N° d'ordre : 177
N° bibliothèque : 00ENSL0177

ÉCOLE NORMALE SUPÉRIEURE DE LYON
Laboratoire de l'informatique du parallélisme

THÈSE

pour obtenir le grade de
Docteur de l'École normale supérieure de Lyon
Spécialité : Informatique
au titre de l'École doctorale de mathématiques et informatique fondamentale

présentée et soutenue publiquement le 23 février 2001
par **Bruno MARTIN**

Automates cellulaires, information et chaos

Après avis de : M. Robert Cori
M. Jean-Paul Delahaye
M. Giancarlo Mauri

Devant la commission d'examen formée de :
M. Nino Boccara
M. Jean-Paul Delahaye
M^lle^ Marianne Delorme (directrice de thèse)
M. Alejandro Maass
M. Giancarlo Mauri
M. Jacques Mazoyer (directeur de thèse)

Sommaire

Remerciements **5**

Introduction **9**

Automates cellulaires . . . 9

La complexité dans les automates cellulaires . . . 11

Le chaos dans les automates cellulaires . . . 12

L'entropie apparente . . . 16

La propagation d'information . . . 17

La sensibilité aux conditions initiales μ-presque partout pour la pseudo-distance de Besicovitch . . . 17

Plan de la thèse . . . 18

1 Définitions **19**

1.1 Les automates cellulaires . . . 19

1.2 Mesure sur l'ensemble des configurations . . . 21

1.3 Automates groupés . . . 26

1.4 Produit d'automates . . . 27

2 Exemples d'automates cellulaires **29**

2.1 La notation de Wolfram pour les ECA . . . 29

2.2 Les règles 210 et 120 . . . 31

2.3 La règle 184 et l'automate T . . . 32

2.4 Les automates additifs . . . 34

2.5 La règle 18 . . . 36

2.6 La règle 54 . . . 37

3 La règle 54 **41**

3.1 Les 8 fonds de la règle 54 . . . 41

3.2 Comprendre les particules à l'aide du groupe . . . 47

4 Les principales classifications 53
4.1 Rapaport 54
4.1.1 La relation d'ordre introduite par I. Rapaport 54
4.1.2 Non-autosimilarité et classification de Rapaport 55
4.1.3 Paraboles et non-autosimilarité 58
4.1.4 La règle 54 n'est pas autosimilaire 60
4.2 Wolfram, Čulik et Yu, Ishii 65
4.2.1 Wolfram 65
4.2.2 Čulik et Yu 69
4.2.3 Ishii 71
4.3 Gilman 74
4.4 Notions de chaos invariantes par décalage 76
4.4.1 Cattaneo *et al.* 76
4.4.2 Formenti 77
4.5 Conclusion 81

5 Entropie apparente 83
5.1 Introduction et définition 83
5.2 Calcul/approximation de l'entropie apparente 88
5.2.1 Calcul exact de l'entropie apparente 89
5.2.2 Approximation de l'entropie apparente 95
5.3 L'entropie apparente et le « chaos » 97
5.4 Les règles « presque chaotiques » 100

6 Vers une définition du chaos 103
6.1 Propagation d'information 103
6.2 Sensibilité μ-presque partout 107
6.3 Comparaisons entre les notions introduites 112
6.4 Retour sur la classification de Wolfram 116

Conclusion 119

Remerciements

Ce n'est pas sans une certaine appréhension que je rédige cette partie qui, selon toute vraisemblance, sera la plus fréquemment lue de ma thèse. Comme aime à le souligner l'un de mes directeurs de thèse, l'exercice qui consiste à remercier des personnes qu'on ne connaît pas est toujours amusant et, n'ayant pas encore soutenu ma thèse au moment où j'écris ces lignes, je n'y échapperai pas. En effet, mes remerciements commencent par les membres de mon jury, comme le veut l'usage, ce qui n'enlève rien à leur sincérité.

Je tiens donc à remercier Robert Cori, Jean-Paul Delahaye et Giancarlo Mauri d'avoir accepté de rapporter sur ma thèse ainsi que Niño Boccara et Alejandro Maass de bien vouloir être membres de mon jury.

Je souhaite remercier ensuite les personnes qui ont eu une influence décisive sur le bon déroulement scientifique de ma thèse.

Comme tout le monde, j'avais entendu dire qu'il était primordial de choisir un bon directeur de thèse et, comme tout le monde, j'ai compris au cours de ma thèse que cela était encore beaucoup plus important que je ne l'avais imaginé. Je dois dire que j'ai eu beaucoup de chance de préparer ma thèse sous la direction de Marianne Delorme et Jacques Mazoyer. Au début de mon stage de DEA, un de leurs précédents thésards m'avait dit qu'il s'agissait probablement de directeurs de thèses parfaits. Si, sur le coup, je n'ai pas compris toute la portée de cette remarque, j'ai pu mesurer ensuite à quel point elle était fondée.

Je tiens à les remercier pour la disponibilité extraordinaire dont ils ont fait preuve. J'ai beaucoup apprécié d'avoir à la fois une très grande liberté quant aux choix de mes sujets de recherche, à l'organisation de mon travail tout en ayant l'assurance d'être toujours soutenu lorsque je rencontrais un problème, qu'il soit de nature scientifique, administratif, social (– Je suis timide et j'ai peur de la secrétaire, vous voulez bien m'accompagner ?), lexical (– Quelle est la différence entre proximité et promiscuité ?) ou très bassement matériel

(– Je me suis enfermé dehors, pouvez-vous m'héberger cette nuit ?). Bien sûr, je souhaite aussi les remercier pour la pertinence des remarques scientifiques qu'ils m'ont faites et des orientations qu'ils m'ont suggérées, les relectures minutieuses de mes travaux, leur patience inébranlable et la bonne humeur de nos (souvent longues) réunions de travail. C'est avec un réel plaisir que j'ai préparé ma thèse sous leur direction et je regrette moi aussi de ne pas avoir pris part aux moments de convivialité que constituait le déjeuner à la cafeteria.

Beaucoup de résultats de cette thèse n'auraient jamais vu le jour sans l'aide précieuse que m'a apportée Sylvie Ruette. Je souhaite la remercier pour tout le temps qu'elle a passé à expliquer de façon claire et accessible à l'informaticien que je suis les connaissances mathématiques sur l'entropie, les systèmes dynamiques, la théorie ergodique et la géométrie non euclidienne qui m'ont été indispensables pour faire aboutir mes recherches. Je la remercie également pour la relecture soigneuse de certaines de mes preuves.

Je tiens également à remercier Bill Allombert pour les nombreuses discussions scientifiques mi-sérieuses, mi-délirantes qui m'ont permis de prendre du recul sur mes recherches. Je le remercie également pour l'aide technique qu'il m'a généreusement apportée tant en théorie des nombres que pour l'administration de mon équipement informatique extravagant. Pour me faire pardonner du manque de disponibilité dont j'ai fait preuve pour jouer au bridge, ce n'est pas un hasard si la chaîne "merci" apparaît exactement 17 fois dans ce texte.

Je remercie aussi Enrico Formenti, Alejandro Maass, Nicolas Ollinger, Ivan Rapaport et surtout François Blanchard pour l'intérêt qu'ils ont porté à mes travaux et pour leur collaboration.

C'est en cherchant à comprendre pourquoi, alors que j'ai toujours nettement préféré les mathématiques à la physique et que rien ne me prédisposait donc lors de mon entrée à l'école à être l'auteur d'une publication en physique à la fin de ma thèse, que j'ai compris la grande influence que mes nombreuses discussions avec Patrick Iglesias sur les bords de Saône ont eue sur mon orientation scientifique. J'ai en effet été impressionné par sa très grande maîtrise de nombreux domaines et surtout par sa capacité à les motiver et les illustrer de façon très pertinente à la fois d'un point de vue mathématique et d'un point de vue physique. Cette thèse lui doit assurément beaucoup.

Je terminerai par tous ceux qui, par leur gentillesse et leur bonne humeur, ont fait de ces trois petites années un moment inoubliable.

Outre Sylvie et Bill déjà cités, je pense en premier lieu à Matthieu, Magalie et Clémence et aux nombreux moments très forts passés avec eux.

Je pense ensuite au personnel administratif de l'école qui a toujours été d'une extrême gentillesse avec moi malgré mon incompétence totale en la matière... merci à Anne-Pascale Botonnet, Nicole Noury, Jean-Alex Losdat, Claire Desecures, Simone Durand et Marie-Jeanne Barrier.

Merci également à mes « vieux » copains de prépas et pièces rapportées, Sébastien, Laurette, Philippe, Mai, Fabien, Laure (les 2), Steeve, JP, Olivier (les 2), Christelle, Florence (et tous ceux que j'oublie) pour m'avoir rappelé qu'il n'y avait pas que la thèse dans la vie (en effet, il y a aussi le ski).

Enfin je n'oublie pas les membres de l'équipe MC2 (Hervé, Olivier et Nicolas, mes sympathiques co-bureaux qu'il m'arrivait même de croiser de temps en temps, Fred, Nathalie, Grégory, Christophe, Pascal, Natacha, Anahi, Éric, Vincent et Ioan pour les réunions d'équipe aussi philosophiques que joyeuses et improvisées au milieu du couloir) et du laboratoire (Claire, Antoine (au fait, cette partie de bridge...), Alice et Vincent au coin café (– Tu bosses ? – Non et toi ? – Moi non plus...), Philippe (– Tu étais au ski ce week-end ? – Oui, et toi ? – Moi aussi...), Vincent (– Au fait, on peut émuler RiscOS sous netBSD ?), Daniel, Olivier et Laurent (– Quelle drôle d'idée d'aller sur le balcon quand il fait froid)).

Je remercie enfin ma maman de m'avoir mis au monde et mon papa d'avoir participé à la conception parce que sinon, rien de tout cela ne serait arrivé. Je remercie plus généralement ma famille qui a supporté mes absences et mes délires mathématico-philosophiques, qui s'est mobilisée pour venir assister à ma soutenance, et m'a apporté en toute circonstance un soutien logistique salvateur.

Introduction

Automates cellulaires

Les automates cellulaires (ou AC) ont été introduits, dans les années 50, par Von Neumann afin d'étudier l'auto-reproduction de systèmes artificiels [50]. Ils se sont révélés des outils précieux en physique et en biologie. D'une part, dans l'esprit de Von Neumann, afin de mieux comprendre certains phénomènes tels que les changements de phase ou la croissance des plantes : en cherchant à définir un automate cellulaire dont le comportement est semblable au phénomène étudié, il est possible de formuler des hypothèses sur les mécanismes fondamentaux qui en sont à l'origine. D'autre part, afin de réaliser des simulations, que ce soit pour vérifier que les sorties de secours d'un bâtiment sont correctement positionnées et en nombre suffisant avant qu'il ne soit construit, ou pour représenter en trois dimensions la propagation de stimulations du cerveau. Les automates cellulaires ont un intérêt propre en informatique. En effet, en 1960, Smith [56] les a définis et étudiés en tant qu'objets calculant des fonctions. Dans cette optique, les automates cellulaires sont considérés comme des modèles bien formalisés du calcul massivement parallèle. Ils permettent, d'une part, d'appréhender des questions de complexité et d'algorithmique parallèle pour une structure donnée et, d'autre part, d'étudier inversement quelles sont les architectures les plus efficaces du point de vue du calcul pour un problème donné.

Les automates cellulaires définissent une classe de systèmes dynamiques discrets. Dans sa forme la plus générale, un AC de dimension d peut être considéré comme un ensemble infini de cellules élémentaires identiques indexées par les éléments de $\mathbb{Z}^d$ (on peut se les représenter comme répartis uniformément dans l'espace à d dimensions aux points de coordonnées entières). À un instant donné (le temps est discret) chaque cellule se trouve dans un certain état pris dans un ensemble fini (le même pour tous). Ces états varient au cours du temps de façon :

- **Synchrone** : Les états de toutes les cellules sont modifiés en même temps. Les états à l'instant $t+1$ ne dépendent que de ceux à l'instant t.

- **Locale** : L'état d'une cellule est déterminé en fonction de ceux des cellules qui se trouvent autour d'elle dans un rayon borné.
- **Uniforme** : La méthode appliquée pour effectuer les mises à jour d'états est la même pour toutes les cellules.

Si les utilisations des automates cellulaires sont pluridisciplinaires, les méthodes impliquées dans leur étude le sont également.

En physique, de nombreux phénomènes sont régis par des interactions locales. Souvent, leur compréhension passe par des bilans (de forces, d'énergies, de matière, ...) sur des unités d'espace dont on fait tendre la taille vers zéro pour obtenir des équations différentielles. Il s'agit ensuite d'essayer de les intégrer en prenant en compte les conditions initiales. La plupart du temps, on ne sait pas - ou il est impossible - d'intégrer ces équations. La prévision de l'évolution de ces phénomènes nécessite alors des simulations. Les automates cellulaires sont fréquemment utilisés dans ce contexte, dès lors qu'il est possible de discrétiser convenablement les interactions locales. C'est alors l'évolution de l'automate sur un grand nombre de cellules qui permet de simuler le phénomène continu. Il est donc naturel de se demander comment les lois de conservation qui permettent de prévoir l'évolution du phénomène physique s'expriment sur l'automate cellulaire qui le modélise. En effet, la mise en évidence de lois de conservation sur un automate cellulaire permet de prévoir partiellement son évolution. Dès lors, les résultats obtenus peuvent servir dans d'autres disciplines.

En tant qu'objets mathématiques, ce sont des systèmes dynamiques particuliers auxquels peuvent s'appliquer des résultats de la théorie ergodique.

Enfin, en tant que modèles de calcul, s'introduisent naturellement des notions de simulation, conduisant à discriminer des classes d'automates regroupant des AC ayant des comportements algorithmiques semblables. Par ailleurs, les problèmes d'algorithmique sur les AC ont nécessité l'introduction de la notion de signal. Très informelle, mais très importante, elle traduit une transmission d'information. Nous l'utiliserons à la fois pour proposer des exemples d'automates et pour interpréter et tenter de comprendre certaines évolutions.

Notre motivation principale est de mieux comprendre et maîtriser les évolutions des automates cellulaires. Lorsqu'un automate cellulaire simule un phénomène naturel, qu'il soit physique ou biologique, il est naturel de se demander comment se traduisent les propriétés de ce système (évolution vers un attracteur, caractère chaotique ou non de son comportement, variations de l'entropie) sur l'automate qui le modélise afin d'obtenir des informations sur cet automate. Ces propriétés, une fois formalisées dans le contexte des AC, permettent d'établir des classifications des automates cellulaires. Le véritable point de départ de ces travaux est un article de Wolfram en 1984 [60] qui propose une classification empirique en quatre classes. Les deux premières classes correspondent à des comportement simples, la troisième à des comportements « chaotiques » et la quatrième à des comportements souvent appelés « complexes ». Les notions d'automates « chaotiques » ou « complexes » ne sont pas formellement définies. Cependant, elles sont souvent utilisées dans la littérature avec des sens relativement précis auxquels nous nous conformerons dans ce travail. Elles sont en effet devenues des outils irremplaçables pour désigner informellement certaines caractéristiques des automates cellulaires. Comme nous le verrons sur des exemples lorsque nous présenterons la classification de Wolfram (section 4.2.1) dans le chapitre consacré aux principales classifications existantes, les automates « chaotiques » sont ceux dont les diagrammes espace-temps sont homogènes et font penser à des points placés au hasard dans le plan, tandis qu'un automate est dit « complexe » lorsque ses évolutions semblent imprévisibles mais que des organisations locales apparaissent. De nombreuses classifications formelles ont suivi celle de Wolfram sans parvenir à définir de façon satisfaisante ces notions de chaos et de complexité. Notre travail s'articule autour de deux axes ; nous nous intéresserons dans un premier temps, à l'étude d'un cas particulier d'automate dit « complexe » pour lequel nous proposerons une méthode permettant de démontrer des résultats mis au jour expérimentalement sur son évolution. Nous nous intéresserons dans un second temps à la question de la classification des AC en fonction de notions de chaos et de complexité. La définition d'une bonne notion de complexité semble plus problématique que celle d'une bonne notion de chaos. De plus, complexité et chaos ne semblent pas liés contrairement aux observations de Wolfram ; nous verrons précisément qu'une règle peut être complexe, chaotique, ou les deux à la fois.

La complexité dans les automates cellulaires

Certaines règles, celles que Wolfram avait mises dans la classe 4, sont dites complexes, par exemple les règles "54" et "110" qui figurent pourtant parmi les règles les plus élémentaires (unidimensionnelles, deux états, plus proches voisins). Ces règles sont étudiées depuis longtemps pour leur aspect compliqué, mais c'est suite à l'article de Cattaneo [11] qui propose une classification des règles élémentaires, à l'exception de ces deux règles qui ne rentrent dans aucune de ses classes, qu'elles sont considérées comme les deux règles les plus compliquées. Ces deux règles ont un fonctionnement un peu similaire qui fait apparaître des structures périodiques dans le diagramme espace-temps séparées par des particules, que l'on appelle aussi parfois défauts. Afin de comprendre la dynamique de ces règles, il est naturel d'identifier les particules qui apparaissent et d'en étudier les relations. Une telle étude à été conduite par Boccara [9] pour la règle 54. Nous proposons, pour la règle 54, une nouvelle méthode qui permet de mieux comprendre l'origine et les propriétés des interactions observées. La méthode consiste à introduire le groupe associé au pavage que définit le motif des structures périodiques. Nous montrerons que les interactions entre les particules sont liées à la structure de ce groupe. Nous en déduisons des résultats sur la règle, qui ont des conséquences à la fois sur sa dynamique et sur ses propriétés algorithmiques. Cette méthode n'est utilisable que dans le cas unidimensionnel, mais dans ce cadre, elle s'applique à toutes les règles faisant apparaître des structures que l'on peut interpréter à l'aide de particules.

Le chaos dans les automates cellulaires

Nous l'avons déjà dit, les automates cellulaires peuvent être utilisés pour modéliser des phénomènes physiques, qui sont régis par des lois d'interaction locales. L'existence, parmi les phénomènes simulés, de systèmes dits chaotiques parce qu'ils sont très sensibles aux conditions initiales suggère que le même type de comportement existe pour les automates cellulaires. Et, en effet, même parmi ceux dont les règles sont les plus élémentaires, certains ont des comportements difficilement prévisibles et (ou) très sensibles aux conditions initiales, qui s'apparentent à des comportements chaotiques. Bien entendu, un automate cellulaire est déterministe et il y a un débat pour savoir s'il est possible de parler de chaos déterministe ou si ce terme doit être réservé aux systèmes régis par des règles aléatoires. Cependant, sans participer au débat, nous utiliserons ce terme comme cela est l'usage dans la littérature consacrée aux AC et en mathématiques à propos des systèmes dynamiques.

Selon [23, 22], la théorie du chaos est née avec les travaux de E. Lorenz. Ce dernier s'est rendu compte par accident de la très grande sensibilité aux conditions initiales du modèle, pourtant déterministe, qu'il avait développé pour simuler les phénomènes météorologiques, ce qu'il appellera l'effet papillon [35, 36, 37]. Lorenz en déduit l'impossibilité de prévision météo à long terme et cherche des modèles plus simples présentant le même comportement. Il s'est en particulier intéressé à la convection dans une tasse de café brûlant. Ce système est signifiant puisque son comportement à long terme est connu : la chaleur du café va se dissiper, la température tendre vers la température de la pièce tandis que le mouvement du liquide va s'arrêter en raison des frottements. Par contre, les mouvements de convection qui se produisent si le café est bien chaud, provoquent des écoulements désordonnés et turbulents. En partant des équations de la convection et en les simplifiant, il obtient un système de trois équations présentant des analogies avec certains systèmes réels. La biologie, ou plus précisément la dynamique des populations, a ensuite joué un rôle important dans l'étude du chaos. En effet, contrairement aux problèmes physiques, l'évolution d'une population suggère l'utilisation de modèles mathématiques discrets plutôt que celle d'équations différentielles. Smale propose de modéliser l'évolution d'une population avec la suite définie par $u_{n+1} = ru_n(1 - u_n)$. Cette suite tend vers un unique point fixe pour de faibles valeurs du paramètre r, mais a un comportement apériodique pour de plus grandes valeurs. Le mathématicien Yorke s'intéressa à cette suite, puis aux articles de Lorenz. Avec May [46], un ami biologiste, il fait connaître les travaux de Smale et de Lorenz à la communauté scientifique.

D'un point de vue systèmes dynamiques (donc mathématique), l'article de Li et Yorke [33] a été le premier à introduire le terme de chaos, mais de façon informelle et pour les transformations de l'intervalle. Une définition précise (« chaos au sens de Li-Yorke ») a ensuite été dérivée de leurs idées. Un système dynamique (X, T) (X est un espace compact et $T : X \longrightarrow X$ une fonction continue) est chaotique au sens de Li-Yorke s'il existe un ensemble non dénombrable S tel que, si x, y sont deux points distincts de S, alors (x, y) est un couple de Li-Yorke, c'est-à-dire que $\limsup d(T^n(x), T^n(y)) > 0$ et $\liminf d(T^n(x), T^n(y)) = 0$. Une autre définition célèbre est celle de Devaney [14] sur laquelle nous reviendrons, un système est alors chaotique s'il est transitif (grossièrement il n'est pas possible de le décomposer en sous-systèmes) et si l'ensemble des points périodiques est dense. En fait, Devaney ajoute une troisième condition qui est la sensibilité aux conditions initiales, mais il a été démontré par la suite qu'il s'agissait d'une conséquence des deux premières [21, 4]. Bien d'autres notions ont été considérées, Auslander et Yorke ont appelé chaotique un système à la fois transitif et sensible aux conditions initiales [1], Glasner et Weiss ont proposé « entropie strictement positive, transitivité et existence d'une mesure de support total » [21], et bien d'autres encore impliquant, outre les notions déjà évoquées, le mélange faible, l'expansivité (nous reparlerons de ces deux notions) ou les systèmes dispersants [17].

En physique, d'autres approches du chaos ont été introduites. En particulier la notion d'attracteur étrange apparaît dans un article de Ruelle et Takens [55] (pour une introduction, cf [49]). Un attracteur peut être représenté dans l'espace des phases et il est dit étrange si cette représentation a une structure fractale. Nous ne nous y intéresserons pas ; par contre, nous nous inspirerons des exposants de Lyapunov (cf par exemple [5, 52]) qui traduisent une dépendance exponentielle aux conditions initiales.

Définir des automates ayant des comportements chaotiques présente un grand intérêt. Cela permet d'étudier la notion de chaos dans un cadre formel. Plusieurs définitions sont considérées, comme c'est d'ailleurs le cas dans le cadre plus général des systèmes dynamiques. Pour définir le chaos, il est très naturel de considérer les définitions en provenance des systèmes dynamiques. Cela a été fait, avec des approches très différentes, et sans que la liste ne soit exhaustive, par Gilman [20] (expansivité presque partout) et Cattaneo [10] (Devaney, forte transitivité, expansivité). Malheureusement, ces définitions donnent des résultats qui ne sont pas satisfaisants. En effet, la topologie la plus naturelle sur l'ensemble des configurations est la topologie produit et, pour celle-ci, il s'agit principalement de savoir s'il est possible de prévoir l'évolution d'une cellule en connaissant la configuration initiale sur un voisinage de cette cellule. Ainsi, dans le cas d'une ligne de cellules, le décalage qui consiste, pour chaque cellule, à recopier l'état de sa voisine de droite, est l'exemple typique du système dynamique chaotique. Cependant, cet automate n'est pas considéré comme chaotique dans le contexte des automates cellulaires. En effet, cela est mal adapté à de nombreuses utilisations possibles (la simulation du trafic routier par exemple), mais surtout, comme les automates cellulaires sont invariants par décalage, il est bon que les définitions sur les AC respectent l'invariance par décalage. De plus, compte tenu de leur utilisation en physique, la composition de l'automate avec un décalage s'interprète comme un changement de référentiel : on passe d'un référentiel fixe à un référentiel ayant un mouvement rectiligne uniforme. Dans ce contexte, il est évident que le changement de référentiel ne change en rien le phénomène étudié et qu'il est inconcevable que ce système soit chaotique dans un cas et pas dans l'autre. Le statut du décalage vis-à-vis du chaos conduit donc à une nouvelle vision de l'espace des phases dans laquelle nous nous plaçons à la suite de Formenti.

Outre le problème du décalage, il est important de préciser quels objets sont classés. Wolfram, comme nous le verrons, ne classifie pas les automates cellulaires mais les diagrammes espace-temps de ces automates construits à partir de configurations « aléatoires ». Čulik et Yu classifient pour leur part l'ensemble des diagrammes espace-temps construits à partir de configurations finies. Cattaneo ou Formenti, classifient l'ensemble des diagrammes espace-temps, ce qui revient à classifier les automates cellulaires. Quant à Gilman et Ishii, ils choisissent une notion d'aléatoire, c'est-à-dire qu'ils se donnent une mesure et classifient l'ensemble des diagrammes espace-temps construits à partir de configurations aléatoires pour cette mesure. Pour notre part, nous choisirons la dernière approche, celle d'Ishii et Gilman. La première raison de ce choix est la pertinence des résultats obtenus. En particulier, c'est l'approche qu'a utilisée Wolfram sans le dire ni le formaliser. En plus, cela correspond au regard que nous portons sur les automates cellulaires : des modèles utilisés pour simuler des comportements physiques. Dans cette perspective, le système discret « automate cellulaire » modélise souvent l'aspect microscopique d'un phénomène, qui d'un point de vue macroscopique est continu. De cela découle trois idées. D'une part, que le phénomène étudié apparaît lorsque l'on effectue un zoom arrière infini et cela justifie l'étude des automates pour des configurations infinies. D'autre part, que les conditions initiales ne sont pas connues avec une très grande précision, on connaît seulement leurs propriétés statistiques et on ne s'intéresse pas du tout aux évolutions particulières que l'on ne rencontrera jamais dans la réalité. Cela justifie notre approche qui consiste à définir des notions pour « presque toutes » les configurations pour une mesure donnée. En effet, la mesure traduit les propriétés statistiques de la configuration initiale de telle sorte que les comportements pathologiques, qui se produisent sur un ensemble non représentatif (de mesure nulle) de configurations initiales, ne sont pas pris en compte. Enfin qu'il existe des exemples de systèmes physiques, comme par exemple l'écoulement d'un fluide, dont le comportement est chaotique ou non selon la vitesse de l'écoulement, et que notre approche permet de dire qu'un automate est chaotique pour certaines mesures et non chaotique pour d'autres, tandis que les définitions topologiques (les classifications de l'ensemble des diagrammes espace-temps) ne donnent qu'une réponse binaire.

Maintenant que nous avons précisé ce que nous voulons classifier et pourquoi, il nous reste à nous demander comment. Nous avons commencé à répondre à cette question en expliquant pourquoi nous exigeons des définitions invariantes par décalage. En fait, l'objectif est d'obtenir une classification qui donne des résultats proches de ceux de Wolfram ; il a en effet soulevé des problèmes importants et non résolus sur l'ensemble des comportements observables parmi les AC. De plus, comme Wolfram, nous cherchons des définitions qui coïncident avec notre intuition. En particulier, nous souhaitons obtenir des définitions invariantes par changement d'échelle, car, comme nous l'avons vu, cette propriété d'invariance par changement d'échelle est très importante lorsque les automates sont utilisés pour modéliser le comportement microscopique d'un phénomène physique. Cette invariance sera formalisée par la notion, et méthode, de groupage.

Afin de mieux cerner et maîtriser les comportements dits chaotiques des AC, nous proposons trois nouvelles notions (l'entropie apparente, la propagation d'information et la sensibilité presque partout) détaillées cidessous, qui ont de bonnes propriétés d'invariance par décalage et par groupage (zoom). Un des objectifs est de pouvoir comparer formellement des représentations intuitives que l'on a du chaos : une évolution qui semble aléatoire ou désordonnée, le fait qu'une toute petite modification peut avoir des conséquences très importantes sur l'évolution du système (effet papillon) et la sensibilité aux conditions initiales.

L'entropie apparente

Parmi les définitions du chaos utilisées en systèmes dynamiques, certaines impliquent que l'entropie de la fonction soit strictement positive. Nous ne donnerons pas la définition de l'entropie d'une fonction en général, car l'utilisation directe de cette notion sur les AC n'est pas très pertinente pour les raisons, liées à la topologie choisie, que nous avons évoquées précédemment. Cependant, nous allons nous inspirer de cette idée que le chaos est lié à l'entropie d'une façon très différente : nous allons observer l'évolution de l'« entropie des configurations ». En effet, étant donné une configuration, on peut considérer la fermeture topologique de l'ensemble de ses translatées. La restriction du décalage à cet ensemble est un sous-shift. On peut ainsi associer à une configuration l'entropie du sous-shift ainsi construit, que nous appellerons l'entropie de la configuration. L'entropie des configurations successives ne peut pas augmenter au cours du temps en raison du caractère déterministe des AC [34], nous donnerons une preuve de ce résultat dans le chapitre 5. Cela peut paraître surprenant que l'on simule des phénomènes physiques dont l'entropie ne peut qu'augmenter à l'aide d'automates cellulaires dont l'entropie ne peut que décroître, mais c'est inévitable avec des modèles discrets et déterministes. Bien que l'entropie n'augmente pas, les configurations semblent de plus en plus désordonnées parce qu'il est de plus en plus difficile de voir qu'elles sont « simples ». En effet, l'entropie mesure la quantité d'information de la configuration. Au bout d'un temps t fini, cette quantité d'information est inférieure à la quantité d'information de la configuration initiale plus celle de la fonction de transition locale et de t. Or l'information de la fonction de transition et de t est finie et donc négligeable devant la quantité infinie d'information de la configuration initiale. Ainsi, pour tout t fini, l'entropie de la configuration après un temps t est inférieure à celle de la configuration initiale. Il est donc nécessaire d'attendre un temps infini pour éventuellement observer une augmentation de l'entropie. Pour rendre cette observation possible, nous définirons l'entropie apparente. Ce faisant, nous passons en fait à la limite, c'est-à-dire que nous faisons le zoom arrière infini qui permet de passer du modèle discret microscopique au phénomène macroscopique continu étudié et pour lequel l'entropie peut augmenter.

La propagation d'information

En physique, la notion de chaos est plutôt associée à celle de grande sensibilité aux conditions initiales, d'où l'introduction de coefficients de Lyapunov, qui suggère qu'un système est chaotique lorsque les trajectoires de deux configurations proches divergent exponentiellement. Cette notion a été adaptée aux AC [2], le problème est que deux configurations initiales proches ne peuvent diverger que polynomialement. En effet, au temps t, seules les cellules situées à distance inférieure à αt (α est une constante) d'une cellule dont l'état diffère entre les deux configurations initiales peuvent avoir des états différents. Or les modèles étudiés étant discrets et euclidiens, le nombre de cellules dans une boule de rayon αt est polynomial en fonction de t. L'idée sous-jacente pour résoudre ce problème est de compter une cellule chaque fois que l'une des différences entre les configurations à l'étape précédente lui fait prendre des états différents, ce qui conduit à compter certaines cellules plusieurs fois et permet alors une divergence maximale exponentielle. En nous inspirant de cette notion, nous nous intéresserons à la notion de propagation d'information, c'est-à-dire au nombre moyen (pour une mesure donnée) de modifications (on ne compte les cellules qu'une seule fois) que le changement d'une seule cellule provoque. Nous montrerons alors qu'un automate dont la propagation d'information est bornée ne peut pas être chaotique.

La sensibilité aux conditions initiales μ-presque partout pour la pseudo-distance de Besicovitch

Les difficultés pour adapter les définitions du monde des systèmes dynamiques proviennent de la topologie produit. Pour le contourner, Formenti [12, 6] a introduit la pseudo-métrique de Besicovitch, qui, contrairement à la distance naturelle sur l'ensemble des configurations, est invariante par le décalage. Il étudie ensuite la notion de sensibilité aux conditions initiales pour cette pseudo-métrique. Cependant, l'espace des configurations n'est pas compact pour la topologie induite par cette pseudo-distance, ce qui ne permet pas de réutiliser directement les résultats des systèmes dynamiques. En fait, la pseudo-distance de Besicovitch est la densité supérieure des différences entre les deux configurations. L'utilisation d'une densité supérieure est nécessaire parce qu'il est impossible de définir une densité sur $\mathbb{Z}$. Ce problème est résolu par l'introduction d'une mesure μ puisque la densité est alors correctement définie sur un ensemble de mesure 1. Nous nous intéresserons donc à la sensibilité μ-presque partout pour Besicovitch et observerons que, bien que sensibles, certains automates ne semblent pas chaotiques.

Plan de la thèse

Dans le premier chapitre, nous définirons les automates cellulaires ainsi que quelques notions de base sur les mesures et le groupage que nous utiliserons dans cette thèse. Le chapitre 2 sera consacré à l'étude de quelques automates cellulaires intéressants qui ont joué un rôle important dans notre recherche et serviront d'exemples par la suite. Nous nous intéresserons particulièrement à l'un de ces exemples, la règle 54 dans le chapitre 3 [41] afin de mieux comprendre les interactions entre ses particules. Nous développerons également des résultats sur cette règle dans le chapitre 4. Celui-ci est consacré aux principales classifications proposées jusque là dans la littérature, que nous présenterons et commenterons. Le chapitre 5 est dédié à la notion d'entropie apparente que nous définissons [39] et dont nous testons la pertinence sur de nombreux exemples. Dans le chapitre 6, nous essayerons de mieux comprendre la notion de chaos pour les AC grâce aux notions que nous introduisons : la propagation d'information et la sensibilité μ-presque partout aux conditions initiales [40].

Chapitre 1

Définitions

1.1 Les automates cellulaires

Afin de ne pas alourdir inutilement les notations, dans cette thèse, nous ne considérerons que des automates cellulaires unidimensionnels (c'est-à-dire que toutes les cellules sont alignées de façon régulière sur une droite). Excepté le deuxième chapitre à propos de la règle 54 (dont les résultats sont intrinsèquement liés aux AC unidimensionnels), les résultats des chapitres 3 et 4 sont topologiques et s'étendent sans problème au cas général des AC en dimension supérieure (cf remarque 6.2.1 page 108 pour plus de détails).

Il existe plusieurs définitions équivalentes des automates cellulaires. Nous allons donner les deux principales : la première est préférée en informatique et en physique parce qu'elle met l'accent sur la localité des interactions. La seconde est préférée en mathématiques où l'on considère les automates cellulaires comme des cas particuliers de systèmes dynamiques, c'est-à-dire de fonctions continues d'un espace (compact) dans lui même.

Définition 1.1.1 *(locale)*
Un automate cellulaire unidimensionnel *de rayon r est un couple (Q, δ) où Q est un ensemble fini d'états et $\delta : Q^{2r+1} \longrightarrow Q$ est une fonction de transition. Une configuration $c \in Q^{\mathbb{Z}}$ de (Q, δ) est une fonction de $\mathbb{Z}$ dans Q et sa fonction de transition globale $G_\delta : Q^{\mathbb{Z}} \longrightarrow Q^{\mathbb{Z}}$ est telle que $(G_\delta(c))(i) = \delta(c(i-r), ..., c(i), ..., c(i+r))$.*

La seconde définition nécessite les définitions suivantes. Notons Q un ensemble fini d'états. Posons $\sigma : Q^{\mathbb{Z}} \longrightarrow Q^{\mathbb{Z}}$ le décalage (ou shift) vers la gauche : $\sigma(c)(i) = c(i+1)$. Par ailleurs, munissons $Q^{\mathbb{Z}}$ de la topologie de Cantor, c'est-à-dire la topologie produit (dite parfois de Tychonoff) induite par la topologie discrète sur Q.

Définition 1.1.2 *(globale)*
Un automate cellulaire unidimensionnel *est un couple* $(Q^{\mathbb{Z}}, F)$ *où* $Q^{\mathbb{Z}}$ *est l'ensemble des configurations et* $F : Q^{\mathbb{Z}} \longrightarrow Q^{\mathbb{Z}}$ *est une fonction continue (pour la topologie produit) shift-invariante (i.e.* $F(\sigma(c)) = \sigma(F(c))$*) appelée fonction globale de transition.*

Notation 1.1.1 *Si* Q *est l'ensemble d'états d'un AC, il est parfois appelé alphabet, c'est la raison pour laquelle* Q^n *est appelé l'ensemble des mots de longueur* n *sur* Q*. Si* $q \in Q$ *est un état, nous noterons* q^* *la configuration uniforme dont tous les états sont* q*. La* i^{th} *coordonnée* $c(i)$ *d'un point* $c \in Q^{\mathbb{Z}}$ *sera aussi notée* c_i *et* $c_{[j,k]} = c_j...c_k \in Q^{k+1-j}$ *est le sous-mot de* c *entre les indices* j *et* k*. Le cylindre défini par la présence du mot* $u \in Q^p$ *à la position* $k \in \mathbb{Z}$ *est l'ensemble*

$$[u]_k = \{c \in Q^{\mathbb{Z}} | c_{[k,k+p-1]} = u\}.$$

L'équivalence entre ces deux définitions est assurée par un théorème de Hedlund dont nous donnons ci-dessous une démonstration.

Théorème 1.1.1 *(Hedlund [27])*
Les définitions, locale et globale, des automates cellulaires sont équivalentes.

Preuve 1.1.1 Les ouverts de $Q^{\mathbb{Z}}$ pour la topologie de Cantor sont engendrés par des ouverts élémentaires, qui sont les cylindres $[u]_k$. Le support d'un ouvert élémentaire $[u]_k$ est l'ensemble (fini) des indices des cellules dont l'état est commun à toutes les configurations de l'ouvert, c'est-à-dire $\{k, k+1, ..., k+p\}$, lorsque $u \in Q^p$.

Il est évident que la définition locale implique la définition globale. En effet, il suffit de montrer la continuité de la fonction globale. Pour cela on montre facilement que l'image réciproque d'un ouvert est un ouvert car l'image réciproque d'un ouvert élémentaire $[u]_k$ avec $u \in Q^p$ est l'union finie des ouverts élémentaires $[v]_{k-r}$ avec $v \in Q^{p+2r}$ tels que $\forall i, 1 \leq i \leq p, \delta(v_{i-r}, ..., v_{i+r}) = u_i$.

Montrons à présent la réciproque. Soit q un état ($q \in Q$), considérons l'ouvert élémentaire $\Omega_q = [q]_0 = \{c \mid c_0 = q\}$. Par continuité de F, $F^{-1}(\Omega_q)$ est un ouvert donc il s'écrit comme l'union d'ouverts élémentaires $F^{-1}(\Omega_q) = \cup_{i \in I_q} \mathcal{O}_i^q$. Or $\cup_{q \in Q} F^{-1}(\Omega_q) = Q^{\mathbb{Z}} = \cup_{i \in \cup I_q} \mathcal{O}_i^q$. Il s'agit d'un recouvrement de $Q^{\mathbb{Z}}$ qui est compact (par Tychonoff, c'est un produit de compacts) par des ouverts (élémentaires). Donc, il existe un sous-recouvrement fini noté $\cup_{i \in \cup I'_q} \mathcal{O}_i^q$ avec I'_q fini. Notons S l'union des supports des ouverts élémentaires $\mathcal{O}_i^q$ pour $q \in Q$ et $i \in I'_q$. Définissons le rayon de l'automate $r = \max\{|j|, j \in S\}$. Par construction, la fonction de transition $\delta : Q^{2r+1} \longrightarrow Q$ est bien définie par $\delta(s_{-r}, ..., s_r) = q$ si et seulement si l'ouvert élémentaire

$$[s_{-r}...s_r]_{-r} \subseteq \cup_{i \in I'_q} \mathcal{O}_i^q.$$

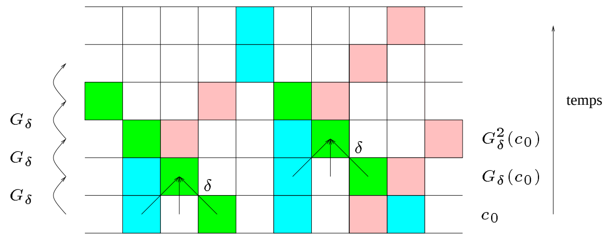


FIG. 1.1 – Exemple de diagramme espace-temps.

On vérifie que F, parce qu'elle est invariante par décalage, est la fonction globale associée à δ. $\square$

Remarque 1.1.1 *Notons que le théorème de Hedlund justifie la composition de deux automates cellulaires. Si $\mathcal{A}$ et $\mathcal{B}$ sont deux automates cellulaires, alors la fonction $G_{\delta_\mathcal{A}} \circ G_{\delta_\mathcal{B}}$ est une fonction continue et invariante par décalage de $Q^\mathbb{Z} \longrightarrow Q^\mathbb{Z}$. Cette fonction est donc la fonction globale de transition d'un automate cellulaire que nous notons $\mathcal{A} \circ \mathcal{B}$.*

Définition 1.1.3 *Un* automate cellulaire élémentaire (ECA) *est un AC unidimensionnel de rayon $r = 1$ et avec deux états (en général* 0 *et* 1*).*

Pour représenter une trajectoire d'un automate cellulaire, nous utiliserons un diagramme espace-temps. Les configurations seront représentées par des lignes de carrés dont la couleur indiquera l'état de la cellule. Les configurations successives de l'évolution seront placées les unes au dessus des autres, le temps allant vers le haut (cf figure 1.1). Pour choisir la configuration initiale des diagrammes espace-temps, il est tentant de vouloir choisir une configuration « aléatoire » afin que le diagramme ne représente pas un cas particulier de trajectoire, mais permette d'avoir une idée générale du comportement de l'automate cellulaire.

1.2 Mesure sur l'ensemble des configurations

Historiquement, Wolfram a observé de nombreux diagrammes espace-temps en partant de configurations « aléatoires »[1] pour établir sa classification. La justification implicite de ce protocole expérimental est d'obtenir un diagramme espace-temps représentatif de « la plupart » des diagrammes espace-temps. Il est alors possible de parler du comportement, de l'évolution

[1] Nous mettons aléatoires entre guillemets parce qu'elles ont été générées par ordinateur et qu'elles ne sont donc pas aléatoires pour la définition que nous donnons par la suite.

ou de la dynamique d'un automate cellulaire en faisant référence au comportement, à l'évolution ou à la dynamique d'une trajectoire typique, c'est-à-dire en partant d'une configuration aléatoire.

Afin d'étudier formellement l'évolution des automates cellulaires en partant des configurations aléatoires, nous devons définir ces dernières. Pour qu'une configuration (c'est-à-dire une suite bi-infinie d'éléments de Q l'ensemble des états) soit aléatoire, il semble nécessaire que tous les mots sur Q (tous les éléments de Q^*) apparaissent dans la suite. Cela paraît raisonnable et est intéressant pour le but que nous poursuivons, car alors, l'évolution d'un automate en partant d'une configuration aléatoire contient toutes les évolutions finies de l'automate et peut donc être vue comme représentative de l'ensemble de ses évolutions. Cependant, cette condition, qui est équivalente au fait que la configuration ait une entropie topologique maximale, n'est pas suffisante pour assurer le caractère aléatoire de la configuration. En effet, si $Q = \{0, 1\}$, nous pouvons écrire en binaire tous les entiers à la suite. Cette suite commence par

$$0\,1\,10\,11\,100\,101\,110\,111\,1000\,1001\,1010...$$

et il est facile de montrer que tous les éléments de Q^* apparaissent. Il suffit de compléter cette suite par symétrie pour un faire une suite bi-infinie qui n'est pas aléatoire. Pourquoi n'est-elle pas aléatoire ? Parce qu'il existe un programme fini qui écrit cette suite. En fait, on connait une définition formelle d'aléatoire. Une suite binaire aléatoire est, par définition, une suite telle qu'il n'est pas possible de décrire pour tout n les n premiers termes de la suite en moins de $n - C$ bits, où C est une constante. Cela se formalise à l'aide de la complexité de Kolmogorov[2]. Sans rappeler toute le théorie qui est derrière la complexité de Kolmogorov (cf [32] pour une introduction sur ce sujet), la complexité de Kolmogorov d'un mot est la taille minimale d'un programme qui écrit ce mot dans un langage de programmation (ou un modèle de calcul) acceptable, c'est-à-dire optimal à une constante près. Ainsi, une suite bi-infinie aléatoire est une suite telle que pour tout n, la complexité de Kolmogorov des $2n + 1$ termes centrés en 0 de la suite est $2n + 1$ à une constante additive près, indépendante de n.

La complexité de Kolmogorov n'a pas été introduite pour définir l'aléatoire mais pour mesurer la quantité d'information d'un objet. Elle a donc

[2]Il existe une autre définition d'une suite aléatoire dont nous ne parlerons pas dans cette thèse. Une suite est aléatoire pour Martin-Löf si elle est *typique*, c'est-à-dire qu'elle appartient à toutes les majorités « raisonnables » [45]. Levin et Schnorr ont démontré que ces deux définitions étaient équivalentes. Le fait que deux approches très différentes aboutissent à la même notion laisse penser qu'il s'agit de la bonne définition d'une suite aléatoire.

été développée dans le cadre de ce qui s'appelle la théorie de l'information. Cette théorie propose notamment une définition axiomatique des fonctions d'information. Ces fonctions définissent intuitivement la quantité d'information d'un objet. Elles sont définies sur un ensemble d'objets et prennent leurs valeurs dans $\mathbb{R}^+$. L'ensemble d'objets doit contenir un objet « nécessairement connu » noté ϵ et il doit exister une opération de couplage qui, à deux objets x et y, associe un nouvel objet noté (x, y). Les fonctions d'information sont celles qui vérifient un certain nombre de conditions intuitives (la fonction est à valeur positive, la quantité d'information du couple (x, x) est la même que celle de x, la quantité d'information du couple (x, y) est inférieure à la somme des quantités d'information...). En fait, la complexité de Kolmogorov est une fonction d'information. Notons qu'une suite est aléatoire si sa quantité d'information est maximale pour la complexité de Kolmogorov. Il nous arrivera ainsi par la suite d'assimiler la quantité d'information d'une configuration à son caractère plus ou moins aléatoire.

En fait, nous utiliserons la complexité de Kolmogorov pour justifier intuitivement certains résultats et certains protocoles expérimentaux. Par contre, nous ne l'utiliserons dans aucun théorème. En effet, les définitions de suite aléatoire et de quantité d'information qui en découlent sont les plus pertinentes sur le plan théorique mais ne sont pas suffisamment paramétrables. Wolfram génère par exemple des configurations dans lesquelles l'état d'une cellule est choisi aléatoirement mais avec des probabilités différentes pour chaque état. L'étude des évolutions des automates cellulaires en partant de ces configurations est très intéressante, bien qu'elles ne soient pas aléatoires pour la définition que nous venons de donner. La complexité de Kolmogorov n'est pas adaptée pour caractériser ces configurations.

Il se trouve qu'avant Kolmogorov, une autre fonction d'information a été définie, il s'agit de l'entropie de Shannon. Puisque la complexité de Kolmogorov ne convient pas, il est naturel de regarder si cette autre fonction d'information est plus adaptée à nos besoins. Contrairement à la complexité de Kolmogorov, qui est algorithmique, l'approche de Shannon est probabiliste. Nous aurons donc besoin de quelques notions de probabilité et en premier lieu de la notion de mesure dont la définition est l'objet de cette section. L'entropie de Shannon est l'entropie métrique pour une mesure donnée, celle-ci sera définie formellement page 85.

La théorie de la mesure a été introduite pour formaliser les probabilités sur des ensembles infinis. En particulier, pour expliquer qu'il y a plusieurs façons de choisir aléatoirement trois points sur un cercle et qu'il est néanmoins possible de faire rigoureusement des probabilités dans ce cas à condition de se fixer une méthode pour choisir les trois points, c'est-à-dire une mesure. La mesure qui accorde la même probabilité à tous les mots de taille donnée, est dite uniforme. Dans notre thèse, nous ne nous contenterons pas de la mesure uniforme, nous utiliserons en fait les mesures de Bernoulli. Dans le cas des suites binaires, la mesure uniforme correspond aux suites de résultats de pile ou face, les autres mesures de Bernoulli sont celles qui correspondent aux suites de résultat de pile ou face avec une pièce plus ou moins biaisée qui retombe plus souvent d'un côté que de l'autre. L'entropie de Shannon est donc plus expressive que la complexité de Kolmogorov puisque les mesures de Bernoulli vont nous permettre d'avoir un cadre formel pour étudier les configurations générées par Wolfram dont les états sont choisis aléatoirement, mais avec des probabilités différentes pour chaque état. En référence aux configurations que Wolfram qualifie d'aléatoires et que nous associons à une mesure de Bernoulli, et par abus de langage, nous parlerons de configurations aléatoires pour une mesure. Cette notion n'est pas définie en théorie ergodique (cf [53, 58] pour une introduction à ce domaine), nous l'utiliserons avec un sens très proche de la notion de configuration générique pour une configuration.

L'adéquation est parfaite entre la complexité de Kolmogorov et la notion d'aléatoire. Ce n'est pas le cas avec l'entropie de Shannon, même s'il y a des liens très forts que nous explicitons ci-dessous. Il se trouve que, pour la mesure uniforme, une suite aléatoire pour Kolmogorov est d'entropie métrique maximale, la réciproque étant fausse car l'exemple précédent (la suite des entiers écrits en binaire) a une entropie métrique maximale. Un autre résultat très intéressant est que l'ensemble des suite aléatoires est de mesure (uniforme) 1, c'est-à-dire que « presque toutes » les suites sont aléatoires. Ainsi, une propriété vérifiée pour les configurations aléatoires est vraie « presque partout ». Réciproquement, une propriété vérifiée « presque partout » est par définition vraie sur un ensemble de mesure 1 dans lequel est inclus un ensemble de configurations aléatoires qui est également de mesure 1. La notion de vérité « presque partout » va donc nous permettre de formaliser les observations que l'on peut faire à partir d'une configuration générée aléatoirement.

Nous supposons ici que l'ensemble Q des états contient au moins deux éléments.

Définition 1.2.1 *Une* mesure de probabilité de Borel sur $Q^{\mathbb{Z}}$ *est une fonction positive μ définie sur les Boréliens (l'ensemble des ouvert clos pour les opérations de complémentaire et d'union, ce qui inclut les fermés, mais pas seulement puisque les intersections infinies d'ouverts sont possibles). Elles est caractérisée par ses valeurs sur les cylindres, vérifie $\mu(Q^{\mathbb{Z}}) = 1$ et pour* *tout* $u \in Q^+$, $k \in \mathbb{Z}$,

$$\sum_{q \in Q} \mu[uq]_k = \mu[u]_k \ \ et \sum_{q \in Q} \mu[qu]_k = \mu[u]_{k+1}$$

Définition 1.2.2 *Une mesure de probabilité est dite de* support total *si et seulement si le seul ensemble fermé de mesure* 1 *est l'espace tout entier.*

Une mesure est de support total si et seulement si tous les ensembles de mesure 1 intersectent tous les ouverts de l'espace, c'est-à-dire d'une certaine façon que la mesure représente l'espace $Q^{\mathbb{Z}}$ tout entier.

Définition 1.2.3 *Une mesure de probabilité* μ *est* σ-invariante *(*σ *est le décalage) si pour tous les ensembles mesurables* Y, $\mu(\sigma^{-1}(Y)) = \mu(Y)$.

Notons que les notions d'invariance, d'ergodicité et de mélange faible sont définies pour toutes les fonctions continues, mais que nous ne les utiliserons que pour le décalage σ.

Définition 1.2.4 *Une mesure* σ*-invariante est* σ-ergodique *si pour tout ensemble mesurable shift-invariant* Y $(\sigma(Y) = Y)$, *soit* $\mu(Y) = 0$ *soit* $\mu(Y) = 1$.

Le fait qu'une mesure μ soit σ-ergodique n'implique pas que $\mu \times \mu$, le produit de la mesure par elle même, soit $\sigma \times \sigma$-ergodique. Cette propriété nous sera utile, elle est équivalente à la notion de mélange faible définie ci-dessous, qui est plus forte que celle de σ-ergodicité. L'exemple classique de mesure ergodique non faiblement mélangeante est la mesure uniforme ν pour la rotation R d'angle irrationnel sur le cercle. En effet, les orbites sont denses ce qui implique l'ergodicité, mais on voit bien que la rotation ne mélange pas du tout les points pour la mesure uniforme et on peut vérifier qu'elle n'est pas R-faiblement mélangeante parce que $\nu \times \nu$ n'est pas ergodique.

Définition 1.2.5 *Les mesures* σ-faiblement mélangeantes *sont caractérisées par l'équivalence suivante.*

Fait 1.2.1 *Soit* μ *une mesure* σ*-invariante, les propriétés suivantes sont équivalentes :*

1. μ *est* σ*-faiblement mélangeante.*
2. $\mu \times \nu$ *est* σ*-faiblement mélangeante pour toute mesure* ν σ*-faiblement mélangeante.*
3. $\mu \times \nu$ *est* σ*-ergodique pour toute mesure* ν σ*-ergodique.*

Les mesures de Bernoulli sont les mesures les plus simples, l'idée est que toutes les cellules sont indépendantes. La probabilité de chaque état q est donnée par une constante p_q. Leur simplicité est la raison pour laquelle nous allons les utiliser dans tous nos exemples bien que les définitions que nous allons donner, et probablement la plupart des propositions, restent vraies pour d'autres mesures σ-faiblement mélangeantes, par exemple les mesures de Markov (avec des relations sur un nombre fini de cellules) ou les mesures telles que les corrélations entre deux états décroissent exponentiellement avec leur distance.

Définition 1.2.6 *Une mesure de Bernoulli est définie par un vecteur de probabilité strictement positif* $(p_q)_{q\in Q}$ *avec* $\sum_{q\in Q} p_q = 1$ *et, pour* $u = u_0...u_{n-1} \in Q^n$, $\mu[u_0...u_{n-1}] = p_{u_0}...p_{u_{n-1}}$.

Pour les AC à deux états, les mesures de Bernoulli seront notées μ_ρ où $\rho = p_1 = 1 - p_0$ est la probabilité qu'un état soit 1.

Nous allons utiliser les résultats classiques suivants [53, 58] :

Fait 1.2.2 *Les mesures de Bernoulli sont de support total et σ-faiblement mélangeantes (elles sont donc également σ-ergodiques).*

1.3 Automates groupés

La notion de groupage est une opération algébrique qui transforme un automate cellulaire en un automate dont chaque état contient k états de l'automate initial et qui simule, en une seule étape, k étapes de calcul de ce dernier. Le nouvel automate a donc beaucoup plus d'états que l'automate initial ($(\#Q)^k$ pour être précis). On transforme un diagramme espace-temps de l'automate initial en un diagramme de l'automate groupé en compressant spatialement les configurations d'un facteur k et en ne conservant que les temps multiples de k. Bien que les opérations spatiales et temporelles soient différentes, les deux opèrent une contraction de rapport $1/k$, ce qui fait globalement penser à un changement d'échelle ou, plus précisément, à un zoom arrière. Nous avons expliqué dans l'introduction que nous souhaitions définir des notions invariantes par changement d'échelle, le groupage va nous permettre de formaliser cela.

Étant donné un automate cellulaire (Q, δ), et une cellule sur son diagramme espace-temps, le graphe de dépendance de son état est un cône (cf figure 1.2 (a)). Nous pouvons donc définir récursivement la fonction $\delta^n : Q^{2nr+1} \longrightarrow Q$ qui calcule l'état de la cellule centrale d'un bloc de $2nr+1$ cellules après n étapes de calcul :

$$\begin{aligned}
\delta^1(x_{-r}, ..., x_0, ..., x_r) &= \delta(x_{-r}, ..., x_0, ..., x_r) \\
\delta^{n+1}(x_{-(n+1)r}, ..., x_{(n+1)r}) &= \delta^n(\delta(x_{-(n+1)r}, ..., x_{-(n-1)r+1}), ..., \\
&\qquad \delta(x_{(n-1)r-1}, ..., x_{(n+1)r}))
\end{aligned}$$

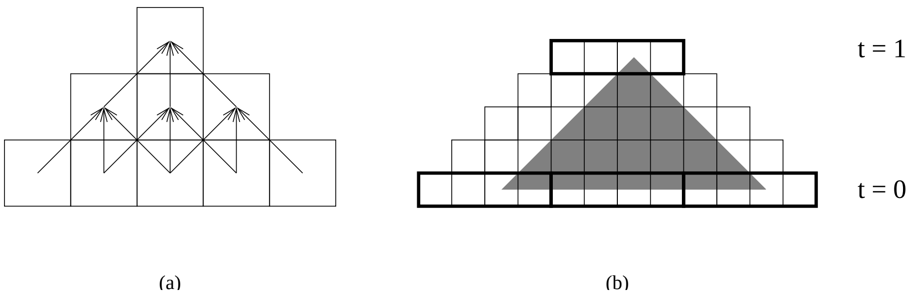


FIG. 1.2 – (a) La relation de dépendance entre les états dans le diagramme espace-temps d'un AC.
(b) Un représentation du groupage (par 4) d'un automate cellulaire élémentaire.

En groupant les états k par k et en faisant k étapes de calcul de l'AC initial en une seule étape (figure 1.2 (b)), les k cellules centrales k étapes plus tard ne dépendent bien que des $2kr+1$ cellules connues à l'étape précédente, nous pouvons définir un nouvel automate cellulaire avec (exponentiellement en fonction de k) plus d'états. Comme nous l'avons vu, sur les diagrammes espace-temps, cette transformation ressemble à un changement d'échelle, plus précisément à un zoom arrière. Il y a formellement un isomorphisme entre les configurations d'un automate et celles de l'automate groupé. Il y a donc une bijection entre les trajectoires des deux automates. De ce fait, ils ont les mêmes propriétés dynamiques.

Définition 1.3.1 *Soit $\mathcal{A}$ un automate cellulaire, l'automate groupé k fois de $\mathcal{A}$ est l'automate noté $\mathcal{A}^k = (Q_\mathcal{A}, \delta_\mathcal{A})^k = (Q_\mathcal{A}^k, \delta_{\mathcal{A}^k})$ tel que*

$$\forall i \in \{1, ..., k\} \quad (\delta_{\mathcal{A}^k}(\vec{x}^{-r}, ..., \vec{x}^{r}))_i = {\delta_\mathcal{A}}^k(x_i^{-r}, ..., x_k^{-r}, ..., x_1^r, ..., x_i^r)$$

avec les conventions suivantes : un élément $(q_1, ..., q_k) \in Q_\mathcal{A}^k$ est noté $\vec{q}$ et $(\vec{q})_i = q_i$ représente la $i^{ème}$ coordonnée de $\vec{q}$.

Notation 1.3.1 *Nous utiliserons la même notation $\mathcal{A}^i$ pour les automates à deux états, ainsi 54^4 représentera l'automate 54 groupé 4 fois et non le nombre.*

1.4 Produit d'automates

Le produit est une autre opération algébrique que nous utiliserons afin de construire des contre-exemples dans le dernier chapitre. Étant donnés deux AC $\mathcal{A}$ et $\mathcal{B}$, l'automate produit est un automate qui effectue en parallèle un calcul de $\mathcal{A}$ et un calcul de $\mathcal{B}$. Dans la suite, il nous arrivera d'interpréter le produit de deux automates comme une opération de superposition de leurs diagrammes espace-temps.

Définition 1.4.1 *Soient $\mathcal{A}$ et $\mathcal{B}$ deux AC, leur automate produit est*

$$\mathcal{A} \times \mathcal{B} = (Q_{\mathcal{A}} \times Q_{\mathcal{B}}, \delta_{\mathcal{A}\times\mathcal{B}}))$$

où $\delta_{\mathcal{A}\times\mathcal{B}}((x_{\mathcal{A}}^{-r}, x_{\mathcal{B}}^{-r}), ..., (x_{\mathcal{A}}^{r}, x_{\mathcal{B}}^{r})) = (\delta_{\mathcal{A}}(x_{\mathcal{A}}^{-r}, ..., x_{\mathcal{A}}^{r}), \delta_{\mathcal{B}}(x_{\mathcal{B}}^{-r}, ..., x_{\mathcal{B}}^{r}))$.

Chapitre 2

Exemples d'automates cellulaires

Ce chapitre est consacré à l'étude de quelques automates cellulaires sur lesquels nous reviendrons souvent par la suite. En effet, l'étude d'un certain nombre d'exemples (principalement les ECA parce qu'une étude exhaustive en est possible, mais pas uniquement) a joué un rôle important dans notre travail. Tout d'abord parce que nous avons cherché à comprendre leur fonctionnement et leur dynamique, ce qui nous a conduit aux résultats du chapitre 3, ensuite parce qu'il nous a semblé que les classifications proposées ainsi que les définitions du chaos ne correspondaient pas vraiment avec la réalité ne serait-ce que dans les cas simples que nous avions étudiés. En fait, nous reviendrons dans le chapitre 4 sur les classifications existantes en les discutant sur des exemples concrets, avant de proposer de nouvelles notions et méthodes dans les deux chapitres suivants.

Nous profitons également de ce chapitre pour introduire des notations que nous serons amenés à utiliser souvent dans la suite de ce document.

2.1 La notation de Wolfram pour les ECA

Nous utiliserons pour les automates à deux états (0 et 1) la notation de Wolfram : ils sont numérotés de 0 à 255 de telle sorte que la fonction de transition de l'automate i dont l'écriture en base 2 est $\overline{a_7a_6a_5a_4a_3a_2a_1a_0}^2$ vérifie :

$$\begin{array}{ll} \delta_i(0,0,0)=a_0 & \delta_i(1,0,0)=a_4 \\ \delta_i(0,0,1)=a_1 & \delta_i(1,0,1)=a_5 \\ \delta_i(0,1,0)=a_2 & \delta_i(1,1,0)=a_6 \\ \delta_i(0,1,1)=a_3 & \delta_i(1,1,1)=a_7 \end{array}$$

Nous parlerons ainsi de l'automate $54 = (\{0,1\}, \delta_{54})$ ou de la règle 54. Notons que des automates ayant des numéros différents peuvent avoir le même comportement en échangeant les états 0 et 1 ou par symétrie. Remarquons que l'on passe d'une règle à celle qui a le même comportement mais en inversant les états 0 et 1, en échangeant les 0 et les 1 dans l'écriture binaire de la règle puis en la lisant de droite à gauche. C'est ainsi que l'on passe de $110 = \overline{01101110}^2$ à $137 = \overline{10001001}^2$ (nous noterons $137 = \overline{110}$). En outre, par symétrie, les règles $124 = \overline{01111100}^2$ (symétrique de 110, nous noterons $124 = \overleftrightarrow{110}$) et $193 = \overline{11000001}^2 = \overleftrightarrow{137}$ ont également le même comportement (cf figure 2.1).

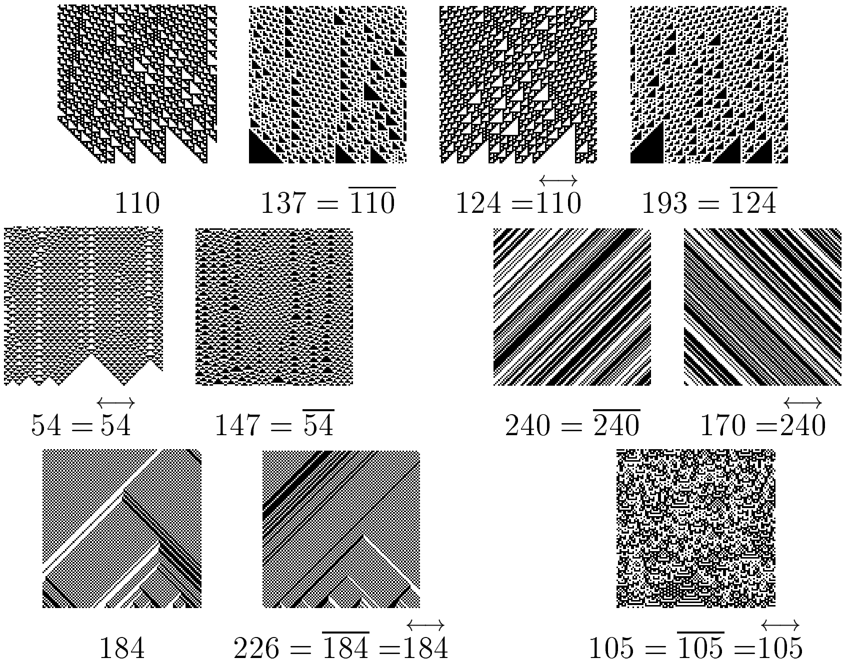
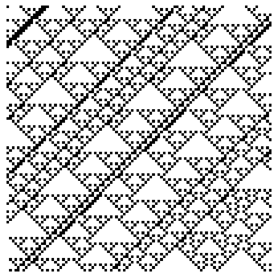


Fig. 2.1 – Il peut y avoir jusqu'à quatre numéros pour le même comportement dans la notation de Wolfram. Il n'y en a parfois que deux ou bien un seul ; en effet, certaines règles sont leur propre symétrique (54, 147), sont invariantes par échange des 0 et des 1 (les décalages 240 et 170), il se peut aussi que leur symétrique coïncide avec l'automate obtenu en échangeant les 0 et les 1 (184, 226), enfin certaines règles sont à la fois leur propre symétrique et invariantes par échange des 0 et des 1 (105).

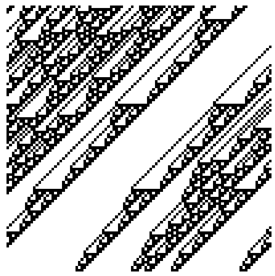
2.2 Les règles 210 et 120

Les règles 210 et 120 sont très intéressantes. En effet, elles sont construites sur le même principe, que nous allons détailler, tandis que leurs comportements sont très différents. Comme nous le montre le tableau ci-dessous, le nouvel état d'une cellule est l'état de la cellule placée à sa gauche à une exception près : sauf si les cellules centrale et de droite sont dans les états 01 pour 210 et 11 pour 120, le nouvel état est alors différent de celui de la cellule de gauche. Elles sont donc toutes les deux permutives à gauche (c'est-à-dire que les fonctions $\delta_{120}(.,x,y)$ et $\delta_{210}(.,x,y)$ sont bijectives pour tous les couples $(x,y) \in \{0,1\}^2$).

	210	120
000	0	0
001	1	0
010	0	0
011	0	1
100	1	1
101	0	1
110	1	1
111	1	0



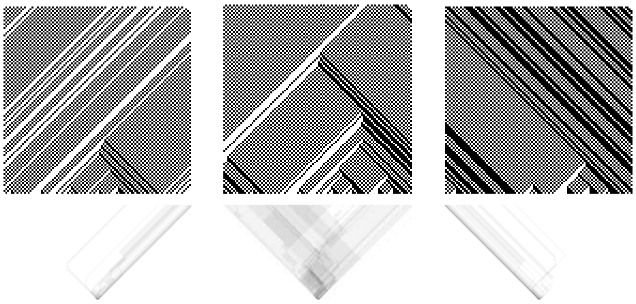
210



120

Malgré leur définition presque identique, leurs comportements sont très différents. En effet, la succession de deux états 1 se comporte comme un mur se translatant vers la droite pour 210, ce qui fait que son aspect général ressemble à un décalage, bien que son évolution ne soit pas la composition d'un décalage avec un AC ayant un comportement périodique. En effet, son évolution se décompose en bandes, mais on démontre que les périodes dans les bandes peuvent être arbitrairement grandes. En ce qui concerne 120, au contraire, aucun mur ne vient entraver la propagation d'information vers la gauche et quand on le fait évoluer à partir de configurations générées aléatoirement pour différentes densités de 1, on observe que ses configurations ressemblent de plus en plus à des configurations aléatoires.

Ces deux exemples montrent bien que la prédiction du comportement d'un AC, connaissant sa fonction de transition locale est un exercice très difficile. En particulier, toutes les approches visant à classer les AC en fonction d'un paramètre calculé à partir de leurs fonctions de transitions [31, 24, 15] se trouvent prises en défaut.



$\rho = 0,4$ $\rho = 0,5$ $\rho = 0,6$

FIG. 2.2 – Nous avons représenté les diagrammes espace-temps de 184 pour des densités de 1 de $0,4$, $0,5$ et $0,6$. Sous les diagrammes espace-temps, nous voyons en niveau de gris la répartition dans l'espace du nombre moyen de modifications provoqué par la modification de la cellule centrale (cf chapitre 6).

2.3 La règle 184 et l'automate T

La règle 184 est l'exemple non trivial le plus simple de ce que Boccara appelle les « number preserving rules » [8], ce sont les AC dont les états sont des entiers, et dont la définition est la suivante.

Definition 2.3.1 *Un AC unidimensionnel $\mathcal{A}$ de rayon r avec $Q_{\mathcal{A}} = \{1, ..., q\}$ est « number preserving » si pour toutes les configurations cycliques de périodes $L \leq 2r + 1$, il vérifie*

$$\delta_{\mathcal{A}}(x_0, x_1, ..., x_{2r}, x_{2r+1}) + \delta_{\mathcal{A}}(x_1, x_2, ..., x_{2r+1}, x_{2r+2}) + ...$$
$$+\delta_{\mathcal{A}}(x_L, x_{L+1}, ..., x_{L+2r-1}, x_{L+2r}) = x_1 + x_2 + ... + x_L.$$

Dans le cas de la règle 184, le nombre de cellules blanches et noires est préservé au cours des évolutions. De plus, les séquences de plus de deux particules blanches se décalent vers la droite tandis que celles de cellules noires se décalent vers la gauche. Lors de la rencontre de ces groupes de cellules que nous pouvons assimiler à des particules, le groupe le moins nombreux disparaît réduisant d'autant le groupe le plus nombreux qui continu sur sa lancée (cf figure 2.2). Ainsi, cette règle compare le nombre de cellules blanches et noires de la configuration initiale puisqu'asymptotiquement, plus aucune séquence de deux cellules blanches n'apparaît si la densité de cellules noires était supérieure initialement, et réciproquement.

Boccara [9] donne une interprétation de cette règle en tant que modèle de trafic routier. En effet, si nous considérons que les cases noires sont occupées par une voiture tandis que les blanches sont libres, le comportement des voitures est d'avancer d'une case vers la droite si celle-ci est libre et de rester

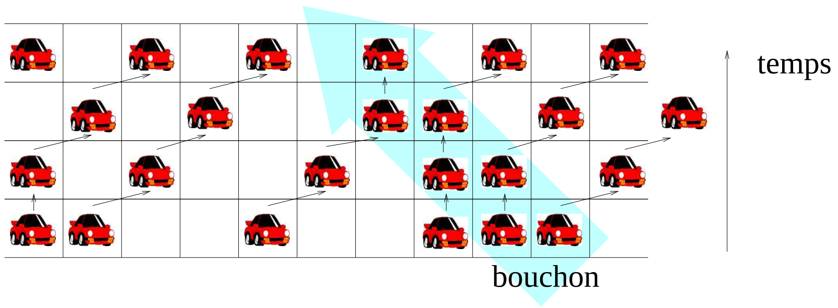


FIG. 2.3 – Le fonctionnement de la règle 184 peut être interprété comme un modèle de trafic routier.

sur place dans le cas contraire (cf figure 2.3). Ainsi, les séquences de cellules noires sont interprétées comme des embouteillages. La première voiture de l'embouteillage (la plus à droite) peut avancer, par définition, ce qui fait qu'à chaque étape, le début de l'embouteillage recule d'une position. En queue de bouchon, si la dernière voiture n'est séparée que par une cellule de la voiture qui la suit, alors cette dernière augmente d'un le nombre de voitures bloquées et la fin de l'embouteillage recule d'une position. Ainsi, tant que la densité des voitures arrivantes est 1/2, le bouchon recule en conservant la même taille, sinon, il va se résorber. On observe exactement le même comportement des bouchons dans la réalité, sauf qu'au delà d'une certaine densité des voitures qui arrivent, sa taille peut aussi augmenter, ce qui n'est pas possible dans le cas de la règle 184. Il suffit de modifier la règle de telle sorte qu'une voiture se met à avancer alors qu'elle était arrêtée si et seulement si les deux positions devant elle sont libres pour observer ce genre de comportement.

Nous allons utiliser ces caractéristiques pour construire un autre AC. Notre objectif est de concevoir un automate dont le comportement change en fonction de la densité des états sur sa configuration initiale de telle sorte qu'il ait un comportement « simple » ou « chaotique » selon les cas. Cette recherche est motivée par des phénomènes naturels tels que le changement de phase solide/liquide ou l'écoulement de fluides dont le comportement change brutalement lorsqu'un paramètre (la température dans le premier cas, la vitesse du fluide dans le second) dépasse une valeur limite. Le trafic routier présente également cette propriété en fonction de la densité des voitures, en ce qui concerne le passage entre une circulation fluide et la présence d'embouteillages. Le problème de la règle 184 est que les voitures sont identiques et donc indiscernables. Pour remédier à cela, nous définissons l'automate T (cf figure 2.4) qui simule un trafic routier de la même façon que la règle 184,

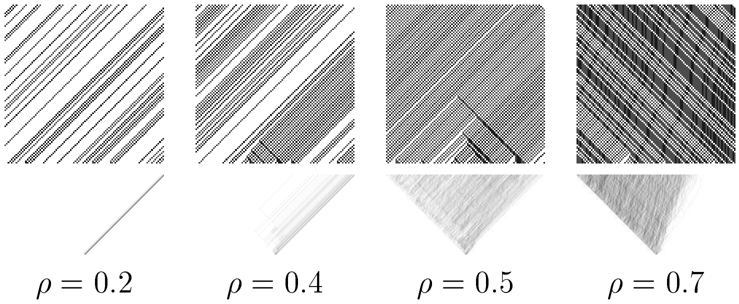


FIG. 2.4 – L'AC T est un modèle très simple de trafic routier qui change de comportement en fonction de la densité des voitures (en raison des embouteillages). Nous voyons que le nombre moyen de modifications dans une configuration donnée est borné pour de faibles densités de voitures (strictement inférieur à 1/2) et non borné sinon.

mais avec deux modèles distincts de voitures, des voitures gris foncé et des voitures gris clair. Comme prévu, le système semble « chaotique » lorsque la densité ρ de voitures est supérieure à $1/2$ en raison des embouteillages et non « chaotique » sinon. Ainsi, dire que cet AC est chaotique ou non n'a pas vraiment de sens puisque cela va dépendre de son utilisation : selon que l'on souhaite modéliser des bouchons ou un trafic fluide. Nous ne pouvons pas parler de son comportement moyen tant que nous ne définissons pas ce qu'est une configuration aléatoire, c'est-à-dire quelle mesure sur l'ensemble des configurations nous choisissons. Si nous supposons que la répartition des voitures est initialement uniforme et qu'il y a le même nombre de voitures gris clair et de voitures gris foncé, nous allons considérer les mesures de Bernoulli μ_ρ^* telles que la probabilité de trouver une voiture gris clair est $\rho/2$ et égale à la probabilité de trouver une voiture gris foncé tandis que la probabilité qu'il n'y ait pas de voiture est $1 - \rho$. Cela permettra de dire (cf chapitre 6) que cet AC est μ_ρ^*-presque partout sensible aux conditions initiales si $\rho \geq 1/2$ tandis qu'il est μ_ρ^*-presque jamais sensible aux conditions initiales sinon. S'il est important de prendre en compte le fait que de nombreux conducteurs prennent leur voiture à la même heure pour aller travailler, d'autres mesures permettent de modéliser des répartitions non uniformes par des mesures de Bernoulli sur des paquets de cellules.

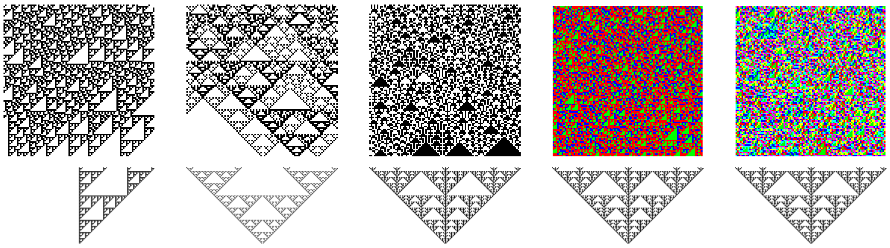
2.4 Les automates additifs

Les automates cellulaires les plus étudiés, sont, outre les automates ayant peu d'états, les familles d'automates se prêtant bien aux méthodes algébriques d'analyse. Parmi celles-ci nous utiliserons la famille des AC additifs.

Dans la définition générale des automates additifs de Wolfram [44, 19], un AC additif est un AC dont l'ensemble d'états est un groupe et qui vérifie le principe de superposition ($\delta(x+x', y+y', z+z') = \delta(x,y,z) + \delta(x',y',z')$). Ces AC sont très intéressants pour fournir des exemples parce que leurs évolutions obéissent à des règles algébriques qu'il est possible d'étudier formellement tandis que leurs diagrammes espace-temps semblent compliqués. Cependant, lorsqu'il s'agira de démontrer des propositions du type « Les AC additifs sont ... », nous utiliserons, comme dans [47, 30], une définition plus restrictive :

Définition 2.4.1 *Un* automate cellulaire additif *est un AC unidimensionnel dont l'ensemble d'états est* $\mathbb{Z}/n\mathbb{Z}$ *et dont la fonction de transition est de la forme :*

$$\delta_{+,n}(x_{-1}, x_0, x_1) = x_0 + x_1 (mod\ n)$$



60 90 150 $\mathbb{Z}/3\mathbb{Z}$ $\mathbb{Z}/6\mathbb{Z}$

FIG. 2.5 – Quelque automates additifs, leur diagramme espace-temps est en fait une superposition de triangles de Pascal.

Par contre, lorsqu'il s'agira de donner des diagrammes espace-temps d'additifs, notamment dans le chapitre suivant (et la figure 2.5), nous utiliserons également des automates dont la fonction de transition est du type $\delta(x_{-1}, x_0, x_1) = x_{-1} + x_0 (\text{mod } n)$, $\delta(x_{-1}, x_0, x_1) = x_{-1} + x_1 (\text{mod } n)$ ou bien $\delta(x_{-1}, x_0, x_1) = x_{-1} + x_0 + x_1 (\text{mod } n)$. En fait, la plupart du temps, ces automates ont les mêmes propriétés, mais la généralisation aux fonction $\delta(x_{-r}, ..., x_r) = P(x_{-r}, ..., x_r)(\text{mod } n)$ (P étant un polynôme) est souvent délicate.

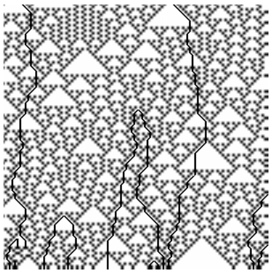
Ce qui rend l'étude des automates cellulaires additifs possible, c'est leur représentation dipolynomiale. Il s'agit d'associer à l'automate de rayon r dont la fonction de transition globale est G_δ, le dipolynôme $\delta(X^{-r}, ..., X^r)$. Ce dipôlynome est bien défini parce que la règle est additive, $\delta(X^{-r}, ..., X^r) = \delta(1, 0, ..., 0)X^{-r} + \delta(0, 1, 0, ..., 0)X^{-r+1} + ... + \delta(0, 0, ..., 1)X^r$. L'intérêt de cette notation est que si on représente la configuration c par la série entière $\sum_{i\in\mathbb{Z}} c_i X^i$ alors la représentation de $G_\delta(c)$ est la série formelle

$$\delta(X^{-r}, ..., X^r) \times \sum_{i\in\mathbb{Z}} c_i X^i.$$

Remarquons que cette représentation peut s'interpréter comme une transformée de Fourier [19]. De plus, il est alors facile de montrer que le dipolynôme associé à l'automate qui fait n étapes de calcul d'un coup (sa fonction de transition globale est G_δ^n) est $\delta(X^{-r}, ..., X^r)^n$. Ce résultat permet d'interpréter les diagrammes des automates additifs comme une superposition de triangles de Pascal [54].

2.5 La règle 18

La règle 18 est définie par :



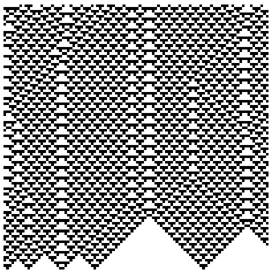
$$\delta_{18}(x,y,z) = \begin{cases} 1 & \text{si } (x,y,z) = \begin{cases} & (0,0,1) \\ \text{ou} & (1,0,0) \end{cases} \\ 0 & \text{sinon} \end{cases}$$

Cette règle ajoute les états de ses deux voisins modulo 2, mais seulement lorsque son propre état est 0. Elle n'est donc pas additive, mais peut présenter localement des comportements additifs. Son évolution est facile à comprendre expérimentalement : son diagramme espace-temps peut être découpé en régions qui sont de deux sortes, celles dont seules les cellules paires sont non nulles et celles dont seules les cellules impaires sont non nulles. Entre ces régions, il est possible de définir des particules comme cela a été remarqué plusieurs fois (par exemple [9]). Le déplacement de ces particules est probablement une sorte de marche aléatoire, sans que nous sachions le démontrer. Le problème pour pouvoir démontrer des résultats sur la règle 18 est précisément de comprendre le déplacement de ces particules. On peut par contre prouver que, lorsque deux particules se rencontrent, elles s'annihilent mutuellement et la région qui était entre les deux particules disparaît de façon irréversible. Ces collisions deviennent de plus en plus rares au fur et à mesure que le nombre de particules décroît, cependant, on peut penser qu'asymptotiquement, le nombre de particules tend vers zéro parce qu'un marcheur aléatoire sur $\mathbb{Z}$ a une probabilité 1 de visiter toutes les cellules. Dans chaque région, la règle 18 se comporte comme la règle 90. Ainsi, sur la moitié des cellules, il semble que la probabilité de 1 et de 0 tend vers 1/2 (c'est une propriété des règles additives qui est une conséquence de [18] et sur laquelle nous reviendrons dans la section 5.1) et donc qu'asymptotiquement la densité de 1 tend vers 1/4.

2.6 La règle 54 et son interprétation à l'aide de particules

La règle 54 est définie par :

$$\delta_{54}(x,y,z) = \begin{cases} 1 \quad \text{si } (x,y,z) = \begin{cases} & (0,0,1) \\ \text{ou} & (1,0,0) \\ \text{ou} & (0,1,0) \\ \text{ou} & (1,0,1) \end{cases} \\ 0 \quad \text{sinon} \end{cases}$$



Parmi les ECA, la règle 54 est l'une des plus célèbres. La principale raison de sa célébrité est l'article de Boccara [9] qui donne une interprétation (que nous allons rappeler ci-dessous) des diagrammes espace-temps de 54 à l'aide de particules. Par ailleurs, la règle 54 est l'une des deux règles (avec 110) ayant un comportement compliqué parmi les ECA (ce sont d'ailleurs les deux seules que Cattaneo ne peut pas classer avec sa méthode dans [11]). Enfin, la question de son universalité pour le calcul (sa capacité à simuler toutes les machines de Turing) est toujours ouverte.

Rappelons à présent l'interprétation de cette règle en termes de particules décrite dans [9]. L'observation d'un diagramme espace-temps de la règle 54 met en évidence un fond (ou background) ayant une période spatiale et temporelle de 4 cellules (voir les figures 2.6 et 2.7). Nous pouvons distinguer trois types de particules. Deux d'entre elles ne se déplacent pas et sont périodiques dans le temps, de période 4. Nous les nommerons g_0 (cf figure 2.6 (a)) et g_e (cf figures 2.6 (b) et 2.7). Il y a aussi une particule w qui peut se déplacer vers la droite ou vers la gauche à une vitesse 1. Nous noterons $\overleftarrow{w}$ les particules w se déplaçant vers la gauche et $\overrightarrow{w}$ celles qui vont vers la droite.

Les interactions entre ces particules sont assez riches, par exemple, comme nous le voyons figure 2.6 (a) nous avons :

$$\begin{array}{ccccc} \overrightarrow{w} & + & \overleftarrow{w} & \longrightarrow & g_0 \\ \overrightarrow{w} & + & g_0 & \longrightarrow & \overleftarrow{w} \\ \overleftarrow{w} & + & g_0 & \longrightarrow & \overrightarrow{w} \end{array}$$

La figure 2.6 (b) représente la collision d'une particule w avec une particule g_e. Les diagrammes (c) et (d) de la figure 2.7 montrent l'annihilation de deux particules g_e côte à côte. Deux particules g_e proches bombardées par une particule w peuvent faire apparaître une autre particule fixe (diagramme (e)) de période 32 qui émet en permanence des particules w (quatre $\overrightarrow{w}$ et quatre $\overleftarrow{w}$ durant chaque cycle). Cette particule peut être annihilée par une particule w qui, bien que retardée par les particules w rayonnées, finit par percuter la particule, ce qui conduit soit à son annihilation (comme sur le diagramme (e)) soit à la création de deux particules g_e.

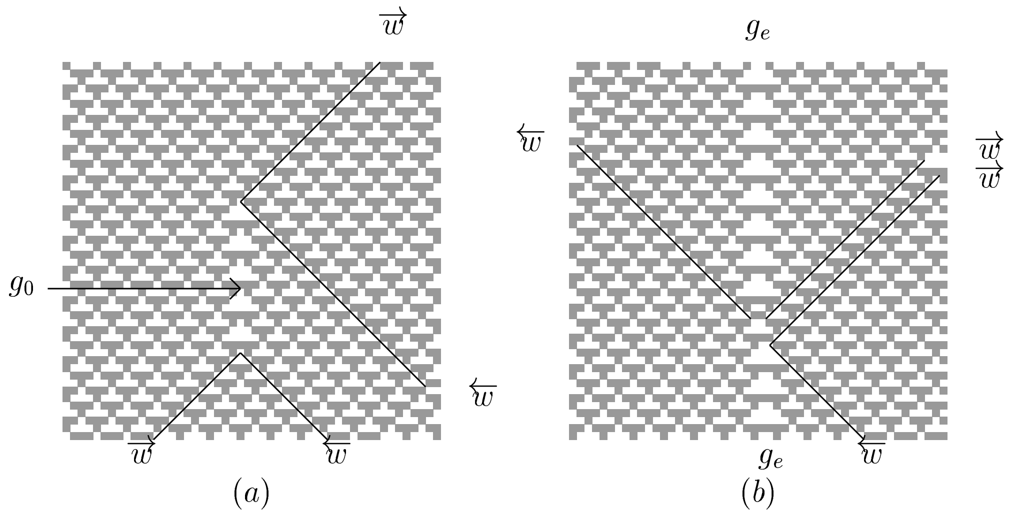


FIG. 2.6 – Les collisions entre une particule w et une particule g_0 ou g_e.

Nous allons étudier cette règle plus en détail dans le chapitre suivant.

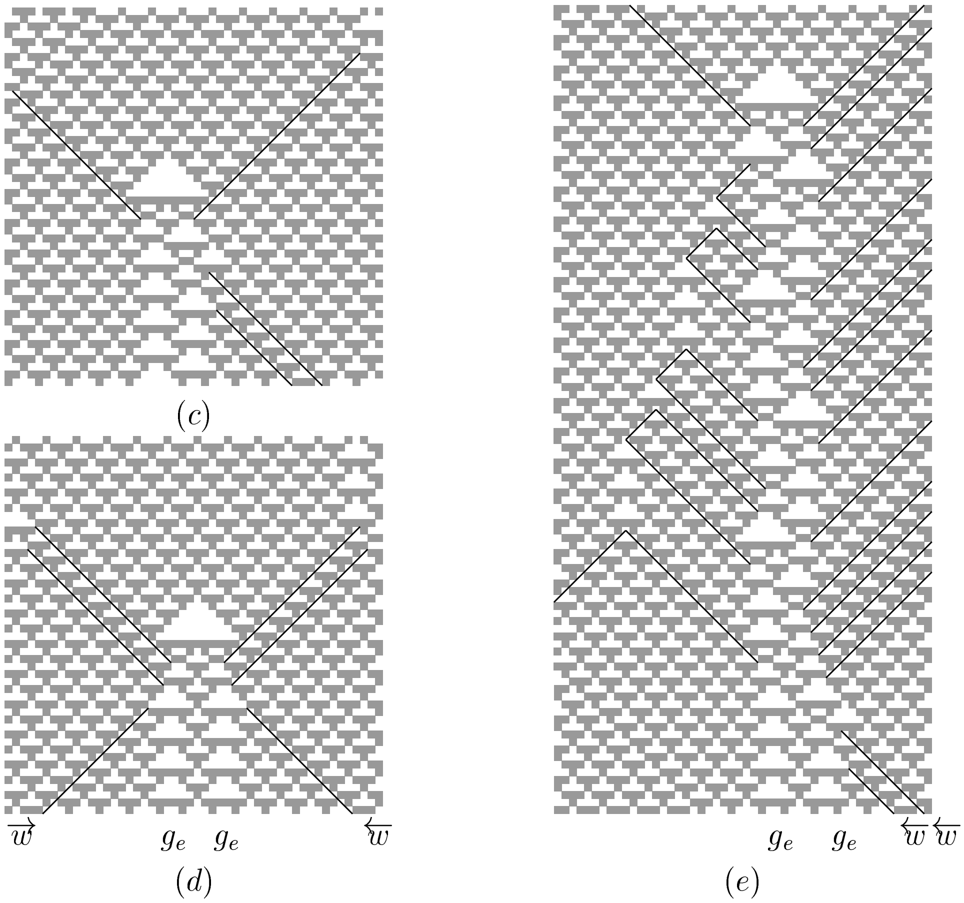


FIG. 2.7 – Annihilation de couples de particules g_e.

Chapitre 3

La règle 54

En observant les diagrammes espace-temps d'un certain nombre d'automates, dont 184, 54 et 110, on voit apparaître quelque chose que l'on interprète comme un fond (background) dans lequel se produisent des collisions de particules. Nous développons dans ce chapitre une méthode générale pour obtenir des résultats sur de telles règles que nous avons mise au point sur la règle 54 (qui est plus simple que 110, tandis que la règle 184 peut se comprendre facilement, indépendamment de la méthode). Dans ce chapitre, nos résultats sont spécifiques au cas unidimensionnel.

Dans un premier temps, nous allons donc étudier plus en détail le fond qui apparaît dans la règle 54, qui est de période 4. Notons que H. V. McIntosh [48] et N. Ollinger ont remarqué que la règle 110 présentait un fond de période 14, ce qui permet d'utiliser notre méthode sur cette règle. Cependant, nous sommes toujours incapable d'expliquer l'apparition quasi-systématique de ces fonds dans les diagrammes espace-temps de ces deux règles. Notons enfin que les régions de la règle 18 (cf section 2.5) ne sont pas composées d'une structure périodique, il n'y a donc pas de fond et la méthode présentée ici n'est pas directement utilisable dans ce cas. Cependant, une généralisation de la méthode permet dans ce cas encore de prouver que l'annihilation de deux particules est irréversible.

3.1 Les 8 fonds de la règle 54

L'interprétation de Boccara de la règle 54 introduit une notion de fond (background) de période spatiale et temporelle 4. Sa démarche a été d'observer des diagrammes de 54 tels que ceux des figures 2.6 et 2.7 et d'appeler fond le motif régulier composé d'un T noir et d'un T inversé blanc. Il a appelé particule les défauts constatés dans ce fond. Afin de pouvoir utiliser cette approche pour démontrer des résultats sur cette règle, il nous faut formaliser ces notions de fond et de particule. Nous allons donc définir, formellement, une seconde notion de fond. Aussi, afin d'éviter les confusions, nous parlerons de fond au sens de Boccara et de fond tout court pour distinguer les deux notions. En fait, la seconde notion de fond correspond à un état stable d'une instance groupée de l'automate.

Définition 3.1.1 *Soit $\mathcal{A}$ un automate cellulaire, un fond de cet AC est un état quiescent de l'automate groupé $\mathcal{A}^k$. c'est-à-dire un état $s \in Q_{\mathcal{A}}^k$ tel que $\delta_{\mathcal{A}^k}(s, ..., s) = s$.*

Remarque 3.1.1 *Notons que, pour cette définition, tous les AC ont des fonds. En effet l'évolution d'un AC à partir d'une configuration périodique est ultimement périodique parce que la période de la configuration initiale est préservée donc l'ensemble des configurations atteignables est fini. Or, d'une configuration spatialement périodique telle que l'évolution de l'AC en partant d'elle soit cyclique, on déduit un état quiescent de l'automate groupé k fois si k est le produit des périodes spatiales et temporelles.*

Nous allons à présent regarder comment exprimer le fond de Boccara avec notre nouvelle définition. Nous adopterons pour cela un point de vue géométrique. En fait, la notion de fond utilisée par Boccara correspond à un ensemble d'états d'une instance groupée, stable par translation et avec une notion de stabilité qu'il est difficile de formaliser. En effet, dans l'automate cellulaire 54^4 (i.e. la règle 54 groupée quatre fois), le fond décrit par Boccara apparaît uniforme (i.e. composé par un seul état groupé). Mais en fonction de l'origine choisie, nous obtenons 8 fonds stables de l'automate 54^4 qui sont également des fonds de 54 :

$$\begin{array}{ll} b_{0,0} = (0,0,0,1) & b_{0,1} = (1,0,1,1) \\ b_{1,0} = (0,0,1,0) & b_{1,1} = (0,1,1,1) \\ b_{2,0} = (0,1,0,0) & b_{2,1} = (1,1,1,0) \\ b_{3,0} = (1,0,0,0) & b_{3,1} = (1,1,0,1) \end{array}$$

Pour des raisons qui seront explicitées plus bas, nous avons indexé les fonds par les éléments du groupe G de présentation $< a, b \mid ab = ba, a^4 = 1, b^2 = a^2 >$[1] (qui a 8 éléments). Pour simplifier, nous notons $b_{k,l}$ le fond associé à l'élément $a^k b^l$ de G. Par exemple, si $k = 4$ et $l = 2$, $b_{4,2} = b_{0,2} = b_{2,0}$. Nous adopterons dans la suite cette dernière représentation (lorsque le premier indice est entre 0 et 3 et que le second est entre 0 et 1).

[1]Ce groupe est fini et abélien, il peut donc s'écrire comme un produit fini de groupes cycliques, ici $G \simeq \mathbb{Z}_4 \times \mathbb{Z}_2$, dont les composantes sont engendrées respectivement par a et ab.

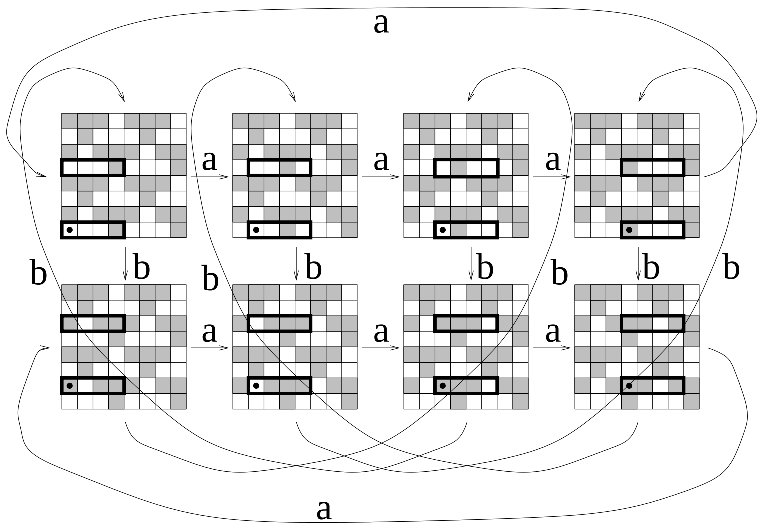


FIG. 3.1 – Le groupe indexant les 8 fonds de 54. Le générateur a correspond à un décalage vers la droite de l'origine du groupage (indiqué par un point) et b à un décalage vers le haut.

Nous allons expliquer l'introduction du groupe G pour indexer les fonds de 54. Considérons un diagramme espace-temps de 54 composé exclusivement du fond au sens de Boccara. En groupant l'automate 4 par 4, ce diagramme se transforme en un diagramme uniforme uniquement composé de l'un des 8 fonds précédemment décrits, supposons qu'il s'agisse de $b_{0,0} = (0,0,0,1)$. En fait, nous venons implicitement de choisir une origine dans le plan (cf figure 3.1). Si nous décalons cette origine vers la droite, le diagramme sera transformé en un diagramme uniforme composé de l'état $b_{1,0} = (0,0,1,0)$. Le décalage vers la droite de l'origine sera associé au générateur a de G. En décalant de nouveau l'origine vers la droite, nous obtenons respectivement $b_{2,0} = (0,1,0,0)$, $b_{3,0} = (1,0,0,0)$ puis à nouveau $b_{4,0} = b_{0,0} = (0,0,0,1)$, ce qui est logique puisque le fond au sens de Boccara est de période spatiale 4, et ce qui justifie la relation $a^4 = 1$. Le générateur b de G représente un décalage de l'origine vers le haut. Bien entendu, un décalage vers le haut puis vers la droite est équivalent à un décalage vers la droite puis vers le haut, nous en déduisons la commutativité de a et b : $ab = ba$. Un premier décalage vers le haut en partant de $b_{0,0} = (0,0,0,1)$ nous amène à un fond nouveau : $b_{0,1} = (1,0,1,1)$. Après un second décalage, nous retrouvons un fond connu $b_{0,2} = b_{2,0} = (0,1,0,0)$, d'où la relation $b^2 = a^2$. Ainsi, le groupe G est-il construit à partir des deux opérations (décalage vers la droite et vers le haut de l'origine du groupage) qui permettent de passer d'un fond de 54^4 à tout autre.

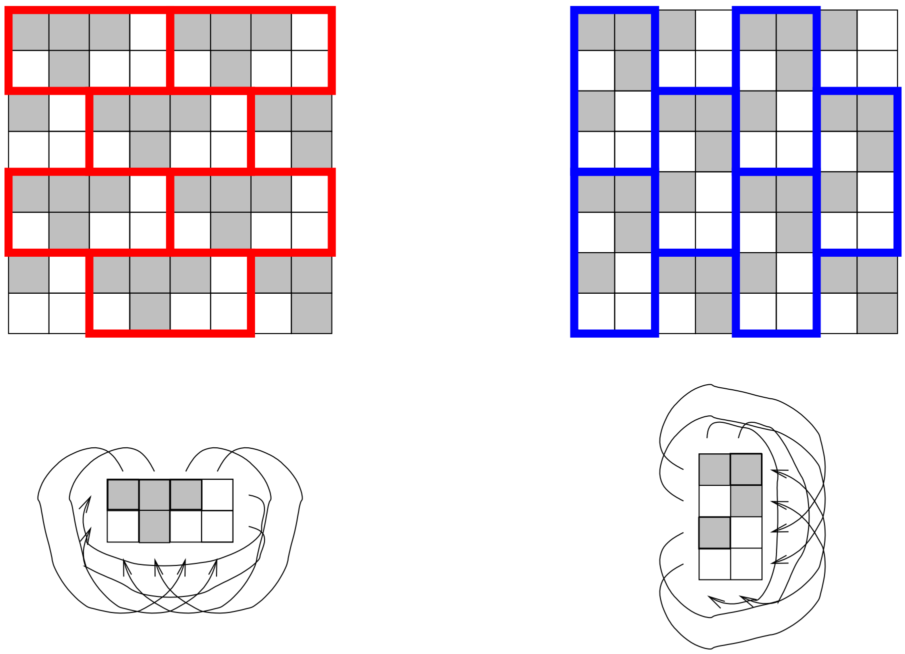


FIG. 3.2 – Deux pavages du fond de 54 avec des tuiles de taille minimale (8 cellules). Dans les deux cas, le groupe associé au pavage (le groupe quotient de $\mathbb{Z}^2$ qui identifie les tuiles) est le groupe G de présentation $< a, b \mid ab = ba, a^4 = 1, b^2 = a^2 >$.

Dans l'explication précédente, l'apparition du groupe G peut paraître un peu magique. Nous allons ici justifier l'existence des 8 fonds de 54 et leur lien avec le groupe G avec un point de vue géométrique. En fait il est possible d'associer à tous les fonds au sens de Boccara un groupe. L'existence de ce groupe est assurée par la signification implicite des fonds au sens de Boccara, leur périodicité spatiale et temporelle. Ainsi pour chacun d'entre eux, nous pouvons définir $\mathcal{T}$, le groupe des translation de $\mathbb{Z}^2$ laissant le diagramme espace-temps associé au fond invariant. Nous pouvons donc associer à chaque fond au sens de Boccara le groupe abélien $\mathbb{Z}^2$ quotienté par le groupe $\mathcal{T}$ et ce groupe est caractéristique du motif qui apparaît sur le diagramme espace-temps du fond.

Si nous considérons le fond au sens de Boccara qui apparaît sur les diagrammes espace-temps de 54 en partant d'une configuration initiale aléatoire, le groupe $\mathcal{T}$ est engendré par la translation horizontale de 4 cellules et la translation diagonale de 2 cellules vers le bas et de deux cellules vers la droite. De plus, en choisissant la notation $b_{k,l}$ avec $0 \leq k \leq 3$ et $0 \leq l \leq 1$

pour les 8 fonds de 54^4, nous décrivons implicitement un pavage de $\mathbb{Z}^2$ avec des tuiles de taille minimale qui respecte le motif. Plus précisément, nous choisissons pas par pas un représentant dans chaque classe d'équivalence de $\mathbb{Z}^2$ quotienté par le groupe $\mathcal{T}$, ce qui nous donne le pavage représenté à gauche sur la figure 3.2. Le choix de cette présentation du groupe G est arbitraire, nous aurions pu choisir n'importe quelle autre présentation, ce qui correspond géométriquement à un autre pavage avec des tuiles de tailles minimales. Un autre pavage est représenté à droite de la même figure. Nous pouvons vérifier que, dans les deux cas, le groupe associé au pavage est le groupe G.

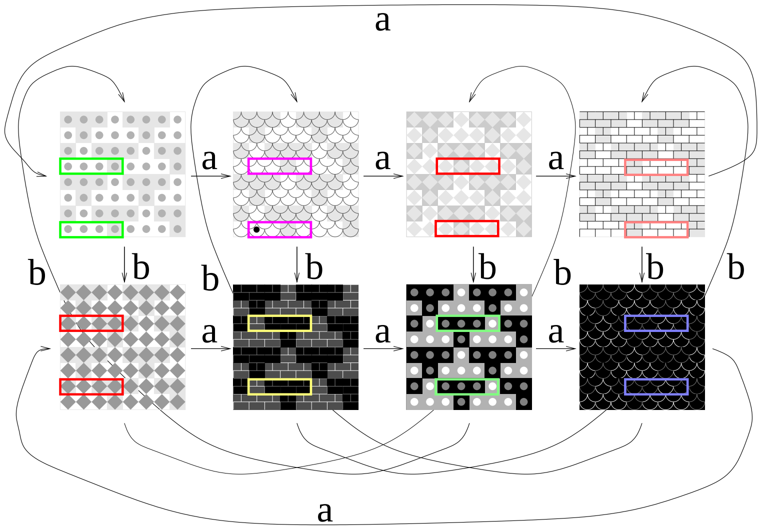


FIG. 3.3 – L'action naturelle du groupe G sur les 8 couleurs.

De plus, les 8 fonds sont indexés par les éléments du groupe G, donc ce groupe agit naturellement sur l'ensemble des 8 fonds. En fait, c'est cette action qui va fournir des informations sur le comportement de la règle puisque l'important est de savoir par quel élément du groupe il faut multiplier l'indice d'une région pour obtenir l'indice de la région voisine. En effet, le comportement de 54 en partant d'une configuration initiale composée de $b_{0,0}$ puis de $b_{1,1}$ est globalement la même qu'en partant d'une configuration composée de $b_{k,l}$ puis de $b_{k+1,l+1}$ puisque cela correspond simplement à une translation (de vecteur (k, l)) du choix de l'origine du groupage dans le diagramme.

Pour illustrer cela, nous allons filtrer (au sens de Hanson et Crutchfield [26]) l'espace-temps de la règle 54. Afin de reconnaître facilement les 8 fonds, nous allons colorier les diagrammes avec 8 couleurs, une par fond. L'idée est la suivante, la couleur d'une cellule ne va pas seulement dépendre de son propre état, mais également de ses deux voisins et de sa position modulo 4 dans l'espace et dans le temps. Soit $(s, t) \in \mathbb{Z} \times \mathbb{N}$ les coordonnées d'une cellule dans le diagramme espace-temps, nous colorions cette cellule avec une des 8 couleurs notées $c_{k,l}, 0 \leq k \leq 3, 0 \leq l \leq 1$ comme cela est indiqué dans le tableau suivant.

100 $c_{0-s-t[4],0-t[2]}$	010 $c_{1-s-t[4],0-t[2]}$	001 $c_{2-s-t[2],0-t[2]}$	000 $c_{3-s-t[4],0-t[2]}$
101 $c_{0-s-t[4],1-t[2]}$	110 $c_{1-s-t[4],1-t[2]}$	111 $c_{2-s-t[4],1-t[2]}$	011 $c_{3-s-t[4],1-t[2]}$

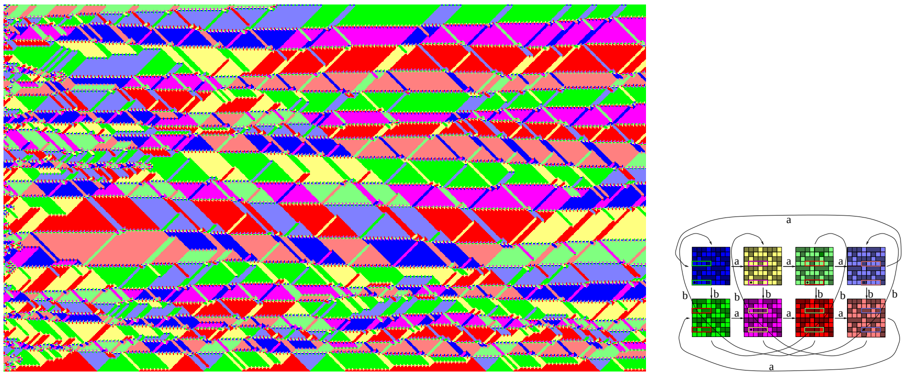


FIG. 3.4 – Une évolution filtrée en partant d'une configuration aléatoire, le temps va vers la droite.

Il est facile de vérifier qu'une région contenant le fond $b_{k,l}$ est remplie par la couleur $c_{k,l}$ dans la configuration filtrée, excepté sur le bord de la région. Ainsi, le groupe G qui agit sur l'ensemble des fonds agit également sur l'ensemble des couleurs, ce que nous avons représenté dans la figure 3.3. La figure 3.4 contient un diagramme filtré en partant d'une configuration aléatoire. Nous voyons que les particule (w, g_0, g_e) n'évoluent pas dans un fond homogène, mais sont en fait les frontières entre les différents fonds. Formellement, nous allons appeler *région* un ensemble 4-connexe de plus de 5 cellules d'un diagramme espace-temps de 54 qui ont la même couleur lorsque l'on filtre le diagramme. Le minimum de 5 cellules est bien entendu arbitraire, mais bien adapté à l'usage que nous ferons de la notion de région. Bien entendu, à chaque région, nous associons le fond qui la constitue (celui qui est associé à la couleur qui la compose dans le diagramme filtré). Nous pouvons à présent définir la trajectoire d'une particule simple : c'est la frontière entre deux régions. Notons que cette frontière peut contenir des cellules n'appartenant à aucune région, mais n'en contient pas nécessairement.

3.2 L'interprétation des particules à l'aide du groupe

Nous allons à présent expliquer comment l'action du groupe sur les fonds permet de comprendre les interactions entre les différentes particules.

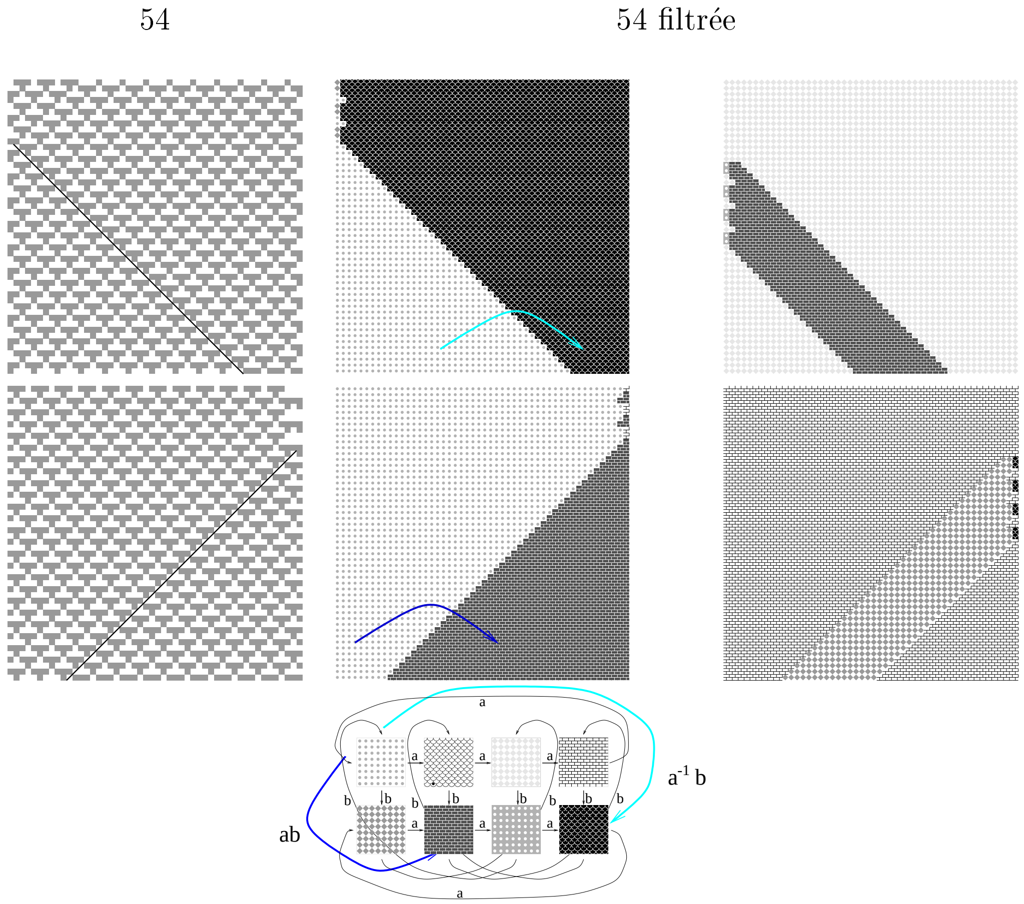


FIG. 3.5 – La particule $\overrightarrow{w}$ est associée à ab et $\overleftarrow{w}$ à $a^{-1}b$.

Sur la figure 3.5, nous voyons que la particule w est en fait la frontière entre deux fonds distincts. De plus, nous obtenons l'indice du fond situé à droite en multipliant celui du fond de gauche par ab pour une particule $\overrightarrow{w}$ et par $a^{-1}b$ pour une particule $\overleftarrow{w}$. Comme nous l'avons expliqué précédemment, cela est indépendant du fond situé à gauche puisque cela correspond à une translation du diagramme espace-temps. Nous pouvons donc associer l'élément ab aux particules $\overrightarrow{w}$ et $a^{-1}b$ aux particules $\overleftarrow{w}$. Nous pouvons aussi observer que lorsque nous avons deux particules w parallèles, les fonds de gauche et de droite sont identiques, ce qui est normal puisque l'ordre des éléments ab et $a^{-1}b$ est 2 dans G (i.e. $(ab)^2 = (a^{-1}b)^2 = 1$).

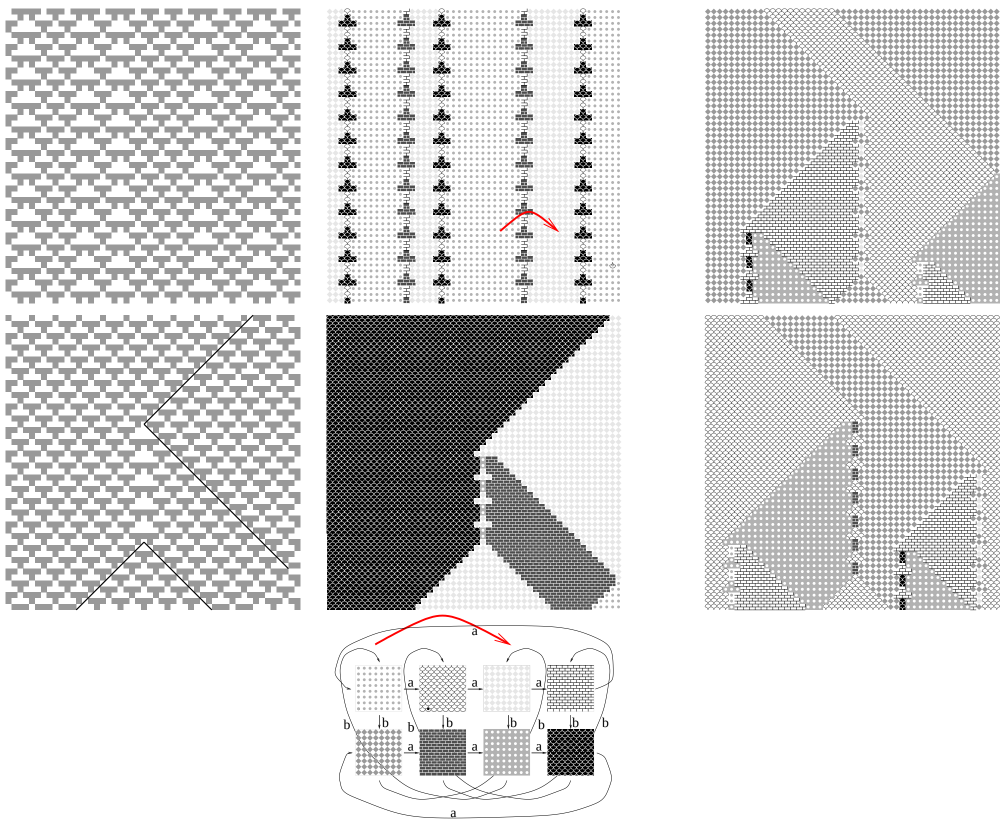


FIG. 3.6 – La particule g_0 est associée à a^2.

Sur la figure 3.6, nous voyons que nous devons associer l'élément a^2 de G à la particule g_0. De nouveau, $(a^2)^2 = 1$ donc lorsque nous avons deux particules g_0 parallèles, le fond de gauche est identique à celui de droite. En ce qui concerne la collision d'une particule g_0 avec une particule $\overleftarrow{w}$, au point d'impact, l'indice du fond de gauche doit être multiplié par a^2 puis ab et donc par $a^3b = a^{-1}b$ pour obtenir celui du fond de droite. Il n'est donc pas très surprenant que la collision crée une particule $\overrightarrow{w}$ sur le diagramme en bas à gauche. Les éléments associés aux particules w et g_0 génèrent un sous-groupe strict $\{1, ab, a^2, a^{-1}b\}$ de G. En fait, le comportement partant de configurations ne contenant que des particules w et g_0 est parfaitement compris. Ainsi, la principale raison de la complexité de 54 nous semble être, à ce point de notre étude, la dernière particule : g_e.

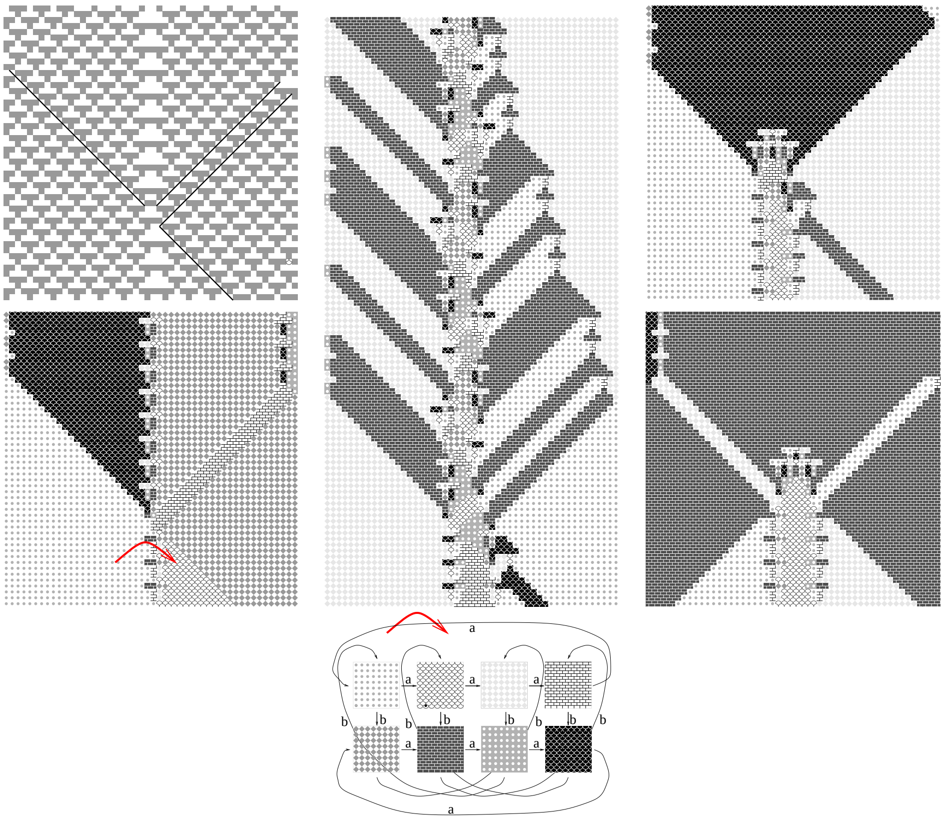


FIG. 3.7 – La particule g_e est associée à a.

Dans la figure 3.7, nous étudions la particule g_e à laquelle nous associons l'élément a de G. Nous avons représenté ses interactions avec une particule w et la « particule » émettant périodiquement des particules w. Il ne s'agit pas d'une particule pour notre définition. En effet, son évolution se décompose en une succession d'interactions entre des particules simples. Nous pouvons à présent expliquer pourquoi il est impossible d'annihiler une seule particule g_e. Supposons que sur la configuration initiale nous n'ayons qu'une particule g_e avec des particules w et g_0, alors nous devons multiplier l'indice du fond le plus à gauche par a, a^{-1}, a^2b ou b afin d'obtenir celui de droite. S'il était possible de supprimer la particule g_e, au bout d'un certain temps, il n'y aurait plus que des particules w et g_0, donc l'indice du fond de droite divisé par celui du fond de gauche appartiendrait à $\{1, a^2, ab, a^{-1}b\}$ ce qui n'est pas le cas. Cependant, il est possible d'annihiler deux (ou plus généralement un nombre pair) de particules g_e comme nous pouvons le voir sur le diagramme de droite (figure 3.7). En fait, le groupe G nous donne une sorte de loi de conservation au cours des interactions entre les particules, ce qui permet de démontrer des résultats négatifs sur le comportement de la règle.

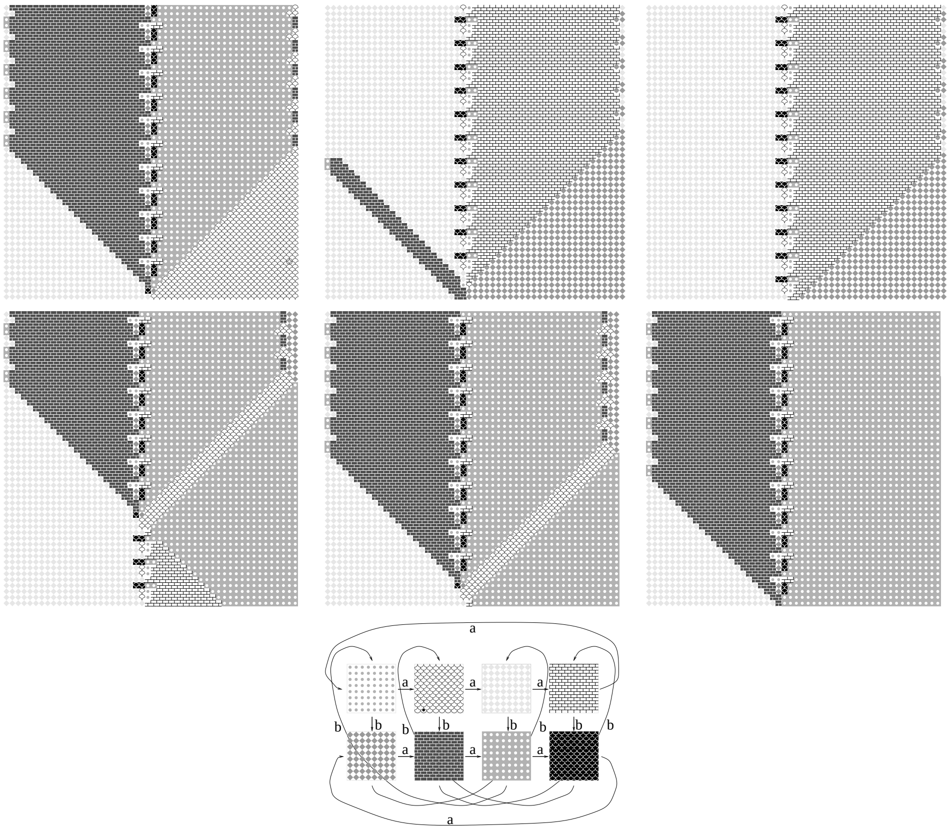


FIG. 3.8 – Les autres interactions entre les fonds, aucune d'entre elles n'est stable.

Nous avons associé quatre éléments du groupe aux quatre particules connues. Il est naturel de se demander ce qu'il advient des trois autres éléments du groupe qui sont distincts de l'unité d'une part et d'autre part s'il existe d'autres particules. La réponse à la première question est partiellement donnée par la figure 3.8. Nous avons mis côte à côte les fonds correspondant aux trois éléments du groupe restants. En ce qui concerne les éléments b et a^2b, il n'y a pas de configuration qui ne soit composée que de deux régions côte à côte, nous avons donc représenté les deux seules possibilités ayant moins de deux cellules entre ces deux régions. Il n'est pas très surprenant, comme $b = a(a^{-1}b) = a^{-1}(ab)$ et $a^2b = a(ab) = a^{-1}(a^{-1}b)$, que nous observions le même comportement que lors de la collision entre une particule g_e et une particule w (que nous rappelons sur le diagramme en bas à gauche). Nous observons donc que pour les trois éléments restants, aucun ne conduit à l'apparition d'une particule stable, en fait, plusieurs particules connues sont générées. Ceci dit, rien ne prouve qu'il n'existe pas une particule (éventuellement très grosse) entre deux de ces fonds. Nous ne pouvons qu'observer qu'une telle particule n'apparaît pas lorsque l'on choisit une configuration initiale aléatoire.

Une première application de l'introduction du groupe est la preuve de l'impossibilité d'annihiler une seule particule g_e, nous utiliserons cette méthode dans le chapitre suivant pour prouver que la règle 54 est située au-dessus d'une chaîne infinie dans la classification de Rapaport.

Chapitre 4

Les principales classifications

La problématique de la classification des automates cellulaires est véritablement posée dans un article de Wolfram en 1984, qui propose une classification empirique. Toute une série d'articles a suivi, la plupart ayant pour objectif de produire une formalisation de la classification de Wolfram. Dans ce chapitre, nous allons décrire et discuter les principales classifications connues. Cependant, nous ne respecterons pas l'ordre historique pour commencer par une classification récente, celle de Rapaport, qui est une classification algébrique, contrairement aux autres qui tentent de classer les automates en fonction de leur dynamique. Si nous commençons par cette classification, c'est qu'elle regroupe dans une même classe des automates qui ont le même comportement à un changement d'échelle près et que donc, en principe, les autres classifications devraient porter sur ces classes algébriques, ce qui n'est pas toujours le cas. Nous décrirons ensuite la classification de Wolfram, celle de Čulik et Yu qui se veut être une formalisation de la précédente et celle d'Ishii, largement inspirée de cette dernière mais qui introduit une mesure sur les configurations. Ces deux tentatives de formalisations souffrent d'un certain nombre de problèmes non résolus car elles s'appuient sur les observations de Wolfram, qui, comme nous le verrons, ne sont pas toujours générales. Nous nous intéresserons ensuite à la classification de Gilman, fondée sur la théorie ergodique, qui est, en fait, antérieure aux deux précédentes, et cependant beaucoup plus pertinente. Son principal défaut, partagé avec les deux premières, c'est que ses classes ne sont pas invariantes par décalage. Nous étudierons donc pour finir deux classifications dont les classes sont invariantes par décalage, bien qu'issues d'approches très différentes, celle de Cattaneo puis celle de Formenti qui introduisent la suite de cette thèse.

Nous éprouverons la pertinence de chaque classification à l'aide d'un jeu d'exemples comprenant ceux décrits dans le premier chapitre.

4.1 Rapaport

La classification de Rapaport [47] est algébrique. En fait, on définit un préordre sur l'ensemble des automates cellulaires unidimensionnels, donc un ensemble de classes d'équivalence qui est alors muni d'un ordre partiel. Dans ce préordre, un AC est plus grand s'il peut simuler l'autre à un changement d'échelle près. Nous avons déjà défini la notion de groupage (qui formalise le changement d'échelle) dans le premier chapitre, nous allons à présent définir l'ordre et étudier ses propriétés.

4.1.1 La relation d'ordre introduite par I. Rapaport

Inspirée de la notion de sous-système en système dynamique, la notion de sous-automate s'introduit naturellement sur les automates cellulaires. Des couleurs distinctes étant assignées à chaque état de $\mathcal{A}$, $\mathcal{A}$ est un sous-automate de $\mathcal{B}$ s'il est possible d'assigner des couleurs distinctes (comprenant les couleurs assignées aux états de $\mathcal{A}$) à chaque état de $\mathcal{B}$ de telle sorte que tous les diagrammes espace-temps de $\mathcal{A}$ soient des diagrammes espace-temps de $\mathcal{B}$. Formellement, cela se traduit par l'existence d'une injection des états de $\mathcal{A}$ dans ceux de $\mathcal{B}$ qui soit compatible avec la fonction de transition opérant sur les états de $\mathcal{A}$ et de $\mathcal{B}$:

Définition 4.1.1 Sous-automate
Étant donnés $\mathcal{A} = (Q_\mathcal{A}, \delta_\mathcal{A})$ et $\mathcal{B} = (Q_\mathcal{B}, \delta_\mathcal{B})$ deux automates cellulaires, $\mathcal{A}$ est un sous-automate de $\mathcal{B}$, ce que nous noterons $\mathcal{A} \subseteq \mathcal{B}$, si et seulement s'il existe une fonction $\iota : Q_\mathcal{A} \longrightarrow Q_\mathcal{B}$ injective telle que $\forall x_{-r}, ..., x_r \in Q_\mathcal{A} \quad \delta_\mathcal{B}(\iota(x_{-r}), ..., \iota(x_r)) = \iota(\delta_\mathcal{A}(x_{-r}, ..., x_r))$.

Notation 4.1.1 *Étant donnés $\mathcal{A} = (Q_\mathcal{A}, \delta_\mathcal{A})$ et $\mathcal{B} = (Q_\mathcal{B}, \delta_\mathcal{B})$ deux automates cellulaires, nous noterons $\mathcal{A} \simeq \mathcal{B}$ si et seulement si il existe une bijection f entre $Q_\mathcal{A}$ et $Q_\mathcal{B}$ telle que $\forall x, y, z \in Q_\mathcal{A} \quad \delta_\mathcal{B}(f(x_{-r}), ..., f(x_r)) = f(\delta_\mathcal{A}(x_{-r}, ..., x_r))$.*

Nous admettrons les propriétés suivantes vis-à-vis du groupage [47] :

Propriété 4.1.1 *Si $\mathcal{A} \subseteq \mathcal{B}$ alors pour tout n, $\mathcal{A}^n \subseteq \mathcal{B}^n$.*

Propriété 4.1.2 *Soit $\mathcal{A}$ un AC, pour tout $n, m \in \mathbb{N}^*$ il existe une bijection entre les états de $(\mathcal{A}^n)^m$ et ceux de $\mathcal{A}^{nm}$ compatible avec leurs fonctions de transition (i.e. $(\mathcal{A}^n)^m \simeq \mathcal{A}^{nm}$).*

Pour comprendre le diagramme espace-temps d'un automate cellulaire, l'œil humain cherche à adapter l'échelle pour faire apparaître les phénomènes principaux sans tenir compte des phases de transition. C'est en partant de cette constatation que [47] relie la notion de sous-automate à celle de groupage pour définir le préordre ci-dessous :

Définition 4.1.2 Préordre introduit dans [47]
Étant donnés $\mathcal{A}$ et $\mathcal{B}$ deux automates cellulaires, $\mathcal{A} \leq \mathcal{B}$ si et seulement s'il existe des entiers i et j tels que $\mathcal{A}^i \subseteq \mathcal{B}^j$, c'est-à-dire que $\mathcal{A}^i$ soit un sous-automate de $\mathcal{B}^j$.

Il s'agit bien d'un préordre car la réflexivité est triviale, quant à la transitivité, elle se déduit des propriétés admises : si $\mathcal{A} \leq \mathcal{B}$ et $\mathcal{B} \leq \mathcal{C}$ alors il existe n_1, m_1, n_2, m_2 tels que $\mathcal{A}^{n_1} \subseteq \mathcal{B}^{m_1}$ et $\mathcal{B}^{n_2} \subseteq \mathcal{C}^{m_2}$ donc $\mathcal{A}^{n_1 n_2} \subseteq \mathcal{C}^{m_1 m_2}$.

Bien entendu, en quotientant l'ensemble des AC unidimensionnels par la relation d'équivalence induite par le préordre $\sim$, nous pouvons définir l'ordre $(CA/\sim, \leq)$.

Remarque 4.1.1 *Notons qu'il est possible de définir d'autres préordres et donc d'autres classifications de cette façon. Par exemple, en remplaçant la notion de sous-automate par celle de facteur. C'est-à-dire en remplaçant l'existence d'une injection par celle d'une surjection des états de $\mathcal{B}^j$ vers les états de $\mathcal{A}^i$ qui soit compatible avec la fonction de transition locale [38].*

4.1.2 Non-autosimilarité et classification de Rapaport

Cet ordre admet un minimum, c'est l'ensemble des AC ayant un seul état. De plus, les automates ayant des propriétés dynamiques globales telles que la périodicité, les décalages ou la nilpotence sont équivalents et situés juste au-dessus du minimum.

Nous allons, jusqu'à la fin de ce chapitre, compléter les résultats démontrés par Rapaport sur sa classification en nous intéressant d'une part aux automates additifs et d'autre part aux automates non-autosimilaires.

Les automates additifs sur $\mathbb{Z}/p\mathbb{Z}$ avec p premier sont également situés juste au-dessus du minimum. En fait, les automates additifs sur $\mathbb{Z}/p^n\mathbb{Z}$ sont à hauteur n [43]. La preuve est plutôt technique et nous ne l'expliciterons pas ici. L'idée est de montrer que l'automate produit d'un additif sur $\mathbb{Z}/p^n\mathbb{Z}$ par lui-même (c'est-à-dire l'automate qui effectue en parallèle deux calculs de l'automate initial) est dans la même classe que l'automate additif. La clef est d'utiliser la représentation dipolynomiale avec la propriété arithmétique suivante.

Fait 4.1.1 *Si $k \geq n$, alors $(1+X)^{p^k} = (1+X^p)^{p^{k-1}}$ dans $\mathbb{Z}/p^n\mathbb{Z}$.*

Ensuite, on montre que tout sous-automate d'une instance groupée d'un additif sur $\mathbb{Z}/p^n\mathbb{Z}$ est dans la même classe que l'un des automates additifs sur $\mathbb{Z}/p^k\mathbb{Z}$ avec $0 \leq k \leq p$.

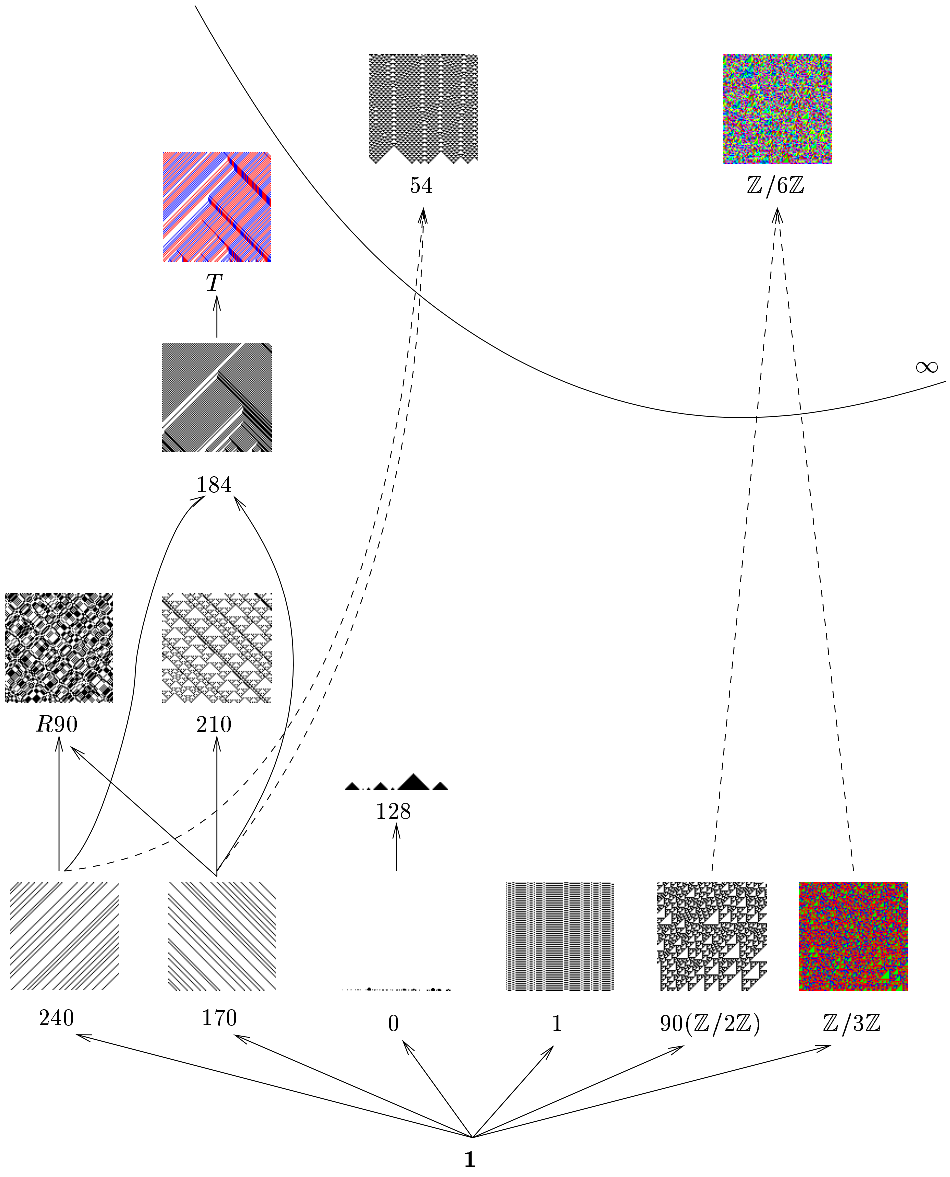


FIG. 4.1 – La classification de Rapaport.

Quant aux autres additifs (sur $\mathbb{Z}/n\mathbb{Z}$ lorsque n n'est pas puissance d'un nombre premier), ils sont situés à une distance infinie du minimum. Pour établir ce dernier résultat, il suffit de montrer que le comportement de ces AC additifs n'est pas autosimilaire.

Le fait d'être à une hauteur infinie dans l'ordre de Rapaport est très lié à la notion de non-autosimilarité comme le montre le résultat ci-dessous (qui nous permettra de démontrer que la règle 54 est au-dessus d'une chaîne infinie à la fin de cette section).

Définition 4.1.3 *Un AC $\mathcal{A}$ est non-autosimilaire si et seulement si pour tout $i \neq j$, $\mathcal{A}^i \nsubseteq \mathcal{A}^j$.*

Cette définition est naturelle puisque les AC qui ne la satisfont pas sont tels qu'il existe i et j tels que $\mathcal{A}^i \subseteq \mathcal{A}^j$, donc l'ensemble de ses diagrammes espace-temps est stable par une homothétie de rapport j/i[1].

Tous les AC que nous connaissons et qui sont au-dessus d'un nombre fini de classes sont autosimilaires. Certains ont des comportements très simples (les AC n'ayant qu'un seul état, les AC périodiques, les décalages, les nilpotents), quant aux additifs sur $\mathbb{Z}/p^n\mathbb{Z}$, leurs diagrammes espace-temps sont autosimilaires puisqu'il s'agit de superpositions de fractales (générées par les triangles de Pascal modulo p^n).

Si on se place dans le cadre de la classification de Rapaport, il existe des AC tels que toute chaîne les reliant au minimum soit finie ($\mathbb{Z}/p^n\mathbb{Z}$, avec p premier) et des AC tels qu'il existe une chaîne infinie entre le minimum et eux-mêmes. La proposition suivante montre qu'une propriété algorithmique implique d'être dans le second cas.

Proposition 4.1.1 *Si $\mathcal{A}$ est un AC non surjectif tel que pour tout $i \neq j$, $\mathcal{A}^i \nsubseteq \mathcal{A}^j$, il existe une suite infinie de classes $(\overline{\mathcal{D}}_n)_n$ distinctes telles que $\mathcal{D}_n < \mathcal{A}$ et $n|m$ implique $\mathcal{D}_n < \mathcal{D}_m$.*

Preuve 4.1.1 L'AC n'étant pas surjectif, il existe un mot u tel que toute configuration contenant ce mot soit un jardin d'Eden (c'est-à-dire une configuration sans antécédent) [16]. Soit k la longueur de ce mot, on considère l'automate $\mathcal{A}^k$ et son sous-automate strict $\mathcal{A}' = (Q_{\mathcal{A}'}, \delta_{\mathcal{A}'})$ avec $Q'_{\mathcal{A}} = Q_{\mathcal{A}^k} \backslash \{u\}$. En effet, u qui est un mot de longueur k, est un état de l'automate $\mathcal{A}^k$. Enfin, $\mathcal{A}'$ est bien défini parce que toutes les configurations de $\mathcal{A}^k$ contenant u sont des jardins d'Eden et ne peuvent donc pas apparaître au cours de l'évolution.

[1]La notion d'autosimilarité n'est pas bien définie pour les automates cellulaires [59]. Un bonne formalisation de cette notion qui permette d'interpréter notre remarque en terme de fractale (et éventuellement de pouvoir associer une dimension fractale aux AC autosimilaires) nécessiterait une longue étude. Le problème est de passer correctement du discret au continu.

Construisons les automates $\mathcal{D}_n = (Q_n, \delta_n)$ dont l'ensemble d'états est $Q_n = Q_{\mathcal{A}'} \times Q_{\mathcal{A}}^{k(n-1)} \not\subseteq Q_{\mathcal{A}}^{kn}$ et dont $\delta_n = G_{\delta_{\mathcal{A}^{kn}}}\big|_{Q_n}$ est la restriction à Q_n de la fonction de transition de $\mathcal{A}^{kn}$. Il est clair que pour tout n, $\mathcal{A}' \leq \mathcal{D}_n \leq \mathcal{A}$ et pour $n|m$, $\mathcal{D}_n < \mathcal{D}_m$. Il suffit en effet de considérer les injections canoniques pour montrer $\mathcal{A}'^n \subseteq \mathcal{D}_n \subseteq \mathcal{A}^{kn}$ et $\mathcal{D}_n^{m/n} \subseteq \mathcal{D}_m$.

Montrons que pour tout $n < m$, $\mathcal{D}_m \not\leq \mathcal{D}_n$. Sinon, il existe i et j tels que $\mathcal{D}_m^i \subseteq \mathcal{D}_n^j$. Ainsi, $\mathcal{A}'^{mi} \subseteq \mathcal{D}_m^i \subseteq \mathcal{D}_n^j \subseteq \mathcal{A}^{knj}$. Il est impossible que $mi \geq nj$ car un sous-automate a nécessairement moins d'états, or $\mathcal{D}_m^i \subseteq \mathcal{D}_n^j$ alors que $\mathcal{D}_m^i$ a $|Q_{\mathcal{A}'}|^i \times |Q_{\mathcal{A}}|^{(m-1)i}$ états, soit plus que $\mathcal{D}_n^j$ qui en a $|Q_{\mathcal{A}'}|^j \times |Q_{\mathcal{A}}|^{(n-1)j}$ (en effet $n < m$ donc $i < j$). Ainsi, $mi < nj$, or $\mathcal{A}'^{mi} \subseteq \mathcal{A}^{knj}$ implique $\mathcal{A}^{kmi} \subseteq \mathcal{A}^{knj}$. En effet, il suffit d'injecter les états contenant u de $\mathcal{A}^{kmi}$ dans les états contenant u de $\mathcal{A}^{knj}$ qui sont plus nombreux. Mais cela est impossible par hypothèse.

Les automates $\mathcal{D}_n$ sont donc tous dans des classes distinctes, il y a donc bien un nombre infini de classes distinctes (et une chaîne infinie) au-dessous de $\mathcal{A}$. $\square$

Réciproquement, on aimerait bien montrer que tout AC non surjectif autosimilaire n'est relié au minimum que par des chaînes finies. En fait, la surjectivité n'est pas l'élément clef de la preuve, on s'attend à ce que tout AC autosimilaire ne soit relié au minimum que par des chaînes finies. Ceci dit, nous n'avons aucune piste pour tenter de démontrer un résultat de ce type.

4.1.3 Paraboles et non-autosimilarité

Nous allons maintenant donner un exemple d'automate non-autosimilaire pour mieux comprendre les conséquences de la proposition précédente sur l'ordre de Rapaport. Pour cela, nous allons utiliser le fait qu'une parabole n'est pas autosimilaire. L'automate $\mathcal{P}$ dessine une parabole à l'aide d'un signal qui dans l'état l va vers la gauche et dans l'état r va vers la droite. Ce signal fait l'aller-retour entre un état fixe $|$ et un état p qui se décale à chaque collision.

Définition 4.1.4 *Soit $\mathcal{P} = (\{|, l, r, P, p, \_\}, \delta_{\mathcal{P}})$ l'automate cellulaire dont la fonction de transition est donnée par :*

$$
\begin{array}{rlrl}
\delta_{\mathcal{P}}(\_,\_,\_) = & \_ & \delta_{\mathcal{P}}(\_,\_,l) = & l \\
\delta_{\mathcal{P}}(\_,|,\_) = & | & \delta_{\mathcal{P}}(|,\_,l) = & l \\
\delta_{\mathcal{P}}(\_,|,l) = & | & \delta_{\mathcal{P}}(|,l,\_) = & r \\
\delta_{\mathcal{P}}(r,\_,\_) = & r & \delta_{\mathcal{P}}(\_,p,\_) = & p \\
\delta_{\mathcal{P}}(r,\_,p) = & r & \delta_{\mathcal{P}}(\_,P,\_) = & l \\
\delta_{\mathcal{P}}(r,p,\_) = & P & \delta_{\mathcal{P}}(P,\_,\_) = & p \\
\delta_{\mathcal{P}}(*,*,*) = & \_ & \textit{dans tous les autres cas} &
\end{array}
$$

Voici par exemple une configuration initiale traçant une parabole :

$$
\begin{array}{cccc|ccccccccc}
\cdots & \_ & \_ & \_ & \_ & \_ & \_ & l & \_ & p & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & \_ & \_ & \_ & \_ & l & p & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & \_ & \_ & \_ & \_ & P & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & \_ & \_ & \_ & r & p & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & \_ & \_ & r & \_ & p & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & \_ & r & \_ & \_ & p & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & r & \_ & \_ & \_ & p & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & l & \_ & \_ & \_ & p & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & \_ & l & \_ & \_ & p & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & \_ & \_ & l & \_ & p & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & \_ & \_ & \_ & l & p & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & \_ & \_ & \_ & P & \_ & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & \_ & \_ & r & p & \_ & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & \_ & r & \_ & p & \_ & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & r & \_ & \_ & p & \_ & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & l & \_ & \_ & p & \_ & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & \_ & l & \_ & p & \_ & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & \_ & \_ & l & p & \_ & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & \_ & \_ & P & \_ & \_ & \_ & \_ & \_ & \cdots \\
\cdots & \_ & \_ & \_ & \_ & r & p & \_ & \_ & \_ & \_ & \_ & \cdots
\end{array}
$$

Proposition 4.1.2 *$\mathcal{P}$ est non-autosimilaire.*

Preuve 4.1.2 Pour montrer la non-autosimilarité de $\mathcal{P}$, nous devons en premier lieu vérifier qu'aucune autre configuration initiale n'engendre une évolution contenant une parabole discrète (ce qui est fastidieux mais facile puisque toutes les autres interactions annihilent les particules concernées). Nous remarquons ensuite que cet AC ne peut donc dessiner qu'une parabole de paramètre fixé pour conclure par l'absurde que, sinon, l'ensemble des diagrammes espace-temps de la règle serait stable par une homothétie, or les paraboles ne sont pas stables par homothétie. □

On déduit donc le corollaire suivant de la proposition 4.1.1.

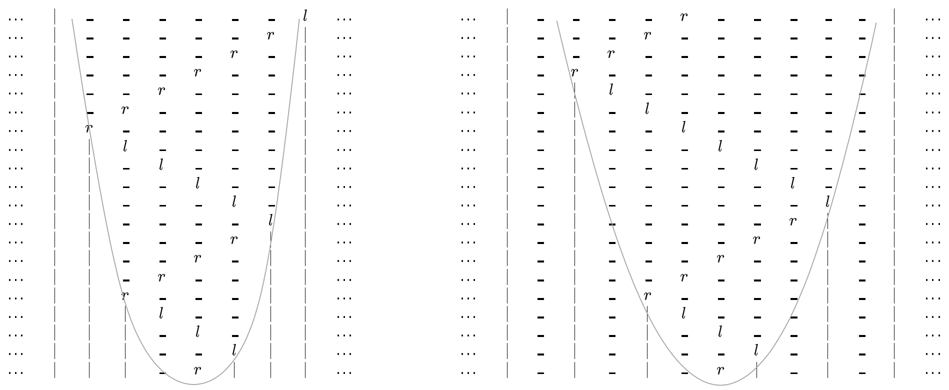
Corollaire 4.1.1 *L'automate $\mathcal{P}$ est à une hauteur infinie.*

Ce raisonnement échouerait si $\mathcal{P}$ pouvait dessiner des paraboles avec différents paramètres, bien qu'il puisse sembler qu'il soit algorithmiquement plus compliqué de dessiner plusieurs types de paraboles. Définissons un AC $\mathcal{P}'$ qui ait cette propriété :

Définition 4.1.5 *Soit $\mathcal{P}' = (\{|, g, d, \_\}, \delta_{\mathcal{P}'})$ un automate cellulaire traçant des paraboles de différents paramètres. La parabole est dessinée par un signal* *(g vers la gauche et d vers la droite) faisant l'aller-retour entre des signaux | qu'il efface à chaque passage. Sa fonction de transition est la suivante :*

$$\begin{array}{rlrl}
\delta_{\mathcal{P}'}(_,_,_) = & _ & \delta_{\mathcal{P}'}(\{_,|,l\},|,\{_,|,l\}) = & | \\
\delta_{\mathcal{P}'}(\{_,|,l\},_,l) = & l & \delta_{\mathcal{P}'}(r,_,\{_,|,r\}) = & r \\
\delta_{\mathcal{P}'}(_,|,l) = & r & \delta_{\mathcal{P}'}(r,|,_) = & l \\
\delta_{\mathcal{P}'}(|,|,l) = & r & \delta_{\mathcal{P}'}(r,|,|) = & l \\
\delta_{\mathcal{P}'}(*,*,*) = & _ & \textit{dans tous les autres cas} &
\end{array}$$

Voici deux configurations initiales dessinant des paraboles :



Nous pouvons observer que cet AC a le même comportement à toutes les échelles, c'est-à-dire que pour tout i, $\mathcal{P}' \subseteq \mathcal{P}'^i$. Pour le montrer, il suffit de considérer l'injection qui insère $i-1$ "$\_$" entre les éléments de $u \in Q^i$: $u \longmapsto (u_1, \_, ..., \_, u_2, \_, ..., \_, ..., u_i, \_, ..., \_)$. $\mathcal{P}'$ semble n'être séparé du minimum que par des chaînes finies dans la classification de Rapaport. On pourrait définir une notion de « complexité » avec la classification de Rapaport en séparant les automates situés au-dessus d'une chaîne infinie des autres. Nous voyons ici que cette complexité serait plus liée à la non-autosimilarité qu'à la puissance algorithmique. En effet, le fait de dessiner une seule parabole ne semble pas algorithmiquement plus compliqué que de dessiner plusieurs paraboles.

4.1.4 La règle 54 n'est pas autosimilaire

Nous allons ici montrer qu'aucune instance groupée de 54 n'est un sous-automate d'une autre instance groupée de la règle. Dans la section précédente, afin de construire l'automate $\mathcal{P}$ ayant ce comportement, nous avons fait en sorte qu'il puisse dessiner des paraboles. Les paraboles n'étant pas autosimilaires, les évolutions des automates groupés $\mathcal{P}^i$ sont donc toutes différentes. La possibilité de dessiner des paraboles avec la règle 54 est donc un indice sérieux de sa non-autosimilarité. Le diagramme 4.2 (le temps va vers la droite) propose une première méthode pour dessiner des paraboles avec la règle 54 : deux particules w rebondissent entre deux particules g_e qui sont écartées à chaque collision.

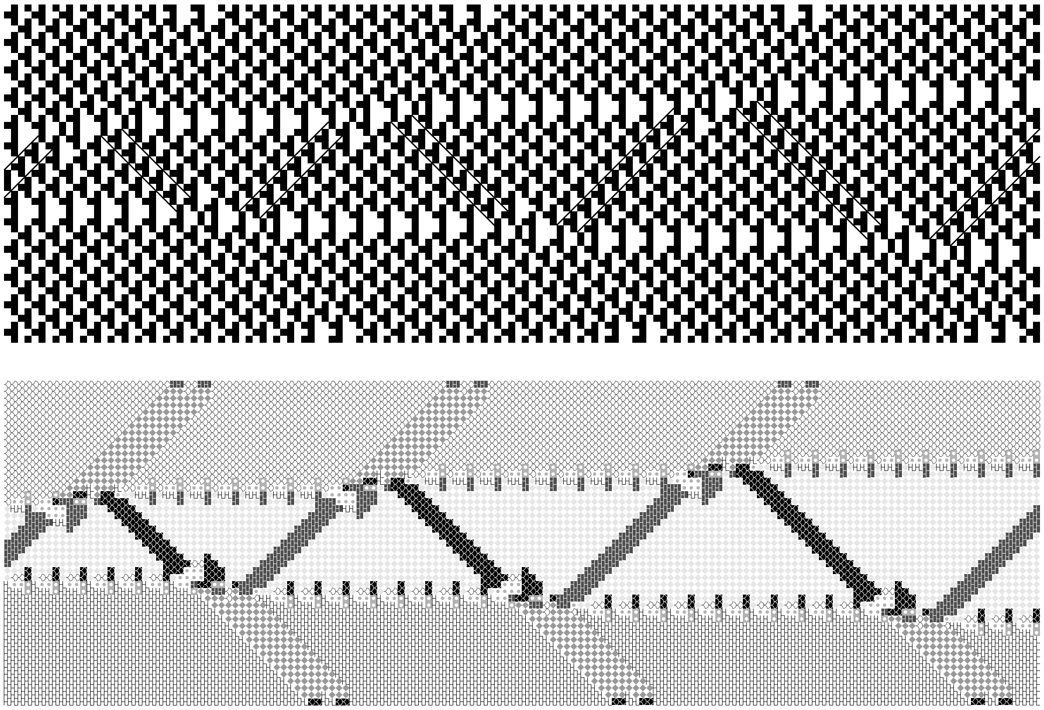


FIG. 4.2 – Premier algorithme dessinant des paraboles : deux particules w entre deux particules g_e. Le diagramme du haut représente la règle 54 tandis que celui du bas représente la règle 54 filtrée (cf chapitre 3) en partant de la même configuration (le temps va vers la droite).

Il y a d'autres algorithmes pour dessiner des paraboles avec la règle 54, par exemple, celui de $\mathcal{P}'$: une seule particule w rebondit sur des particules g_0 qui disparaissent après la collision (cf figure 4.3).

Si nous écartons les particules g_e, nous obtenons des paraboles de paramètres différents. Ainsi, l'argument utilisé pour l'automate $\mathcal{P}$ n'est pas réutilisable pour la règle 54. Donc la preuve de la non-autosimilarité de 54 est plutôt technique, l'idée de la preuve est basée sur le premier algorithme (cf figure 4.2) qui dessine une parabole de paramètre fixe. Nous allons montrer que les seules injections possibles de 54^i dans 54^j $(i < j)$ simulent les particules $\overrightarrow{w}$, $\overleftarrow{w}$, g_0 et g_e par elles-mêmes (i.e. l'image par l'injection d'un état de 54^i composé d'une particule entourée deux fonds (backgrounds) est un état

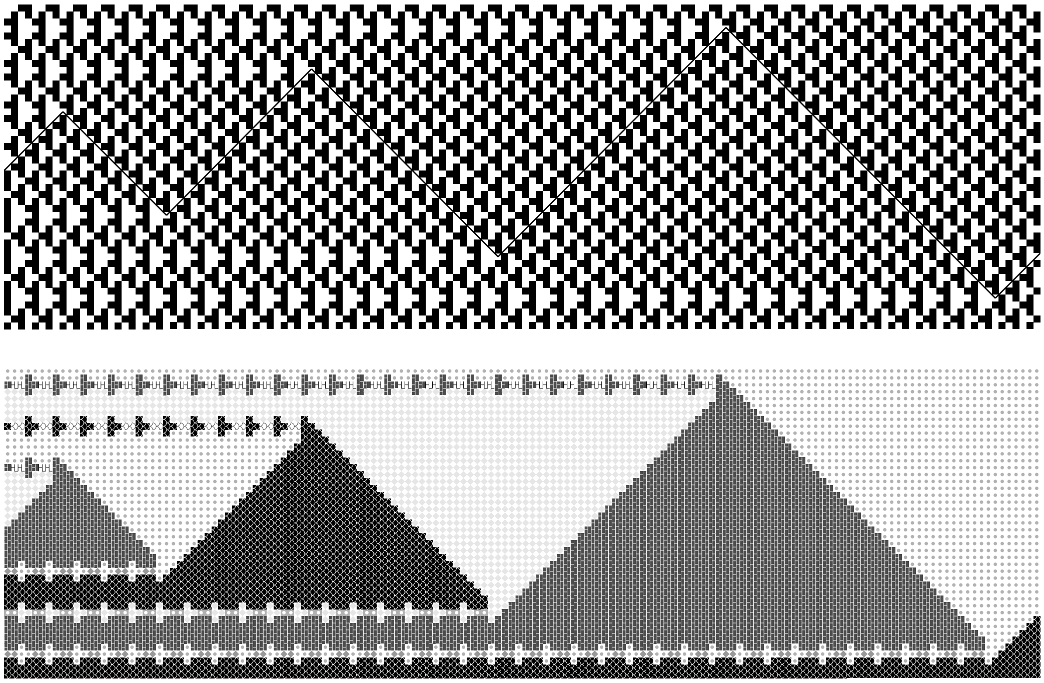


FIG. 4.3 – Second algorithme dessinant des paraboles : une particule w efface des particules g_0.

de 54^j composé de la même particule entourée par des fonds (éventuellement différents)).

Proposition 4.1.3 *La règle* 54 *est non-autosimilaire.*

Corollaire 4.1.2 *La règle* 54 *est située au-dessus d'une chaîne infinie dans l'ordre de Rapaport.*

Preuve 4.1.3 (de la proposition)
L'objectif est de montrer que l'ensemble des fonds de l'automate ne peut s'injecter que dans lui-même tout en respectant sa structure de groupe et donc que les particules w, g_0 et g_e sont simulées par elles-mêmes. Pour cela, nous allons travailler sur l'automate groupé 4 fois et montrer dans le lemme suivant que pour tout i et j $(i < j)$, $54^{4i} \not\subseteq 54^{4j}$, ce qui montre le résultat souhaité puisque $54^i \subseteq 54^j \Longrightarrow 54^{4i} \subseteq 54^{4j}$. $\square$

Lemme 4.1.1 *Pour tout $i \neq j$, $54^{4i} \not\subseteq 54^{4j}$.*

Preuve 4.1.4 Le résultat est évident si $i > j$ parce que 54^{4i} a plus d'états que 54^{4j}. Nous supposons donc $i < j$.

Montrons que $54^{4i} \nsubseteq 54^{4j}$ par l'absurde et notons ι l'injection de $\{0,1\}^{4i}$ dans $\{0,1\}^{4j}$. Montrons tout d'abord, en reprenant les notations définies dans la section 3.1, qu'il existe un élément $a^{k'}b^{l'}$ de G tel que pour tout $a^kb^l \in G$,

$$\iota(\underbrace{b_{k,l}, b_{k,l}, ..., b_{k,l}}_{i}) = (\underbrace{b_{k+k',l+l'}, b_{k+k',l+l'}, ..., b_{k+k',l+l'}}_{j})$$

La première étape est de montrer qu'un fond uniforme $(b_{k,l}, ..., b_{k,l})$ de 54^{4i} est nécessairement envoyé sur un fond uniforme de 54^{4j}. La caractéristique principale de ces fonds est qu'ils permettent la transmission de signaux de vitesse maximale dans les deux directions (i.e. ils peuvent être traversés par les particules $\overrightarrow{w}$ et $\overleftarrow{w}$). Une étude détaillée de la fonction de transition de 54 montre que seuls les triplets 00α et 01α (resp. $\alpha00$ et $\alpha10$) dépendent du troisième état (le fait que 00α dépend du troisième état signifie que $\delta_{54}(0,0,0) \neq \delta_{54}(0,0,1)$, c'est-à-dire que $\delta_{54}(0,0,.)$ est bijective). Il s'agit donc des seules possibilités de transmettre un signal de vitesse maximale vers la gauche (resp. vers la droite). Ainsi, le diagramme espace-temps de 54^{4j} ayant pour configuration initiale la configuration cyclique composée de $\iota(b_{k,l})$ doit contenir une diagonale vers la gauche composée de 00 ou de 01 et une diagonale vers la droite composée de 00 ou de 10 pour permettre la transmission d'une particule à la vitesse maximale. En fait, les diagonales sont nécessairement composées d'une seule séquence (00, 01 ou 10) : les séquences 01 et 10 imposent que la séquence suivante sur la diagonale soit respectivement 01 et 10 parce que $\delta_{54}(0,0,1) = \delta_{54}(1,0,1) = \delta_{54}(1,0,0) = 1$. De plus, une diagonale composée de 00 ne peut croiser qu'une diagonale composée également de 00 ce qui limite les possibilités à étudier à deux cas : deux diagonales composées de 00 ou une diagonale (gauche) composée de 01 et une diagonale (droite) composée de 10. La figure 4.4 représente les deux possibilités dans le cas $j = 3$.

Comme $\delta_{54^{4j}}(\iota(b_{k,l}, ..., b_{k,l}), \iota(b_{k,l}, ..., b_{k,l}), \iota(b_{k,l}, ..., b_{k,l})) = \iota(b_{k,l}, ..., b_{k,l})$, la présence de deux diagonales permet de déduire $\iota(b_{k,l}, ..., b_{k,l})$, ainsi dans le cas de deux diagonales composées de 00,

$$\iota(\underbrace{b_{k,l}, ..., b_{k,l}}_{i}) = (\underbrace{0, 0, ..., 0}_{4j})$$

et dans le cas contraire, nous voyons qu'il existe k' et l' (selon le choix des positions des diagonales) tels que

$$\iota(\underbrace{b_{k,l}, ..., b_{k,l}}_{i}) = (\underbrace{b_{k',l'}, b_{k',l'}, ..., b_{k',l'}}_{j}).$$

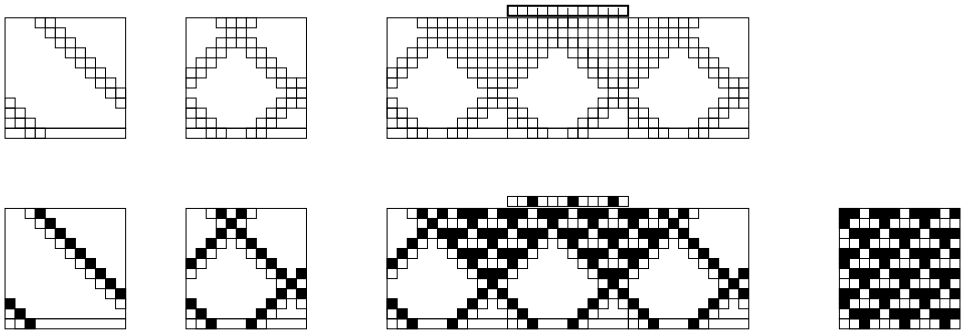


FIG. 4.4 – Les états stables de 54^{12} permettant la propagation de signaux de vitesse maximale dans les deux directions :
De la gauche vers la droite, selon la composition de la diagonale gauche (00 ou 01), la composition de la seconde diagonale est fixée (respectivement 00 et 10). Connaissant les deux diagonales, par stabilité, nous pouvons déduire l'état, respectivement $(0, ..., 0)$ et $(b_{k,l}, b_{k,l}, b_{k,l})$ pour des valeurs de k et de l qui dépendent des positions des diagonales, dans notre exemple $(b_{1,0}, b_{1,0}, b_{1,0})$. Sur le diagramme de droite, nous montrons le diagramme complet correspondant au fond $(b_{1,0}, b_{1,0}, b_{1,0})$.

Il est clair que $\iota(b_{k,l}, ..., b_{k,l}) \neq (0, 0, ..., 0)$ parce qu'alors, comme ι est injective, $\iota(b_{k+2,l}, ..., b_{k+2,l}) = (b_{k',l'}, ..., b_{k',l'})$ pour des valeurs bien choisies de k' de l' mais

$$\delta_{54^{4i}}((b_{k,l}, ..., b_{k,l}), (b_{k,l}, ..., b_{k,l}), (b_{k+2,l}, ..., b_{k+2,l})) = (b_{k,l}, ..., b_{k,l})$$

$$\delta_{54^{4j}}((0, ..., 0), (0, ..., 0), (b_{k',l'}, ..., b_{k',l'})) = (b_{k',l'}, ..., b_{k',l'}).$$

Enfin, en prenant en compte les spécificités des interactions entre les différents fonds, montrons que si $\iota(b_{0,0}, ..., b_{0,0}) = (b_{k',l'}, ..., b_{k',l'})$, alors $\iota(b_{k,l}, ..., b_{k,l}) = (b_{k+k',l+l'}, ..., b_{k+k',l+l'})$ (i.e. que ι préserve l'action du groupe sur les fonds). Il est facile de voir que si $\iota(b_{0,0}, ..., b_{0,0}) = (b_{k',l'}, ..., b_{k',l'})$, alors $\iota(b_{1,1}, ..., b_{1,1}) = (b_{k'+1,l'+1}, ..., b_{k'+1,l'+1})$ parce que $b_{k'+1,l'+1}$ est le seul fond dont l'interaction avec $b_{k',l'}$ se comporte comme un décalage vers la gauche (en raison de la création d'une particule $\overrightarrow{w}$, cf figure 3.5). En appliquant le même raisonnement aux autres interactions, comme nous pouvons à présent différencier les particules g_0 et g_e en les bombardant avec des particules $\overrightarrow{w}$, une étude exhaustive permet de montrer que l'action du groupe G sur les fonds de 54^i est préservée par ι.

Dans une seconde étape, en juxtaposant les fonds, nous pouvons créer toutes les particules usuelles de la règle 54 sur l'automate 54^i et nous savons que par ι elles sont simulées par les mêmes particules sur l'automate 54^{4j}. Pour conclure la preuve, nous remarquons que, dans la règle 54, les particules $\overleftarrow{w}$ issues de la collision entre une particule $\overleftarrow{w}$ et une particule g_e ont exactement une cellule de retard par rapport à la particule $\overleftarrow{w}$ incidente (cf figure 2.6 (b)). Donc, après i collisions, elle a exactement une cellule de retard pour l'automate 54^i (et sépare donc deux fonds) mais qu'elle n'est pas décalée d'une cellule complète dans 54^{4j} et donc il y a une cellule dont l'état n'est pas un fond (celui qui contient la particule $\overleftarrow{w}$) ce qui contredit l'hypothèse de l'existence de ι. □

4.2 Wolfram, Čulik et Yu, Ishii

4.2.1 Wolfram

Dans un article de 1984 [60], S. Wolfram propose de classer les automates cellulaires en 4 classes supposées distinctes. Il fonde cette classification sur l'observation de nombreux diagrammes espace-temps de configurations « aléatoires » ainsi que sur de nombreuses mesures d'entropie. Les « définitions » des 4 classes sont les suivantes :

- « Class 1 cellular automata evolve after a finite number of time steps from almost all initial states to a unique homogeneous state, in which all sites have the same value. [...] Their evolution completely destroys any information on the initial state. »
- « Class 2 cellular automata serve as "filters" which generate separated simple structures from particular (typically short) initial site value sequences. The density of appropriate sequences in a particular initial state therefore determines the statistical properties of the final state into which it evolves. (There is therefore no unique large-time (invariant) probability measure on the set of possible configurations.) Changes of site values in the initial state almost always affect final site values only within a finite range, typically of order r. » (r est le rayon de l'automate).
- « Evolution of infinite class 3 cellular automata from almost all possible initial states leads to aperiodic ("chaotic") patterns. After sufficiently many time steps, the statistical properties of these patterns are typically the same for almost all initial states. In particular, the density of nonzero sites typically tends to a fixed nonzero value (often close to $1/k$). » (k est le nombre d'états) « Small changes in initial states thus almost always lead to increasingly large changes in later states. »

– À propos des AC de la classe 4, Wolfram écrit : « Evolution leads to complex localized structures, sometimes long-lived. [...] An important feature, shared by other class 4 cellular automata, is the presence of propagating structures. By arranging for suitable reflections of these propagating structures, final states with any cycle lengths may be obtained. [Les structures apparaissant sur les diagrammes espace-temps de ces automates] are strongly reminiscent of the two-dimensional (essentially totalistic) cellular automaton known as the "Game of Life". The Game of Life has been shown to have the important property of computational universality. Cellular automata may be viewed as computers, in which data represented by initial configurations is processed by time evolution. Computational universality implies that suitable initial configurations can specify arbitrary algorithmic procedures. The system can thus serve as a general purpose computer, capable of evaluating any (computable) function. Given a suitable encoding, the system may therefore in principle simulate any other system, and in this sense may be considered capable of arbitrarily complicated behaviour. » Notons que dans la littérature, ce passage de l'article de Wolfram à été interprété comme une conjecture de l'universalité des automates de la classe 4.

Les « définitions » ci-dessus ne sont pas formelles. Ainsi, lorsque l'on prend des exemples d'automates cellulaires, il est facile de les mettre dans l'une des classes pour les uns et très difficile pour d'autres. Dans ce dernier cas, il est possible que plusieurs classes ou bien qu'aucune ne conviennent. Nous allons proposer ci-dessous une classification qui respecte dans la mesure du possible celle de Wolfram. Cependant, lorsqu'un problème surviendra, nous ferons des choix personnels tout en adoptant la même méthode, empirique, que Wolfram.

La classe 1 pose déjà problème. En effet, la règle 128 semble devoir y appartenir, d'une part, en raison de sa grande simplicité (une cellule évolue vers l'état blanc dès qu'elle-même ou que l'une de ses voisines est blanche), d'autre part parce que très rapidement elle évolue vers la configuration uniformément blanche. D'ailleurs, Wolfram donne des exemples de règles ayant le même comportement et qu'il positionne dans la classe 1. Malheureusement, d'après la définition donnée par Wolfram, cette règle n'appartient pas à la classe 1. En effet, depuis presque toute configuration pour une mesure non triviale de Bernoulli, l'évolution tend bien asymptotiquement vers la configuration uniformément blanche, mais pas en un temps fini. Pire, pour tout n, après n étapes de calcul, la configuration n'est presque jamais la configuration uniformément blanche. Le problème, c'est que pour tout n, la probabilité

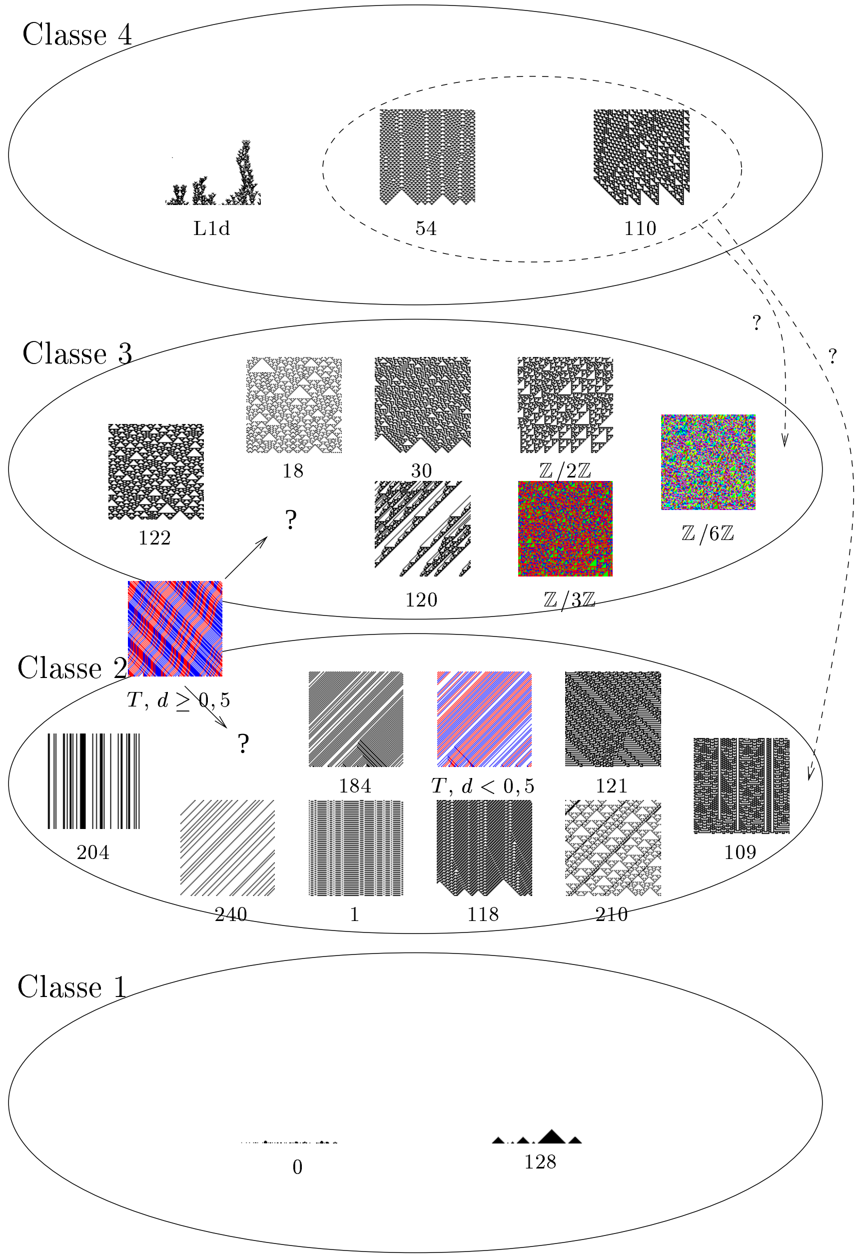


FIG. 4.5 – La classification de Wolfram.

de trouver $2n+1$ cellules noires côte à côte est 1.

J'ai choisi, comme la plupart des articles récents sur ce sujet, de placer les règles 54 et 110 dans la classe 4. Historiquement, Wolfram les a plutôt mises dans la classe 3. Le problème, c'est que les diagrammes espace-temps n'ont pas l'air « chaotiques » puisque l'on constate l'apparition de structures locales. Par ailleurs, la modification d'un petit nombre d'états modifie un nombre toujours croissant de cellules, ce qui semble les écarter de la classe 2. Nous avons également représenté L1d, un des exemples donnés par Wolfram d'AC de la classe 4, il s'agit d'une version unidimensionnelle du jeu de la vie, c'est-à-dire que les cellules naissent ou meurent en fonction du nombre de cellules vivantes dans leur voisinage (de rayon 2).

Dans la classe 3 se trouvent les règles additives (notées $\mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/3\mathbb{Z}$ et $\mathbb{Z}/6\mathbb{Z}$) ainsi que quelques ECA dont le diagramme semble aléatoire.

Pour décider de l'appartenance à la classe 2, nous avons en fait procédé par élimination et nous y avons placé toutes les règles qu'il n'était pas possible de placer ailleurs. Ainsi, le décalage n'est ni chaotique, ni aléatoire et ne présente pas de structures locales compliquées, il n'y a donc d'autre choix que de le mettre dans la classe 2, alors que les structures engendrées ne sont ni stables, ni périodiques. Par contre l'identité ou la règle 109 satisfont à la définition, quant à la règle 210, elle ne peut être dans une classe différente de celle du décalage. De plus, en reproduisant le raisonnement fait pour le décalage, le produit (ou la superposition) des deux décalages se trouve nécessairement dans cette classe.

Enfin, se pose le problème de la classifiabilité des règles. En effet, que dire d'un automate cellulaire, tel que T (le modèle de trafic routier décrit précédemment), qui a des comportements différents selon la densité de sa configuration initiale ? Bien sûr, lorsque la densité des voitures est inférieure à $1/2$, nous voulons placer T dans la classe 2, par contre pour une densité supérieure, il correspond plutôt à la classe 3. Nous verrons que dans les tentatives de formalisation proposées, s'opposent deux points de vue : certains proposent de classer les règles, tandis que les autres introduisent des mesures (ce qui est suggéré par les mentions « pour presque toute configuration » de Wolfram). Bien sûr, dans le second cas, bien que cela soit rarement explicite, nous n'avons plus une classification des automates, mais soit une classification pour une mesure donnée, soit une classification des couples (AC, mesure). Notons que le même problème se pose si l'on considère des configurations finies et non infinies [3].

4.2.2 Čulik et Yu

Dans leur article [13], Čulik et Yu essayent de proposer une formalisation de Wolfram. Leur classification a donc également quatre classes dont les définitions sont les suivantes :

- Un AC est dans la classe 1 s'il existe un état quiescent q tel que toutes les configurations q-finies évoluent vers une configuration homogène s^*.
- Un AC est dans la classe 2 s'il existe un état quiescent q tel que toutes les configurations q-finies ont une évolution ultimement périodique.
- Un AC est dans la classe 3 s'il existe un état quiescent q tel que pour toutes paires de configurations q-finies c_1 et c_2, le fait que c_1 évolue ou non vers c_2 est décidable.
- La classe 4 contient tous les automates cellulaires.

La première classe correspond plutôt bien à celle de Wolfram, les auteurs affirment qu'elle coïncide sur les automates à deux états et $r = 2$ et qu'elle serait en fait plus grande, mais sans expliquer pourquoi.

Ils font d'ailleurs le même commentaire à propos de leur classe 2, privée de leur classe 1, vis-à-vis de la classe 2 de Wolfram. Cependant, leur classe 2 est beaucoup plus restrictive que cette dernière. En effet, le décalage fait indiscutablement partie de la classe 2 de Wolfram, mais l'évolution des configurations q-finies ne sont pas ultimement périodiques. Outre le problème du décalage, les définitions à base de configurations finies sont sensibles à des effets de bord. Par exemple, la règle 109 est dans la classe 2 parce que 00 est un « mur » et que 0 est également un état quiescent de la règle. Au contraire, dans le cas de la règle $82 \circ \sigma$ (la composition de 82 avec le décalage), la séquence 11 est un mur tandis que 1 n'est pas un état quiescent (seul 0 l'est), du coup cet automate n'est pas dans la classe 2 en raison des effets de bords. Remarquons que σ composé avec la règle $210 = 82 + 128$ (ce qui signifie que les fonctions locales de transitions sont identiques sauf que $\delta_{210}(1,1,1) = 1$ alors que $\delta_{82}(1,1,1) = 0$) est dans la classe 2 ou dans la classe 3 selon l'état quiescent choisi.

Nous retrouvons donc dans la classe 3, privée de la classe 2, beaucoup d'automates dont le comportement n'a rien de chaotique, dont le décalage (240) et la superposition des décalages (170×240). Plus surprenant, on y trouve aussi des automates tels que 54 et 110 ainsi que des automates universels pour le calcul. En effet, en ce qui concerne 54 et 110 (quel que soit l'état quiescent choisi), la taille des mots est nécessairement croissante, ce qui donne un algorithme trivial pour décider si une configuration évolue ou non vers une autre : il suffit de calculer les itérés de la première tant que la taille

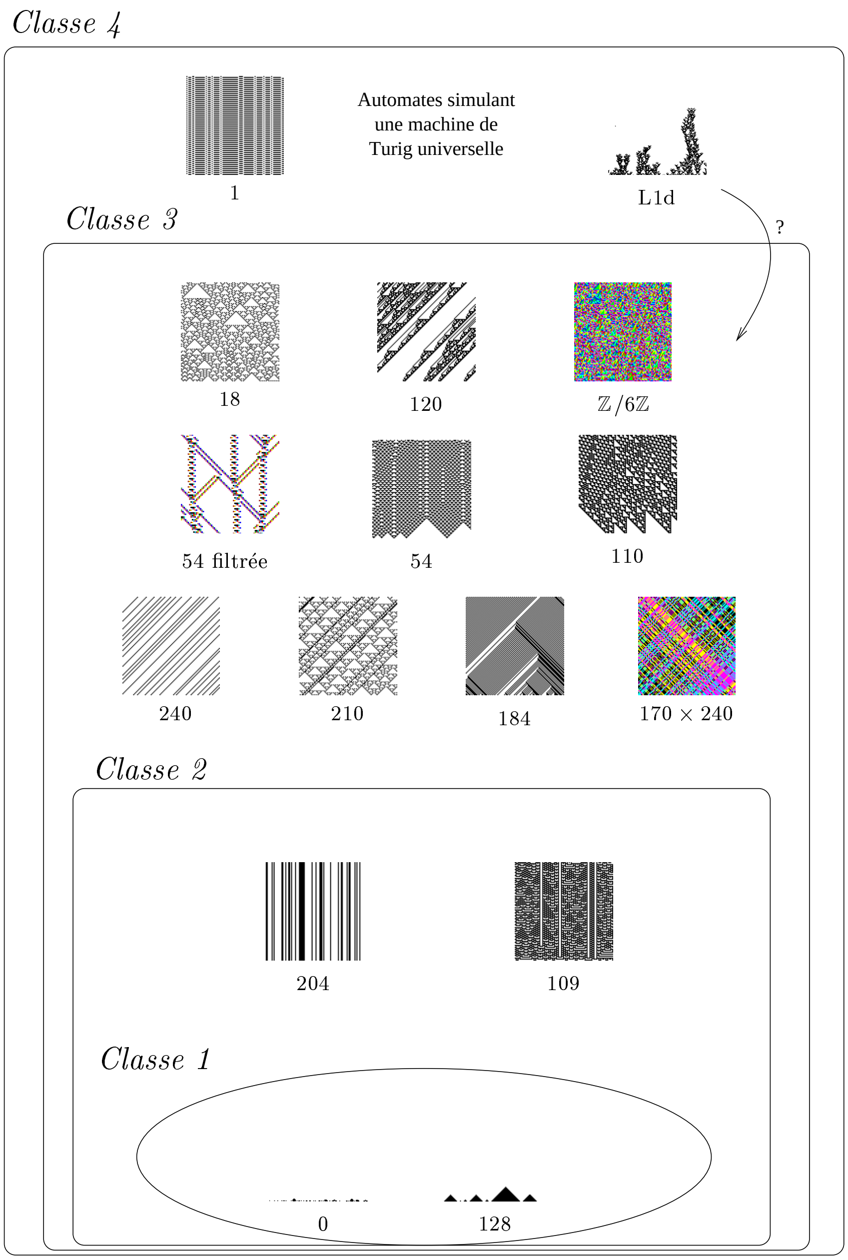


FIG. 4.6 – La classification de Čulik et Yu.

de la configuration obtenue est inférieure à celle de la configuration cherchée. De même, de nombreux AC universels vont se trouver dans la classe 3 de Čulik et Yu.

Quant à la classe 4 privée de la classe 3, outre des automates cellulaires simulant une machine de Turing universelle, on y trouve tous les automates n'ayant pas d'état quiescent... dont l'automate 1 dont l'évolution est des plus simples.

Cette tentative de classification n'est donc pas particulièrement concluante. Si le choix d'utiliser des configurations finies pour définir les classes est bien adapté à la classe 1, on voit bien sur les exemples 109, $82 \circ \sigma$ et $210 \circ \sigma$ (qui ont le même comportement global mais pas le même comportement sur les configurations finies) qu'il est particulièrement difficile de déduire le comportement global d'un automate en regardant uniquement ses configurations finies. De plus, ce choix exclut arbitrairement toutes les règles sans état quiescent du champ d'étude. Enfin, [3] montre qu'il n'est pas possible de proposer une classification basée sur des configurations finies en remarquant qu'il est possible de construire des AC ayant un comportement donné avec une probabilité p/q et un comportement différent sinon.

4.2.3 Ishii

La classification proposée par Ishii [28] est largement inspirée par celle de Čulik et Yu mais implique des configurations infinies et l'introduction d'une mesure ce qui permet (pas toujours à bon escient) de considérer le comportement des AC « pour presque toutes configurations », comme Wolfram.

- Un AC est dans la classe I si presque toutes ses configurations évoluent en un temps fini vers la même configuration uniforme.
- Un AC est dans la classe II si presque toutes ses configurations évoluent vers un attracteur cyclique modulo un décalage en un temps fini et s'il n'est pas dans la classe I.
- Un AC est dans la classe III si presque toutes ses configurations évoluent asymptotiquement vers un attracteur cyclique modulo un décalage et s'il n'est pas dans la classe II.
- La classe IV contient tous les autres automates cellulaires.

Ishii utilise en fait des mesures de Bernoulli. Il est facile de vérifier que les mesures de Bernoulli sont de support total, c'est-à-dire que la mesure des ouverts non vides (pour la topologie de Cantor considérée par Ishii) n'est pas nulle, ce qui implique que tout ensemble de mesure 1 est dense (il intersecte

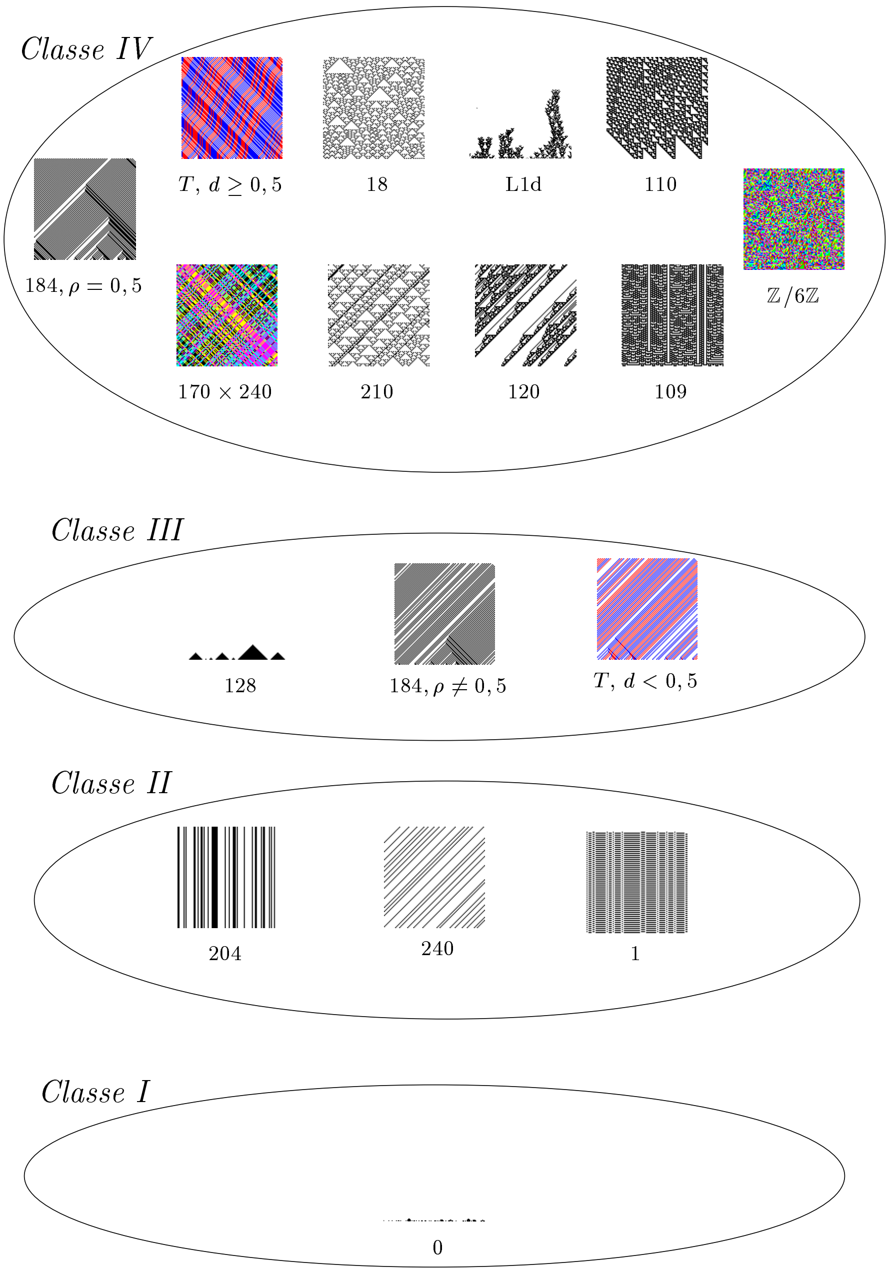


FIG. 4.7 – La classification de Ishii.

tous les ouverts). D'ailleurs, même si l'on souhaite étendre ces définitions à d'autres mesures, c'est une propriété que l'on souhaiterait conserver. Le seul problème, c'est que, dans ce cas, si presque toutes les configurations d'un automate évoluent en un temps fini vers une configuration uniforme, alors toutes les configurations évoluent en un temps fini vers une configuration uniforme. La preuve est topologique et assez simple. Les fonctions de transitions des AC sont continues et l'espace $Q^{\mathbb{Z}}$, et plus généralement $Q^{\mathbb{Z}^n}$, est compact. Donc si toutes les configurations d'un ensemble E de mesure 1 tendent en un temps fini t vers un singleton $\{c_0\}$ qui est fermé, $G_\delta^t(E) = \{c_0\}$ donc $E = G_\delta^{-t}(\{c_0\})$, E est fermé car image réciproque d'un fermé par une application continue, or un espace dense fermé est nécessairement égal à l'ensemble des configurations. Le même raisonnement s'applique à la définition de la classe II. Il se trouve que si l'automate tend presque partout en un temps fini t vers un attracteur cyclique modulo les décalages, alors il tend presque partout vers un attracteur cyclique de période p fixée modulo σ^k avec k fixé. En effet, pour p et k fixés, l'ensemble des configurations cycliques de période p modulo σ^k est shift-invariant, donc son image réciproque par G_δ^t est également invariante par décalage, la mesure de cette image réciproque est donc soit 0, soit 1 (les mesures de Bernoulli sont σ-ergodiques). Il existe donc un unique couple (p, k) pour lequel cette mesure soit non nulle. Or l'ensemble des configurations cycliques de période p module σ^k est fermé puisque c'est l'image réciproque de 0^* par l'application continue $c \longmapsto \sigma^k \circ G_\delta^p(c) - c$ (dès lors qu'une définition raisonnable de l'opérateur binaire $-$ sur $Q^{\mathbb{Z}}$ est donnée). Comme précédemment, son image réciproque par G_δ^t est un fermé de mesure 1, c'est donc l'ensemble tout entier.

Ainsi, la classe I de Ishii est plus restrictive que celle de Čulik et Yu puisque la règle 128 s'en trouve exclue. Sa classe II est constituée des automates dont l'évolution est périodique au bout d'un temps fini à un décalage près ; comme nous l'avons vu précédemment, elle contient le shift, mais ni la superposition des deux shifts, ni la règle 184, ni la règle 210. Sa classe III récupère donc la règle 128 ainsi que les règles dont le comportement est ultimement un décalage comme 184 pour μ_ρ avec $\rho \neq 1/2$ ou T si la densité est inférieure strictement à $1/2$. Par contre, les règles localement périodiques mais qui ne sont pas globalement périodiques (modulo un décalage) telles que les règles 109 et 210 ne se trouvent pas non plus dans la classe III. Enfin, le produit d'un décalage vers la gauche et d'un décalage vers la droite n'est pas dans cette classe, en effet, même composé par un décalage, le comportement n'est pas ultimement périodique.

Nous voyons donc que le problème du décalage est plus profond qu'il n'y paraît, nous verrons qu'il est vraiment lié à la topologie de Cantor (cf section 4.4.1).

4.3 Gilman

Gilman [20], qui munit l'ensemble des configurations de la topologie de Cantor, propose, pour une mesure donnée, de partitionner l'ensemble des AC en trois classes à l'aide des définitions suivantes :

Définition 4.3.1 *Un AC est* équicontinu *en x si les trajectoires des points proches de x restent proches de la trajectoire de x, c'est-à-dire*

$$\forall \eta > 0, \exists \epsilon > 0, \forall y \in Q^{\mathbb{Z}} \quad d(x,y) \leq \epsilon \Longrightarrow \forall n \in \mathbb{N} \quad d(G_\delta^n(x), G_\delta^n(y)) \leq \eta.$$

Définition 4.3.2 *Un AC est* expansif *si les trajectoires de deux points divergent nécessairement, ce qui s'écrit*

$$\exists M > 0, \forall\, x \neq y \in Q^{\mathbb{Z}}, \exists n \in \mathbb{N} \quad d(G_\delta^n(x), G_\delta^n(y)) \geq M.$$

- Un AC est dans la classe $\mathcal{A}$ s'il est équicontinu en certains points.
- Un AC est dans la classe $\mathcal{B}$ s'il est presque partout équicontinu mais n'est pas dans $\mathcal{A}$.
- Un AC est dans la classe $\mathcal{C}$ s'il est presque partout expansif.

Gilman démontre que ces trois classes constituent une partition de l'ensemble des AC pour une mesure de Bernoulli donnée. Se trouvent dans la classe $\mathcal{A}$ les automates cellulaires dont l'évolution fait apparaître des « murs », c'est-à-dire des séquences d'états telles que les cellules situées à leur droite n'ont aucune influence sur les cellules situées à leur gauche et réciproquement (les séquences interdisant toute propagation vers la gauche et celles les interdisant vers la droite peuvent être différentes, mais les deux doivent exister). En formalisant cette notion de mur (ou mot bloquant), on obtient une caractérisation de l'ensemble des AC unidimensionnels de la classe $\mathcal{A}$ de Gilman (cette propriété se trouve de façon implicite dans [29], un énoncé explicite se trouve dans [7]). Par exemple, l'état 0 est un mur de la règle 128 qui se trouve donc dans la classe $\mathcal{A}$, il en va de même pour la séquence 00 pour la règle 109.

Dans la classe $\mathcal{B}$, nous trouvons les automates qui ont un comportement ultimement périodique mais sans « murs ». Par exemple, la composition de la règle 184 avec un décalage σ lorsque la densité de 1 est strictement inférieure à 1/2. Ce même automate, lorsque la densité de 1 est supérieure à 1/2 se comporte ultimement comme un décalage et est donc presque partout expansif et dans la classe $\mathcal{C}$. Cela illustre bien que cette classification porte sur les couples (AC, mesure).

Nous retrouvons tous les autres AC dans la classe $\mathcal{C}$, en particulier le décalage (shift). En fait, le fait que le décalage soit présumé « chaotique »

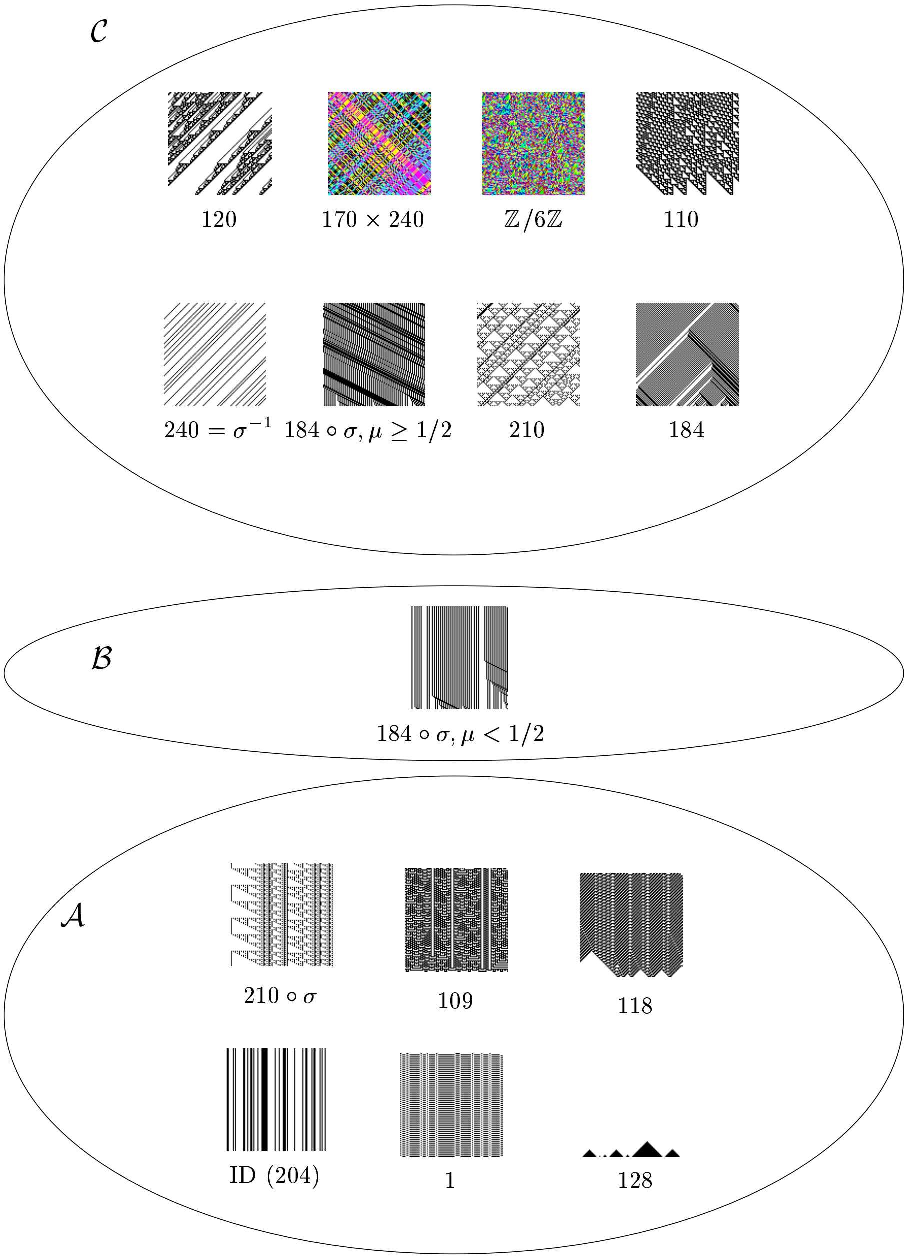


FIG. 4.8 – La classification de Gilman.

puisque Gilman suggère que sa classe $\mathcal{C}$ correspond à la classe 3 de Wolfram est presque inhérent au choix de la topologie de Cantor puisque c'est l'exemple type du système dynamique chaotique. En effet, la distance qui est associée à Cantor est

$$d(c, c') = \sum_{i \in \{i \in \mathbb{Z} \mid c_i \neq c'_i\}} \frac{1}{2^{|i|}}$$

ce qui fait qu'un automate va être sensible aux conditions initiales dès lors qu'un changement très loin de l'origine dans la configuration initiale peut changer la cellule 0 au cours de l'évolution. Or c'est bien entendu le cas du décalage. Nous allons voir ci-dessous qu'il n'est pas simple de proposer une définition de chaos telle que le décalage ne soit pas chaotique.

En fait, la classification de Gilman, bien qu'étant la première et malgré le problème du décalage, est l'une des plus intéressantes qui ait été proposées pour formaliser Wolfram.

On constate que, contrairement aux deux précédentes classifications, celle de Gilman repose sur les travaux effectués en systèmes dynamiques et en théorie ergodique. Sans doute est-ce cela qui explique que cette classification est beaucoup plus robuste et pertinente que les deux précédentes. Remarquons que toutes les propositions de définitions que nous donnons dans la suite, que ce soit celles de Cattaneo, celles de Formenti ou les nôtres, sont inspirées par les définitions proposées en systèmes dynamiques.

4.4 Notions de chaos invariantes par décalage

Notons que si une notion est invariante par décalage, alors, dès lors que l'identité n'est pas chaotique (ce qui est raisonnable), le décalage n'est pas chaotique non plus.

4.4.1 Cattaneo *et al.*

L'objectif de la classification de Cattaneo [10] que nous allons présenter ici est de proposer une définition de chaos pour laquelle le décalage ne soit pas chaotique. Cette définition est basée sur la topologie de Cantor, en remarquant que le shift n'est pas expansif. En fait, très peu d'AC sont expansifs puisque les additifs au sens auquel nous les avons définis ne le sont pas non plus, alors que ceux qui sont permutifs à gauche et à droite, c'est-à-dire qui font la somme d'au moins une cellule située plus à gauche et d'au moins une cellule située plus à droite de la cellule dont on calcule le nouvel état, le sont. Précisons les définitions nécessaires. Nous les donnons dans le cadre des AC puisque c'est l'objet de notre étude, elles sont bien entendu toutes définies dans le cadre plus général des systèmes dynamiques en remplaçant G_δ par une fonction continue F d'un espace compact $\mathcal{E}$ dans lui-même.

- Rappelons qu'un AC est expansif s'il existe $M > 0$ tel que pour tout couple x, y de configurations, $\exists n \in \mathbb{N}$ tel que $d(G_\delta^n(x), G_\delta^n(y)) \geq M$.
- un AC est chaotique (au sens de Cattaneo) s'il est fortement transitif (pour tout sous-ensemble non vide U, $\cup G_\delta^n(U) = Q^{\mathbb{Z}}$) et si l'ensemble des points périodiques est dense.
- un AC est chaotique au sens de Devaney [14] s'il est transitif (pour tout couple d'ensembles non vides U, V, $\exists n, G_\delta^n(U) \cap V \neq \emptyset$) et si l'ensemble des points périodiques est dense (cette définition de chaos implique la sensibilité aux conditions initiales et la surjectivité dans le cas des AC).

On constate que (même dans le cas général des systèmes dynamiques) toutes ces définitions impliquent la surjectivité, ce qui n'est pas gênant dans le contexte système dynamique. Pour étudier une fonction F non surjective sur $\mathcal{E}$, il suffit d'étudier la restriction de la fonction sur $\cap_n F^n(\mathcal{E})$ (cet espace est compact car intersection de compacts, en effet, l'image directe d'un compact par une application continue est compact) qui est surjective. La situation est tout autre pour les AC puisque la restriction indiquée ci-dessus n'est en général plus un automate cellulaire, ces définitions sont donc très restrictives. En particulier, les règles 54 et 110 n'étant pas surjectives, sont exclues du champ d'étude. Les relations entre ces définitions représentées sur la figure 4.9 sont spécifiques aux automates cellulaires. Cattaneo remarque de plus que les règles chaotiques (pour sa définition) ne sont pas injectives (donc le décalage n'est pas chaotique). En fait, si cette notion de chaos a l'avantage d'être shift-invariante (ce que l'expansivité n'est pas), et de ne pas considérer le shift comme chaotique, les seuls exemples concrets qui la satisfont sont les automates additifs (à un sens assez large, mais seulement les surjectifs).

4.4.2 Formenti

Le problème inhérent aux travaux de Gilman et de Cattaneo provient de la topologie choisie. En effet, avec cette topologie, il s'agit de regarder si l'on peut prévoir l'état d'une cellule en ne connaissant que son état et l'état des cellules voisines dans la configuration initiale. Du coup, le décalage est nécessairement chaotique puisque l'état d'une cellule dépend de l'état initial d'une cellule de plus en plus éloignée. Pour résoudre le problème, E. Formenti [12, 6] propose d'utiliser une autre topologie, celle de Besicovitch. Afin que la métrique associée donne la même importance à toutes les cellules, l'idée est que la distance entre deux configurations soit la « densité » des différences entre les deux configurations. Cependant, il est impossible

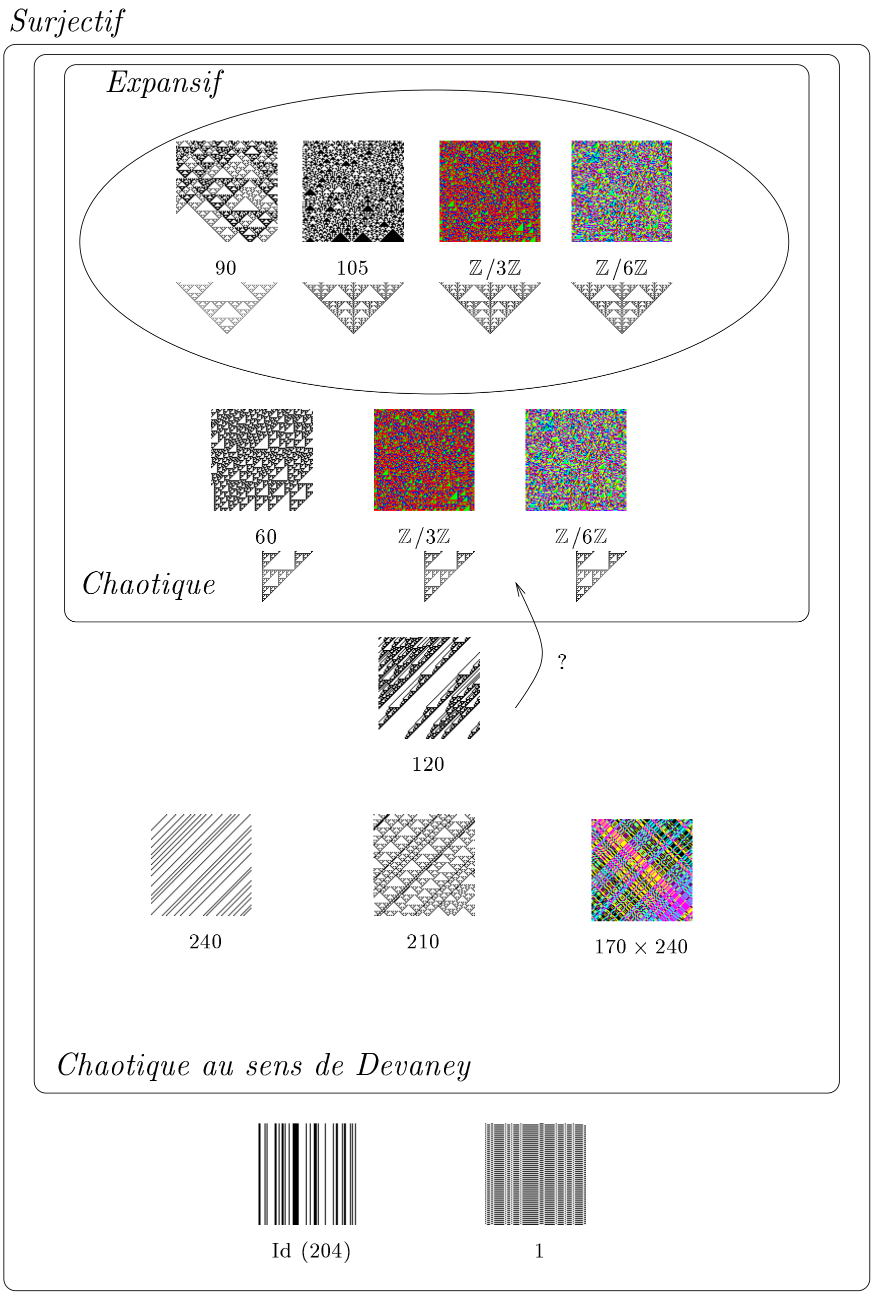


FIG. 4.9 – Le chaos d'après Cattaneo.

de définir la « densité » d'un sous-ensemble de $Q^{\mathbb{Z}}$. Dans le cas de Besicovitch, il s'agit d'une densité supérieure qui introduit une pseudo-distance sur l'ensemble des configurations. La topologie induite est shift-invariante mais l'ensemble des configurations n'est pas compact. Il est donc impossible de réutiliser directement les résultats de la théorie ergodique (qui étudie les comportements chaotiques des systèmes dynamiques), en effet cette dernière suppose en général que l'espace est compact.

Définition 4.4.1 *La* pseudo-distance de Besicovitch *sur $Q^{\mathbb{Z}}$ est définie par*

$$d(x,y) = \limsup_{l \longrightarrow +\infty} \frac{\#\{j \in [-l,l] | x_j \neq y_j\}}{2l+1} \qquad \text{où } \# \text{ représente le cardinal.}$$

Ce n'est pas une distance parce que la pseudo-distance entre deux configurations égales partout, excepté pour un nombre fini de cellules, est nulle. Il en va de même pour la pseudo-distance entre deux configurations identiques excepté pour les cellules aux positions 2^n, bien qu'il y ait un nombre infini de différences.

Propriété 4.4.1 *La topologie de Besicovitch, c'est-à-dire la topologie sur $Q^{\mathbb{Z}}$ quotienté par la relation $\sim$, $x \sim y \iff d(x,y) = 0$, induite par d, est métrique, connexe par arc, de dimension infinie, complète, ni séparable, ni localement compact. De plus, $x \sim y \Longrightarrow G_\delta(x) \sim G_\delta(y)$ et les fonctions de transition des AC sont des fonctions continues de $Q^{\mathbb{Z}}/\sim$ dans lui même [6].*

La dernière propriété de continuité est très importante et justifie l'utilisation de notions de théorie ergodique dans ce contexte.

La notion d'expansivité utilisée par Gilman et Cattaneo apparaît beaucoup trop restrictive dans la contexte de Besicovitch. C'est donc à la sensibilité aux conditions initiales que Formenti s'intéresse.

Définition 4.4.2 *Un AC $\mathcal{A}$ est sensible aux conditions initiales pour une pseudo-distance d s'il n'est équicontinu nulle part, ce qui s'écrit*

$$\forall c \in Q^{\mathbb{Z}}, \exists M > 0, \forall \epsilon > 0, \exists c' \in Q^{\mathbb{Z}}, \exists n$$
$$d(c,c') < \epsilon \text{ et } d(G^n_{\delta_{\mathcal{A}}}(c), G^n_{\delta_{\mathcal{A}}}(c')) \geq M.$$

Il est très facile de vérifier que l'identité et les décalages sont équicontinus. Cette fois, le produit des décalages est également continu. Nous voyons que le changement de topologie a résolu le problème du décalage plus profondément parce que, cette fois, la distance associée à la topologie est shift-invariante ($d(\sigma(x),\sigma(y)) = d(x,y)$), ce qui n'était pas le cas avec Cantor. Les seuls automates pour lesquels il est prouvé qu'ils sont sensibles aux conditions initiales sont les additifs. Nous retrouvons parmi les règles « ni équicontinues,

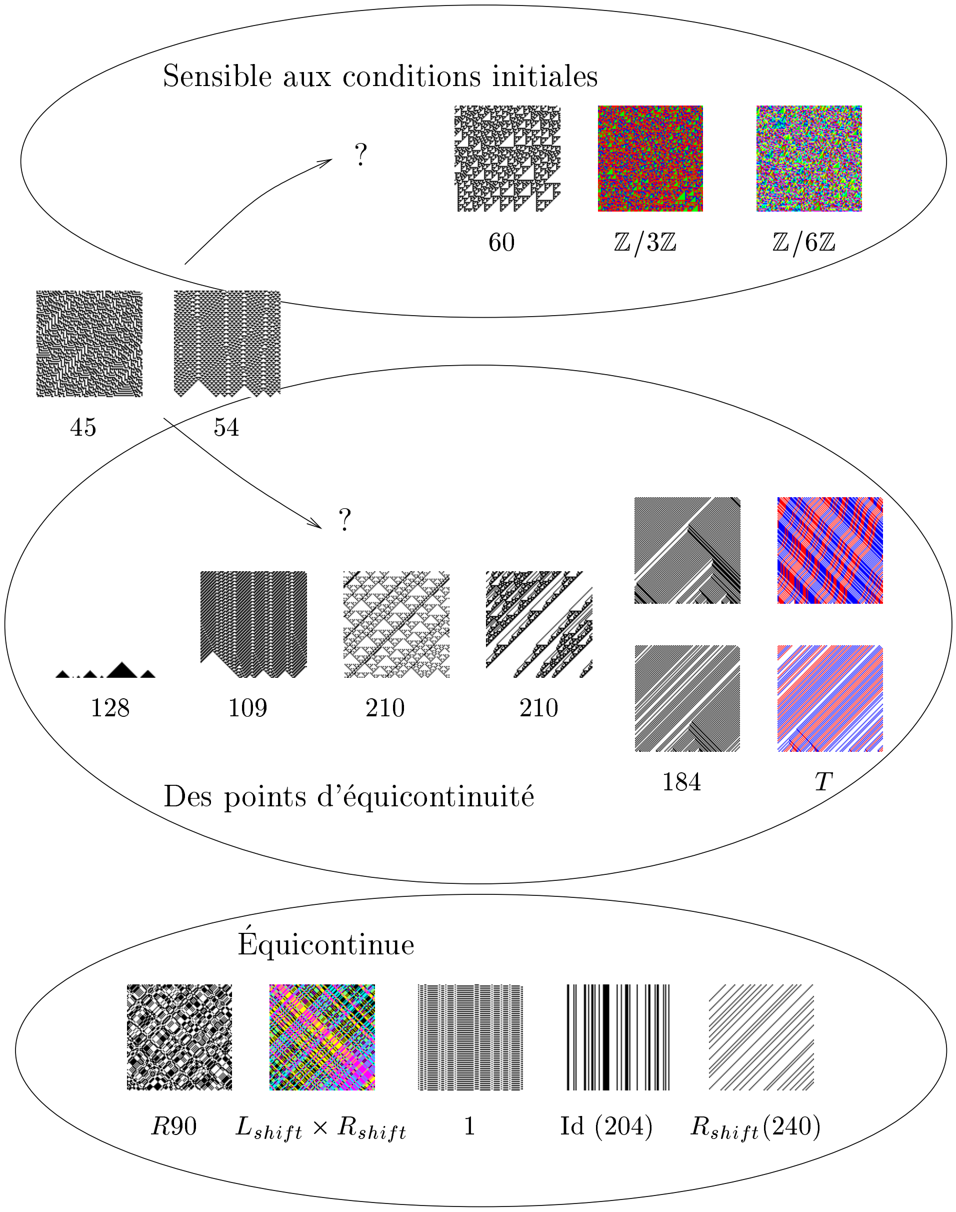


FIG. 4.10 – Le chaos d'après Formenti.

ni sensibles », des règles telles que 128 qui est sensible en 1^*, 109 et 210 en raison de 0^*. Nous trouvons enfin, comme prévu, les automates 184 et T dont le comportement dépend des conditions initiales. Ainsi, le principal problème est la grande classe de AC ni sensibles, ni équicontinus, qui présente des comportements très variés. C'est pour cela que nous utiliserons cette notion en nous munissant d'une mesure dans le chapitre 6.

4.5 Conclusion

Nous venons de faire le tour des principales classifications proposées dans la littérature. Cette étude a eu une grande influence sur nos travaux concernant la notion de chaos sur les automates cellulaires. En particulier, nous avons compris la nécessité de nous appuyer sur les théories du chaos développées en mathématiques (systèmes dynamiques, théorie ergodique) pour obtenir des définitions suffisamment robustes. Nous avons également compris l'origine topologique de la difficulté de définir des notions invariantes par décalage. Enfin, nous retenons trois idées qui sont à l'origine des résultats de notre dernier chapitre. La première est d'essayer de formaliser les remarques de Wolfram sur la propagation des défauts (propagation bornée dans sa classe 2 et non bornée dans sa classe 3). La deuxième est d'introduire une mesure comme dans les classifications d'Ishii et de Gilman. La troisième est d'utiliser la pseudo-distance de Besicovitch introduite par Formenti en raison de son invariance par décalage.

Chapitre 5

Entropie apparente

5.1 Introduction et définition

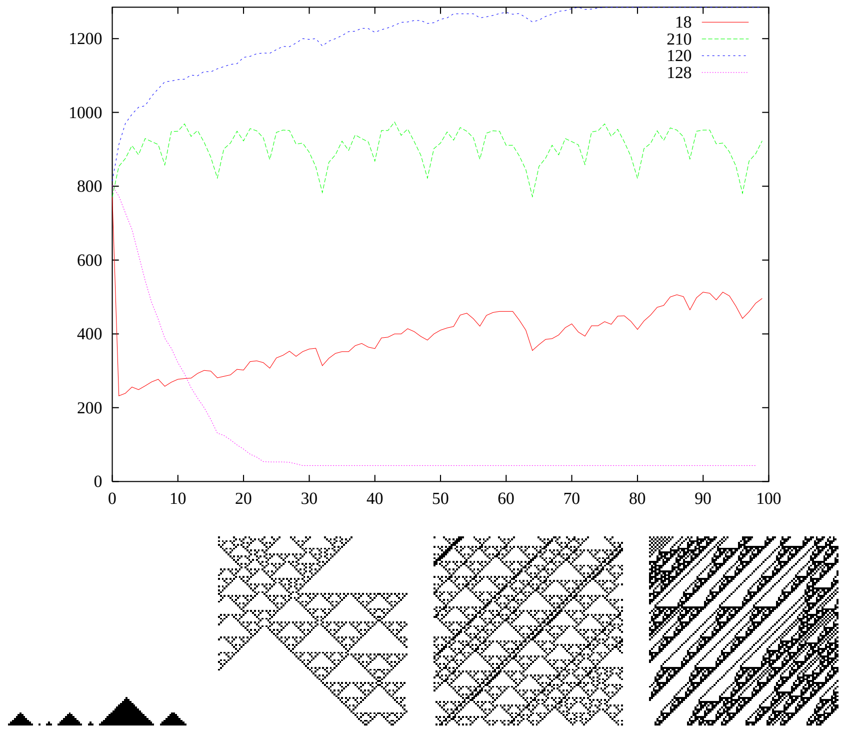
Il peut paraître surprenant que, malgré de nombreuses tentatives, aucune ne parvienne à vraiment rendre compte de la classification de Wolfram. Celle-ci est pourtant intuitive : à regarder un diagramme espace-temps, nous « voyons » du premier coup d'œil s'il est simple, complexe, ou s'il a l'air aléatoire. Dans un premier temps, nous allons donc essayer de comprendre ce que nous « voyons », de comprendre pourquoi tel diagramme semble simple et tel autre plus compliqué.

Wolfram, pour établir sa classification, s'est aidé de nombreuses mesures, en particulier, il a fait de nombreux calculs d'entropie. En effet une idée naturelle est de penser que les diagrammes qui ont l'air plus compliqués sont en fait plus désordonnés, ce qui suggère qu'il est possible de mesurer cela par un calcul d'entropie. Si les physiciens associent plutôt l'entropie au désordre, nous adopterons un point de vue plus informatique. Nous interpréterons donc l'entropie en termes de quantité d'information, puisque, comme nous l'avons déjà dit, l'entropie (dite de Shannon) est une fonction d'information. Rappelons que l'autre fonction d'information très célèbre est la complexité ou entropie de Kolmogorov. Cette complexité est algorithmique. La complexité de Kolmogorov d'un mot est la taille minimale d'un programme qui écrit ce mot dans un langage de programmation (ou un modèle de calcul) acceptable, c'est-à-dire optimal à une constante près. Les liens entre l'entropie de Shannon et la complexité de Kolmogorov sont très étroits puisqu'un résultat important de Romashchenko [25] montre que les formules linéaires vérifiées par l'entropie de Shannon et par la complexité de Kolmogorov sont les mêmes. Ce résultat justifie que, informellement, nous donnerons les idées des preuves en utilisant la complexité de Kolmogorov, tandis que, pour simplifier, nous ferons directement les preuves avec l'entropie de Shannon. De plus, nous utiliserons une propriété de la complexité de Kolmogorov (elle est approximable par dessus) pour proposer des protocoles expérimentaux d'approximation de notions liées à l'entropie.

Une toute première idée serait de considérer la quantité d'information du diagramme espace-temps tout entier en espérant que cela coïncide avec l'impression de complexité que nous avons en le regardant. Il est facile de vérifier que cela ne marche pas du tout parce que la quantité d'information du diagramme tout entier est exactement la même que la quantité d'information de la configuration initiale choisie, et ce, quel que soit l'automate cellulaire considéré. Cela a été remarqué plusieurs fois, cf [60, 15]. En effet, du point de vue de la complexité de Kolmogorov, un programme très simple peut construire le diagramme espace-temps en connaissant la configuration initiale et la table de transition (qui est finie). Une deuxième idée naturelle est de considérer l'évolution de la quantité d'information des configurations successives du diagramme espace-temps. Il y a alors deux attitudes possibles : celle du mathématicien qui prend une feuille de papier et un stylo pour définir la quantité d'information à l'aide de l'entropie et essaye de démontrer des résultats sur son évolution ou celle de l'informaticien qui se souvient qu'une méthode rapide (et imparfaite) de calculer l'entropie d'un objet est de mesurer le taux de compression obtenu avec un programme tel que `gzip` et qui fait un rapide programme pour voir expérimentalement ce que cette idée donne. C'est cette seconde attitude que nous allons adopter.

Le résultat de cette expérience est représenté sur la figure 5.1. Il semble plutôt encourageant puisque la taille des fichiers compressés correspondant à la règle 120 qui semble chaotique augmente jusqu'à ce que la compression devienne impossible tandis qu'elle reste stable pour la règle 210 qui ressemble à un shift et qu'elle tend vers une constante très faible pour la règle 128 qui tend vers la configuration uniforme. Quant à la règle 18 (cf section 2.5) qui semble chaotique mais n'est pas surjective, la taille décroît rapidement à la première étape (une cellule ne passe dans l'état 1 que si le mot 100 ou 001 apparaît sur la configuration initiale, ce qui est très peu probable puisque la densité de 1 de cette configuration est $0,9$) avant de remonter vers une valeur intermédiaire.

L'idée selon laquelle le désordre des configurations successives d'un automate cellulaire chaotique augmente est séduisante mais elle est fausse. En effet, comme l'a remarqué Wolfram [60] et démontré Lindgren [34], l'entropie des configurations des AC ne peut que décroître. L'explication est simple grâce à la complexité de Kolmogorov : il suffit de connaître la fonction de transition (finie) et la configuration au temps précédent pour calculer une configuration, celle-ci ne peut donc pas contenir plus d'information. Il s'agit



128 18 210 120

FIG. 5.1 – Pour chacune des règles 18, 120 et 128, nous avons choisi une configuration aléatoire de densité $0,9$ et stocké dans un fichier différent chacune des 100 premières configurations de leur évolution (chaque fichier contient 10000 cellules). Nous avons fait la même opération avec la règle 210 mais en choisissant une densité de $0,1$. Le diagramme représente la taille (en octets) des fichiers compressés avec `gzip`.

donc à présent de comprendre pourquoi l'œil humain d'une part et les compresseurs d'autre part ont la même « impression » de complexité concernant la règle 120 qu'ils n'ont pas vis-à-vis de la règle 210 alors que d'un point de vue théorique l'entropie reste constante dans les deux cas.

Définition 5.1.1 *Soit μ une mesure σ-ergodique. Soit $(\mathcal{A}, \delta)$ un AC, l'*entropie métrique *de sa configuration après t itérations est définie comme suit :*

$$S_\mu^{(t)}(\mathcal{A}) = \lim_{n\to+\infty} \frac{-\sum_{u\in Q^n} p_u^{(t)} \log(p_u^{(t)})}{n}$$

avec la convention usuelle $0 \times \log(0) = 0$ et où p_u est la probabilité d'apparition du mot u dans la configuration c : $p_u^{(t)} = G_\delta^t(\mu)([u]_0)$ sachant que

$G_\delta^t(\mu)$ représente la mesure définie par $G_\delta^t(\mu)(X) = \mu(G_\delta^{-t}(X))$. En langage mathématique, $S_\mu^{(t)}(\mathcal{A})$ est l'entropie métrique de σ pour la mesure $G_\delta^t(\mu)$.

Proposition 5.1.1 *Soit $\mathcal{A}$ un AC et μ une mesure sur l'ensemble de ses configurations. Si $t_1 \leq t_2$ alors $S_\mu^{(t_1)}(\mathcal{A}) \geq S_\mu^{(t_2)}(\mathcal{A})$.*

Preuve 5.1.1 [34] Afin de simplifier les notations, posons

$$S_{\mu,n}^{(t)} = - \sum_{u \in Q^n} p_u^{(t)} \log(p_u^{(t)})$$

où $p_u^{(t)}$ est la probabilité que le mot u apparaisse dans la configuration après t étapes de calcul. Les probabilités d'apparition aux temps t et $t+1$ sont liées par la relation

$$p_u^{(t+1)} = \sum_{\{v | \delta_{\mathcal{A}}^{n+2r \rightarrow r}(v) = u\}} p_v^{(t)}$$

ce qui permet d'exprimer l'entropie au temps $t+1$

$$S_{\mu,n}^{(t+1)} = - \sum_{u \in Q^n} p_u^{(t+1)} \log(p_u^{(t+1)}) \leq S_{\mu,n+2r}^{(t)}.$$

L'augmentation de l'entropie métrique en une étape de calcul est

$$\Delta S_\mu^{(t)}(\mathcal{A}) = S_\mu^{(t+1)}(\mathcal{A}) - S_\mu^{(t)}(\mathcal{A}).$$

Le calcul explicite donne

$$\begin{aligned} \Delta S_\mu^{(t)}(\mathcal{A}) &= \lim_{n \longrightarrow +\infty} \left(\frac{1}{n} S_{\mu,n}^{(t+1)}(\mathcal{A}) - \frac{1}{n+2r} S_{\mu,n+2r}^{(t)}(\mathcal{A}) \right) \\ &= \lim_{n \longrightarrow +\infty} \left(\frac{1}{n+2r} (S_{\mu,n}^{(t+1)}(\mathcal{A}) - S_{\mu,n+2r}^{(t)}(\mathcal{A})) \right. \\ &\qquad \left. + \left(\frac{1}{n} - \frac{1}{n+2r} \right) S_{\mu,n}^{(t+1)}(\mathcal{A}) \right) \leq 0 \end{aligned}$$

le premier terme dans la limite est négatif et le second tend vers zéro lorsque $n \longrightarrow +\infty$. Ainsi l'entropie métrique ne peut pas augmenter lorsqu'une règle déterministe est appliquée. $\square$

Nous voyons que l'entropie ne peut pas augmenter, mais que la limite qui la définit peut converger de plus en plus lentement dans les configurations successives. Cela signifie qu'il est nécessaire de regarder des mots de plus en plus longs pour détecter les corrélations existantes entre les états des

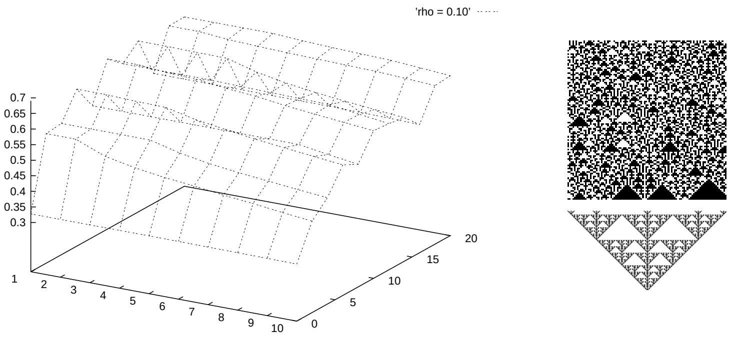


FIG. 5.2 – Ce diagramme représente, pour les 20 premières configurations de la règle additive 150 en partant d'une configuration aléatoire pour $\mu_{0,9}$, le i^{e}, $1 \leq i \leq 10$, terme de la limite qui définit l'entropie métrique. C'est-à-dire $-\sum_u p_u \log(p_u)$ si $u \in \{0,1\}^i$ et p_u est la probabilité d'apparition du mot u de longueur i. Nous voyons que pour tout i, la limite supérieure est $\log(2)$, ce qui signifie que les probabilités d'apparition tendent à être égales pour tous les mots de longueur i.

configurations. Cela a été remarqué dans [34] et représenté pour un AC additif unidimensionnel (la règle 150). La figure 5.2 représente les corrélations sur les mots de taille 1 à 10 pour les 20 premières configurations de l'évolution de la règle 150 partant d'une configuration initiale aléatoire dont la densité de 1 est 1/10. Nous voyons que très rapidement il n'y a presque plus de corrélation entre les mots de taille inférieure à 10, à quelques exceptions près qui sont spécifiques aux AC additifs aux temps 2^n.

Remarquons que cette observation explique pourquoi les AC que nous souhaitons qualifier de « chaotiques » ont un diagramme espace-temps compliqué, ou plus exactement des configurations compliquées la plupart du temps. En fait, puisque nous ne pouvons représenter qu'une partie finie du diagramme, nous sommes incapables de détecter ces corrélations et nous ne pouvons donc pas les distinguer d'une configuration aléatoire.

Afin de quantifier la limite sup de la complexité des configurations d'un diagramme, nous allons définir l'entropie apparente d'un AC qui est obtenue en inversant les limites.

Définition 5.1.2 *Soit $\mathcal{A}$ un AC et μ une mesure de Bernoulli, nous définissons l'*entropie apparente *de $\mathcal{A}$ de la façon suivante :*

$$S^a_\mu(\mathcal{A}) = \lim_{n\to+\infty} \limsup_{t\to+\infty} \sum_{u\in Q^n} p_u^{(t)} \log(p_u^{(t)}) = \lim_{n\to+\infty} \limsup_{t\to+\infty} S^{(t)}_{\mu,n}$$

où $p_u^{(t)}$ est la probabilité que le mot u apparaisse dans la configuration après t itérations.

Observons que :

– La limite $\lim_{n\to+\infty} \limsup_{t\to+\infty} S_{\mu,n}^{(t)}$ existe parce que la suite
$$\left(\limsup_{t\to+\infty} S_{\mu,n}^{(t)}\right)_n$$
est sous-additive.

– En mathématiques, la définition équivalente suivante de l'entropie apparente serait probablement préférée. Considérons la topologie associée à la limite simple sur l'ensemble des mesures, c'est-à-dire, telle que la suite de mesures $(\mu_n)_n$ converge si et seulement si pour tout mot u, la suite $(\mu_n(u))_n$ converge. Il est clair que l'ensemble des mesures est compact pour cette topologie (c'est un produit de compacts). Donc la suite de mesures $(G_\delta^n(\mu))_n$ a des points d'adhérence. L'entropie apparente est la limite sup des entropies des points d'adhérence.

Cette définition ressemble à celle de Ferrari *et al.* [18] qui considèrent, pour certains AC additifs (lorsque le nombre d'états est p^s avec p premier), la moyenne de Cesàro des distributions des configurations successives en partant d'une mesure de Markov μ
$$\mathcal{M}_\mu = \lim_{M\longrightarrow+\infty} \frac{1}{M} \sum_{m=0}^{M-1} \mu \circ G_\delta^{-m}.$$
Ils démontrent que cette limite existe et que c'est le produit de mesures uniformes (c'est-à-dire la mesure d'entropie maximale). Bien sûr, ce résultat implique que l'entropie apparente des AC additifs est maximale (i.e., $\log(|Q|)$). De plus, cela montre que la plupart des configurations vont avoir l'air complètement aléatoires. En fait, il semble plus intéressant de définir l'entropie apparente comme l'entropie de la moyenne de Cesàro puisque cela représente la complexité de presque toutes les configurations plutôt que sa limite sup. Le problème est que nous ne savons pas si la limite est définie pour tous les AC et que la définition que nous proposons est plus adaptée aux notions que nous introduirons par la suite.

5.2 Calcul exact et approximation de l'entropie apparente

Nous ne connaissons pas de méthode générale qui permette de calculer l'entropie apparente d'une règle quelconque. Cependant, un certain nombre d'automates ont des propriétés grâce auxquelles un calcul exact de l'entropie apparente est possible. Dans un premier temps, nous donnerons l'entropie apparente de ces règles et nous proposerons une nouvelle méthode qui permet de majorer l'entropie apparente des AC ayant un « attracteur ». Nous indiquerons ensuite des méthode expérimentales pour obtenir des approximations de l'entropie apparente des autres règles. Ces méthodes utilisent l'approximabilité par le haut de la complexité de Kolmogorov.

5.2.1 Calcul exact de l'entropie apparente

La bonne connaissance du comportement de certaines règles permet d'en déduire l'entropie apparente. Parmi elles, se trouvent les règles pour lesquelles nous pouvons démontrer que la configuration tend μ-presque sûrement vers une configuration uniforme (les règles 0, 1, 128, ...). Dans ce cas, l'entropie apparente est égale à 0 pour μ. Pour l'identité ou les décalages, l'entropie apparente est bien entendu égale à l'entropie de la configuration initiale. Enfin, nous avons vu que les règles additives ont une entropie apparente maximale pour les mesures de Bernoulli non triviales lorsque le nombre d'états est p^s avec p premier (c'est une conséquence de [18]).

Il est également possible de démontrer que l'entropie apparente est nulle lorsque l'automate évolue asymptotiquement vers un ensemble de configurations périodiques. En fait, il est possible de formaliser cette notion d'évolution vers des configurations périodiques à l'aide des notions plus générale de sous-shift μ-attracteurs définies dans [30]. Une de ces définitions de μ-attracteurs s'applique aux sous-shifts de type fini dont la définition est la suivante.

Définition 5.2.1 *Un* sous-shift *est un sous-ensemble $\Sigma \subseteq Q^{\mathbb{Z}}$ qui est σ-invariant et fermé dans la topologie produit. Le langage $L(\Sigma)$ d'un sous-shift $\Sigma \subseteq Q^{\mathbb{Z}}$, est l'ensemble des mots finis de Σ. Un* sous-shift de type fini *(SFT) est un sous-shift tel qu'il existe un entier positif p appelé ordre, vérifiant pour tout $c \in Q^{\mathbb{Z}}$,*

$$c \in \Sigma \iff \forall i \in \mathbb{Z}, c_{[i,i+p-1]} \in L(\Sigma).$$

Définition 5.2.2 *Considérons un SFT $\Sigma \subseteq Q^{\mathbb{Z}}$ d'ordre p et $x \in Q^{\mathbb{Z}}$, définissons la* densité supérieure de défauts d_D (D pour défaut, nous reprenons ainsi la notation de [30]) par rapport à Σ *dans x par*

$$d_D(x,\Sigma) = \limsup_{l \longrightarrow +\infty} \frac{\#\{i \in [-l,l] \mid x_{[i,i+p-1]} \notin L(\Sigma)\}}{2l+1}.$$

Notation 5.2.1 *Si d est une pseudo-distance, nous pouvons naturellement définir la pseudo-distance entre un élément x et un ensemble comme étant* *l'inf de la pseudo-distance entre x et les éléments de l'ensemble :*

$$d(x, \Sigma) = \inf_{y \in \Sigma} d(x, y).$$

Remarque 5.2.1 *Notons que $d_D(x, \Sigma)$ n'est associée à aucune pseudo-distance parce qu'en général $d_D(x, \{y\}) \neq d_D(y, \{x\})$.*

Définition 5.2.3 *Soit μ une mesure σ-ergodique, un sous-shift Σ est un Bμ-attracteur (resp. un Dμ-attracteur) si*

$$\mu(\{c \in Q^{\mathbb{Z}} \mid \lim_{n \to +\infty} \mathbf{d}(G_\delta^n(c), \Sigma) = 0\}) = 1$$

si $\mathbf{d}$ est la pseudo-distance d de Besicovitch (cf définition 4.4.1) (resp. $\mathbf{d} = d_D$).

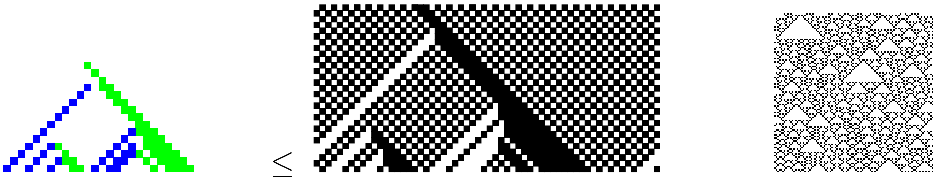
Remarque 5.2.2
- *Nous ne nous intéresserons dans la suite qu'aux Bμ-attracteurs finis. En effet, l'existence d'un Bμ-attracteur infini n'apporte pas d'information par rapport à l'existence d'un Dμ-attracteur.*
- *Comme l'ensemble $\{c \in Q^{\mathbb{Z}} \mid lim_{n \to +\infty} \mathbf{d}(G_\delta^n(c), \Sigma) = 0\}$ est shift-invariant, sa mesure est soit 0 soit 1.*
- *Nous pouvons extraire un Bμ-attracteur minimal de tout Bμ-attracteur (fini) en considérant les configurations c' de Σ telles que*

  $$\mu(\{c \in Q^{\mathbb{Z}} \mid \liminf d(G_\delta^n(c), c') = 0\}) = 1.$$

  Et lorsque μ est une mesure de σ-faiblement mélengeante, cela implique que pour tout c', il existe $q \in Q$ tel que $d(c', q^*) = 0$.
- *De plus, si un sous-shift est un Bμ-attracteur, alors c'est un Dμ-attracteur.*

La définition d'un Bμ-attracteur est très naturelle : un sous-shift est un Bμ-attracteur si pour presque toutes les configurations (vis-à-vis de μ), la pseudo-distance (de Besicovitch) entre les configurations successives et le sous-shift tend vers 0. Nous avons vu dans la dernière remarque que, dans ce cas, presque toutes les évolutions tendent vers des configurations uniformes lorsque μ est une mesure de Bernoulli. La règle 128 (voir la figure 6.1) ou les automates $\mathcal{S}$ et 184^2 pour certaines mesures ont des Bμ-attracteurs (cf figure 5.3).

La définition des Dμ-attracteurs est plus topologique parce qu'elle ne dépend que du langage $L(\Sigma)$ sans s'occuper du nombre d'occurrences des mots. Ainsi, les Dμ-attracteurs sont plus expressifs et permettent de mettre en évidence des processus d'homogénéisation vers des configurations périodiques. Par exemple (cf [30]), le sous-shift $\{(01)^*, (10)^*\}$ est un B$\mu_{1/2}$-attracteur de la règle 184. En observant la règle 18 (cf section 2.5), nous voudrions étendre la définition de Dμ-attracteur aux sous-shifts qui ne sont pas



$\mathcal{S}$ $\leq$ 184 18

Considérons l'AC $\mathcal{S}$ qui, outre l'état blanc, B, dans lequel évoluent les autres états, a 2 états, un se déplaçant vers la droite et un se déplaçant vers la gauche, qui, lorsqu'ils se rencontrent, s'annihilent. Si μ est une mesure telle que les états allant vers la gauche et ceux allant vers la droite ont la même probabilité de présence sur la configuration initiale alors $\{B^*\}$ est un Bμ-attracteur (et donc un Dμ-attracteur). Comme $\mathcal{S}$ est un sous-automate de 184^2 (l'automate 184 dont les états sont groupés deux par deux), nous avons le même résultat sur 184^2 pour la mesure telle que $(1,0)$ a la probabilité 0 d'apparaître et que les probabilités d'apparition de $(1,1)$ et $(0,0)$ sont égales. Cette mesure sur $(\{0,1\} \times \{0,1\})^{\mathbb{Z}}$ correspond à une mesure sur $\{0,1\}^{\mathbb{Z}}$ qui n'est pas shift-invariante. En fait, 184 n'a pas de Bμ-attracteur lorsque μ est une mesure shift-invariante non triviale. En effet, pour $\mu_{1/2}$ (telle que les particules allant vers la gauche et vers la droite ont les mêmes probabilités), le sous-shift $\{(01)^*, (10)^*\}$ est un Dμ-attracteur, mais asymptotiquement, les configurations sont à la pseudo-distance $1/2$ des deux configurations. La règle 18 pourrait être un exemple d'AC avec un Dμ-attracteur qui ne soit pas de type fini s'il était possible d'étendre la définition et n'a pas de Bμ-attracteur.

FIG. 5.3 – Exemples d'ensembles Bμ-attracteurs et Dμ-attracteurs.

de type fini afin de pouvoir dire que $\{c \in \{0,1\}^{\mathbb{Z}} | \forall i, c(2i) = 0\} \cup \{c \in \{0,1\}^{\mathbb{Z}} | \forall i, c(2i+1) = 0\}$ est un Dμ-attracteur (cf figure 5.3).

L'introduction des Dμ-attracteurs nous permet d'obtenir des résultats exacts sur l'entropie apparente. En effet on va établir le résultat suivant.

Proposition 5.2.1 *Si Σ est un SFT et un Dμ-attracteur de $\mathcal{A}$, alors l'entropie apparente $S^a_\mu(\mathcal{A})$ est inférieure à l'entropie topologique de Σ.*

Définition 5.2.4 *L'entropie topologique d'un sous-shift Σ est*

$$h_\Sigma(\sigma) = \limsup_{n \longrightarrow +\infty} \frac{\log(\#\{u \in Q^n \mid \exists c \in \Sigma, c_{[0,n-1]} = u\})}{n}.$$

Lemme 5.2.1 *Soit Σ un SFT d'ordre p. Si $k \geq p$, définissons*

$$d_D^{(k)}(x, \Sigma) = \limsup_{l \longrightarrow +\infty} \frac{\#\{i \in [-l, l] \mid x_{[i,i+k-1]} \notin L(\Sigma)\}}{2l+1}.$$

Alors $d_D^{(k)}(x,\Sigma) \leq (k-p)d_D(x,\Sigma)$.

Preuve 5.2.1 Comme Σ est un SFT, $x_{[i,i+k-1]} \in L(\Sigma)$ si et seulement si $\forall j, 0 \leq j \leq k-p, x_{[j,j+p-1]} \in L(\Sigma)$. Donc pour tout $i \in [-l,l]$ tel que $x_{[i,i+p-1]} \notin L(\Sigma)$, nous avons $x_{[j,i+k-1]} \notin L(\Sigma)$ pour $i-k+p \leq j \leq i$. Ainsi dans le pire cas, $\#\{i \in [-l,l] \mid x_{[i,i+k-1]} \notin L(\Sigma)\} \leq (k-p)\#\{i \in [-l,l] \mid x_{[i,i+p-1]} \notin L(\Sigma)\}$. Nous en déduisons que

$$d_D^{(k)}(x,\Sigma) \leq (k-p)d_D(x,\Sigma).$$

$\square$

Preuve 5.2.2 (de la proposition 5.2.1). La preuve est une conséquence du fait que l'entropie métrique est inférieure à l'entropie topologique.

Par définition,

$$S_\mu^a(\mathcal{A}) = \lim_{n\to+\infty} \limsup_{t\to+\infty} S_{\mu,n}^{(t)} = \lim_{n\to+\infty} \limsup_{t\to+\infty} \sum_{u\in Q^n} p_u^{(t)} \log(p_u^{(t)})$$

où $p_u^{(t)} = \mu((G_\delta^t)^{-1}([u]))$.

Nous allons démontrer que pour tout $\epsilon > 0$, cette limite est inférieure à $h_\Sigma(\sigma) + o(\epsilon)$ où o est une fonction de $\mathbb{R} \longrightarrow \mathbb{R}$ telle que $o(\epsilon) \longrightarrow_{\epsilon\to 0} 0$.

Par définition de $h_\Sigma(\sigma)$, il existe K tel que pour tout $k \geq K$

$$\frac{\log(\#\{u \in Q^k \mid \exists c \in \Sigma, c_{[0,k-1]} = u\})}{k} \leq h_\Sigma(\sigma) + \epsilon.$$

Si un SFT Σ est un Dμ-attracteur de $(Q^{\mathbb{Z}}, G_\delta)$, par définition

$$\mu(\{c \in Q^{\mathbb{Z}} \mid \lim_{n\to+\infty} d_D(G_\delta^n(c),\Sigma) = 0\}) = 1$$

donc $\mu(\{c \in Q^{\mathbb{Z}} \mid \lim_{n\to+\infty} d_D^{(k)}(G_\delta^n(c),\Sigma) = 0\}) = 1$ d'après le lemme 5.2.1.

Donc, pour tout $\alpha > 0$, il existe M tel que pour $t \geq M$,

$$\mu(\{c | d_D^{(k)}(G_\delta^t(c),\Sigma) \leq \epsilon\}) \geq 1-\alpha.$$

Comme σ est ergodique, la mesure de cet ensemble est 1.

Dans ce cas, nous avons

$$\begin{aligned} S_{\mu,k}^{(t)} &= \frac{-\sum_{u\in Q^k} p_u^{(t)} \log(p_u^{(t)})}{k} \\ &= \frac{-\sum_{u\in Q^k\cap L(\Sigma)} p_u^{(t)} \log(p_u^{(t)})}{k} + \frac{-\sum_{u\in Q^k\setminus L(\Sigma)} p_u^{(t)} \log(p_u^{(t)})}{k} \end{aligned}$$

$$\begin{aligned}
&\leq \frac{\log(\#\{u \in Q^k \mid \exists c \in \Sigma, c_{[0,k-1]} = u\})}{k} \\
&\qquad\qquad + \frac{-\sum_{u \in Q^k \setminus L(\Sigma)} p_u^{(t)} \log(p_u^{(t)})}{k} \\
&\leq h_\Sigma(\sigma) + \epsilon + \frac{-\sum_{u \in Q^k \setminus L(\Sigma)} p_u^{(t)} \log(p_u^{(t)})}{k}
\end{aligned}$$

à présent, pour majorer le dernier terme, notons $\epsilon_t(c) = d_D^{(k)}(G_\delta^t(c), \Sigma)$ et remarquons que pour $t \geq M$ et μ-presque tout c, $\epsilon_t(c) \leq \epsilon$.

Notre premier objectif est de montrer que $\sum_{u \in Q^k \setminus L(\Sigma)} p_u^{(t)} \leq \epsilon$. Définissons

$$f \left| \begin{array}{ccl} Q^{\mathbb{Z}} & \longrightarrow & \mathbb{R} \\ c & \longmapsto & \begin{array}{l} 1 \text{ if } c_{[0,k-1]} \in Q^k \setminus L(\Sigma) \\ 0 \text{ sinon} \end{array} \end{array} \right.$$

et $g_t = f \circ G_\delta^t$. Nous avons $\int g_t d\mu = \sum_{u \in Q^k \setminus L(\Sigma)} p_u^{(t)}$. f est mesurable (en fait continue) donc g_t est mesurable également et nous pouvons appliquer le théorème ergodique de Birkhoff (cf par exemple [53, 58]) avec μ et σ : si $S_n(g_t)(c) = 1/(2n+1) \sum_{i=-n}^{n} g_t \circ \sigma^i(c)$, pour μ-presque tout c,

$$\lim_{n \to \infty} S_n(g_t)(c) = \int g_t d\mu.$$

Par ailleurs,

$$\begin{aligned}
\lim_{n \to \infty} S_n(g_t)(c) &= \lim_{n \to \infty} \frac{\#\{i \in [-n,n] | G_\delta^t(c)_{[i,i+k-1]} \notin L(\Sigma)\}}{2n+1} \\
&= d_D^{(k)}(G_\delta^t(c), \Sigma).
\end{aligned}$$

Ainsi, pour μ-presque tout c,

$$\epsilon_t(c) = d_D^{(k)}(G_\delta^t(c), \Sigma) = \int g_t d\mu = \sum_{u \in Q^k \setminus L(\Sigma)} p_u^{(t)} \leq \epsilon.$$

De plus, $-\sum p_u^{(t)} \log(p_u^{(t)})$ est maximal lorsque tous les $p_u^{(t)}$ sont égaux à $\epsilon / |Q^k \setminus L(\Sigma)|$, donc dans le pire cas,

$$S_{\mu,k}^{(t)} = \frac{-\sum_{u \in Q^k} p_u^{(t)} \log(p_u^{(t)})}{n}$$

$$\begin{aligned}
&\leq h_\Sigma(\sigma) + \epsilon + |Q^k \backslash L(\Sigma)| \left(\frac{-1}{k} \frac{\epsilon}{|Q^k \backslash L(\Sigma)|} \log \left(\frac{\epsilon}{|Q^k \backslash L(\Sigma)|} \right) \right) \\
&\leq h_\Sigma(\sigma) + \epsilon + \frac{-\epsilon}{k} \log \left(\frac{\epsilon}{|Q^k \backslash L(\Sigma)|} \right) \\
&\leq h_\Sigma(\sigma) + \epsilon + \frac{-\epsilon}{k} \log \left(\frac{\epsilon}{|Q|^k} \right) \\
&\leq h_\Sigma(\sigma) + \epsilon - \frac{\epsilon}{k} \log(\epsilon) + \epsilon \log(|Q|).
\end{aligned}$$

Cette inégalité est vraie pour tout $t \geq M$ et pour tout $k \geq K$, donc, $S^a_\mu(\mathcal{A}) = \lim_n \limsup_t S^{(t)}_{\mu,n} \leq h_\Sigma(\sigma) + \epsilon - \lim_k \epsilon/k \log(\epsilon) + \epsilon \log(|Q|) = h_\Sigma(\sigma) + \epsilon + \epsilon \log(|Q|)$ et, comme $\epsilon + \epsilon \log(|Q|) \longrightarrow 0$, $S^a_\mu(\mathcal{A}) \leq h_\Sigma(\sigma)$. $\square$

Donnons quelques applications de ce résultat. Comme nous l'avons signalé ci-dessus, le sous-shift $\{(01)^*, (10)^*\}$ est un B$\mu_{1/2}$-attracteur pour la règle 184, donc $S^a_{\mu_{1/2}}(184) = 0$. De plus, si $0 < \rho < 1/2$ (resp. $1/2 < \rho < 1$) alors le sous-shift de type fini engendré par 00 (resp. 11), 01, et 10 est un Bμ_ρ-attracteur de la règle 184. Or l'entropie topologique d'un SFT est donnée par la formule

$$h_\Sigma(\sigma) = \limsup_{n \longrightarrow +\infty} \frac{\log(\#\{u \in Q^n \mid \exists c \in \Sigma, c_{[0,n-1]} = u\})}{n}.$$

Pour le SFT engendré par 00 (resp. 11), 01, et 10,

$$\#\{u \in Q^n \mid \exists c \in \Sigma, c_{[0,n-1]} = u\} = F_{n+1}$$

où F_n est le n^{e} terme de la suite de Fibonacci. En effet, il y a trois mots de longueur 2 qui apparaissent dans Σ dont deux se terminent par un 0 (resp. par un 1). Or par récurrence, s'il y a F_{n+1} mots de longueur n apparaissant dans Σ dont F_n se terminant par un 0 (resp. par un 1), alors les mots de longueur $n+1$ apparaissant dans Σ sont des mots de longueur n auxquels on a ajouté un 0 (resp. un 1) en dernière position plus les mots de longueur n qui ne se terminent pas par un 0 (resp. par un 1) auxquels on a rajouté un 1 (resp. un 0). On trouve donc $F_{n+2} = F_{n+1} + F_n$ mots de longueur $n+1$ dont F_{n+1} se terminent par un 0 (resp. par un 1). Par ailleurs, nous savons que

$$F_n = \frac{5+\sqrt{5}}{10} \left(\frac{1+\sqrt{5}}{2} \right)^n + \frac{5-\sqrt{5}}{10} \left(\frac{1-\sqrt{5}}{2} \right)^n,$$

donc $S^a_{\mu_\rho}(184) \leq \log((1+\sqrt{5})/2)$.

5.2.2 Approximation de l'entropie apparente

L'objectif de cette section est de décrire et discuter la méthode expérimentale que nous avons utilisée pour obtenir des approximations de l'entropie apparente lorsque nous ne pouvons pas la calculer exactement.

Les résultats que nous donnons ont été obtenus en calculant les corrélations sur les mots de taille 10 (c'est-à-dire le 10$^{\text{e}}$ terme de la limite qui définit l'entropie de la configuration) après 100 itérations. Le problème est qu'il est presque impossible de calculer les corrélations sur des mots beaucoup plus longs (c'est possible pour des mots de taille 20, mais plus pour des mots de taille 30). En dessinant les diagrammes tridimensionnels des corrélations en fonction du nombre d'itérations et de la taille des mots, nous pouvons constater que cela constitue une bonne approximation pour la plupart des règles, mais pas pour toutes les règles. Pour certaines règles (par exemple 54 ou 110), le nombre d'itérations est nettement insuffisant, nous avons alors calculé 1000 itérations ce qui reste d'ailleurs insuffisant : il est nécessaire d'utiliser d'autres outils pour deviner l'entropie apparente de ces règles. Parfois le nombre d'itérations est suffisant, mais pas la taille des mots considérés. Cela se produit principalement pour les règles presque chaotiques (voir la section 5.4). Dans ce cas-là, nous n'avons pas essayé de calculer les corrélations sur des mots plus longs, il faut donc garder à l'esprit que les résultats donnés sont parfois assez éloignés de l'entropie apparente. En particulier, ces erreurs ont deux conséquences : elles expliquent, d'une part, que dans nos diagrammes, l'entropie apparente tend toujours vers zéro lorsque ρ tend vers 0 ou 1 et, d'autre part, que dans certains cas l'entropie apparente peut en fait être significativement inférieure à la valeur représentée sur nos diagrammes (par exemple pour les règles telles que 54 et 110 ou la règle 184 avec $\rho = 1/2$, dont l'entropie apparente devrait être zéro).

Notons que pour avoir une idée rapide de l'entropie apparente d'une règle, nous pouvons utiliser les compresseurs usuels sur un fichier contenant une configuration après 100 ou 1000 itérations. Les expériences que nous avons faites avec `gzip` donnent des résultats comparables à ceux obtenus avec la compression de Huffman sur des mots de longueur $l = 10$ pour des configurations de 10000 cellules et $l = 20$ pour des configurations de 100000 cellules. En fait, bien qu'ils utilisent d'autres algorithmes, les autres compresseurs sont soumis aux mêmes contraintes de mémoire et ne peuvent détecter des corrélations sur des mots beaucoup plus longs. L'algorithme de Huffman consiste à calculer les probabilités des mots d'une longueur donnée pour affecter, de façon optimale, des codes plus courts aux mots revenant le plus fréquemment. Asymptotiquement, le taux de compression sur des mots de longueur l est le l^{e} terme de la suite qui définit l'entropie métrique. Il est impossible de compresser pour de grandes valeurs de l puisqu'il faut calculer les fréquences des 2^l mots de longueur l, ce qui est vite impossible : pour $l = 30$ il faudrait calculer les fréquences de plus d'un milliard de mots. `gzip` est basé sur l'algorithme LZ77 qui utilise une fenêtre flottante de 4000 octets comme dictionnaire. Les 17 octets suivants sont lus et l'algorithme cherche le plus long préfixe de ce mot qui apparaît dans la fenêtre afin de coder ce facteur par sa position et sa longueur. Pour détecter des corrélations sur des mots relativement longs, il faut qu'ils aient une probabilité raisonnable d'apparaître dans la fenêtre et donc que celle-ci soit suffisamment grande. Or l'augmentation de la taille de la fenêtre pose de nombreux problèmes, d'une part la taille pour coder la position du facteur trouvé augmente, d'autre part, le temps de calcul et la mémoire nécessaire augmentent également. La taille du dictionnaire est une première limite à la taille des corrélations prises en compte puisqu'il faut que les mots les plus fréquents de cette longueur ait une probabilité suffisante d'être dans le dictionnaire. La seconde limite vient de la taille des configurations à compresser. Comme nous le voyons sur la figure 5.2.2, les taux de compressions sont meilleurs avec des configurations de 100000 cellules (12,5ko) qu'avec seulement 10000 cellules (1,25ko) il faut dire que dans ce dernier cas, la taille des fichiers à compresser est inférieure à la taille de la fenêtre glissante.

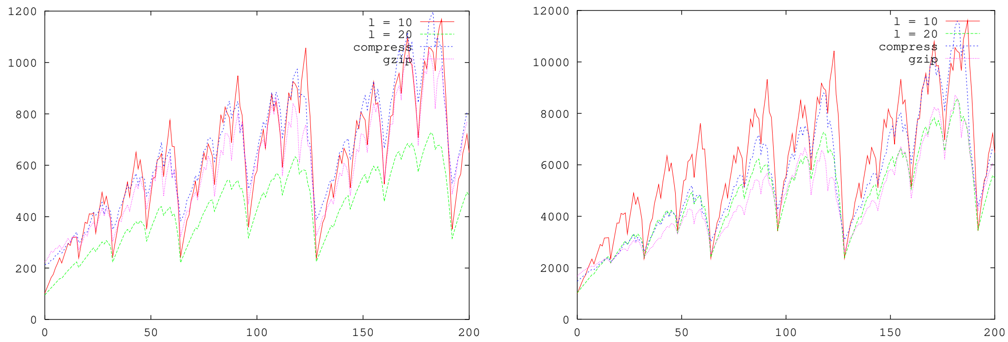


FIG. 5.4 – Nous voyons les taux de compressions obtenus avec `gzip` et `compress` comparés aux 10e et 20e termes de la suite définissant l'entropie. Sur le diagramme de gauche, les configurations contiennent 10000 cellules, contre 100000 pour le diagramme de droite.

Nous avons également essayé l'algorithme `compress` basé sur l'algorithme LZ78. Dans cet algorithme, le dictionnaire est construit au fur et a mesure. On cherche le plus petit facteur qui ne soit pas dans le dictionnaire du mot qu'il reste à coder. Par définition ce facteur privé de sa dernière lettre est dans le dictionnaire. Cette entrée du dictionnaire plus la dernière lettre permettent donc de coder le facteur tandis qu'il est ajouté au dictionnaire. Lorsque le dictionnaire est plein, soit il est vidé, soit il peut être conservé en vérifiant qu'il est toujours adapté à la compression. Les résultats obtenus avec `compress` sont supérieurs à ceux obtenus avec `gzip`, surtout dans le cas de 100000 cellules, ce qui ne veut bien entendu strictement rien dire sur leur efficacité respective puisque les configurations considérées sont très petites (1 à 12 ko) et que ces programmes sont optimisés pour compresser des fichiers beaucoup plus gros. Contrairement à `gzip`, `compress` n'est pas plus efficace pour 100000 cellules que pour 10000, en fait, comme `gzip`, les corrélations détectées sont dans nos expériences limitées par la taille très faible des configurations, mais seraient de toute façon limitées par la taille du dictionnaire utilisé.

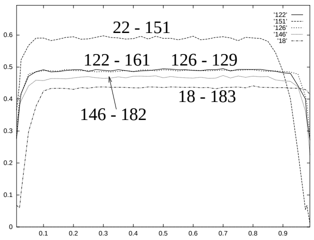
En fait, nous voyons sur la figure 5.2.2 que l'apparence des courbes est la même quelle que soit la méthode utilisée et quelle que soit la taille des configurations. En utilisant les propriétés spécifiques de chaque compresseur, il doit être possible de construire des contre-exemples pour lesquels les résultats sont très différents en fonction de la méthode. En pratique, cela n'est pas un problème, en effet, on ne trouve jamais un tel contre-exemple par hasard. Ainsi, ces méthodes sont en pratique très efficaces pour avoir une approximation de la complexité des configurations.

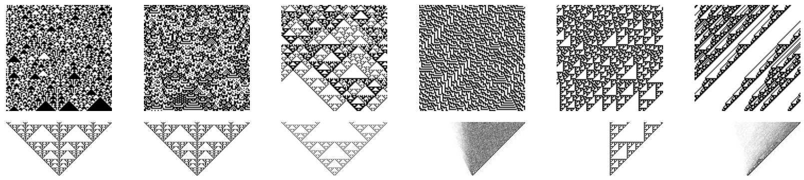
5.3 L'entropie apparente et le « chaos »

Nous nous demandons dans cette section s'il est possible, comme cela est suggéré par Wolfram, de dire qu'un AC est chaotique si son entropie apparente pour les mesures de Bernoulli μ_ρ pour $0 < \rho < 1$ est constante. Dans la définition de la classe 3 de Wolfram, il écrit en effet : « Évolution of infinite class 3 cellular automata from almost all possible initial states leads to aperiodic ("chaotic") patterns. After sufficiently many time steps, the statistical properties of these patterns are typically the same for almost all initial states. In particular, the density of nonzero sites typically tends to a fixed nonzero value. » Nous supposons donc que, par définition, un AC est chaotique si son entropie apparente est constante pour évaluer la pertinence de cette définition.

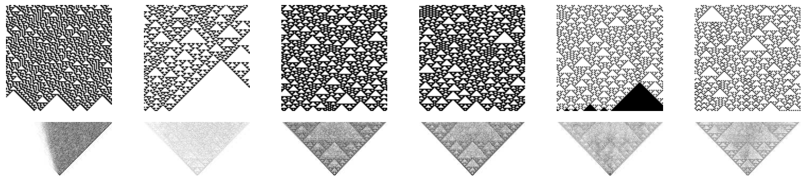
Une étude exhaustive nous a amené à penser que les ECA chaotiques sont les 30 règles représentées sur la figure 5.5. Excepté pour les automates additifs (les règles 150, 105, 90, 165, 60, 195, 102 et 153), on ne sait pas

	r	$\overline{r}$	$\overleftrightarrow{r}$	$\overleftrightarrow{\overline{r}}$
log(2)	150			
	105			
	90	165		
	45	75	101	89
	60	195	102	153
	120	225	106	169
	30	135	86	149
0.59	22	151		
0.49	122	161		
	126	129		
0.47	146	182		
0.43	18	183		





150 105 90 45 60 120



30 22 122 126 146 18

FIG. 5.5 – Tous les comportements chaotiques parmi les ECA

démontrer que leur entropie apparente est constante. Nous avons regroupé les règles ayant le même comportement à une symétrie près ou en échangeant les états 0 et 1. L'AC représenté montre pour chaque classe le diagramme espace-temps de la règle dont le numéro (dans la notation de Wolfram) est le plus petit.

Une première observation est que les entropies apparentes, qui sont donc constantes pour $0 < \rho < 1$, ne sont pas nécessairement maximum. Par exemple, la règle 18 (cf section 2.5 et figure 5.6). Sans que nous puissions le prouver, nous avons expliqué pourquoi il était naturel de penser qu'apparaissent des régions de plus en plus grandes dont une cellule sur deux est nulle et l'autre a une probabilité d'être nulle qui tend vers 1/2. De ce fait, l'entropie apparente de la règle 18 est probablement log(2)/2. Cela est

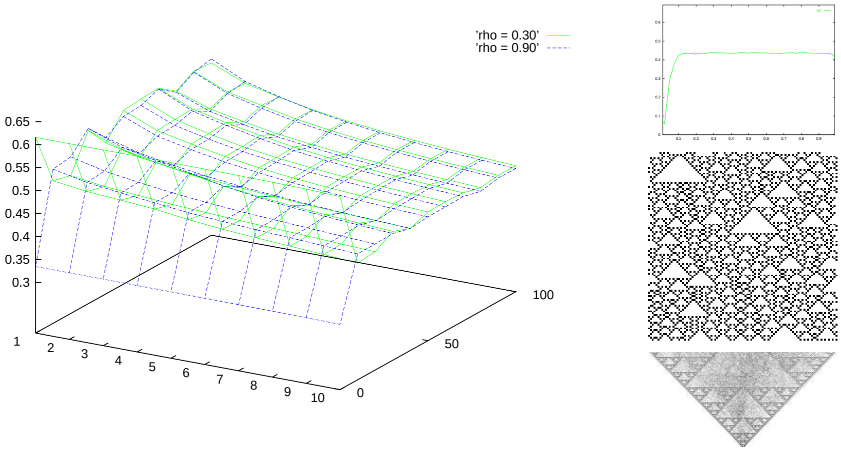


FIG. 5.6 – La règle 18 est une règle chaotique qui a très probablement une entropie apparente de $\log(2)/2$ pour toutes les mesures de Bernoulli non triviales (l'approximation de l'entropie apparente pour les mesures de Bernoulli avec $0 < \rho < 1$ est représentée sur le diagramme en haut à droite). Nous voyons sur le diagramme de gauche que les corrélations sur les mots de longueur inférieure à 10 convergent très rapidement vers la même valeur en partant de deux densités différentes de 1 (0,1 et 0,7).

cohérent avec notre souhait d'étendre la notion de Dμ-attracteur à des sous-shift qui ne sont pas de type fini (cf figure 5.3) puisque l'entropie topologique du sous-shift en question est précisément $\log(2)/2$. Nous observons que nos expériences donnent des résultats supérieurs en raison des imprécisions des mesures discutées ci-dessus. Nous voyons que la règle 146 (excepté au tout début où elle efface les longues séquences de 1) a le même comportement que la règle 18 ($146 = 128 + 18$). Ainsi, leur entropie apparente est la même bien que les expériences donnent des valeurs différentes.

Si cette « définition » de chaos donne des résultats convaincants pour les ECA, il y a des AC ayant plus d'états qui, bien que n'ayant pas une entropie apparente constante, ont un comportement qui semble « chaotique ». Il suffit par exemple de faire le produit d'une règle chaotique au sens précédent avec l'identité pour obtenir une telle règle. De plus, nous avons construit l'automate T de telle sorte qu'il soit « chaotique » lorsque la densité de voitures est grande, or nous observons sur la figure 5.7 que son entropie apparente n'est pas constante, même pour $\rho \geq 1/2$. Ainsi, si la propriété d'avoir une entropie apparente constante non nulle semble liée à la notion de « chaos » puisque les AC la vérifiant que nous connaissons ont un comportement « chaotique »,

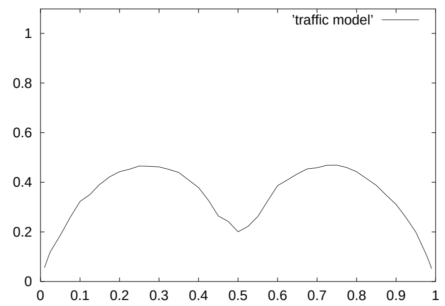


FIG. 5.7 – Ce diagramme représente l'entropie apparente de T, le modèle très simple de trafic routier, pour la mesure μ_ρ^* pour $0 \leq \rho \leq 1$.

sans qu'il soit clair que ce soit toujours le cas, à l'évidence, de nombreux comportements « chaotiques » ne vérifient pas cette propriété.

5.4 Les règles « presque chaotiques »

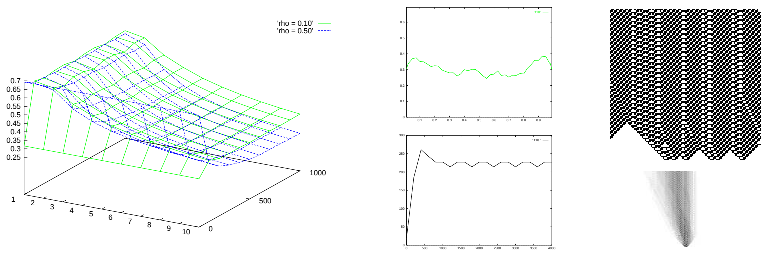


FIG. 5.8 – Pour tout $0 < \rho < 1$, la règle 118 a une entropie apparente qui n'est pas constante, mais qui est relativement indépendante de ρ comme le montre le diagramme au centre en haut. En fait, une seule modification d'état sur la configuration initiale ne modifie presque sûrement qu'un nombre fini de cellules (cf le diagramme au centre en bas). Le diagramme de gauche montre qu'en partant de deux densités différentes, on arrive à des entropie apparentes différentes, mais assez proches.

Outre les règles précédentes, d'autres règles ont une entropie apparente qui est presque constante (cf figure 5.8). Malheureusement, ces automates ne semblent pas partager une propriété spécifique qui permettrait de les définir formellement. Dans le cas des ECA, les règles qui nous intéressent sont celles dont l'entropie apparente ne tend pas vers 0 lorsque ρ tend vers 0 ou 1. Nous constatons que toutes (parmi les ECA) ont des diagrammes espace-temps dans lesquels nous identifions des particules. Afin d'avoir une entropie apparente très supérieure à l'entropie métrique de sa configuration initiale, il est nécessaire « mélanger » les informations d'un grand nombre de cellules de la configuration initiale. En fait, contrairement aux règles ayant une entropie apparente constante et qui mélangent asymptotiquement les informations d'un nombre non borné de cellules, les règles presque chaotiques mélangent un nombre borné d'états pour une mesure donnée, mais cette borne tend vers $+\infty$ lorsque l'entropie de la mesure tend vers 0.

Il serait intéressant de trouver une caractérisation des règles « presque chaotiques ». Nous verrons dans le chapitre suivant qu'il est possible de les distinguer des règles ayant une entropie apparente constante ; par contre, nous ne connaissons pas de définition qui permette de les distinguer des règles telles que l'identité ou 210. Pourtant, l'identité ou 210 ne font que conserver l'information de la configuration initiale sans augmenter l'entropie apparente, tandis que les règles « presque chaotiques » ont une entropie apparente très supérieure à l'entropie de la configuration initiale lorsque celle-ci est très faible.

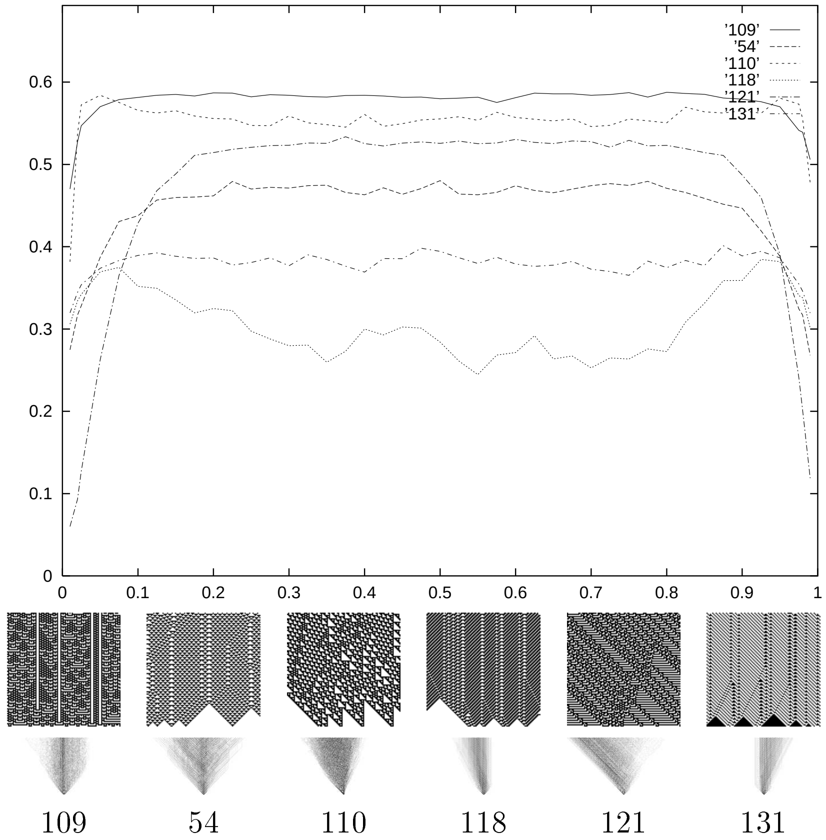


FIG. 5.9 – Quelques comportements presque chaotiques parmi les ECA. Nous observons sur le diagramme du haut que les mesures expérimentales des entropies apparentes pour les mesures de Bernoulli μ_ρ pour $0 \leq \rho \leq 1$, sans être constantes varient peu. Sous les diagrammes espace-temps des règles, les diagrammes indiquant les nombres moyens de modifications suggèrent que les modifications engendrées par un seul défaut (notion que nous définissons formellement dans la section 6.1 ci-dessous) sont de plus en plus nombreuses au cours de l'évolution des règles 54, 110 et 121, mais bornées au cours du temps pour les règles 109, 118 et 131. Cependant, dans ce dernier cas, d'autres expériences montrent que la borne n'est pas uniforme en fonction de ρ.

Chapitre 6

Vers une définition du chaos

Le chapitre précédent et l'entropie apparente nous ont permis de mieux comprendre d'où venait l'aspect plus ou moins compliqué des diagrammes espace-temps bien que les fonctions de transitions soient déterministes. Cependant, excepté lorsque l'entropie apparente tend vers zéro, ce qui correspond à l'effacement de toute l'information contenue dans la configuration initiale, la notion d'entropie apparente ne nous a pas permis de poser une définition vraiment pertinente quant à la « complexité » des règles considérées.

Le but de ce chapitre est donc, par d'autres méthodes, de poser des définitions intéressantes pour comprendre les évolutions des AC « chaotiques ». Compte tenu de notre étude d'un certain nombre d'autres classifications, nous avons les exigences suivantes : nos définitions doivent être invariantes par décalage et par le produit d'automates. En effet, il ne nous paraît pas normal que la composition par un décalage change un automate de classe, ni que le produit de deux automates « non chaotiques » soit chaotique. Nous introduisons deux notions, la propagation d'information et le sensibilité presque partout aux conditions initiales pour Besicovitch qui satisfont à nos exigences. Nous étudierons les relations qui les lient à l'entropie apparente et entre elles ainsi que leur pertinence, à l'aide (la plupart du temps) de résultats démontrés mais aussi de résultats expérimentaux.

6.1 Propagation d'information

Inspiré des exposants de Lyapunov [2] et motivé par les remarques de Wolfram (« The density of appropriate sequences in a particular initial state therefore determines the statistical properties of the final state into which it evolves. » pour la classe 2 et « Small changes in initial states thus almost always lead to increasingly large changes in later states. » pour la classe 3), nous définissons la notion de propagation d'information d'un AC pour une mesure. L'idée est de compter le nombre moyen de modifications que le changement d'une seule cellule de la configuration initiale va provoquer pour une mesure donnée. La principale différence avec les exposants de Lyapunov est que nous ne comptons que les modifications « effectives » induites par la modification d'une seule cellule de la configuration initiale. Par exemple, si la modification de l'état conduit à deux modifications au temps $t-1$ et que chacune de ces modifications changerait un état au temps t mais que l'action des deux laisse l'état inchangé, alors, plutôt que de compter 2 modifications comme dans le cas des exposants de Lyapunov, nous ne comptons aucune modification parce que l'état n'a pas changé. Il semble que la règle 210 (cf figure 6.2) a un exposant de Lyapunov strictement supérieur à 1 (donc le nombre de modifications est exponentiel si une même cellule peut être comptée plusieurs fois) tandis que le nombre de modifications de cet AC est borné du point de vue de la propagation d'information. Nous allons à présent définir cela formellement.

Définition 6.1.1 *Soit $\mathcal{A}$ un AC et ${Q_\mathcal{A}}^\mathbb{Z}$ l'ensemble de ses configurations. Soit μ une mesure de Bernoulli sur ${Q_\mathcal{A}}^\mathbb{Z}$. Si $c \in {Q_\mathcal{A}}^\mathbb{Z}$, définissons*

$$c_{p\leftarrow s} \left| \begin{array}{ccl} \mathbb{Z} & \longrightarrow & Q_\mathcal{A} \\ i & \longmapsto & \begin{cases} c(i) \text{ si } i \neq p \\ s \text{ sinon} \end{cases} \end{array} \right.$$

c'est-à-dire la configuration dont l'état de la cellule p est remplacé par s. Définissons à présent les coefficients de dépendance $\alpha_{\mu,\mathcal{A},t,p}$ qui indiquent la probabilité (selon μ) que l'état de la cellule 0 après t étapes de calcul change lorsque nous remplaçons l'état à la position p par q avec la probabilité $\pi_q = \mu([q])$. Formellement,

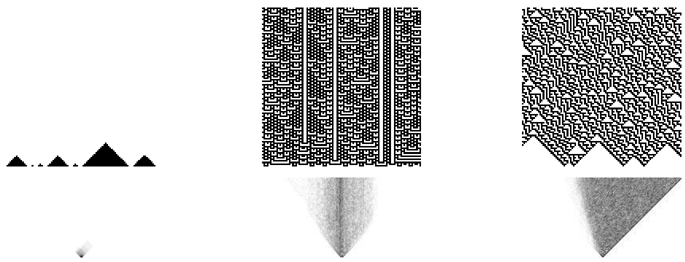
$$\alpha_{\mu,\mathcal{A},t,p} = \sum_{q\in Q_\mathcal{A}} \pi_q \mu(\{c \in {Q_\mathcal{A}}^\mathbb{Z} \mid (G^t_{\delta_\mathcal{A}}(c))_0 \neq (G^t_{\delta_\mathcal{A}}(c_{p\leftarrow q}))_0\})$$

Remarque 6.1.1
- *Lorsqu'aucune confusion à propos de la mesure et de l'AC n'est possible, nous noterons simplement les coefficients de dépendance $\alpha_{t,p}$.*
- *Si r est le rayon de l'AC, $p > r \times t \Longrightarrow \alpha_{t,p} = 0$ parce que l'état de la cellule 0 après t étapes de calcul est indépendant de l'état de la cellule p de la configuration initiale.*

Définition 6.1.2 *La* propagation d'information *d'un AC $\mathcal{A}$ pour la mesure μ est la suite infinie de réels positifs*

$$\left(\sum_{p\in\mathbb{Z}} \alpha_{\mu,\mathcal{A},t,p}\right)_{t\in\mathbb{N}}.$$

Cette séquence est bien définie pour tout t grâce à la remarque précédente : les $(\alpha_{\mu,\mathcal{A},t,p})_{p\in\mathbb{Z}}$ sont presque tous nuls.



128 109 30

FIG. 6.1 – La règle 128 appartient à $[\mathcal{D} \stackrel{\mu}{\rightarrow}^{+} 0]$, 109 à $[\mathcal{D} \stackrel{\mu}{\rightarrow}^{+} a]$ et 30 à $[\mathcal{D} \stackrel{\mu}{\rightarrow}^{+} +\infty]$. Les diagrammes du bas représentent en niveau de gris la probabilité pour chaque cellule d'être modifiée par la modification de la cellule centrale.

Cette notion permet de définir la classe $[\mathcal{D} \stackrel{\mu}{\rightarrow}^{+} 0]$ des AC dont la propagation d'information tend vers zéro, la classe $[\mathcal{D} \stackrel{\mu}{\rightarrow}^{+} +\infty]$ des AC dont la propagation d'information n'est pas bornée (sa limite sup est $+\infty$) et la classe $[\mathcal{D} \stackrel{\mu}{\rightarrow}^{+} a]$ lorsque la limite sup de la propagation d'information est finie mais non nulle. La figure 6.1 montre un exemple dans chaque classe. Il est clair que ces trois classes définissent une partition de l'ensemble de AC pour une mesure donnée.

Notons que souvent, tout au long de cette thèse, nous avons représenté au dessous des diagrammes espace-temps, un diagramme indiquant combien de fois en moyenne une cellule du diagramme est modifiée lorsque la cellule centrale de la configuration initiale est changée. Ces diagrammes permettent d'avoir immédiatement une indication expérimentale de la classe (parmi celles que nous venons de définir) à laquelle appartient l'automate. D'autres informations peuvent être déduites de leur observation. Ainsi, lorsque ce diagramme représente des triangles de Pascal, alors la règle est additive. Parfois, on reconnait des triangles de Pascal mais grisés ou bruité, ce qui suggère que la règle (comme 18, cf section 2.5) est additive sur un sous-ensemble des configurations initiales. Enfin, ces diagrammes permettent d'avoir une idée de la direction des propagations d'informations au cours de l'évolution des règles.

Proposition 6.1.1 *Les automates additifs ont une propagation non bornée pour toute mesure invariante non triviale, ils sont dans la classe $[\mathcal{D} \stackrel{\mu}{\rightarrow}^{+} +\infty]$.*

Les deux preuves suivantes ont été faites par B. Allombert.

Lemme 6.1.1 *Soit p un nombre premier, soient a et n des entiers positifs et $k < p^n$, alors*

$$\binom{ap^n - 1}{k} = \pm 1 (mod\ p) \ .$$

Preuve 6.1.1 (du lemme 6.1.1)
Dans $\mathbb{F}_p[X] = \mathbb{Z}/p\mathbb{Z}[X]$, nous avons

$$\begin{array}{rcl} (X-1)^{ap^n-1} & = & \frac{(X-1)^{ap^n}}{X-1} \\ & = & \frac{(X^{p^n}-1)^a}{X-1} \\ & = & \frac{(X^{p^n}-1)(X^{p^n}-1)^{a-1}}{X-1} \\ \frac{X^{p^n}-1}{X-1} & = & \sum_{j=0}^{p^n-1} X^j \\ (X^{p^n}-1)^{a-1} & = & (-1)^{a-1} + eX^{p^n} + \dots \text{ où } e \in \mathbb{F}_p \end{array}$$

Donc le produit est égal à $\sum_{j=0}^{p^n-1}(-1)^{a-1}X^j + eX^{p^n} + \dots$ et les p^n premiers coefficients sont ± 1. Nous finissons la preuve en remarquant que les valeurs absolues des coefficients de $(X-1)^{ap^n}$ sont les coefficients binomiaux. □

Lemme 6.1.2 *Pour $N \geq 2$, pour tout M, il existe m tel que pour tout $0 < k < M$ nous avons*

$$\binom{m}{k} \in (\mathbb{Z}/N\mathbb{Z})^* \textit{ (i.e. il est inversible modulo } N\textit{)}$$

Preuve 6.1.2 (du lemme 6.1.2)
Pour tout p divisant N, p premier, choisissons e_p tel que $p^{e_p} \geq M$. Définissons

$$m = -1 + \prod_{p|N} p^{e_p}.$$

Nous utilisons le lemme 6.1.1 pour p, $n = e_p$ et $a = \prod_{q|N, q \neq p} q^{e_q}$. Donc pour tout $k < M \leq p^n$

$$\binom{m}{k} = \pm 1 (\text{mod } p)$$

Donc $\binom{m}{k}$ est inversible modulo tous les diviseurs premiers de N et donc modulo N. □

Preuve 6.1.3 (de la proposition 6.1.1)
C'est une conséquence directe du second lemme en utilisant la représentation dipolynomiale des AC additifs. Soit $(\mathbb{Z}/\mathbb{Z}_n{}^{\mathbb{Z}}, G_\delta)$ un AC additif. Par additivité, l'évolution depuis une configuration c lorsque nous remplaçons l'état

de la cellule 0 par l'état q est la superposition des évolutions en partant de c et en partant de $0^*_{0\leftarrow q}$ (i.e. la configuration telle que tous les états sont 0 excepté l'état de la cellule à l'origine qui est q). Formellement, $\forall t, G^t_\delta(c_{0\leftarrow q}) = G^t_\delta(c) + G^t_\delta(0^*_{0\leftarrow q})$ si $+$ est l'opération de groupe. La représentation dipolynomiale nous indique que $(G^t_\delta(0^*_{0\leftarrow q}))_i$ (l'état de la cellule i de la configuration après t étapes de calcul depuis la configuration nulle excepté la cellule 0 qui est dans l'état q) est q fois le coefficient de degré i de $(1+X)^t$ modulo n. D'après le second lemme, quel que soit $q \neq 0$, pour tout M, il existe m tel que, pour tout $0 < k < M$, $(G^m_\delta(0^*_{0\leftarrow q}))_k \neq 0$. Donc pour toute mesure non triviale, les AC additifs dont dans $[\mathcal{D} \overset{\mu}{\rightarrow}^+ +\infty]$. $\square$

6.2 Sensibilité aux conditions initiales μ-presque partout

Nous nous inspirons ici de la notion introduite par Formenti de sensibilité aux conditions initiales pour la pseudo-distance de Besicovitch. Comme nous l'avons vu (cf section 4.4.2), la pseudo-distance de Besicovitch a été introduite pour obtenir des définitions invariantes par décalage. La raison principale de l'introduction de la sensibilité aux conditions initiales μ-presque partout est l'étude de cas particuliers. D'une part, la règle 120 (cf figure 6.2) apparaît sensible aux conditions initiales, mais il semble difficile de démontrer qu'il n'existe pas de configuration particulière pour lesquelles elle serait équicontinue. D'autre part, la règle 210 ne semble pas sensible, mais elle n'est pas équicontinue en 0^*. L'idée est donc de pouvoir dire que 120 est μ-presque partout sensible, tandis que 210 est μ-presque jamais sensible.

Par ailleurs, l'introduction d'une mesure a également un intérêt théorique important. Nous avons en effet remarqué qu'il était impossible de définir la densité d'un sous-ensemble de $\mathbb{Z}$ et que nous avions donc choisi une densité supérieure pour définir une pseudo-métrique. La densité supérieure de l'ensemble A est définie par la formule suivante.

$$\limsup_{l\longrightarrow+\infty} \frac{\#([-l,l]\cap A)}{2l+1}.$$

D'autres choix étaient possibles, par exemple la topologie de Weyl qui approche la densité de façon différente.

$$\limsup_{l\longrightarrow+\infty} \max_{k\in\mathbb{N}} \frac{\#([k+1,k+l]\cap A)}{l}$$

Le fait que la pseudo-distance de Besicovitch ne soit pas un choix canonique devrait nous dissuader de l'utiliser. Mais il se trouve que l'introduction d'une mesure résout ce problème. En effet, dès lors que la mesure choisie est σ-ergodique, le théorème ergodique de Birkhoff nous indique que la densité d'une configuration est bien définie avec probabilité 1. Cela signifie que la sensibilité μ-presque partout pour Besicovitch est équivalente à la sensibilité μ-presque partout pour Weyl ou toute autre topologie basée sur une approximation de la densité. Cette définition est donc canonique, bien qu'un choix arbitraire ait été fait pour la poser.

Remarque 6.2.1 *Notons que cette propriété joue un rôle important dans l'extension des résultats au cas général des AC en dimension supérieure. En effet, l'extension de la pseudo-distance de Besicovitch à des grilles de dimensions supérieures $Q^{\mathbb{Z}^n}$ est :*

$$d(x,y) = \limsup_{l \longrightarrow +\infty} \frac{\#\{i \in B(0,l) \subset \mathbb{Z}^n | x_i \neq y_i\}}{\#B(0,l)}$$

où $B(0,l)$ est la balle centrée à l'origine de rayon 1 dans $\mathbb{Z}^n$ pour une distance choisie arbitrairement sur $\mathbb{Z}^n$, par exemple $\overline{d}_1((a_1,...,a_n),(b_1,...,b_n)) = |b_1 - a_1| + ... + |b_n - a_n|$ ou $\overline{d}_2((a_1,...,a_n),(a_1,...,a_n)) = ((b_1 - a_1)^2 + ... + (b_n - a_n)^2)^{1/2}$. Remarquons que si nous choisissons des distances équivalentes sur $\mathbb{Z}^n$, nous obtenons la même topologie sur $Q^{\mathbb{Z}^n}$. Cependant, des distances différentes peuvent conduire à des notions de sensibilité aux conditions initiales pour Besicovitch différentes, mais toutes les distances usuelles sur $\mathbb{Z}^n$ sont équivalentes pour la sensibilité presque partout parce que les pseudo-distances de Besicovitch qui leurs sont associées ne diffèrent que sur un ensemble de mesure nulle.

Cela constitue la seule difficulté pour étendre les résultats topologiques de cette thèse puisque les concepts de mesure et d'ergodicité sont correctement définis pour toutes les grilles de dimensions finies et que l'extension de la définition de propagation de défauts est directe.

Afin de définir la sensibilité aux conditions initiales presque partout, nous pourrions simplement remplacer « pour toutes les configurations c » par « pour μ-presque toute les configurations c » dans la définition de sensibilité 4.4.2. Nous allons donner une définition plus restrictive, sans d'ailleurs que nous sachions si cette définition est strictement plus restrictive ou finalement équivalente. L'intérêt de cette définition, c'est la possibilité de montrer plus facilement la non μ-sensibilité, tandis que les preuves de μ-sensibilité ne sont en pratique pas plus difficiles. En particulier, nous montrerons qu'une propagation d'information bornée implique la non sensibilité presque partout, ce qui, comme nous le verrons, est équivalent à la « sensibilité presque nulle part ».

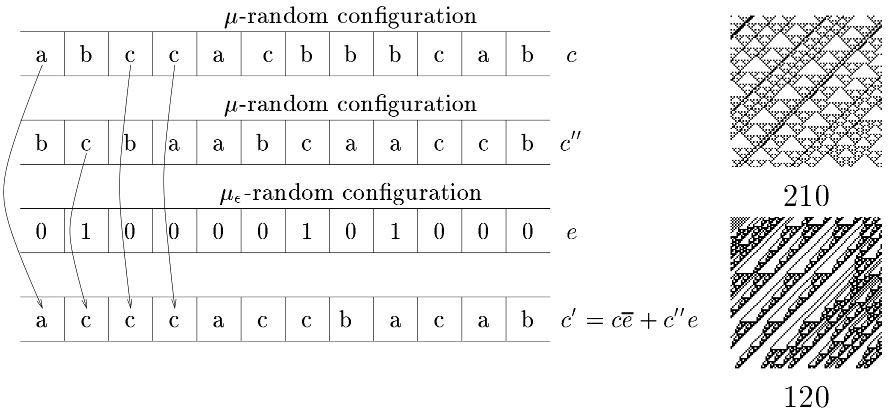


FIG. 6.2 – La configuration $c' = c\overline{e} + c''e$ est la configuration dont l'état d'une cellule donnée est égal à l'état de la même cellule dans c lorsque l'état de cette cellule dans e est 0 et à l'état de la même cellule dans c' sinon.

Définition 6.2.1 *Soit μ une mesure de Bernoulli. Un AC est μ-*presque partout sensible aux conditions initiales (pour la topologie de Besicovich) *s'il existe $M > 0$ tel que pour tout $\epsilon_0 > 0$, il existe $\epsilon < \epsilon_0$ tel que si c et c'' sont deux configurations μ-aléatoires, si e est une configuration μ_ϵ-aléatoire et si $c' = c\overline{e} + c''e$ est la configuration dont l'état à la position i est égal à c_i lorsque e_i est egal à 1 et à c''_i sinon (cf figure 6.2), alors avec probabilité 1 (pour $\mu \times \mu \times \mu_\epsilon$) il existe n tel que $d(G^n_{\delta_\mathcal{A}}(c), G^n_{\delta_\mathcal{A}}(c')) \geq M$.*

Observons que :

- Si $c' = c\overline{e} + c''e$ et que la répartition de c et c'' (resp. e) sont données par μ (resp. μ_ϵ), il est facile de vérifier que la loi de répartition de c' est donnée par μ ;
- Cette propriété implique qu'il existe M tel que pour μ-presque toutes les configurations c et pour tout $\epsilon > 0$, il existe c' et n avec $d(c, c') < \epsilon$ et $d(G^n_{\delta_\mathcal{A}}(c), G^n_{\delta_\mathcal{A}}(c')) \geq M$. Elle implique donc la μ-sensibilité, et même une « μ-sensibilité uniforme » puisqu'alors la constante M ne dépend pas de c ;
- Avec la topologie produit, le résultat précédent impliquerait la sensibilité aux conditions initiales. L'argument est que les mesures de Bernoulli sont de support total (i.e. les ensembles ouverts non vides ont une mesure non nulle), mais ce n'est pas le cas sur $Q^\mathbb{Z}/\sim$ avec la topologie de Besicovitch ;
- L'ensemble des triplets de configurations (c, c'', e) tels que $c' = c\overline{e} + c''e$ et qu'il existe n satisfaisant $d(G^n_{\delta_\mathcal{A}}(c), G^n_{\delta_\mathcal{A}}(c')) \geq M$ est clairement shift-invariant dans $(Q \times Q \times \{0, 1\})^\mathbb{Z}$. Comme les mesures de Bernoulli sont faiblement mélangeantes, $\mu \times \mu \times \mu_\epsilon$ est σ-ergodique, dont la mesure

de l'ensemble est soit 1 soit 0. Ainsi un AC est soit μ-presque partout sensible aux conditions initiales, soit « μ-presque jamais sensible aux conditions initiales » : pour tout η, il existe ϵ tel que si nous construisons c, c' comme précédemment, pour tout n, $d(G^n_{\delta_\mathcal{A}}(c), G^n_{\delta_\mathcal{A}}(c')) \leq \eta$ $\mu \times \mu \times \mu_\epsilon$-presque partout.

La sensibilité aux conditions initiales μ-presque partout a du sens. En effet certains automates ne le sont pas (clairement les règles 0, 128, les décalages, etc. ne le sont pas) et nous allons démontrer que les AC additifs le sont.

Proposition 6.2.1 *Les AC additifs sont μ-presque partout sensibles aux conditions initiales pour toute mesure de Bernoulli μ non triviale.*

Preuve 6.2.1 Soient c, c'' deux configurations μ-aléatoires, e une configuration μ_ϵ-aléatoire et $c' = c\overline{e} + c''e$. La i$^{\text{e}}$ cellule de $G^k_{\delta_\mathcal{A}}(c')$ a un état différent de celui de la cellule correspondante de $G^k_{\delta_\mathcal{A}}(c)$ si et seulement si $[G^k_{\delta_\mathcal{A}}(e(c''-c))]_i \neq 0$ parce que $c' = c\overline{e} + c''e = (c - ce) + c''e = c + e(c'' - c)$. Le nombre d'éléments inversibles de la ligne t du triangle de Pascal est non borné d'après le lemme 6.1.2. Comme $\mathbb{Z}/n\mathbb{Z}^*$ est fini, il existe $f \in \mathbb{Z}/n\mathbb{Z}^*$ tel que le nombre x_t d'éléments de la ligne t du triangle de Pascal égaux à f modulo n n'est pas borné. $[G^t_{\delta_\mathcal{A}}(e(c''-c))]_0 = r_t + fS_t$

$$\text{avec } S_t = \sum_{\left\{i \mid \binom{t}{i} = f (\text{mod } n)\right\}} e_i(c''_i - c_i)(\text{mod } n)$$

$$\text{et } r_t = \sum_{g \in \mathbb{Z}/n\mathbb{Z}, g \neq f} g \sum_{\left\{i \mid \binom{t}{i} = g (\text{mod } n)\right\}} e_i(c''_i - c_i).$$

Nous voulons à présent calculer la probabilité que la somme S_t soit égale à $g \in \mathbb{Z}/n\mathbb{N}$. Pour cela, nous pouvons considérer le processus de Markov sur $\mathbb{Z}/n\mathbb{Z}$ partant de l'état 0 et qui, à chaque étape, ajoute $h \in \mathbb{Z}/n\mathbb{Z}$ modulo n avec la probabilité p'_h définie par $p'_0 = \epsilon p_0 + 1 - \epsilon$ et $p'_h = \epsilon p_h$ si $h \neq 0$. Nous observons que la probabilité que S_t soit égal à $g \in \mathbb{Z}/n\mathbb{N}$ est exactement la probabilité qu'après x_t étapes le processus de Markov soit dans l'état g (une étape du processus correspond à un terme de la somme, les probabilités étant indépendantes parce que les mesures sont de Bernoulli). Remarquons que cette probabilité peut être obtenue en utilisant le développement de $(p'_0 + p'_1X + p'_2X^2 + \ldots + p'_{n-1}X^{n-1})^{x_t} = \sum_{i=0}^{(n-1)x_t} a_iX^i$: c'est la somme des coefficients de degrés k pour $k \in g + n\mathbb{Z}$, c'est-à-dire $\sum_{k=g \text{ mod } (n)} a_k$. Cette probabilité est donc donnée par la formule suivante :

$$\pi_g(t) = \frac{\sum_{k=0}^{n-1} \xi^{-kg}(p'_0 + \xi^k p'_1 + \ldots + (\xi^{n-1})^k p'_{n-1})^{x_t}}{n}$$

avec $\xi = e^{\frac{2i\pi}{n}}$. Tous les termes, excepté pour $k = 0$, tendent vers 0 lorsque $x_t \longrightarrow +\infty$ parce que $|p'_0 + \xi^k p'_1 + ... + (\xi^{n-1})^k p'_{n-1}| < 1$ et le premier terme ($k = 0$) est toujours 1 puisque $p'_0 + ... + p'_{n-1} = 1$. Donc la probabilité que la somme soit égale à $g \in \mathbb{Z}/n\mathbb{N}$ tend vers $1/n$ sur une sous-suite t_k telle que $x_{t_k} \longrightarrow +\infty$.

Par compacité, sur une sous-suite de t_k (que nous notons toujours t_k), la probabilité $\pi'_h(t_k)$ que $r_{t_k} = h \in \mathbb{Z}/n\mathbb{N}$ converge vers une valeur que nous notons π'_h. Remarquons à présent que pour tout t, S_t et r_t sont indépendants parce qu'ils dépendent de cellules différentes et que μ et μ_ϵ sont des mesures de Bernoulli. Ainsi la probabilité que $[G^{t_k}_{\delta_\mathcal{A}}(e(c'' - c))]_0 = i \in \mathbb{Z}/n\mathbb{Z}$ est

$$\sum_{h+fg=i \mod (n)} \pi'_h(t_k)\pi_g(t_k)$$

qui tend vers $1/n$ lorsque k tend vers $+\infty$. Pour tout $M < 1 - 1/n$ (notons que M ne dépend pas de ϵ) il existe t tel que l'état de la cellule 0 de $G^t_{\delta_\mathcal{A}}(c)$ et celui de la cellule 0 de $G^t_{\delta_\mathcal{A}}(c')$ soient différents avec une probabilité supérieure à M pour $\mu \times \mu \times \mu_\epsilon$.

Pour achever la preuve, il ne nous reste plus qu'à remarquer que, presque sûrement, la pseudo-distance de Besicovitch entre $G^t_{\delta_\mathcal{A}}(c)$ et $G^t_{\delta_\mathcal{A}}(c')$ est exactement la probabilité que $[G^t_{\delta_\mathcal{A}}(c)]_0$ et $[G^t_{\delta_\mathcal{A}}(c')]_0$ soient différents. C'est une conséquence de la σ-ergodicité de $\mu \times \mu \times \mu_\epsilon$ grâce au théorème ergodique de Birkhoff. Formellement, définissons

$$f_t \left| \begin{array}{ccl} Q^\mathbb{Z} \times Q^\mathbb{Z} \times \{0,1\}^\mathbb{Z} & \longrightarrow & \mathbb{R} \\ (c, c'', e) & \longmapsto & \begin{array}{l} 1 \text{ if } [G^t_{\delta_\mathcal{A}}(c)]_0 \neq [G^t_{\delta_\mathcal{A}}(c')]_0 \\ 0 \text{ sinon.} \end{array} \end{array} \right.$$

La probabilité que $[G^t_{\delta_\mathcal{A}}(c)]_0$ et $[G^t_{\delta_\mathcal{A}}(c')]_0$ soient différents est $\int f_t d(\mu\times\mu\times\mu_\epsilon)$, donc $\int f_t d(\mu \times \mu \times \mu_\epsilon) \geq M$.

Comme f_t est mesurable (elle est en fait continue), nous pouvons appliquer le théorème ergodique de Birkhoff avec $\mu \times \mu \times \mu_\epsilon$ et σ : si

$$S_n(f_t)(c, c'', e) = \frac{1}{n}\sum_{i=0}^{n-1} f_t \circ \sigma^i(c, c'', e),$$

pour $\mu \times \mu \times \mu_\epsilon$-presque tout (c, c'', e),

$$\lim_{n\to\infty} S_n(f_t)(c, c'', e) = \int f_t d(\mu \times \mu \times \mu_\epsilon) \geq M.$$

Notons que

$$\begin{aligned}S_n(f_t)(c,c'',e) &= \frac{1}{n}\sum_{i=0}^{n-1} f_t\circ\sigma^i(c,c'',e)\\ &= \frac{1}{n}\#\{i\in[0,n-1]|[G^t_{\delta_{\mathcal{A}}}(c)]_i\neq[G^t_{\delta_{\mathcal{A}}}(c')]_i\}.\end{aligned}$$

Afin d'obtenir une somme sur $[-n,n]$, nous appliquons Birkhoff à nouveau avec μ et σ^{-1}, nous définissons $S'_{n-1}(f_t)$ et ajoutons :

$$\lim_{n\to+\infty}((n+1)S_{n+1}(f_t)(c,c'',e)+nS'_n(f_t)(\sigma^{-1}(c,c'',e)))/(2n+1)$$

$$= d(G^t_{\delta_{\mathcal{A}}}(c),G^t_{\delta_{\mathcal{A}}}(c')).$$

Par σ-invariance de $\mu\times\mu\times\mu_\epsilon$ et G_δ, $\int f_t\circ\sigma d(\mu\times\mu\times\mu_\epsilon)=\int f_t d(\mu\times\mu\times\mu_\epsilon)$. Donc, pour $\mu\times\mu\times\mu_\epsilon$-presque tout (c,c'',e), $d(G^t_{\delta_{\mathcal{A}}}(c),G^t_{\delta_{\mathcal{A}}}(c'))=\int f_t d(\mu\times\mu\times\mu_\epsilon)\geq M$. $\square$

6.3 Comparaisons entre l'entropie apparente, la propagation d'information et la μ-sensibilité aux conditions initiales

Nous allons à présent nous intéresser aux relations entre l'entropie apparente et la propagation d'information d'un AC.

Proposition 6.3.1 *Soit μ une mesure de Bernoulli de coefficients $(p_q)_{q\in Q}$. Si la μ-propagation d'information de $\mathcal{A}$ tend vers 0 alors il existe un ensemble de configurations uniformes qui est un $B\mu$-attracteur.*

Preuve 6.3.1 Soit Q^∞ le sous-ensemble d'états qui apparaissent après un nombre arbitraire d'itérations de la fonction globale de transition. Soient c et c' deux configurations μ-aléatoires. Fixons $t>0$. Nous considérons l'état de la cellule 0 après t étapes de calcul. Imaginons que nous remplaçons les états des cellules d'indice $-t$ à t de c par les états de c' un par un : soit $c^0=c$ et $c^{i+1}=c^i{}_{-t-1+i\leftarrow c'(-t-1+i)}$. Clairement, $G^t_\delta(c^{2t+1})_0=G^t_\delta(c')_0$ et la probabilité que $G^t_\delta(c^i)_0\neq G^t_\delta(c^{i+1})_0$ est $\alpha_{t,-t-1-i}$ par définition. Donc, le nombre moyen de modifications de l'état de la cellule 0 après t itérations pendant ce processus est $\sum_p\alpha_{t,p}\longrightarrow_t 0$; donc la probabilité que la cellule reste toujours dans le même état (i.e. $\forall k, 0\leq k\leq 2t+1, [G^t_\delta(c^k)]_0=[G^t_\delta(c)]_0$) est $1-\sum\alpha_{t,p}\longrightarrow 1$. Ainsi pour tout t, il existe q tel que $\mu(\{c|[G^t_\delta(c)]_0=q\})\geq 1-\sum\alpha_{t,p}$. Nous déduisons que pour tout ϵ il existe M tel que pour $t \geq M$, il y a un état $q \in Q^\infty$ (q dépend de t) tel que $\mu(\{c|[G_\delta^t(c)]_0 = q\}) \geq 1-\epsilon$. D'après le théorème ergodique de Birkhoff, avec probabilité 1, $\mu(\{c|[G_\delta^t(c)]_0 \neq q\}) = d(G_\delta^t(c), q^*)$ parce que μ est σ-ergodique (nous utilisons le même argument que dans la preuve 6.2.1 avec la fonction $f_t : c \longmapsto 1$ si $[G_\delta^t(c)]_0 \neq q$ et 0 sinon). Donc $d(G_\delta^t(c), \{q^* \mid q \in Q^\infty\}) \leq \epsilon$ pour tout $t \geq M$. Par conséquence $\{q^* \mid q \in Q^\infty\}$ est un Bμ-attracteur. $\square$

Corollaire 6.3.1 *Si la μ-propagation d'information de $\mathcal{A}$ tend vers 0 alors son entropie apparente tend vers zéro.*

Réciproquement, il y a des AC ayant un Bμ-attracteur qui ne sont pas dans $[\mathcal{D} \overset{\mu}{\to}^+ 0]$: la règle $\mathcal{S}$ (cf figure 5.3) a un Bμ-attracteur et est dans $[\mathcal{D} \overset{\mu}{\to}^+ a]$. De plus, nous allons décrire plus tard (après la proposition 6.3.2) un AC qui semble expérimentalement être dans $[\mathcal{D} \overset{\mu}{\to}^+ +\infty]$ mais tend vers une configuration uniforme.

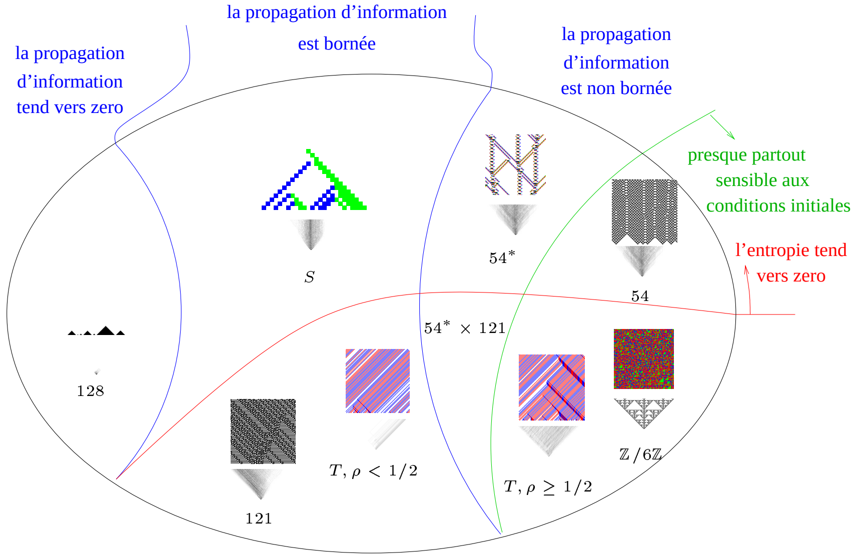
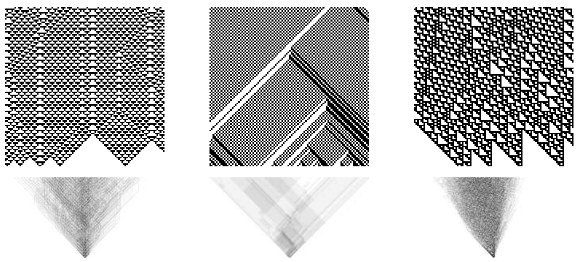


FIG. 6.3 – Relations entre l'entropie apparente, la propagation d'information et la sensibilité aux conditions initiales μ-presque partout.

Intéressons-nous à présent aux relations entre propagation d'information et sensibilité μ-presque partout. L'idée est de prendre une configuration μ-aléatoire c. Nous construisons ensuite c' à partir de c : pour toutes les cellules, l'état de c' est égal à l'état de c avec une probabilité $1-\eta$, et est égal à q avec une probabilité ηp_q. Avec probabilité 1, la pseudo-distance de Besicovitch $d(c, c')$ entre c et c' est inférieure à η. Cela nous permet de démontrer la proposition suivante.

Proposition 6.3.2 *Soit μ une mesure σ-faiblement mélangeante. Être μ-presque partout sensible aux conditions initiales implique une propagation d'information non bornée.*

Preuve 6.3.2 Lorsque la propagation d'information dans un AC $\mathcal{A}$ est bornée, démontrons que pour tout η, il existe ϵ tel que pour tout n $d(G_\delta^n(c), G_\delta^n(c')) \leq \eta$ où c et c'' sont deux configurations μ-aléatoires, e est une configuration μ_ϵ-aléatoires et $c' = c\overline{e} + c''e$. Soit M une borne sur la propagation d'information dans $\mathcal{A}$. Imaginons, comme dans la preuve précédente, que nous remplaçons les états de c entre les indices $-t$ et t par les états de c' un par un : soit $c^0 = c$ et $c^{i+1} = c^i{}_{-t-1+i \leftarrow c'(-t-1+i)}$. Clairement $G_\delta^t(c^{2t+1})_0 = G_\delta^t(c')_0$. Cette fois, nous remplaçons un état par q avec une probabilité de ϵp_q et conservons l'état initial sinon (avec une probabilité de $(1-\epsilon)$), donc, la probabilité que $G_\delta^t(c^i)_0 \neq G_\delta^t(c^{i+1})_0$ est $\epsilon\alpha_{t,-t-1-i}$. Au cours de cette expérience le nombre moyen de changements de l'état de la cellule 0 après t itérations est $\epsilon \sum \alpha_{t,p} \leq \epsilon M$ qui tend vers 0 quand $\epsilon \longrightarrow 0$. La probabilité que l'état ne change jamais est au moins $1 - \epsilon \sum \alpha_{t,p} \geq 1 - \epsilon M$, donc $\mu \times \mu \times \mu_\epsilon(\{(c, c'', e) | G_\delta^t(c)_0 = G_\delta^t(c\overline{e} + c''e)_0\}) \geq 1 - \epsilon M$. À nouveau (cf la preuve 6.2.1), nous utilisons la σ-ergodicité de $\mu \times \mu \times \mu_\epsilon$ (car μ et μ_ϵ sont σ-faiblement mélangeantes) et le théorème ergodique de Birkhoff pour démontrer qu'avec une probabilité 1, $d(G_\delta^t(c), G_\delta^t(c')) = \mu \times \mu \times \mu_\epsilon(\{(c, c'', e) | G_\delta^t(c)_0 \neq G_\delta^t(c')_0\}) \leq \epsilon M$. Ainsi pour un ϵ suffisamment petit, la pseudo-distance entre $G_\delta^t(c)$ et $G_\delta^t(c')$ est inférieure à η avec probabilité 1. □



54 184 ($\rho = 1/2$) 110

FIG. 6.4 – Expérimentalement, 54 et 184 pour $\mu_{1/2}$ sont dans $[\mathcal{D} \overset{\mu}{\to}{}^{+} +\infty]$ et ont une entropie apparente nulle, de nouveaux résultats de N. Ollinger sur la règle 110 nous laissent penser que cette règle a le même comportement.

Il est expérimentalement facile à observer mais semble difficile de prouver que la règle 184 est $\mu_{1/2}$-presque partout sensible aux conditions initiales. Pour comprendre ce qui se passe, considérons un marcheur aléatoire sur $\mathbb{Z}$

partant de 0 et lisant la configuration initiale vers la droite (resp. vers la gauche). Lorsqu'il lit un 1, il se déplace vers la droite et lorsqu'il lit un 0, il va vers la gauche. Si nous changeons l'état de la cellule 0 de la configuration initiale en 1 (resp. en 0), cela provoque au moins une modification sur chacune des configurations successives de l'évolution de l'automate. La position de l'état modifié se déplace vers la droite lorsque le marcheur aléatoire est sur $\mathbb{Z}_+^*$, vers la gauche lorsqu'il est sur $\mathbb{Z}_-^*$ et reste à la même position lorsqu'il est à la position 0. Parfois, en raison de cette modification, toute une région contenant des états 0 et 1 en quinconce est décalée et donc tous les états de cette région sont changés. Si nous regardons maintenant l'évolution globale de l'AC, on peut démontrer qu'elle génère des régions de plus en plus grandes composées de 0 et 1 alternés (elles sont de la forme $(01)^*$ ou $(10)^*$). Évidemment, en partant de deux configurations génériques, il n'y a aucune raison que les régions se trouvent aux mêmes positions, du coup, la pseudo-distance entre les configurations tend asymptotiquement vers 1/2.

Il semble que la règle 54 soit également presque partout sensible tandis que son entropie apparente est nulle. En fait, la disparition des particules g_e est un phénomène irréversible (nous avons démontré l'irréversibilité de leur disparition dans le chapitre 2) qui va se produire de plus en plus rarement au fur et à mesure que le nombre de particules g_e va diminuer, mais le nombre de particules g_e tend vers 0 comme le confirment les expériences de [60, 9, 26]. Comme toutes les interactions entre des particules w et g_0 font disparaître une des particules, nous pouvons supposer que le sous-shift $\{(0001)^*, (1110)^*\}$ est un $\mathrm{D}\mu_\rho$-attracteur pour $0 < \rho < 1$.

Les propositions précédentes soulèvent une question naturelle : les AC avec une propagation d'information non bornée sont-ils μ-presque partout sensibles aux conditions initiales ?

Il semble que non. Définissons l'automate 54^* qui simule la règle 54 mais en remplaçant le fond périodique par un fond uniforme. Un tel AC peut être défini formellement grâce au filtre de Hanson et Crutchfield [26] qui est un AC, mais seulement sur un sous-ensemble des configurations. Nous pouvons considérer ici une extension de cet AC avec une mesure telle que les configurations sont correctes avec probabilité 1. Expérimentalement, le nombre de particules décroît comme $t^{-1/2}$ et tend donc vers 0. Ainsi, cet AC, pour les mesures décrites précédemment, tend presque sûrement vers une configuration uniforme. En raison de la faible décroissance du nombre de particules, cet AC semble avoir une propagation d'information non bornée : avec une probabilité non nulle, une seule modification crée ou supprime une ou plusieurs particules. À chaque interaction de ces particules avec d'autres particules, elles sont dupliquées. La probabilité de telles collisions est liée à la densité de particules, donc le nombre moyen de particules ajoutées ou supprimées par la modification de la configuration initiale doit évoluer comme $\int t^{-1/2} = 2\sqrt{t}$ qui tend vers $+\infty$ lorsque $t \longrightarrow +\infty$. Ainsi cet AC semble avoir une propagation d'information non bornée (ce que les expériences confirment). De plus, la proposition suivante suggère que cet AC n'est pas presque partout sensible aux conditions initiales.

Notons une certaine similitude entre l'évolution de 54^* et le refroidissement d'une tasse de café brûlant dont nous avons parlé dans l'introduction. En effet ces systèmes sont prévisibles à long terme, mais pas à court terme. Cependant, les mécanismes à l'origine de ces comportements sont très différents.

Proposition 6.3.3 *Soit $\mathcal{A}$ un AC qui tend μ-presque toujours vers une configuration uniforme c_0, alors $\mathcal{A}$ n'est pas μ-presque partout sensible aux conditions initiales.*

Preuve 6.3.3 Pour toute constante α, soit $\epsilon' < \alpha/2$, il existe M tel que pour presque toutes configurations c, $n > M$ implique $d(c_0, G_\delta^n(c)) \leq \epsilon'$. Cela permet de borner la distance entre les orbites après M itérations. À présent, pour la borner pour les M premières configurations, nous utilisons le fait qu'une seule modification génère au plus $2t+1$ modifications après t itérations. Soit $\epsilon \leq \frac{\alpha}{2Mr+1}$. Si la répartition de $e \in \{0,1\}^{\mathbb{Z}}$ est donnée par μ_ϵ et si $c' = c\overline{e} + c''e$, c et c'' étant deux configurations de $\mathcal{A}$, alors μ_ϵ-presque partout $d(c, c') \leq \epsilon$. Donc presque partout, pour $n \leq M$, $d(G_\delta^n(c), G_\delta^n(c')) \leq \epsilon(2Mr+1) \leq \alpha$ et pour $n > M$, $d(G_\delta^n(c), G_\delta^n(c')) \leq d(c_0, G_\delta^n(c)) + d(c_0, G_\delta^n(c')) \leq 2\epsilon' \leq \alpha$. Donc $\mathcal{A}$ n'est pas presque partout sensible aux conditions initiales. $\square$

En fait, on montre qu'un AC ayant un Bμ-attracteur si μ est une mesure de Bernoulli n'est pas μ-presque partout sensibles aux conditions initiales. Le problème est que si l'existence d'un Bμ-attracteur assure celle d'un Bμ-attracteur composé de configurations uniformes, ce dernier n'est pas forcément un singleton. Cependant, à un temps donné (suffisamment grand), toutes les configurations sont proches d'une seule et même configuration de l'attracteur en raison de la σ-ergodicité de μ.

6.4 Retour sur la classification de Wolfram

Les AC μ-presque partout sensibles dont l'entropie apparente est non nulle parmi les ECA apparaissent expérimentalement comme étant exactement ceux qui ont une entropie apparente constante dont nous avons parlé dans le chapitre précédent et qui sont représentés sur la figure 5.5. Clairement, ces observations ne se généralisent pas du tout aux AC ayant plus d'états. En effet, il suffit de superposer 54^* et l'identité pour obtenir un AC μ-presque partout sensible dont l'entropie apparente est non nulle (mais l'entropie apparente n'est pas constante), cet AC a donc les mêmes propriétés que la superposition de la règle 30 avec l'identité. Bien entendu, le premier AC qui est le produit de deux AC intuitivement non chaotiques n'est pas chaotique, tandis que le second, produit d'un AC chaotique avec un non chaotique doit être considéré comme chaotique. En fait, nous constatons que la sensibilité μ-presque partout est une condition nécessaire mais non suffisante pour être chaotique. Ainsi, d'une certaine manière, nous avons démontré l'observation de Wolfram à propos des AC chaotiques selon laquelle « the value of a particular site depends on an ever-increasing number of initial site values ». Nous avons également constaté que réciproquement, cette condition n'est même pas suffisante pour impliquer la sensibilité presque partout.

Nous ne nous interesserons pas à la classe 4 de Wolfram. En effet, elle semble plus motivée par des considération algorithmiques que dynamiques. Or nous nous intéressons ici seulement aux propriétés dynamiques et celles-ci semblent totalement indépendantes des propriétés algorithmiques. En effet, il existe des AC universels pour Turing qui tendent presque partout vers une configuration uniforme[1], donc par superposition, dans toutes les classes dynamiques presque partout que nous pouvons définir. En fait, un AC est généralement universel sur un ensemble de mesure nulle. Cela suggère que deux types de classifications coexistent, les classifications dynamiques auxquelles nous nous sommes principalement intéressé et les classifications algorithmiques, dont les plus intéressantes sont sans doute celle de Rapaport que nous avons évoquée et la variante proposée pas Ollinger [51] qui utilise une notion de groupage plus générale (rectangle est non plus seulement carré). Cette dernière est très semblable à la classification de Rapaport (tous les résultats que nous mentionnons semblent demeurer exacts) mais elle admet un maximum global lié à la notion d'universalité intrinsèque.

Nous avons déjà mis en évidence la contradiction entre la définition de la classe 1 et la présence de 128 dans cette classe (qui ne tend pas vers une configuration uniforme en un temps fini). Nous avons en fait plusieurs candidats pour donner une définition de cette classe. Une propagation d'information bornée, la présence d'un Bμ-attracteur composé de configurations uniformes ou une entropie apparente nulle. La première définition implique la seconde

[1] Il suffit par exemple de prendre l'automate universel décrit par Smith [57] et d'y ajouter un nouvel état q_0 qui se propage comme une tache d'encre : une cellule passe (et reste) dans cet état dès qu'une de ses voisine y est. Comme dans le cas de 128, dès que la probabilité que q_0 apparaisse sur la configuration initiale est non nulle, $\{q_0^*\}$ est un Bμ-attracteur.

dans le cas de mesures de Bernoulli, quant à la seconde, elle implique la troisième dans tous les cas. La définition de Wolfram serait plutôt l'existence d'un Bμ-attracteur composé d'une seule configuration uniforme. Cependant, cette définition n'est pas stable par groupage, certains automates ne vérifiant pas cette propriété ont des automates groupés la vérifiant, ce qui est quand même un problème. En fait, puisque Wolfram fait des remarque à la fois sur la propagation bornée des défauts et sur la destruction de l'information initiale, le choix de l'une ou l'autre est plus philosophique qu'autre chose. Le plus important est de comprendre en quoi ces trois notions sont différentes, de connaître les relations qui les lient et, qu'à elles trois, elles permettent de caractériser assez finement le comportement des règles qui tendent vers des configurations homogènes ou périodiques.

Deux principales questions demeurent, obtenir une définition (ou plus probablement un ensemble de définitions) qui soit plus proche de la notion de chaos telle que la conçoivent les physiciens, et trouver un moyen de séparer (ou comprendre pourquoi cela est impossible) les AC « presque chaotiques » comme 109 ou 118 des AC tels que l'identité ou 210. Probablement, le fait d'avoir une propagation d'information bornée pour toute mesure, mais non uniformément bornée pour toutes les mesures est une condition nécessaire pour être « presque chaotique », mais elle n'est a priori pas suffisante puisque vérifiée par 210.

Conclusion

Nous nous sommes intéressé à la notion de chaos sur les automates cellulaires. Nous avons vu que cette notion était assez difficile à formaliser et en particulier, que le comportement d'un automate cellulaire pouvait être très différent selon la méthode utilisée pour choisir une configuration aléatoire. Ainsi il semble plus pertinent de proposer des notions de chaos pour une mesure donnée. De ce fait, nous avons introduit trois nouvelles définitions. L'entropie apparente permet de mieux comprendre l'impression de complexité ou d'aléatoire que nous avons en regardant certains diagrammes espaces-temps. De plus, nous pouvons à présent formellement quantifier cette impression. Les notions de propagation d'information et de sensibilité aux conditions initiales presque partout sont des conditions nécessaires, mais non suffisantes de chaos. En fait, ces notions permettent de mettre en évidence la difficulté de classer les automates en fonction de leur comportement. En effet, les différentes remarques faites par Wolfram sur les AC de ses classes sont souvent partiellement vérifiées, mais prêtent également à confusion. Par exemple, comment devons nous classer un AC tel que la règle 54^* dont la propagation d'information est non bornée, comme les AC chaotiques, mais dont l'entropie apparente est nulle, ce qui signifie qu'elle efface toutes les informations de sa configuration initiale, et tend vers une configuration uniforme, comme les AC les plus simples de la classe 1 ?

Ces derniers automates, ceux qui ont des comportements compliqués bien qu'ils détruisent l'information de leur configuration initiale, ont des comportements qui ressemblent à des interactions entre des particules. Pour étudier ces AC dans le cas unidimensionnel, nous avons introduit une nouvelle méthode que nous avons utilisée dans le cas de la règle 54. Outre les résultats algorithmiques sur cette règle (sa hauteur infinie dans l'ordre de Rapaport) et des indications quant à son comportement dynamique, ce travail permet de mieux comprendre comment des comportements compliqués peuvent apparaître dans des AC ayant très peu d'états. En effet, l'utilisation d'un fond ou background permet d'avoir une grande variété de particules avec seulement deux états alors qu'il est très difficile de simuler des particules avec un AC sans attribuer un état (voire beaucoup plus) à chaque particule.

La première question soulevée par notre dernier chapitre est de comprendre quelles sont les relations entre la notion naïve de sensibilité presque partout que nous avons écartée, la notion naïve uniforme, plus restrictive, et la définition de μ-sensibilité que nous avons choisie. Par ailleurs, nous nous sommes restreints aux mesures de Bernoulli dans cette thèse, les mesures σ-invariantes les plus simples. Cependant, il semble que les résultats que nous avons obtenus, ou tout au moins une grande partie d'entre eux, soient vrais pour des mesures plus générales. Si le mélange faible est nécessaire à la notion de sensibilité presque partout (nous n'utilisons pour les autres définitions que l'ergodicité), il faut probablement ajouter une condition de localité. Ainsi, il semble naturel d'étudier le cas des mesures de Markov (les dépendances sont finies) ou des mesures dont les dépendances décroissent exponentiellement avec la distance. Comme nous l'avons vu, nous avons obtenu des définitions intéressantes, mais qui ne prétendent pas être des définitions du chaos, seulement des conditions nécessaires. D'ailleurs, les définitions proposées en systèmes dynamiques impliquent la sensibilité mais sont toutes plus restrictives. Cependant, nous obtenons des résultats naturels qui correspondent bien à notre intuition sur le chaos. Ces résultats suggèrent une nouvelle approche de ce problème, il s'agit d'adapter, lorsque cela est possible, les notions définies et étudiées en systèmes dynamiques à la topologie de Besicovitch (malgré la non compacité de l'espace) pour une mesure donnée. De plus, il ne faut pas espérer proposer une seule définition du chaos, mais c'est d'un ensemble de définitions et des relations qui les relient que progressera notre connaissance sur les comportements chaotiques des automates cellulaires.

À propos de la règle 54, nous ne savons toujours pas expliquer pourquoi le fond apparaît lorsque cet automate évolue en partant d'une configuration aléatoire, ce qui pourrait permettre de caractériser les règles ayant ce type de comportement et peut-être de démontrer des résultats sur leur comportement dynamique presque partout. La question de l'universalité pour le calcul est toujours ouverte. En fait, bien que nous ayons des outils pour démontrer des résultats négatifs sur cette règle et que son comportement ne semble pas assez compliqué pour simuler une machine de Turing universelle, démontrer que 54 n'est pas universelle nécessite notamment une définition formelle de l'universalité.

Bibliographie

[1] J. Auslander and J. Yorke. Interval maps, factors of maps and chaos. *Tohoku Math. J.*, 32:177–188, 1980.

[2] F. Bagnoli, R. Rechtman, and S. Ruffo. Lyapunov exponents for cellular automata. In M. López de Haro and C. Varea, editors, *Lectures on Thermodynamics and Statistical Mechanics*. World Scientific, 1994.

[3] J. T. Baldwin and S. Shelah. On the classifiability of cellular automata. *Theoretical Computer Science*, 230:117–129, 2000.

[4] J. Banks, J. Brooks, G. Cairns, G. Davis, and P. Stacey. On Devaney's definition of chaos. *Amer. Math. Monthly*, 99:332–334, 1992.

[5] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn. Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems: A method for computing all of them. *Meccanica*, pages 9– and 21–, 1980.

[6] F. Blanchard, E. Formenti, and P. Kůrka. Cellular automata in the Cantor, Besicovitch and Weyl topological spaces. *Complex Systems*, 11(2), 1999.

[7] F. Blanchard and A. Maass. Dynamical behaviour of Coven's cellular automata. *Theoretical Computer Science*, 163:291–302, 1996.

[8] N. Boccara and H. Fukś. Number-conserving cellular automaton rules. *Fundamenta Informaticae*, 2000, to appear.

[9] N. Boccara, J. Nasser, and M. Roger. Particle-like structures and their interactions in spatio-temporal patterns generated by one-dimensional deterministic cellular-automata rules. *Physical Review A*, 44:866–875, 1991.

[10] G. Cattaneo, E. Formenti, L. Margara, and G. Mauri. On the dynamical behavior of chaotic cellular automata. *Theoretical Computer Science*, 217:31–51, 1999.

[11] G. Cattaneo, E. Formenti, L. Margara, and G. Mauri. Topological chaos and CA. In M. Delorme and J. Mazoyer, editors, *Cellular Automata, a Parallel Model*, pages 213–259. Springer Verlag, 1999.

[12] G. Cattaneo, E. Formenti, L. Margara, and J. Mazoyer. Shift invariant distance on $S^{\mathbb{Z}}$ with non trivial topology. In *Proceedings of MFCS' 97*. Springer Verlag, 1997.

[13] K. Čulik and S. Yu. Undecidability of CA classification schemes. *Complex System*, 2:177–190, 1988.

[14] R. Devaney. *Introduction to Chaotic Dynamical Systems (Second Edition)*. Addison-Wesley, 1989.

[15] J.-C. Dubacq, B. Durand, and E. Formenti. Kolmogorov complexity and cellular automata classification. *Theoretical Computer Science*, 2000. À paraître.

[16] B. Durand. *Cellular automata: reversibility and complexity.* PhD thesis, École Normale Supérieure de Lyon, 1994.

[17] A. Maass F. Blanchard, B. Host. Topological complexity. *Ergod. Th. & Dynam. Sys.*, 20:641–662, 2000.

[18] P. A. Ferrari, Alejandro Maass, and Servet Martinez. Cesàro mean distribution of group automata starting from Markov measures. *preprint*, 2000.

[19] M. Garzon. *Models of massive parallelism.* ETACS. Springer, 1995.

[20] R. H. Gilman. Classes of linear automata. *Ergod. Th. & Dynam. Sys.*, 7:105–118, 1987.

[21] E. Glasner and B. Weiss. Sensitive dependence on initial conditions. *Nonlinearity*, 6:1067–1075, 1993.

[22] J. Gleick. *Chaos.* The Viking Press, New York, 1987.

[23] J. Gleick. *La théorie du chaos, vers une nouvelle science.* Flammarion, 1991.

[24] H. A. Gutovitz and C. G. Langton. Mean field theory of the edge of chaos. In *Proceedings of the Fourth European Conference on Artificial Life*, 1995.

[25] H. Hammer, A. Romashchenko, A. Shen, and N. Vereshchagin. Inequalities for Shannon entropy and Kolmogorov complexities. In *Twelfth Annual IEEE Conference on Computational Complexity*, pages 13–23, Ulm, Germany, June 1997.

[26] J. E. Hanson and J. P. Crutchfield. Computational mechanics of cellular automata: an example. *Physica D*, 103:169–189, 1997.

[27] G. A. Hedlund. Endomorphism and automorphism of the shift dynamical system. *Mathematical System Theory*, 3:320–375, 1969.

[28] S. Ishii. Measure theoretic approach to the classification of cellular automata. *Discrete Applied Mathematics*, 39:125–136, 1992.

[29] P. Kůrka. Languages, equicontinuity and attractors in cellular automata. *Ergod. Th. & Dynam. Syst.*, 217:417–433, 1997.

[30] P. Kůrka and A. Maass. Stability of subshifts in cellular automata. *preprint*, 1999.

[31] C. G. Langton. Computation at the edge of chaos. *Physica D*, 22:120–149, 1986.

[32] M. Li and P. Vitányi. *An Introduction to Kolmogorov Complexity and its Applications*. Graduate texts in computer science. Springer Verlag, 1997.

[33] T. Y. Li and J. A. Yorke. Period three implies chaos. *Amer. Math. Monthly*, 82(no. 10):985–992, 1975.

[34] K. Lindgren. Correlation and random information in cellular automata. *Complex System*, 1:529–543, 1987.

[35] E. Lorenz. Determinist nonperiodic flow. *Journal of the Athmospheric Sciences*, 20:130–141, 1963.

[36] E. Lorenz. The mecanics of vacillation. *Journal of the Athmospheric Sciences*, 20:448–464, 1963.

[37] E. Lorenz. The problem of deducing the climate from the governing equations. *Tellus*, 16:1–11, 1964.

[38] B. Martin. Méthodes algébriques de construction d'automates cellulaires. Mémoire de DEA Informatique, École Normale Supérieure de Lyon, 1998.

[39] B. Martin. Apparent entropy of cellular automata. *Complex Systems*, 12(2), 2000.

[40] B. Martin. Damage spreading and μ-sensitivity on cellular automata. In *Proceedings of IFIP TCS2000*, pages 226–242, Sendai, Japan, August 2000. Submitted to *Ergod. Th. & Dyn. Sys.*

[41] B. Martin. A group interpretation of particles generated by one dimensional cellular automaton, 54 wolfram's rule. *Int. Journ. of Mod. Phys. C*, 11(1):101–123, 2000.

[42] B. Martin. Simulations between cellular automata on Cayley graphs. *Fondamenta Informaticae*, to appear, 2000.

[43] B. Martin and I. Rapaport. Abelian cellular automata according to the preorder $(CA, \leq)$, preprint, 2000.

[44] O. Martin, O. Odlyzko, and S. Wolfram. Algebraic properties of cellular automata. *Communication in Mathematical Physics*, 93:219–258, 1984.

[45] P. Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–619, 1966.

[46] R. M. May. Simple mathematical models with very complicated dynamics. *Nature*, 261:645–647, 1974.

[47] J. Mazoyer and I. Rapaport. Inducing an order on cellular automata by a grouping operation. In *STACS'98*, volume 1373, pages 116–127. Lecture Notes in Computer Science, 1998.

[48] H. V. McIntosh. Rule 110 as it relates to the presence of gliders, 2000. http://delta.cs.cinvestav.mx/~mcintosh/newweb/rule110/rule110.html.

[49] J. Miles. Strange attractor in fluid dynamics. *Advances in Applied Mechanics*, 24:189–214, 1984.

[50] J. Von Neumann. *Theory of Self-reproducing Automata.* University of Illinois Press, Chicago, 1966.

[51] N. Ollinger. Automates cellulaires, (re)groupage. Mémoire de DEA Informatique, École Normale Supérieure de Lyon, 2000.

[52] E. Ott. *Chaos in Dynamical Systems*. Cambridge University Press, 1993.

[53] K. Petersen. *Ergodic theory*. Cambridge University Press, 1983.

[54] N. Reimen. *Une contribution à l'étude des automates cellulaires*. PhD thesis, Université Paris VII, 1993.

[55] D. Ruelle and F. Takens. On the nature of turbulance. *Communication in Mathematical Intelligencer*, 2:167–192, 1971.

[56] A. Smith III. Cellular automata theory. Technical Report 2, Stanford University, 1960.

[57] A. Smith III. A simple computation-universal cellular space. *Journal of Assoc. Comp. Mach.*, 18(3):233–253, 1971.

[58] P. Walters. *An Introduction to Ergodic Theory*. Springer-Verlag, 1982.

[59] S. Wilson. Cellular automata can generate fractals. *Discrete applied mathematics*, 8:91–99, 1984.

[60] S. Wolfram. Universality and complexity in cellular automata. *Physica D*, 10:1–35, 1984.