics*, 43(2) :645–646, 1972. (Cité page 30.)

[McKay *et al.*, 1979] Michael D McKay, Richard J Beckman, and William J Conover. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2) :239–245, 1979. (Cité page 35.)

[Mitchell, 1991] Don P Mitchell. Spectrally optimal sampling for distribution ray tracing. *ACM SIGGRAPH Computer Graphics*, 25(4) :157–164, 1991. (Cité page 34.)

[Niederreiter, 1987] Harald Niederreiter. Point sets and sequences with small discrepancy. *Monatshefte für Mathematik*, 104(4) :273–337, 1987. (Cité page 43.)

[Niederreiter, 1992] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*. Society for Industrial and Applied Mathematics, 1992. (Cité pages 42, 57, 79, 81, 83 et 84.)

[Olenko, 2006] A Ya Olenko. Upper bound on $\sqrt{x}jv(x)$ and its applications. *Integral Transforms and Special Functions*, 17(6) :455–467, 2006. (Cité page 101.)

[Ostromoukhov *et al.*, 2004] Victor Ostromoukhov, Charles Donohue, and Pierre-Marc Jodoin. Fast hierarchical importance sampling with blue noise properties. In *ACM Transactions on Graphics (TOG)*, volume 23, pages 488–495. ACM, 2004. (Cité page 41.)

[Ostromoukhov, 2007] Victor Ostromoukhov. Sampling with polyominoes. In *ACM Transactions on Graphics (TOG)*, volume 26, page 78. ACM, 2007. (Cité page 41.)

[Öztireli and Gross, 2012] A. Cengiz Öztireli and Markus Gross. Analysis and synthesis of point distributions based on pair correlation. *ACM Trans. Graph.*, 31(6) :174 :1–174 :6, 2012. (Cité pages 39 et 77.)

[Penrose and Yukich, 2002] Mathew D Penrose and JE Yukich. Limit theory for random sequential packing and deposition. *Annals of Applied probability*, pages 272–301, 2002. (Cité page 32.)

[Penrose, 1974] Roger Penrose. The rôle of aesthetics in pure and applied mathematical research. *The Institute of Mathematics and its Applications Bulletin*, 10(7/8) :266–271, 1974. (Cité page 41.)

[Peyrot *et al.*, 2013] Jean-Luc Peyrot, Frédéric Payan, and Marc Antonini. Feature-preserving direct blue noise sampling for surface meshes. In *Eurographics 2013*, page 4 pages, May 2013. (Cité page 32.)

[Pilleboue *et al.*, 2015] Adrien Pilleboue, Gurprit Singh, David Coeurjolly, Michael Kazhdan, and Victor Ostromoukhov. Variance analysis for monte carlo integration. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 34(4) :124 :1–124 :14, 2015. (Cité pages 8 et 98.)

[Rudin, 1987] Walter Rudin. *Real and complex analysis*. Tata McGraw-Hill Education, 1987. (Cité pages 12, 13, 19 et 67.)

[Schlömer *et al.*, 2011] Thomas Schlömer, Daniel Heck, and Oliver Deussen. Farthest-point optimized point sets with maximized minimum distance. In *Proc. Symp. High Performance Graphics '11*, pages 135–142. ACM, 2011. (Cité pages 39, 54 et 108.)

[Schlömer and Deussen, 2010] Thomas Schlömer and Oliver Deussen. Towards a standardized spectral analysis of point sets with applications in graphics, 2010. (Cité page 100.)

[Shirley and Wang, 1994] Kenneth Chiu Peter Shirley and Changyaw Wang. Multi-jittered sampling. *Graphics gems IV*, 4 :370, 1994. (Cité page 35.)

[Singh, 2015] Gurprit Singh. *Sampling and Variance Analysis for Monte Carlo Integration in Spherical Domain*. PhD thesis, Université Claude Bernard Lyon 1, 09 2015. (Cité page 99.)

[Sobol, 1967] Ilya M Sobol. On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational mathematics and mathematical physics*, (7) :86–112, 1967. (Cité page 43.)

[Solomon, 1967] Herbert Solomon. Random packing density. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 3 : Physical Sciences*, pages 119–134, Berkeley, Calif., 1967. University of California Press. (Cité page 32.)

[Strichartz, 2003] Robert S Strichartz. *A guide to distribution theory and Fourier transforms*. World Scientific, 2003. (Cité page 93.)

[Subr and Kautz, 2013] Kartic Subr and Jan Kautz. Fourier analysis of stochastic sampling strategies for assessing bias and variance in integration. *ACM Trans. on Graphics*, 32(4) :128 :1–128 :12, 2013. (Cité pages 7, 8, 80, 86, 87 et 88.)

[Thomson, 1904] J.J. Thomson. Xxiv. on the structure of the atom : an investigation of the stability and periods of oscillation of a number of corpuscles arranged at equal intervals around the circumference of a circle ; with application of the results to the theory of atomic structure. *Philosophical Magazine Series 6*, 7(39) :237–265, 1904. (Cité page 34.)

[Torquato and Stillinger, 2003] Salvatore Torquato and Frank H Stillinger. Local density fluctuations, hyperuniformity, and order metrics. *Physical Review E*, 68(4) :041113, 2003. (Cité pages 58 et 113.)

[Torquato *et al.*, 2006] S Torquato, OU Uche, and FH Stillinger. Random sequential addition of hard spheres in high euclidean dimensions. *Physical Review E*, 74(6) :061308, 2006. (Cité pages 32 et 104.)

[Torquato, 2002] Salvatore Torquato. *Random heterogeneous materials : microstructure and macroscopic properties*, volume 16. Springer Science & Business Media, 2002. (Cité page 30.)

[Uche *et al.*, 2006] OU Uche, FH Stillinger, and S Torquato. On the realizability of pair correlation functions. *Physica A : Statistical Mechanics and its Applications*, 360(1) :21–36, 2006. (Cité page 39.)

[Ulichney, 1987] Robert Ulichney. *Digital halftoning*. MIT press, 1987. (Cité pages 7 et 112.)

[Wachtel *et al.*, 2014] Florent Wachtel, Adrien Pilleboue, David Coeurjolly, Katherine Breeden, Gurprit Singh, Gaël Cathelin, Fernando de Goes, Mathieu Desbrun, and Victor Ostromoukhov. Fast tile-based adaptive sampling with user-specified fourier spectra. *ACM Trans. Graph.*, 33(4), 2014. (Cité pages 9, 23 et 41.)

[Wachtel, 2015] Florent Wachtel. *Système de pavage pour l'échantillonnage*. PhD thesis, Université Claude Bernard Lyon 1, 5 2015. À paraitre. (Cité page 41.)

[Wieczorek and Simons, 2005] Mark A. Wieczorek and Frederik J. Simons. Localized spectral analysis on the sphere. *Geophysical Journal International*, 162(3) :655–675, 2005. (Cité page 73.)

[Yan *et al.*, 2014] Dong-Ming Yan, Jianwei Guo, Xiaohong Jia, Xiaopeng Zhang, and Peter Wonka. Blue-noise remeshing with farthest point optimization. In *Computer Graphics Forum*, volume 33, pages 167–176. Wiley Online Library, 2014. (Cité page 39.)

[Zhou *et al.*, 2012] Yahan Zhou, Haibin Huang, Li-Yi Wei, and Rui Wang. Point sampling with general noise spectrum. *ACM Transactions on Graphics (TOG)*, 31(4) :76, 2012. (Cité pages 39 et 77.)