THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *23/06/2023* **par :**

Léo Gayral

Complexité et robustesse des pavages avec perturbations aléatoires

JURY

Mireille Bousquet-Mélou	Directrice de Recherche CNRS	Examinatrice
Cédric Boutillier	Maître de Conférences	Examinateur
Jean-René Chazottes	Directeur de Recherche CNRS	Président
Pascal Maillard	Professeur des Universités	Examinateur
Irène Marcovici	Maîtresse de Conférences	Examinatrice
Tom Meyerovitch	Associate Professor	Rapporteur
Andrei Romashchenko	Chargé de Recherche CNRS	Rapporteur
Mathieu Sablik	Professeur des Universités	Directeur de thèse

École doctorale et spécialité :
MITT : Domaine Mathématiques : Mathématiques fondamentales
Unité de Recherche :
Institut de mathématiques de Toulouse – IMT (UMR 5219)
Directeur de Thèse :
Mathieu Sablik
Rapporteurs :
Jean-René Chazottes, *Tom Meyerovitch* et *Andrei Romashchenko*

Complexité et robustesse des pavages avec perturbations aléatoires

Léo Gayral

Remerciements

Merci à mon encadrant, Mathieu Sablik, pour sa patience pendant ces si longues et si courtes années. Merci également à Siamak Taati et Anthony Quas pour le chaleureux accueil qu'ils m'ont réservé lors de mes séjours dans leurs universités respectives.

Merci aux rapporteurs de ce manuscrit pour leur relecture attentive, et au reste du jury pour le temps qu'elles et ils me dédient dans leur emploi du temps sans aucun doute *outrement* bien rempli.

Enfin, merci à toustes les adelphes, les camarades, les anonymes et pseudonymes, ainsi que les artistes avec qui j'entretiens des relations parasociales, vous êtes ce qui me donne la force d'avancer jour après jour.

Préliminaires

Ce chapitre est une traduction de l'introduction faite au Chapitre 1.

Le sujet principal de ma thèse, comme annoncé dans le titre, est la robustesse des pavages (généralement choisis au hasard, mais pas nécessairement uniformément distribués) aux perturbations aléatoires. Ici, les notions de robustesse et de perturbations sont volontairement vaguement définies, et prendront des significations formelles différentes selon le contexte. Le dernier ingrédient clé, la *complexité*, peut être compris de différentes manières. Tout d'abord, de façon informelle, il fait référence aux structures hiérarchiques apériodiques complexes qu'on retrouve dans la plupart des pavages que j'étudierai. Deuxièmement, il peut être compris comme une complexité computationnelle, soit des ensembles de pavages aléatoires que j'introduirai, soit de la notion même de stabilité en tant que problème de décision.

La saveur générale de mon travail dans ce manuscrit est donc un entrelacs de physique statistique, de complexité computationnelle et de théorie ergodique, qui repose sur des éléments de dynamique symbolique. Je commencerai par donner un aperçu général de la bibliographie existante sur des sujets liés, puis j'exposerai la structure de mon travail dans les chapitres suivants.

I. Contexte

I.i. Structures complexes dans les pavages apériodiques

Dans ce contexte, un pavage est constitué de tuiles carrées alignées sur la grille $\mathbb{Z}^d$, avec des « règles locales » qui déterminent quelles tuiles sont autorisées les unes à côté des autres, à l'instar de pièces de puzzle. Formellement, lorsque le nombre de tuiles différentes (qui formeront un *alphabet*) et de telles règles (les *motifs interdits*) est fini, l'ensemble des pavages correspondants est appelé un sous-décalage de type fini, un SFT (pour *Subshift of Finite Type* en anglais).

En particulier, si chaque tuile utilise une couleur sur chaque bord, et que l'on exige de deux tuiles voisines qu'elles aient la même couleur sur leur bord en commun, on obtient le formalisme des tuiles de Wang, illustré dans la Figure 1.1. Ce formalisme a été introduit par Wang en 1961 [Wan61], en conjecturant que tout jeu de tuiles qui pave le plan doit admettre une configuration périodique. En 1966 [Ber66], son doctorant Berger a prouvé que la conjecture était fausse en exhibant un jeu de tuiles apériodique utilisant un très grand nombre de tuiles, et a également prouvé que la question de la pavabilité du plan est algorithmiquement indécidable. Ce résultat a été rapidement amélioré par Robinson [Rob71], qui a diminué le nombre total de tuiles de Wang dans l'alphabet de plusieurs milliers à quelques dizaines (de 20426 à 56 pour être précis). Ce processus d'optimisation s'est poursuivi jusqu'à la dernière décennie, où Jeandel et Rao [JR21] ont prouvé que tout jeu de tuiles de Wang qui couvre le plan en utilisant dix tuiles ou moins doit contenir une configuration périodique, et ont présenté un jeu de tuiles apériodiques avec onze tuiles (illustré en Figure 1.1). La plupart de ces jeux de tuiles apériodiques reposent sur une structure autosimilaire hiérarchique (y compris le pavage de Jeandel-Rao [Lab21], ou la construction de Labbé à 19 tuiles de Wang [Lab19]), à une exception notable près : le pavage de Kari-Culik [Kar96] est prouvé apériodique, et simule une transformation continue sur les nombres réels induisant des trajectoires apériodiques, mais avec une entropie positive donc sans structure auto-similaire [DGG17]. L'un des points les plus remarquables de ces constructions est que les structures apériodiques complexes peuvent émerger de règles locales très simples.

Une problématique liée est celle des pavages géométriques, où chaque tuile a une forme (généralement polygonale), et où un pavage admissible du plan est entièrement constitué de tuiles dont les intérieurs ne s'intersectent pas (c'est d'ailleurs ainsi que se comportent les pièces de puzzle en réalité). Notamment, les bords des tuiles de Wang peuvent être transformés en formes géométriques qui encodent les règles locales entre les tuiles voisines, ce cadre généralise donc le précédent. Dans certaines situations, les formes des tuiles les forceront à s'aligner sur une grille périodique (comme $\mathbb{Z}^2$ pour les dominos, ou le réseau triangulaire pour les dimères), alors que dans d'autres cas, il n'y a pas de réseau sous-jacent (comme pour le pavage de Penrose). Dans ce contexte, la question de l'existence d'une monotuile apériodique a récemment reçu une réponse positive [Smi+23] en utilisant de nouvelles idées élégantes, prouvant une fois de plus que des structures complexes peuvent émerger d'une unique forme simple.

Même en utilisant une grille carrée, comme nous allons le voir, de nombreux comportements hautement complexes peuvent se produire, en particulier lorsque des perturbations aléatoires sont introduites dans le modèle. En fonction du chapitre, je considérerai deux sources principales de perturbations aléatoires : les perturbations indépendantes localisées, plus proches de la philosophie de la théorie de l'information, et les mesures de Gibbs, qui correspondent à une approche de physique statistique. Dans les deux cas, cependant, ces perturbations doivent être comprises comme une faible probabilité d'ignorer une règle locale, d'avoir des tuiles mal assorties. Par conséquent, la robustesse fait référence au fait qu'un pavage perturbé aléatoire est proche d'une mesure sur les pavages sans erreurs (dont le SFT est le support), par rapport à une certaine propriété ou dans une certaine topologie. Dans un cas, nous verrons qu'une question de robustesse utilisant des perturbations localisées est indécidable, qu'un algorithme ne peut pas dire si un jeu de tuiles induit un système stable. Dans l'autre cas, en utilisant des perturbations thermodynamiques, nous verrons une classe de jeux de tuiles induisant des systèmes chaotiques, où la complexité réside dans la structure de l'ensemble des mesures limites.

I.ii. Physique statistique, quasicristaux et chaos

L'une des principales motivations pour l'étude de ces structures apériodiques a été la découverte par les chimistes Shechtman *et al.* en 1984 [She+84] des *quasi-cristaux*, des matériaux hautement structurés dont les propriétés sont similaires à celles des cristaux, mais avec des preuves empiriques d'une symétrie d'ordre dix incompatible avec les réseaux périodiques connus. Ces matériaux ont d'abord été obtenus en laboratoire, mais ont depuis été découverts à l'état naturel dans certains sites météoritiques [Bin+15].

Les liens entre la physique statistique et la dynamique symbolique étaient déjà connus à l'époque, les tuiles pouvant représenter des arrangements d'atomes, avec les motifs interdits modélisant les contraintes sur la manière dont ces atomes peuvent s'assembler (en utilisant des interactions énergétiques à portée finie entre les particules [LS84]). Cependant, à l'époque, il n'y avait pas de résultats concernant les structures apériodiques permettant d'expliquer théoriquement comment les quasi-cristaux peuvent se former, notamment en présence de bruit et d'imperfections.

En général, ces structures (périodiques ou apériodiques) émergent lors du refroidissement des matériaux. L'exemple historique le plus notable est le modèle (Lenz-)Ising [Nis05], avec une transition de phase d'une phase désordonnée à haute température à une phase magnétisée à basse température, établie dans le cas bidimensionnel par Peierls [Pei36], qui correspond à la transition de la phase paramagnétique à la phase ferromagnétique dans les matériaux réels. En ce qui concerne la robustesse des structures apériodiques, Miękisz a utilisé une variante du pavage de Robinson pour prouver la présence d'une phase ordonnée dans un modèle bidimensionnel avec des interactions à portée finie, avec des comportements périodiques à des échelles de plus en plus grandes dans la structure autosimilaire à mesure que la température décroît [Mię90, Mię97, Mię98]. À la même époque, il a contribué à un un modèle tridimensionnel avec des interactions à portée infinie qui, à des températures basses mais positives, impose la suite apériodique de Thue-Morse le long d'une direction [EMZ98].

Parfois, la complexité ne réside pas seulement dans les structures asymptotiques, mais aussi dans la façon dont le système converge vers elles lorsque la température diminue. Plus précisément, la *dépendance chaotique en la température* fait référence au phénomène de non-convergence des mesures de Gibbs, aux variations volatiles de l'état macroscopique du système lorsque la température s'approche d'une certaine valeur. On peut par exemple imaginer une solution chimique dont la couleur alternerait entre le bleu et le rouge lorsque sa température s'approche du point de congélation, sans jamais se stabiliser. La possibilité d'un tel comportement a été mathématiquement démontrée dans un certain nombre de modèles de pavages aléatoires.

Historiquement, le premier exemple de modèle avec une dépendance chaotique en la température au voisinage de 0 a été découvert par van Enter et Ruszel [ER07]. Leur modèle consistait en des interactions à portée finie, mais basées sur un système de *spins continus* (à valeurs sur le cercle unité et pas seulement dans un alphabet fini) en utilisant un potentiel avec une paire de puits infiniment imbriqués l'un dans l'autre. Ils ont montré que, lorsque la température décroît vers 0, les mesures de Gibbs subissent une suite infinie de transitions de phase, alternant entre des états ferromagnétiques et antiferromagnétiques.

Le second exemple, qui est la principale inspiration du Chapitre 5, est dû à Chazottes et Hochman [CH10]. Il s'agit d'un modèle *unidimensionnel* avec des spins discrets (*i.e.* l'alphabet binaire $\{0, 1\}$) et à *portée infinie* mais *décroissance exponentielle* (*i.e.* un potentiel Lipschitzien) pour lequel les mesures de Gibbs invariantes ne convergent pas à basse température. Avec un argument de dynamique symbolique multidimensionnelle [Hoc09], ils ont pu « simuler » leur modèle unidimensionnel à portée infinie dans un SFT tridimensionnel, obtenant ainsi un modèle avec un alphabet *fini* et des interactions à *portée finie* pour lequel les mesures de Gibbs invariantes ne convergent pas lorsque la température décroît vers 0. Cette construction a ensuite été adaptée aux modèles bidimensionnels [CS20, Bar+22] en utilisant de nouvelles méthodes de simulation [DRS10, AS13]. D'autres résultats similaires de dépendance chaotique ou sensible à la température ont été obtenus ces dernières années [CR15, BGT18, CR19].

En utilisant précautionneusement les mêmes outils thermodynamiques que Chazottes et Hochman, ainsi que des résultats de simulation plus traditionnels, mes co-auteurs et moi-même avons pu obtenir une caractérisation des ensembles limites pour les systèmes *uniformément* chaotiques (pour lesquels le comportement ne dépend pas de la trajectoire) en fonction de leur complexité computationnelle [GST23].

I.iii. Complexité computationnelle des problèmes de pavages

Les machines de Turing [Tur36] sont un cadre de travail naturel pour étudier formellement des questions de calculabilité. Certains problèmes peuvent être résolus par un ordinateur en temps fini, auquel cas l'approche naturelle consiste à étudier la complexité temporelle de la tâche. Cependant, certains problèmes ne peuvent tout simplement pas être résolus en temps fini, auquel cas on les qualifie d'*indécidables*. La *hiérarchie arithmétique* est apparue comme un moyen de classifier la complexité relative de ces tâches indécidables. Notamment, le premier échelon de cette hiérarchie correspond au problème de l'arrêt, qui consiste à savoir si un calcul donné se terminera en temps fini ou non.

Les interactions entre les systèmes dynamiques et des questions de calculabilité ne sont pas nouvelles, et datent pratiquement des premiers balbutiements de la dynamique symbolique. En effet, l'une des principales caractéristiques des pavages de Berger et Robinson était leur capacité à simuler des machines de Turing, établissant ainsi l'indécidabilité du problème *domino*, qui consiste à savoir si un jeu de tuiles donné peut paver le plan, en l'assimilant au problème de l'arrêt [Ber66, Rob71].

Ces interactions ont été largement étudiées au cours de la dernière décennie. En particulier, les valeurs que peuvent prendre certains invariants de conjugaison ont été caractérisées, comme les entropies (dimensionnelles) possibles [HM10, Mey11], les périodes possibles [JV15a], les ensembles μ-limite [Boy+15], certaines classes de SFT [Wes17], les ensembles de mesures limite des automates cellulaires itérés sur une mesure initiale [HS18], les ensembles limite génériques d'automates cellulaires [ENT23]... Une autre façon de mettre en évidence la complexité des pavages consiste à situer la complexité de problèmes indécidables apparentés dans la hiérarchie arithmétique (ou analytique). Dans la hiérarchie arithmétique, le problème du domino est Π_1-complet [Rob71], le problème de conjugaison est Σ_1-complet et le problème de factorisation est Σ_3-complet [JV15b]... En ce qui concerne la hiérarchie analytique, décider si un pavage a une entropie topologique complètement positive ou non est Π^1_1-complet [Wes22], en dimension $d \geq 4$ le problème du domino apériodique is Π^1_1-complete [CH22]... Pour obtenir ces résultats, les preuves impliquent toujours l'encodage et la simulation de machines de Turing dans des structures hiérarchiques autosimilaires (et apériodiques). Une direction apparentée notable a ainsi été la possibilité de simuler des systèmes *complexes* unidimensionnels en utilisant des règles plus simples en dimensions supérieures [Hoc09, DRS12, AS13].

Par conséquent, la quête de comportements robustes est naturelle dans le contexte des pavages considérés comme des modèles de calcul, où l'on voudrait idéalement que les calculs survivent en présence d'une certaine quantité de bruit. De tels systèmes robustes ont déjà été étudiés dans le cas des automates cellulaires [Gác01], des machines de Turing [AC05] ou même d'autres pavages *simulants* (*i.e.* capables de simuler des modèles de calcul) [DRS12]. Dans d'autres cas, comme pour les simulations via des transformations calculables sur les nombres réels [BGR12], les mesures limites ne sont pas calculables, mais l'introduction d'une quantité quelconque de bruit provoque un effondrement de la complexité, auquel cas les mesures limites du système perturbé deviennent calculables. En suivant une philosophie similaire, les jeux de tuiles simulants étudiés au Chapitre 4 seront « stabilisés » à des échelles de plus en plus grandes de leur structure autosimilaire à mesure que la quantité de bruit diminue [GS23a], fournissant ainsi des calculs fiables à des horizons temporels de plus en plus lointains. Ce résultat a ensuite permis de prouver que la notion correspondante de stabilité est indécidable [GS23b].

Quelques résultats existent sur les interactions entre la physique statistique et la calculabilité, notamment à propos de la calculabilité de la fonction de pression sur certaines classes de sous-shifts unidimensionnels [Spa08, Bur+22], ou de certains invariants thermodynamiques comme l'entropie résiduelle [BW20], mais cette interface représente encore un domaine peu exploré dans l'ensemble.

II. Structure du manuscrit

Les prochains chapitres constituent une introduction au cadre existant et au folklore sur lesquels mon travail est basé. Dans le Chapitre 2, j'introduirai des idées de base sur la dynamique symbolique, la théorie ergodique, et j'étudierai plusieurs variantes du pavage de Robinson. Dans le Chapitre 3, en partant des machines de Turing, je détaillerai certains aspects clés de la calculabilité tels que les pavages simulants, la hiérarchie arithmétique et l'analyse calculable. Le formalisme thermodynamique sera introduit au sein du Chapitre 5, lorsque le besoin s'en fera sentir.

Pendant ma thèse, j'ai exploré deux directions principales, en utilisant deux types de perturbations aléatoires et deux notions de robustesse. Chaque axe fera l'objet d'une attention particulière dans un chapitre dédié.

Dans le Chapitre 4, je considèrerai principalement des bruits indépendants, avec une fréquence d'erreurs ε, et je comparerai des pavages (aléatoires) dans la topologie de Besicovitch, qui quantifie la fréquence de leurs différences. Cette topologie est forte, donc dans ce contexte, la robustesse doit être comprise globalement, on cherche à savoir si oui ou non le système converge lorsque la fréquence des erreurs tend vers 0. D'abord, en suivant les traces de mon premier article publié [GS23a], je prouverai qu'une variante du pavage de Robinson est stable. Ensuite, en suivant les traces d'un article exploratoire [Gay21] et de mon deuxième article publié [GS23b], j'utiliserai ce pavage simulant comme fondation pour prouver que cette notion de stabilité est indécidable, située entre Π_2 et Π_4 dans la hiérarchie arithmétique.

Le Chapitre 5 mettra le projecteur sur le formalisme thermodynamique pour les perturbations aléatoires, et sur la topologie faible-$*$ pour comparer les mesures. En raison de sa nature plus faible, cette topologie n'est pas adaptée à l'étude de la même notion de stabilité (le cas échéant, tous les modèles seraient stables), c'est pourquoi la notion de robustesse étudiée dans ce chapitre est celle de la chaoticité des trajectoires de mesures perturbées, lorsque la température (et donc la fréquence des erreurs) tend vers 0. En suivant les traces de l'article correspondant [GST23], j'établirai d'abord une borne supérieure Π_2 sur la complexité de l'ensemble d'accumulation (contenant toutes les valeurs d'adhérence à température 0 le long de toutes les trajectoires de mesures de Gibbs) dans le cas de modèles uniformes (lorsque l'ensemble d'accumulation est le même quelle que soit la trajectoire) avec des interactions calculables et à portée finie. Ensuite, à l'aide d'un jeu de tuiles simulant soigneusement conçu, je prouverai que tout ensemble connexe Π_2-calculable de mesures de probabilité peut être réalisé comme un ensemble d'accumulation (à homéomorphisme affine calculable près), en utilisant des interactions bidimensionnelles à portée finie.

Enfin, dans le Chapitre 6, je parlerai de certains points laissés en suspens dans les chapitres précédents et de perspectives ouvertes qui n'auraient pas leur place ailleurs.

Table of Contents

CHAPTER 1

Introduction

The main topic of my PhD, as announced in the title, is the robustness of tilings (usually chosen at random, but not necessarily uniformly distributed) *to* random perturbations. Here, both the notions of robustness and of perturbations are purposely loosely defined, and will take different formal meanings depending on the context. The last key ingredient, *complexity*, can be understood in various ways. First, informally, it refers to the complex aperiodic hierarchical structures embedded in most of the tilings I will study. Second, it can be here understood as the *computational* complexity, either of the sets of random tilings I will introduce or of the notion of stability as a decision problem.

The general flavour of my work in this manuscript is thus a mix of statistical physics, computational complexity and ergodic theory, thrown onto a base layer of symbolic dynamics. I will first give a broad overview of the literature on related topics, and then discuss the structure of my work in the following chapters.

1.1 Context

1.1.1 Complex Structures in Aperiodic Tilings

In this context, a tiling is made of square tiles aligned on a grid $\mathbb{Z}^d$, with “local rules” that determine which tiles are allowed next to each other, like puzzle pieces. Formally, when the number of different tiles (that will form an *alphabet*) and of such rules (*i.e. forbidden patterns*) is finite, the set of corresponding tilings is called a Subshift of Finite Type, an SFT.

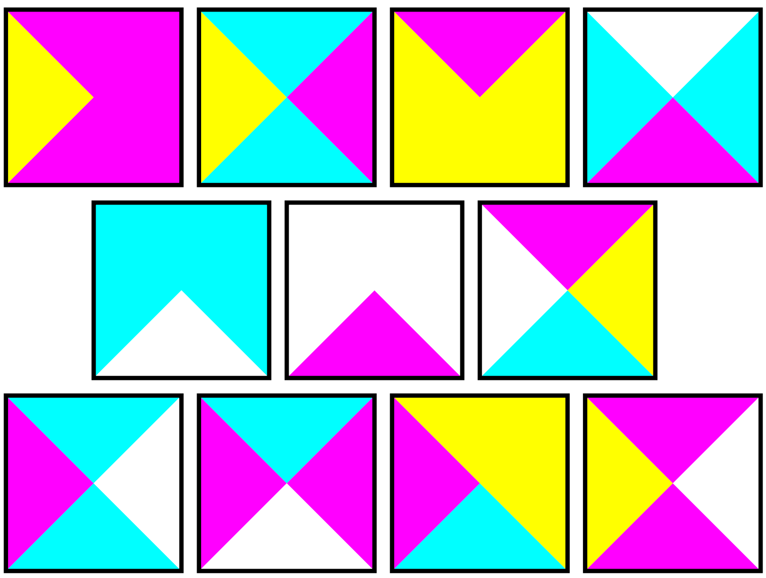

Figure 1.1: Jeandel-Rao set of eleven Wang tiles inducing only aperiodic tilings, using four colours for the interfaces.

In particular, if each tile uses one colour on each side, and we want neighbouring tiles to use the same colour on their common interface, then we obtain the formalism of Wang tiles, illustrated in Figure 1.1. This formalism was introduced by Wang in 1961 [Wan61], while conjecturing that any tileset that tiles the plane must admit a periodic configuration. In 1966 [Ber66], his PhD student Berger proved the conjecture wrong by exhibiting an aperiodic tileset using a very large number of tiles, and also proved that this question of tileability of the plane cannot be decided by an algorithm. This result was quickly improved by Robinson [Rob71], who brought down the total number of Wang tiles from several thousands to dozens (from 20426 to 56 to be precise). This optimisation process was settled during the last decade, by Jeandel and Rao [JR21], who proved that any Wang tileset that tiles the plane using ten tiles or less must contain a periodic configuration, and exhibited an aperiodic tileset with eleven tiles, shown in Figure 1.1. Most such aperiodic tilesets rely on a hierarchical self-similar structure (including the Jeandel-Rao tiling [Lab21], and Labbé's construction using 19 Wang tiles [Lab19]), with one notable exception: the Kari-Culik tiling [Kar96] is proven to be aperiodic, and simulates a continuous transformation on real numbers inducing aperiodic trajectories, but with positive entropy so without a self-similar structure [DGG17]. One of the most notable parts of these constructions is that the highly *complex* aperiodic structures can emerge from really simple local rules.

A related framework is that of geometric tilings, where each prototile has a (usually polygonal) shape, and a valid tiling of the plane is entirely covered by prototiles with non-overlapping interiors (which is incidentally how real-life puzzle pieces behave). Notably, the borders of Wang tiles can be turned into geometric shapes that enforce the local rules between neighbouring tiles, so this framework generalises the previous one. In some situations, the shapes of the prototiles will force them to align on a periodic grid (like $\mathbb{Z}^2$ for dominos, or the triangle lattice for dimers), while in other cases there is no underlying lattice (such as for the Penrose tiling). In this context, the question of existence of an aperiodic monotile was recently positively settled [Smi+23] using some new elegant ideas, proving once again that complex structures can arise from even one single simple shape.

Even when using a square grid, as we will see, many highly complex behaviours can occur, in particular when random perturbations are thrown into the mix. Depending on the chapter, I will consider two main sources of random perturbations: localised independent ones, closer to the philosophy of information theory, and Gibbs measures, that fit a statistical physics approach. In both cases, though, these perturbations are to be understood as a low probability of breaking a local rule, of having mismatched tiles. Consequently, robustness refers to the fact that a random perturbed tiling is close to a measure on tilings without mistakes (*i.e.* supported by the SFT), with respect to some property or in some topology. In one case, we will see that a question of robustness using localised perturbations is undecidable, that an algorithm cannot tell if a tileset induces a stable system. In the other case, using thermodynamic perturbations, we will exhibit a class of tilesets inducing chaotic systems, where the complexity lies in the structure of the set of limit measures.

1.1.2 Statistical Physics, Quasicrystals and Chaos

One major motivation for the study of such aperiodic structures was the discovery by the chemists Shechtman *et al.* in 1984 [She+84] of *quasicrystals*, highly structured materials with properties similar to those of crystals, but with empirical evidence for a complex tenfold symmetry incompatible with known periodic lattices. Such materials were initially obtained in lab conditions, but have since then been found naturally occurring in some meteoritic sites [Bin+15].

Links between statistical physics and symbolic dynamics were already known at the time, with the tiles representing atom clusters, and forbidden patterns modelling constraints on the way these atoms can fit together (*e.g.* using finite range energetic interactions between them [LS84]). However, at the time, there were no results related to aperiodic structures to better explain how quasicrystals may form, notably in the presence of noise and imperfections.

Usually, such structures (either periodic or aperiodic) emerge when cooling down materials. The most notable example historically is the (Lenz-)Ising model [Nis05], with a phase transition from the disordered high-temperature phase to the magnetised low-temperature phase established in the two-dimensional case by Peierls [Pei36], which corresponds to the transition from the paramagnetic to the ferromagnetic phase in real-life materials. Regarding the robustness of aperiodic structures, Miękisz used a variant of the Robinson tiling to prove the presence of an ordered phase in a two-dimensional model with finite-range interactions, with periodic behaviours at increasingly larger scales in the self-similar structure as the temperature decreases [Mię90, Mię97, Mię98]. Around the same time, he contributed to a three-dimensional model with infinite-range interactions which, at low but positive temperature, enforces the aperiodic Thue-Morse sequence along one direction [EMZ98].

In some situations, complexity lies not just in the limit structures, but also in the way the system converges to them as the temperature decreases. More precisely, *chaotic dependence on temperature* refers to the phenomenon of divergence of Gibbs measures, of volatile variations in the macroscopic state of the system as the temperature approaches a certain value. One may for instance picture a chemical solution whose colour keeps alternating between blue and red as its temperature nears a freezing point. The possibility of such a behaviour has been mathematically demonstrated in a number of lattice models.

The first example of a model with chaotic temperature dependence near zero temperature was found by van Enter and Ruszel [ER07]. Their model consisted of finite-range interactions, but on a *continuous spins* system (taking values from the unit circle and not just a finite alphabet) using a potential with a pair of infinitely nested wells. They showed that, as the temperature goes to zero, the Gibbs measures undergo an infinite sequence of phase transitions, alternating between ferromagnetic and antiferromagnetic states.

The second example, which is the main inspiration for Chapter 5, is due to Chazottes and Hochman [CH10]. They constructed a *one-dimensional* model with discrete spins (in the binary alphabet $\{\mathtt{0},\mathtt{1}\}$) and *infinite-range* but *exponentially decaying* interactions (*i.e.* a Lipschitz-continuous potential) for which the shift-invariant Gibbs measures don't converge at low temperatures. Using a technique from multidimensional symbolic dynamics [Hoc09], they were able to "simulate" their long-range one-dimensional model with a *three-dimensional* model, thus obtaining a model with a *finite* alphabet and *finite-range* interactions for which the invariant Gibbs measures do not converge. This construction was then adapted to *two-dimensional* models [CS20, Bar+22] using new simulation methods [DRS10, AS13]. Other similar results of chaotic or sensitive dependence on temperature were given in recent years [CR15, BGT18, CR19].

By carefully using the same thermodynamic tools as Chazottes and Hochman, along with more traditional simulation results, my co-authors and I were able to obtain a characterisation of the limit sets of *uniformly* chaotic systems (for which the behaviour doesn't depend on the trajectory) depending on their computational complexity [GST23].

1.1.3 Computational Complexity of Tiling Problems

Turing machines [Tur36] are in many ways the default framework for formal computability questions. Some problems *can* be solved by a computer in finite time, in which case the natural approach is to study the time complexity of the task. However, some problems simply *cannot* be solved in finite time, which we call *undecidable*. The *arithmetical hierarchy* emerged as a way to classify the relative complexity of such undecidable tasks. Notably, the first ladder of this hierarchy corresponds to the *halting* problem, of whether a given computation will end in finite time or not.

The interactions between dynamical systems and computability are not new, and were practically baked into symbolic dynamics from the start. Indeed, one of the main features of Berger and Robinson's hierarchical tilesets was their ability to simulate Turing machines, thus establishing the undecidability of the *domino* problem, *i.e.* whether a given tileset tiles the plane, by equating it with the halting problem [Ber66, Rob71].

These links have been widely studied in the last decade. Notably, the values that can be taken by some conjugacy invariants have been characterised, such as the possible (dimension) entropies [HM10, Mey11], the possible periods [JV15a], the μ-limit sets [Boy+15], some classes of SFTs [Wes17], the sets of limit measures of cellular automata iterated on an initial measure [HS18], the generic limit sets of cellular automata [ENT23], etc. Another way to highlight the complexity of tilings has been to locate the complexity of related undecidable problems in the arithmetical (or analytical) hierarchy. Regarding the arithmetical hierarchy, the domino problem is Π_1-complete [Rob71], the conjugacy problem is Σ_1-complete and the factorisation problem is Σ_3-complete [JV15b], etc. Regarding the analytical hierarchy, deciding whether a tiling has a completely positive topological entropy or not is Π_1^1-complete [Wes22], in dimension $d \geq 4$ the aperiodic domino problem is Π_1^1-complete [CH22], etc. To obtain all these results, the proofs always involve the embedding of Turing machines into *complex* hierarchical self-similar (and aperiodic) structures. A related direction of significant interest has thus been the possibility of simulating one-dimensional *complex* systems using simpler rules in higher dimensions [Hoc09, DRS12, AS13].

Hence, the possibility of robust behaviours is natural in the context of tilings seen as a model of computation, where we ideally want the computations to survive in the presence of some amount of noise. Such robust systems were already studied in the case of cellular automata [Gác01], Turing machines [AC05] or even other simulating tilings [DRS12]. In other cases, such as simulations through computable transformations on real numbers [BGR12], the limit measures are non-computable, but the introduction of any amount of noise causes a collapse of the complexity and the limit measures become computable. Following a similar philosophy, the simulating tilesets studied in Chapter 4 will be "stabilised" at increasingly large scales of their self-similar structure as the amount of noise decreases [GS23a], thus providing reliable computations at increasingly large time horizons. This in turn allowed to prove that the corresponding notion of stability is undecidable [GS23b].

A few existing results concern the interactions between statistical physics and computability, such as the computability of the pressure function on some classes of one-dimensional subshifts [Spa08, Bur+22], or of some thermodynamic invariants like the residual entropy [BW20], but this intersection still represents a budding field overall.

1.2 Structure of the Dissertation

The following two chapters are an introduction to the existing framework and folklore on which my work is based. In Chapter 2, I will introduce basic ideas about symbolic dynamics, ergodic theory, and discuss several variants of the Robinson tiling. In Chapter 3, starting from Turing machines, I will detail some key aspects of computability such as simulating tilesets, the arithmetical hierarchy and computable analysis. The thermodynamic formalism will be introduced within Chapter 5 as the need arises.

During my PhD, I have explored two main directions, using different kinds of random perturbations and different notions of robustness. Each will be the focal point of its own chapter.

In Chapter 4, I will mostly use independent noises with a frequency ε of mistakes, and compare (random) tilings using the Besicovitch topology, that quantifies the frequency of their differences. This topology is strong, so in this context, robustness is to be understood globally, we are trying to know whether the system converges at all as the frequency of mistakes goes to 0. First, following the tracks of my first published article [GS23a], I prove that a variant of the Robinson tiling is stable. Then, following the tracks of an exploratory paper [Gay21] and my second published article [GS23b], I use this simulating tiling as the bedrock to prove that this notion of stability is undecidable, located between Π_2 and Π_4 in the arithmetical hierarchy.

In Chapter 5, I move onto the thermodynamic formalism for the random perturbations, and the weak-$*$ topology to compare measures. Due to its weaker nature, this topology is ill-fitted to study the same notion of stability (everything would be stable here), so the notion of robustness studied in this chapter is that of chaoticity (or stability) of the cooling trajectories, as the temperature (and thus the frequency of mistakes) goes to 0. Following the tracks of the corresponding article [GST23], I first prove a Π_2 upper bound on the complexity of the set of ground states (*i.e.* the set of all accumulation points of all cooling trajectories) in the case of uniform models (when the accumulation set is the same regardless of the trajectory) with computable finite-range interactions. Then, using a carefully crafted simulating tileset, I prove that any connected Π_2 set of probability measures can be realised as a set of ground states (up to a computable affine homeomorphism), using two-dimensional finite-range interactions.

Lastly, in Chapter 6, I discuss some loose ends and open perspectives from the previous chapters that wouldn't fit well anywhere else.

References

[AC05] Eugène Asarin and Pieter Collins. "Noisy Turing Machines". In: *ICALP*. Vol. 3580. Lecture Notes in Computer Science. Springer, 2005, pp. 1031–1042. 10.1007/11523468_83.

[AS13] Nathalie Aubrun and Mathieu Sablik. "Simulation of Effective Subshifts by Two-Dimensional Subshifts of Finite Type". In: *Acta Applicandae Mathematicae* 126 (2013), pp. 35–63. 10.1007/s10440-013-9808-5.

[Bar+22] Sebastián Barbieri, Rodrigo Bissacot, Gregório Dalle Vedove, and Philippe Thieullen. *Chaos in Bidimensional Models With Short-Range*. 2022. 10.48550/arXiv.2208.10346.

[Ber66] Robert Berger. *The Undecidability of the Domino Problem*. Memoirs of the American Mathematical Society 66. AMS, 1966. 10.1090/memo/0066.

[BGR12] Mark Braverman, Alexander Grigo, and Cristóbal Rojas. "Noise vs Computational Intractability in Dynamics". In: *ITCS*. ACM, 2012, pp. 128–141. 10.1145/2090236.2090247.

[BGT18] Rodrigo Bissacot, Eduardo Garibaldi, and Phillipe Thieullen. "Zero-Temperature Phase Diagram for Double-Well Type Potentials in the Summable Variation Class". In: *Ergodic Theory and Dynamical Systems* 38.3 (2018), pp. 863–885. 10.1017/etds.2016.57.

[Bin+15] Luca Bindi, Nan Yao, Chaney Lin, Lincoln Hollister, Christopher Andronicos, Vadim Distler, Michael Eddy, Alexander Kostin, Valery Kryachko, Glenn MacPherson, William Steinhardt, Marina Yudovskaya, and Paul Steinhardt. "Natural Quasicrystal With Decagonal Symmetry". In: *Scientific Reports* 5.1 (2015). 10.1038/srep09111.

[Boy+15] Laurent Boyer, Martin Delacourt, Victor Poupet, Mathieu Sablik, and Guillaume Theyssier. "μ-Limit Sets of Cellular Automata from a Computational Complexity Perspective". In: *Journal of Computer and System Sciences* 81.8 (2015), pp. 1623–1647. 10.1016/j.jcss.2015.05.004.

[Bur+22] Michael Burr, Suddhasattwa Das, Christian Wolf, and Yun Yang. "Computability of Topological Pressure on Compact Shift Spaces Beyond Finite Type". In: *Nonlinearity* 35.8 (2022), pp. 4250–4282. 10.1088/1361-6544/ac7702.

[BW20] Michael Burr and Christian Wolf. "Computability at Zero Temperature". In: *Nonlinearity* 33.11 (2020), pp. 6157–6175. 10.1088/1361-6544/ab9c71.

[CH10] Jean-René Chazottes and Michael Hochman. "On the Zero-Temperature Limit of Gibbs States". In: *Communications in Mathematical Physics* 297.1 (2010), pp. 265–281. 10.1007/s00220-010-0997-8.

[CH22] Antonin Callard and Benjamin Hellouin de Menibus. "The Aperiodic Domino Problem in Higher Dimension". In: *STACS*. Vol. 219. 2022, 19:1–19:15. 10.4230/LIPIcs.STACS.2022.19.

[CR15] Daniel Coronel and Juan Rivera-Letelier. "Sensitive Dependence of Gibbs Measures at Low Temperatures". In: *Journal of Statistical Physics* 160.6 (2015), pp. 1658–1683. 10.1007/s10955-015-1288-8.

[CR19] Daniel Coronel and Juan Rivera-Letelier. "Sensitive Dependence of Geometric Gibbs States at Positive Temperature". In: *Communications in Mathematical Physics* 368.1 (2019), pp. 383–425. 10.1007/s00220-019-03350-6.

[CS20] Jean-René Chazottes and Mao Shinoda. *On the Absence of Zero-Temperature Limit of Equilibrium States for Finite-Range Interactions on the Lattice* $\mathbb{Z}^2$. 2020. 10.48550/arXiv.2010.08998.

[DGG17] Bruno Durand, Guilhem Gamard, and Anaël Grandjean. "Aperiodic Tilings and Entropy". In: *Theoretical Computer Science* 666 (2017), pp. 36–47. 10.1016/j.tcs.2016.12.013.

[DRS10] Bruno Durand, Andrei Romashchenko, and Alexander Shen. "Effective Closed Subshifts in 1D Can Be Implemented in 2D". In: *Fields of Logic and Computation*. Vol. 6300. Lecture Notes in Computer Science. Springer, 2010, pp. 208–226. 10.1007/978-3-642-15025-8_12.

[DRS12] Bruno Durand, Andrei Romashchenko, and Alexander Shen. "Fixed-Point Tile Sets and Their Applications". In: *Journal of Computer and System Sciences* 78.3 (2012), pp. 731–764. 10.1016/j.jcss.2011.11.001.

[EMZ98] Aernout van Enter, Jacek Miękisz, and Miloš Zahradník. "Nonperiodic Long-Range Order for Fast-Decaying Interactions at Positive Temperatures". In: *Journal of Statistical Physics* 90.5-6 (1998), pp. 1441–1447. 10.1023/A:1023299916107.

[ENT23] Solène Esnay, Alonso Núñez, and Ilkka Törmä. "Arithmetical Complexity of the Language of Generic Limit Sets of Cellular Automata". In: *Journal of Computer and System Sciences* 134 (2023), pp. 20–41. 10.1016/j.jcss.2023.01.002.

[ER07] Aernout van Enter and Wioletta Ruszel. "Chaotic Temperature Dependence at Zero Temperature". In: *Journal of Statistical Physics* 127.3 (2007), pp. 567–573. 10.1007/s10955-006-9260-2.

[Gác01] Peter Gács. "Reliable Cellular Automata With Self-Organization". In: *Journal of Statistical Physics* 103.1-2 (2001), pp. 45–267. 10.1023/A:1004823720305.

[Gay21] Léo Gayral. "The Besicovitch-Stability of Noisy Tilings is Undecidable". In: Automata. July 2021. hal.science/hal-03233596.

[GS23a] Léo Gayral and Mathieu Sablik. "On the Besicovitch-Stability of Noisy Random Tilings". In: *Electronic Journal of Probability* 28 (2023), pp. 1–38. 10.1214/23-EJP917.

[GS23b] Léo Gayral and Mathieu Sablik. "Arithmetical Hierarchy of the Besicovitch-Stability of Noisy Tilings". In: *Theory of Computing Systems* (2023). 10.1007/s00224-023-10142-y.

[GST23] Léo Gayral, Mathieu Sablik, and Siamak Taati. *Characterisation of the Set of Ground States of Uniformly Chaotic Finite-Range Lattice Models*. 2023. 10.48550/arXiv.2302.07326.

[HM10] Michael Hochman and Tom Meyerovitch. "A Characterization of the Entropies of Multidimensional Shifts of Finite Type". In: *Annals of Mathematics* 171.3 (2010), pp. 2011–2038. 10.4007/annals.2010.171.2011.

[Hoc09] Michael Hochman. "On the Dynamics and Recursive Properties of Multidimensional Symbolic Systems". In: *Inventiones mathematicae* 176.1 (2009), pp. 131–167. 10.1007/s00222-008-0161-7.

[HS18] Benjamin Hellouin de Menibus and Mathieu Sablik. "Characterization of Sets of Limit Measures of a Cellular Automaton Iterated on a Random Configuration". In: *Ergodic Theory and Dynamical Systems* 38.2 (2018), pp. 601–650. 10.1017/etds.2016.46.

[JR21] Emmanuel Jeandel and Michaël Rao. "An Aperiodic Set of 11 Wang Tiles". In: *Advances in Combinatorics* (2021). 10.19086/aic.18614.

[JV15a] Emmanuel Jeandel and Pascal Vanier. "Characterizations of Periods of Multi-Dimensional Shifts". In: *Ergodic Theory and Dynamical Systems* 35.2 (2015), pp. 431–460. 10.1017/etds.2013.60.

[JV15b] Emmanuel Jeandel and Pascal Vanier. "Hardness of Conjugacy, Embedding and Factorization of Multidimensional Subshifts". In: *Journal of Computer and System Sciences* 81.8 (2015), pp. 1648–1664. 10.1016/j.jcss.2015.05.003.

[Kar96] Jarkko Kari. "A Small Aperiodic Set of Wang Tiles". In: *Discrete Mathematics* 160 (1996), pp. 259–264. 10.1016/0012-365X(95)00120-L.

[Lab19] Sébastien Labbé. "A Self-Similar Aperiodic Set of 19 Wang Tiles". In: *Geometriae Dedicata* 201 (2019), pp. 81–109. 10.1007/s10711-018-0384-8.

[Lab21] Sébastien Labbé. "Substitutive Structure of Jeandel-Rao Aperiodic Tilings". In: *Discrete & Computational Geometry* 65.3 (2021), pp. 800–855. 10.1007/s00454-019-00153-3.

[LS84] Dov Levine and Paul Steinhardt. "Quasicrystals: A New Class of Ordered Structures". In: *Physical Review Letters* 53 (26 1984), pp. 2477–2480. 10.1103/PhysRevLett.53.2477.

[Mey11] Tom Meyerovitch. "Growth-Type Invariants for $\mathbb{Z}^d$ Subshifts of Finite Type and Arithmetical Classes of Real Numbers". In: *Inventiones Mathematicae* 184 (3 2011), pp. 567–589. 10.1007/s00222-010-0296-1.

[Mię90] Jacek Miękisz. "A Microscopic Model With Quasicrystalline Properties". In: *Journal of Statistical Physics* 58.5–6 (1990), pp. 1137–1149. 10.1007/BF01026568.

[Mię97] Jacek Miękisz. "Stable Quasicrystalline Ground States". In: *Journal of Statistical Physics* 88.3–4 (1997), pp. 691–711. 10.1023/B:JOSS.0000015168.25151.22.

[Mię98] Jacek Miękisz. "An Ultimate Frustration in Classical Lattice-Gas Models". In: *Journal of Statistical Physics* 90.1–2 (1998), pp. 285–300. 10.1023/A:1023264004272.

[Nis05] Martin Niss. "History of the Lenz-Ising Model 1920–1950: From Ferromagnetic to Cooperative Phenomena". In: *Archive for History of Exact Sciences* 3 (2005), pp. 267–318. 10.1007/s00407-004-0088-3.

[Pei36] Rudolph Peierls. "On Ising's Model of Ferromagnetism". In: *Proceedings of the Cambridge Philosophical Society* 32 (1936), pp. 477–481.

[Rob71] Raphael Robinson. "Undecidability and Nonperiodicity for Tilings of the Plane". In: *Inventiones mathematicae* 12 (1971), pp. 177–209. 10.1007/BF01418780.

[She+84] Dan Shechtman, Ilan Blech, Denis Gratias, and John Cahn. "Metallic Phase With Long-Range Orientational Symmetry and No Translational Symmetry". In: *Physical Review Letters* 53 (20 1984), pp. 1951–1953. 10.1103/PhysRevLett.53.1951.

[Smi+23] David Smith, Joseph Myers, Craig Kaplan, and Chaim Goodman-Strauss. *An Aperiodic Monotile*. 2023. 10.48550/arXiv.2303.10798.

[Spa08] Christoph Spandl. "Computability of Topological Pressure for Sofic Shifts With Applications in Statistical Physics". In: *Journal of Universal Computer Science* 14.6 (2008), pp. 876–895. 10.3217/jucs-014-06-0876.

[Tur36] Alan Turing. "On Computable Numbers, With an Application to the Entscheidungsproblem". In: *Proceedings of the London Mathematical Society. Second Series* 42.3 (1936), pp. 230–265. 10.1112/plms/s2-42.1.230.

[Wan61] Hao Wang. "Proving Theorems by Pattern Recognition — II". In: *The Bell System Technical Journal* 40.1 (1961), pp. 1–41. 10.1002/j.1538-7305.1961.tb03975.x.

[Wes17] Linda Westrick. "Seas of Squares With Sizes From a Π^0_1 Set". In: *Israel Journal of Mathematics* 222.1 (2017), pp. 431–462. 10.1007/s11856-017-1596-6.

[Wes22] Linda Westrick. *Topological Completely Positive Entropy Is No Simpler in $\mathbb{Z}^2$-SFTs*. 2022. 10.48550/arXiv.1904.11444.

Chapter **2**

On Symbolic Dynamics

The objective of this chapter is to provide a mostly self-contained introduction to symbolic dynamics, and how it relates to ergodic theory for group actions in particular. The following sections can roughly be grouped in three themes.

In the broadest sense, symbolic dynamics can be seen the study of discrete dynamical systems within a (full) shift space, where configurations are labellings of a lattice G, with the canonical shift action $G \curvearrowright \mathcal{A}^G$ by translation. I will thus introduce the basic notions relating to symbolic spaces in Section 2.1. Some special attention will be given to the notions of periodicity (Section 2.2) and conjugacy (Section 2.3), both of which will play a role in Chapter 4. Then, the related notion of invariant measures will follow in Section 2.4, with the well-known weak-$*$ topology. These sections are heavily inspired by the book by Lind and Marcus [LM21], which provides a broad study of symbolic dynamics for the case of the lattice $G = \mathbb{Z}$. However, as I will need more general notions on $G = \mathbb{Z}^d$ later on, I chose here to introduce the formalism for a general group G. Readers interested in two-dimensional tilings may also refer to the book by Grünbaum and Shephard [GS87].

After this, I move onto ergodic theory, which provides a practical toolset to study these invariant measures. The broader framework for ergodic theory is that of *amenable* groups, for which interested readers may refer to [Kre85, Chapter 6] of Krengel's book, among other references [Mou85, Tem92]. In Section 2.5, I introduce the main results for $G = \mathbb{Z}^d$ instead, following [Kel98, Chapter 2] of Keller's book (which focuses on the semigroup action of $\mathbb{N}^d$ but translates almost verbatim for the group $\mathbb{Z}^d$). Then, using these tools, I will introduce another *stronger* topology on measures, notably discussed in the book by Glasner [Gla03], which will be central in Chapter 4.

I conclude the chapter with Section 2.7, which provides an extended overview of an example of subshift (or rather several related subshifts), that of the Robinson tiling [Rob71]. Notably, this tiling exhibits a self-similar hierarchical aperiodic structure, which allows it to simulate other tilings within itself. We will also see how this aperiodic tiling, and a variant thereof, relate to periodicity.

2.1 Shift Spaces

Definition 2.1 (Full Shift)**.** Let $(G, +)$ be a discrete (finite or) countable group, and $\mathcal{A}$ a finite alphabet. The *full shift* corresponds to the natural group action $G \overset{\sigma}{\curvearrowright} \mathcal{A}^G$, such that $(\sigma^g(\omega))_h := \omega_{h-g}$.

When there is no ambiguity on the group used (*e.g.* $G = \mathbb{Z}^d$ in most of the rest of this dissertation), we will denote this space $\Omega_{\mathcal{A}} = \mathcal{A}^G$. ■

Definition 2.2 (Cylinder Set)**.** Let $p \in \mathcal{A}^I$ be a pattern on a *finite* window $I \Subset G$. We define the *cylinder set* $[p] = \{\omega \in \Omega_{\mathcal{A}}, \omega_I = p\}$. The family $([p])_{I \Subset G, p \in \mathcal{A}^I}$ constitutes a countable base of open sets for the product (discrete) topology on $\Omega_{\mathcal{A}}$. Notice that cylinder sets are actually clopen (*i.e. both* open and closed). ■

The most studied and understood structure for symbolic dynamics is the case of the lattice $G = \mathbb{Z}^d$, and in particular the one-dimensional case $G = \mathbb{Z}$ (the definition can also be adapted for semigroups such as $G = \mathbb{N}$). However, in the last decade, some studies have shown active interest in SFTs on various other structures [Jea15, Bar23], such as free groups [BC21] or Baumslag-Solitar groups [AK13, EM22, ABH23]. While my work in later chapters focuses on the lattice $\mathbb{Z}^d$, some questions may naturally translate to the broader framework, such as the stability of periodic SFTs in Chapter 4 that may translate to other groups where similar percolation arguments hold.

Definition 2.3 (Subshift)**.** A subshift X of $\Omega_{\mathcal{A}}$ is a closed subset stable under the shift action of G. Let $\mathcal{F}$ be a collection of *forbidden patterns*, *i.e.* elements $p \in \mathcal{A}^I$ defined on finite windows $I(p) \Subset G$. We can then define the induced subshift $X_{\mathcal{F}}$ as

$$X_{\mathcal{F}} = \left\{\omega \in \Omega_{\mathcal{A}}, \forall p \in \mathcal{F}, \forall g \in G, \sigma^g(\omega)_{I(p)} \neq p\right\} ,$$

i.e. the set of configurations that contain no translation of any forbidden pattern (with ω_I denoting the restriction of the configuration ω to the window I). ■

For a simple example, consider the one-dimensional case (*i.e.* $G = \mathbb{Z}$), on the binary alphabet $\mathcal{A} = \{\mathtt{0}, \mathtt{1}\}$, with $\mathcal{F} = \{\mathtt{10}\}$. For any configuration $\omega \in X_{\mathcal{F}}$, we have only 0 symbols up to a rank, and only 1 after that rank (which may be a position $k \in \mathbb{Z}$ but also $\pm\infty$). The effect of the shift action on $X_{\mathcal{F}}$ is graphically illustrated in Figure 2.1.

$$\begin{aligned} \sigma^2(\omega) &= \cdots \; \mathtt{0} \; \mathtt{0} \; \mathtt{0} \; \mathtt{0} \; \mathtt{1} \; \mathtt{1} \; \mathtt{1} \cdots \\ \sigma^1(\omega) &= \cdots \; \mathtt{0} \; \mathtt{0} \; \mathtt{0} \; \mathtt{1} \; \mathtt{1} \; \mathtt{1} \; \mathtt{1} \cdots \\ \omega &= \cdots \; \mathtt{0} \; \mathtt{0} \; \mathtt{1} \; \mathtt{1} \; \mathtt{1} \; \mathtt{1} \; \mathtt{1} \cdots \end{aligned}$$

Figure 2.1: Effect of the shift action on $\omega \in \Omega_{\{\mathtt{0},\mathtt{1}\}}$.

To avoid repeating "up to translation" on and on and on, when convenient, we may treat shifted windows $I + g$ as the window I itself, and thus identify the congruent patterns in the sets $\mathcal{A}^{I+g}$ and $\mathcal{A}^I$. In particular, following this logic, $X_{\mathcal{F}}$ is simply the set of configurations that contain no forbidden pattern from $\mathcal{F}$.

Remark 2.4 (Wang Tiles)**.** There is a distinction to be made between the *tiles* (*i.e.* the alphabet $\mathcal{A}$), the *tileset* (*i.e.* the alphabet $\mathcal{A}$ with a given set of local rules/forbidden patterns $\mathcal{F}$), and the *tilings* (*i.e.* the configurations in $\Omega_{\mathcal{A}}$ that may or may not abide by given rules).

While the distinction between the tileset $(\mathcal{A}, \mathcal{F})$ and the induced space of tilings $X_{\mathcal{F}} \subseteq \Omega_{\mathcal{A}}$ is impossible to erase as it relates deeply to how simple rules can induce complex structures (see Section 2.7 for example), the distinction between the tiles and the tileset is less rigid, in particular when the alphabet $\mathcal{A}$ itself has some added structure that intuitively induces local rules.

This is notably the case of Wang tiles on $G = \mathbb{Z}^2$, where we have a set of colours C, and $\mathcal{A} \subseteq C^4$ (with one colour for each cardinal direction). Naturally, in this context, we want matching North-South and East-West interfaces:

$$\mathcal{F} := \left\{\begin{pmatrix} b \\ a \end{pmatrix} \in \mathcal{A}^2, \mathrm{North}(a) \neq \mathrm{South}(b) \in C\right\} \sqcup \left\{(a\,b) \in \mathcal{A}^2, \mathrm{East}(a) \neq \mathrm{West}(b) \in C\right\} .$$

An example of Wang tileset can be seen in Figure 1.1. In such contexts, we may thus simply define the alphabet, the set of tiles, and deduce the local rules, the tileset, from it. Hence, we will mostly indiscriminately consider the two notions when appropriate. ■

Definition 2.5 (Admissible Patterns)**.** Let $x \in \mathcal{A}^I$ be a pattern on a (not necessarily finite) window $I(x) \subseteq G$, and $\mathcal{F}$ a family of forbidden patterns. We say that x is *locally admissible* (with respect to $\mathcal{F}$) if it contains no forbidden pattern $p \in \mathcal{F}$ (*i.e.* $\sigma^g(x)_{I(p)} \neq p$ for any $g \in G$ such that $I(p) + g \subseteq I(x)$).

We say that x is *globally admissible* (with respect to a subshift X) if $x = \omega_I$ for some $\omega \in X$, and we will denote $L_I(X)$ the set of globally admissible words in $\mathcal{A}^I$, which we will also call the *language* of X. ■

Remark 2.6. Naturally, globally admissible patterns of $X_{\mathcal{F}}$ are also locally admissible for $\mathcal{F}$. However, the converse is not true. For example, consider $\mathcal{A} = \{\mathtt{0}, \mathtt{1}\}$, $G = \mathbb{Z}$ and let $\mathcal{F} = \{\mathtt{01}, \mathtt{10}, \mathtt{11}\}$. These forbidden patterns tell us that the symbol 1 cannot be encountered next to either a 0 or a 1 (*i.e.* the whole alphabet), so $X_{\mathcal{F}} = \{\mathtt{0}^\infty\}$ is a singleton, and 0 is the only globally admissible pattern on the window $\{0\} \Subset \mathbb{Z}$. However, all by itself, the symbol 1 contains no forbidden pair from $\mathcal{F}$, so it is still locally admissible. ■

Lemma 2.7. *Let $X \subseteq \Omega_{\mathcal{A}}$ be a subshift. Then there exists a countable set $\mathcal{F}$ such that $X = X_{\mathcal{F}}$.*

Proof. Let $\mathcal{P}_{\Subset}(G)$ be the family of *finite subsets* of G. As G is countable, so is $\mathcal{P}_{\Subset}(G)$. For $I \in \mathcal{P}_{\Subset}(G)$, as $\mathcal{A}^I$ is finite, so is the complementary set $L_I(X)^c$ of the language on I. Thus, consider the following countable set:

$$\mathcal{F} := \bigcup_{I \in \mathcal{P}_{\Subset}(G)} L_I(X)^c .$$

Naturally, by definition of $\mathcal{F}$, we have the inclusion $X \subseteq X_{\mathcal{F}}$. Now, let I_n be an increasing sequence of finite sets, such that $G = \bigcup I_n$. If $\omega \in X_{\mathcal{F}}$, then $\omega_{I_n} \in L_{I_n}(X)$ necessarily, so we have an element $\omega_n \in X$ that matches ω on I_n. Then, in the product discrete topology, we have $\omega_n \to \omega$, and X is closed, so $\omega \in X$, which concludes the proof. □

If there is a *finite* set of forbidden patterns $\mathcal{F}$, then X is a *Subshift of Finite Type* (which will be shortened as SFT in the rest of this dissertation). In particular, the full shift $\Omega_{\mathcal{A}}$ itself is the SFT without any forbidden patterns, where $\mathcal{F} = \emptyset$. More broadly, if there exists a *recursively enumerable* set $\mathcal{F}$ (see Chapter 3 for a proper definition of what it means for a set to be enumerable by a Turing machine) such that $X = X_{\mathcal{F}}$, then X is said to be *effective*.

Usually, in the one-dimensional case $G = \mathbb{Z}$, SFTs have a simple *automatic* structure that allows for a complete characterisation of their properties and behaviours. This will be the case here, both in Chapter 4 where we will obtain a decidable characterisation of stability in the one-dimensional case and Chapter 5 where one-dimensional finite-range potentials cannot induce a chaotic model.

2.2 Periodicity

In Chapter 4, I will often refer to the notion of *periodicity* of configurations or subshifts. While this notion makes some amount of intuitive sense, some of the related results are not obvious, hence the need for a formal introduction.

Definition 2.8 (Periodic and Aperiodic Configurations)**.** The configuration $\omega \in \Omega_{\mathcal{A}}$ is said to be (strongly) *periodic iff* its orbit $\{\sigma^g(x), g \in G\}$ for the shift action is finite. Conversely, $\omega \in \Omega_{\mathcal{A}}$ is said to be (strongly) *aperiodic iff* there is no $g \in G$ such that $\sigma^g(\omega) = \omega$. ■

The finite orbit of a periodic configuration is in particular an SFT.

Proposition 2.9 (Adapted from [BDJ08, Theorem 3.8])**.** *Let $G = \mathbb{Z}^d$, and X a subshift (not necessarily of finite type* a priori*). The following properties are equivalent:*

1. *X is a finite set,*
2. *every element $\omega \in X$ is periodic,*
3. *there is a window $I_n := [\![0, n-1]\!]^d$ such that any element $\omega \in X$ is a labelling of I_n repeated periodically in each direction (*i.e. *it is invariant under the action of the subgroup $(n\mathbb{Z})^d$).*

Proof. The implication ($1 \Rightarrow 2$) follows from the fact that the orbit of ω is included in X, thus finite too. Likewise, the implication ($3 \Rightarrow 1$) directly follows from the fact that the set $\mathcal{A}^{I_n}$ is finite.

For (1&2 $\Rightarrow$ 3), consider first some periodic element $\omega \in X$. Because its shift orbit is finite, it means that in each direction of the canonical base $(e_i)_{1\leq i\leq d}$ of $\mathbb{Z}^d$ we have some integer $n_i(\omega)$ such that $\sigma^{n_i e_i}(\omega) = \omega$. In particular, by taking the least common multiple $n(\omega) = \mathrm{lcm}\,(n_1, \ldots, n_d)$, then ω is invariant under $(n(\omega)\mathbb{Z})^d$. As X itself is finite, we can then consider $n = \mathrm{lcm}\,(n(\omega), \omega \in X)$, such that any element $\omega \in X$ is invariant under $(n\mathbb{Z})^d$.

Finally, let's prove ($2 \Rightarrow 1$) by contradiction. Assume X is infinite, but only contains periodic elements. As the set of tilings invariant under $(n\mathbb{Z})^d$ is finite (each tiling corresponds to a pattern in $\mathcal{A}^{I_n}$), X must contain periodic configurations with arbitrarily large minimal periods (*i.e.* n such that ω is invariant under $(n\mathbb{Z})^d$ but not invariant under $(m\mathbb{Z})^d$ for $m < n$). Let a sequence ω_k with minimal periods $n(k) \to \infty$. Up to extraction, by compactness of X, assume $\omega_k \to \omega \in X$. As ω *must* be periodic, it is invariant under $(n\mathbb{Z})^d$ for some $n \in \mathbb{N}$ (even though the elements ω_k are not after a rank), and we identify it to $x = \omega_{I_n} \in \mathcal{A}^{I_n}$. By definition of the product discrete topology on $\Omega_{\mathcal{A}}$, we must have a maximal window $I_{n\times q(k)}$ where ω matches ω_k, with $q(k) \to \infty$. What's more, when $n(k) > n$ (thus in particular $\omega_k \neq \omega$), we have a finite upper bound on $q(k)$, which we now assume is maximal. On the (top or right) boundary of $I_{n\times q(k)}$, we must then have a translation of I_n that doesn't contain the pattern x. We define ω'_k as the translation of ω_k that brings this mismatching block inside the actual window I_n. Up to extraction, $\omega'_k \to \omega' \in X$. On one hand, inside $]\!]-\infty, -1]\!]^d$, ω' and ω coincide, which implies $\omega = \omega'$ as they are both periodic. On the other hand, we know that $\omega'_{I_n} \neq \omega_{I_n}$, hence a contradiction. □

Remark 2.10. A naive compactness argument would be insufficient to prove ($2 \Rightarrow 1$) here, as we can have a sequence (ω_k), with minimal periods $n(k) \to \infty$, that still converges to a periodic tiling. ■

In particular, by simply enumerating the (finitely many) labellings of I_{n+1} that do not occur in any element of X (*i.e.* the complementary of the language $L_{I_{n+1}}(X)$ in $\mathcal{A}^{I_{n+1}}$), we obtain $\mathcal{F}$ such that $X = X_{\mathcal{F}}$, *i.e.* X is an SFT. We will call such SFTs *periodic*.

Remark 2.11. The reason why we *must* use the window I_{n+1} instead of I_n is not immediately obvious. We in fact need this extra margin to add some redundancy in the patterns, which relates to the notion of *reconstruction function* which we will detail in greater depth in Chapter 4.

For now, let's just prove that using I_n *wouldn't* work with a counter-example on $G = \mathbb{Z}$. Let $\mathcal{A} = \{*, \mathtt{0}, \mathtt{1}\}$, and $X = \{(*\mathtt{1})^\infty, (\mathtt{1}*)^\infty, (*\mathtt{0})^\infty, (\mathtt{0}*)^\infty\}$. This set is the union of two 2-periodic orbits, so $n = 2$ is a good window in the previous proposition. However, with $\mathcal{F} := L_{I_2}(X)^c = \{**, \mathtt{00}, \mathtt{11}, \mathtt{01}, \mathtt{10}\}$, we obtain an infinite SFT $X_{\mathcal{F}} = \{(*\mathtt{0}|*\mathtt{1})^\infty, (\mathtt{0}*|\mathtt{1}*)^\infty\}$ instead. In other words, a configuration in X encodes one single bit in $\{\mathtt{0}, \mathtt{1}\}$ repeated infinitely, but the patterns in $\mathcal{F}$ are too small so each bit of $X_{\mathcal{F}}$ can be chosen freely instead. ■

Lemma 2.12. *Let $G = \mathbb{Z}^2$, $X_{\mathcal{F}}$ an SFT and assume $X_{\mathcal{F}}$ contains an element ω that is not (strongly) aperiodic. Then it contains a (strongly) periodic element.*

Proof. As $X_{\mathcal{F}}$ is an SFT, we suppose $\mathcal{F}$ is a finite set. Up to translation of each forbidden pattern to embed them in the north-eastern quarter plane, let a window size r such that $I(p) \subseteq I_r$ for any $p \in \mathcal{F}$.

Let $\omega \in X_{\mathcal{F}}$ be a non-aperiodic configuration, with a non-zero vector $k = (k_1, k_2) \in \mathbb{Z}^2$ such that $\sigma^k(\omega) = \omega$. Without loss of generality, we may assume that $k_2 \geq r - 1$.

Consider the sliding window $(I + ie_1)_{i \in \mathbb{Z}}$ with $I = [\![0, r + k_1 - 1]\!] \times [\![0, k_2 - 1]\!]$. On each such window, in ω, we observe a pattern in the finite set $\mathcal{A}^I$, hence the existence of distinct positions $i < j$ such that $\omega_{I+ie_1} = \omega_{I+je_1}$ (we can even require disjoint windows, *i.e.* $\ell := j - i \geq k_1 + r$). Let $p = \omega_J$ the restriction on the rectangle $J = [\![i, j-1]\!] \times [\![0, k_2 - 1]\!]$. As $\mathbb{Z}^2 = \bigsqcup_{(x,y) \in \mathbb{Z}^2} (J + x\ell e_1 + yk)$, define ω' by using the pattern p on *every* shifted window. By construction, ω and ω' always match in the union of rectangles $\bigsqcup_{y \in \mathbb{Z}} (J' + yk)$ with $J' = [\![i, j + k_1 + r - 1]\!] \times [\![0, k_2 - 1]\!]$.

By (strong) periodicity of ω', to prove that $\omega' \in X_{\mathcal{F}}$, we just need to check for forbidden patterns inside windows $I_r + k'$ with $k' \in J$. Now, we always have either $I_r + k' \subseteq J' \sqcup (J' + k)$ or $I_r + k' + \ell e_1 \subseteq J' \sqcup (J' + k)$, where it matches $\omega \in X_{\mathcal{F}}$, so the window $I_r + k'$ of ω' is globally admissible in any case. In conclusion, we have a (strongly) periodic configuration $\omega' \in X_{\mathcal{F}}$. □

The same argument, in a simpler way, can be used to prove that *any* non-empty one-dimensional SFT (on $G = \mathbb{Z}$) must contain a periodic configuration. However, in the higher-dimensional case, it is known [Ber66, Rob71, Kar96] that there exists SFTs that contain only (strongly) aperiodic configurations. This is the case of the Robinson tiling in particular, which we will introduce at the end of the current chapter, and whose hierarchical aperiodic structure will allow us to perform complex tasks in the next chapters.

Even though there is a spectrum of "degrees" of (a)periodicity in the general case, for the lattice $\mathbb{Z}^2$ which is my main interest here, an SFT must either contain only aperiodic configurations or include at least one periodic element. Hence, for most of the rest of this dissertation, we will consider SFTs that are either *aperiodic* (*i.e.* all their elements are (strongly) aperiodic) or *periodic* (*i.e.* finite) and forget about the other less structured behaviours in-between.

2.3 Conjugacy

Another standard notion we refer to in Chapter 4, and need to introduce beforehand, is that of *conjugate* subshifts.

Definition 2.13 (Morphism)**.** Let $X \subseteq \Omega_{\mathcal{A}}$ and $Y \subseteq \Omega_{\mathcal{B}}$ be two subshifts, defined using the same group G. A morphism is a continuous map $\theta : X \to Y$ that commutes with the shift action (*i.e.* $\theta \circ \sigma^g = \sigma^g \circ \theta$ for any $g \in G$). ■

One natural way of defining morphisms is to use a *cellular automaton* (sometimes called a *sliding block code*). By this, we mean that we let a map $\theta : \mathcal{A}^I \to \mathcal{B}$ on a finite window $I \Subset G$, and extend it as $\overline{\theta} : \Omega_{\mathcal{A}} \to \Omega_{\mathcal{B}}$ so that $\overline{\theta}(\omega)_g = \theta(\omega_{I+g})$.

Proposition 2.14 (Adapted from [Hed69, Theorem 3.4])**.** *Any morphism $\theta : X \to Y$ is the restriction to X of a cellular automaton $\overline{\theta} : \Omega_{\mathcal{A}} \to \Omega_{\mathcal{B}}$.*

Proof. For each letter $b \in \mathcal{B} \cong \mathcal{B}^{\{0\}}$, we have a corresponding cylinder set $[b]$, so that $\Omega_{\mathcal{B}} = \bigsqcup_{b\in\mathcal{B}}[b]$. It follows that we have a partition $X = \bigsqcup_{b\in\mathcal{B}} U_b$ made of clopen sets $U_b = \theta^{-1}([b])$.

Let J_n an increasing sequence of finite windows such that $\bigcup J_n = G$. Let us prove that, for n large enough, the partition $([p])_{p\in\mathcal{A}^{J_n}}$ of X is a refinement of $(U_b)_{b\in\mathcal{B}}$. If that is not the case, then we have a sequence of patterns $p_n \in \mathcal{A}^{J_n}$ such that each $[p_n]$ (here a *non-empty* subset of X) intersects two sets of $(U_b)_{b\in\mathcal{B}}$. By compactness of X, up to extraction, we have $\bigcap_n [p_n] = \{\omega\}$, and two letters $b, b' \in \mathcal{B}$ such that $\omega_n \in [p_n] \cap U_b \neq \emptyset$ and $\omega'_n \in [p_n] \cap U_{b'} \neq \emptyset$. By continuity of θ, as $\omega_n \to \omega$ and $\theta(\omega_n) \in [b]$, it follows that $\theta(\omega)_0 = b$. The same argument with the sequence (ω'_n) gives $\theta(\omega)_0 = b'$ instead, thus a contradiction.

Now, let n such that $([p])_{p\in\mathcal{A}^{J_n}}$ is a refinement of $(U_b)_{b\in\mathcal{B}}$. For any $p \in \mathcal{A}^{J_n}$, there is a unique $b(p) \in \mathcal{B}$ such that $[p] \subseteq U_b$ (if $[p] = \emptyset$ in X, we may simply arbitrarily chose a value for $b(p)$). This map $p \in \mathcal{A}^{J_n} \mapsto b(p) \in \mathcal{B}$ naturally extends as a cellular automaton $\overline{\theta}$, and it is easy to check that when $\omega \in X$, $\overline{\theta}(\omega) = \theta(\omega)$. ☐

Hence, we may indiscriminately identify morphisms, cellular automata and "local" maps $\mathcal{A}^I \to \mathcal{B}$ later on. So far, we have only introduced morphisms that do a *one-way* trip. In this context, a conjugacy is expectedly an isomorphism of subshifts, that goes back-and-forth:

Definition 2.15 (Conjugacy)**.** A conjugacy $\theta : X \to Y$ is a bijective morphism. Note that, by compactness of the shift spaces, θ is then a homeomorphism, so θ^{-1} is a morphism (of subshifts) too. ■

Naturally, we say that two subshifts are *conjugate* if there exists a conjugacy between them.

Remark 2.16. Let a morphism $\overline{\theta} : X \to Y$ induced by $\theta : \mathcal{A}^I \to \mathcal{B}$. In-between these two scales, we can also denote $\overline{\theta}(p) \in \mathcal{B}^J$ the image of a pattern $p \in \mathcal{A}^{I+J}$. Then, for any window $J \Subset G$, we have $\overline{\theta}\left(L_{I+J}(X)\right) \subseteq L_J(Y)$. What's more, the morphism is a conjugacy, then this inclusion is an equality.

Indeed, for any $p \in L_{I+J}(X)$, we have $\omega \in X$ such that $\omega_{I+J} = p$, thus $\overline{p} = \overline{\theta}\left(\omega_{I+J}\right) = \overline{\theta}(\omega)_J \in L_J(Y)$ as $\overline{\theta}(\omega) \in Y$ (and the equality follows from the surjectivity of $\overline{\theta} : X \to Y$ for a conjugacy). ■

Proposition 2.17. *Let X and Y be conjugate subshifts. If X is an SFT, then so must be Y.*

Proof. Let $\theta : X \to Y$ be a conjugacy, with $\overline{\theta} : \mathcal{A}^I \to \mathcal{B}$ and $\overline{\theta^{-1}} : \mathcal{B}^J \to \mathcal{A}$ corresponding cellular automata. Without loss of generality, up to enlarging the windows I and J, suppose that $X = X_{\mathcal{F}}$ has forbidden patterns $\mathcal{F} \subseteq \mathcal{A}^I$, and $0_G \in I \cap J$ (so that $g \in J + I + g$ for any $g \in G$).

Define the set of forbidden patterns $\mathcal{F}' = L_{J+I}(Y)^c \subseteq \mathcal{B}^{J+I}$. We want to prove that $Y = X_{\mathcal{F}'}$. Naturally, the inclusion $Y \subseteq X_{\mathcal{F}'}$ holds because, for any $g \in G$, $\omega_{J+I+g} \in L_{J+I}(Y)$.

Now, consider $\omega \in X_{\mathcal{F}'}$, and let $\omega' = \overline{\theta} \circ \overline{\theta^{-1}}(\omega)$. First, notice that $\omega' \in Y$, because $\overline{\theta^{-1}}(\omega) \in X$. Indeed, for any window $I + g$ we have $\overline{\theta^{-1}}(\omega)_{I+g} = \overline{\theta^{-1}}\left(\omega_{J+I+g}\right)$ and $\omega_{J+I+g} \in L_{J+I}(Y)$ (*i.e.* the complementary set of $\mathcal{F}'$) so following the previous remark $\overline{\theta^{-1}}(\omega)_{I+g} \in L_I(X)$ (itself a subset of the complementary of $\mathcal{F}$). Second, $\omega' = \omega$. Indeed, let $g \in G$ and $\omega^g \in Y$ such that $\omega^g_{J+I+g} = \omega_{J+I+g} \in L_{J+I}(Y)$. As $g \in J + I + g$:

$$\omega_g = \omega^g_g = \theta \circ \theta^{-1}\left(\omega^g\right)_g = \overline{\theta} \circ \overline{\theta^{-1}}\left(\omega^g_{J+I+g}\right) = \overline{\theta} \circ \overline{\theta^{-1}}\left(\omega_{J+I+g}\right) = \overline{\theta} \circ \overline{\theta^{-1}}\left(\omega\right)_g = \omega'_g \,.$$

Finally, $\omega = \omega' \in Y$, so we have indeed $Y = X_{\mathcal{F}'}$ an SFT. ☐

We will say that the property of "being an SFT" is a conjugacy invariant. Other known invariant properties, that characterise the dynamical behaviour of the subshift, include unique ergodicity (introduced in Section 2.5), minimality (briefly discussed in Section 2.7), transitivity (*i.e.* the existence of a dense orbit), etc. In a similar fashion, some numerical characteristics of a subshift are conjugacy invariants, which allow us to prove that two subshifts are *not* conjugate. This includes the period of periodic subshifts, the entropy (which will play an important role in Chapter 5), the entropic dimension (which allows us to distinguish zero-entropy systems), etc. In particular, in Chapter 4, we will see that the notion of stable SFTs we introduce, which *a priori* depends on the choice of forbidden patterns $\mathcal{F}$, is a conjugacy invariant, hence characterises the SFT itself.

2.4 Measures on Shift Spaces

Now that the general context for subshifts is introduced, we can formally introduce invariant measures, which will allow us to study the average and/or typical properties of the tilings in these spaces.

Definition 2.18 (Weak-$*$ Topology)**.** Let $\mathcal{M}(X)$ be the set of probability measures on a subshift X. We define the weak-$*$ topology as the minimal topology such that $\mu \mapsto \mu([p])$ is a continuous function for any pattern $p \in \mathcal{A}^I$ with $I \Subset G$.

For any increasing sequence J_n of finite windows such that $\bigcup J_n = G$, we can define a (non-canonical) distance $d^*(\mu, \nu) = \sum_{n\in\mathbb{N}} \frac{1}{2^{n+1}} \sum_{p\in\mathcal{A}^{J_n}} |\mu([p]) - \nu([p])|$ that induces the weak-$*$ topology. ■

Because the weak-$*$ topology only looks at what happens inside finite windows $I \Subset G$, where we have only finitely many patterns $\mathcal{A}^I$, it is in particular adapted to study sets of measures through the lens of *computable analysis*, which will be introduced in the next chapter.

Define $\sigma_*^g(\mu) := \mu \circ \sigma^{-g}$ the *pushforward* measure shifted by $g \in G$.

Definition 2.19 (Invariant Measures)**.** Let $\mathcal{M}_\sigma(X)$ be the set of shift-invariant probability measures, *i.e.* measures $\mu \in \mathcal{M}(X)$ such that $\sigma_*^g(\mu) = \mu$ for any $g \in G$ (and it is enough to require/check the property for a family of generators of G, such as the canonical basis $(e_i)_{1\leq i\leq d}$ for $\mathbb{Z}^d$). ■

Proposition 2.20. *Let $G = \mathbb{Z}^d$ and X a non-empty subshift. Then $\mathcal{M}_\sigma(X) \neq \emptyset$.*

Proof. As $X \neq \emptyset$, $\mathcal{M}(X)$ is not empty either (it contains at the very least a Dirac measure δ_ω with $\omega \in X$). Hence, let $\mu \in \mathcal{M}(X)$, and define $\mu_n = \frac{1}{(2n+1)^d} \sum_{k\in[\![-n,n]\!]^d} \sigma_*^k(\mu) \in \mathcal{M}(X)$. Then, by compactness of $\mathcal{M}(X)$, let μ_∞ an adherence value of the sequence (μ_n). Using the same extraction for $(\sigma_*^{e_i}(\mu_n))_n$, we naturally converge to $\sigma_*^{e_i}(\mu_\infty)$. Now, notice that:

$$\sigma_*^{e_i}(\mu_n) - \mu_n = \frac{1}{(2n+1)^d} \left(\sum_{k\in[\![-n,n]\!]^d, k_i=n} \sigma_*^{k+e_i}(\mu) - \sum_{k\in[\![-n,n]\!]^d, k_i=-n} \sigma_*^k(\mu) \right).$$

It follows that (regardless of the choice of J_n) we have $d^*(\sigma_*^{e_i}(\mu_n), \mu_n) \leq \|\sigma_*^{e_i}(\mu_n) - \mu_n\|_\infty \leq \frac{1}{n} \underset{n\to\infty}{\longrightarrow} 0$, hence at the limit $\sigma_*^{e_i}(\mu_\infty) = \mu_\infty$. As this holds for any element e_i of the generating family of $\mathbb{Z}^d$, we conclude that the result holds for any vector $k \in \mathbb{Z}^d$, *i.e.* $\mu_\infty \in \mathcal{M}_\sigma(X)$ is shift-invariant. □

This result quite naturally adapts to the broader case of *amenable* groups. By proving the existence of such measures, we are actually laying the bedrock of ergodic theory, and while we will introduce it in the specific context of $G = \mathbb{Z}^d$ in the next section, most of it generalises well to amenable groups. Note that, for non-amenable groups such as free groups, even the existence of invariant measures in $\mathcal{M}_\sigma(X)$ is not easily guaranteed [ST16, Ber17].

2.5 Ergodic Theory of $\mathbb{Z}^d$-Actions

The key idea of ergodic theory is to study dynamical systems $G \overset{\sigma}{\curvearrowright} X$ through their typical behaviour, either *on average* or *generically* (*i.e.* almost surely in some sense) by looking at the invariant measures in $\mathcal{M}_\sigma(X)$. In particular, the key concept to understand is that of *ergodicity*, of ergodic measures, each of those encoding a typical behaviour occurring in X. As initially announced, let us focus on the case of $G = \mathbb{Z}^d$, and let's introduce the formalism of ergodic theory in this context.

Definition 2.21 (Invariant Function)**.** Let X be a subshift and $\mu \in \mathcal{M}_\sigma(X)$. The function $f \in \mathrm{L}^1(X, \mu)$ is invariant (with respect to μ) if, for any $k \in \mathbb{Z}^d$, we have $f \circ \sigma^k = f$ (μ-a.s.). ■

For a non-negative or integrable function f, denote $\mu(f) = \int f(x)\mathrm{d}\mu(x)$ its integral.

Theorem 2.22 (Birkhoff's Pointwise Ergodic Theorem)**.** *Let X be a subshift and $\mu \in \mathcal{M}_\sigma(X)$. For each function $f \in \mathrm{L}^1(X, \mu)$, the limit*

$$\overline{f} = \lim_{n\to\infty} \frac{1}{(2n+1)^d} \sum_{k\in[\![-n,n]\!]^d} f \circ \sigma^k = \lim_{n\to\infty} \frac{1}{n^d} \sum_{k\in I_n} f \circ \sigma^k$$

is μ-a.s. well-defined, with convergence in $\mathrm{L}^1(X, \mu)$. What's more, $\overline{f} \in \mathrm{L}^1(X, \mu)$ is invariant and $\mu\left(\overline{f}\right) = \mu(f)$.

Remark 2.23. More broadly, we can obtain the same limit $\overline{f}$ by using any increasing sequence of rectangles whose length goes to ∞ in all dimensions, or even by taking some limits first (using a lower-dimensional ergodic theorem). ■

Definition 2.24 (Ergodic Measure)**.** The measure μ is *ergodic* if all invariant functions are constant. There are several other equivalent characterisations of ergodicity, such as:

$$\forall f \in \mathrm{L}^1(X,\mu), \frac{1}{n^d}\sum_{k\in I_n} f\circ\sigma^k \overset{\mu\text{-a.s.}}{\longrightarrow} \mu(f)\,.$$

This matter-of-factly implies that the μ-probability of an event in a finite window must then be equal to the frequencies of occurrences of said event in sliding windows. ■

Theorem 2.25 (Ergodic Decomposition)**.** *Let $\mu \in \mathcal{M}_\sigma(X)$. For almost every $\omega \in X$, we have an ergodic weak-$*$ limit $\mu_\omega = \lim \frac{1}{n^d}\sum_{k\in I_n}\delta_{\sigma^k(\omega)} \in \mathcal{M}_\sigma(X)$. For any integrable function $f \in \mathrm{L}^1(X,\mu)$, for almost every $\omega \in X$, $\frac{1}{n^d}\sum_{k\in I_n} f\circ\sigma^k(\omega) \to \mu_\omega(f)$. What's more, $\omega \mapsto \mu_\omega(f)$ is measurable, integrable and $\int \mu_\omega(f)\mathrm{d}\mu(\omega) = \mu(f)$.*

In other words, this means that ergodic measures are the extremal points of the convex set $\mathcal{M}_\sigma(X)$, and that any measure $\mu \in \mathcal{M}_\sigma(X)$ can be arbitrarily well approached (in the weak-$*$ topology) by convex combinations of (finitely many) ergodic measures, using the Krein-Milman theorem. Consequently, we say that X is *uniquely ergodic* if it admits a unique ergodic measure, *i.e.* if $\mathcal{M}_\sigma(X)$ is a singleton.

2.6 Besicovitch Distance

Now that I have introduced the standard framework of ergodic theory, we have the necessary tools to define another topology on $\mathcal{M}_\sigma(X)$. The Besicovitch distance has been quite used in the recent research literature, but it was already introduced in earlier works, *e.g.* as the Ornstein distance $\overline{d}$ in [Gla03, Chapter 15] of Glasner's book. In that initial context, the interest of this topology was that it turns the Kolmogorov-Sinai entropy $\mu \mapsto h(\mu)$ into a continuous function (see Chapter 5 for a proper definition of h, in the context of Gibbs measures).

The main difference with the weak-$*$ topology is that the Besicovitch distance compares configurations *globally*, instead of looking at (arbitrarily large) finite windows.

Definition 2.26 (Hamming Distance)**.** On a finite window $I \Subset \mathbb{Z}^d$, we define the Hamming distance between two finite patterns $p, q \in \mathcal{A}^I$ as $d_I(x,y) = \frac{1}{|I|}\left|\{k \in I, p_k \neq q_k\}\right|$. We can then define the Hamming-Besicovitch pseudometric d_H on $\Omega_\mathcal{A}$, such that $d_H(\omega,\omega') = \limsup_{n\to\infty} d_{I_n}\left(\omega_{I_n},\omega'_{I_n}\right)$. ■

Definition 2.27 (Besicovitch Distance)**.** A measure $\lambda \in \mathcal{M}_\sigma(X\times Y)$ is said to be a coupling (or a joining) between the measures $\mu \in \mathcal{M}_\sigma(X)$ and $\nu \in \mathcal{M}_\sigma(Y)$ if $\pi^X_*(\lambda) = \mu$ and $\pi^Y_*(\lambda) = \nu$ (with π^X and π^Y the canonical projections of $X \times Y$ onto X and Y). For two measures $\mu, \nu \in \mathcal{M}_\sigma(X)$ we define their Besicovitch distance as:

$$d_B(\mu,\nu) := \inf_{\lambda \text{ a coupling}} \int d_H(x,y)\mathrm{d}\lambda(x,y)\,.$$

Note that we can always consider the independent coupling $\mu \otimes \nu$, so the set of couplings is non-empty and d_B is well-defined. By the pointwise ergodic theorem, d_H is obtained by averaging $\mathbb{1}_{x_0\neq y_0}$ over translations, so $\lambda(d_H) = \lambda([x_0 \neq y_0])$ for any coupling. ■

Remark 2.28. Instead of couplings with projections, we can more generally define d_B using invariant measures $\lambda \in \mathcal{M}_\sigma(Z)$ with measurable maps $\psi^1 : Z \to X$ and $\psi^2 : Z \to X$, such that $\psi^1_*(\lambda) = \mu \in \mathcal{M}_\sigma(X)$ and $\psi^2_*(\lambda) = \nu \in \mathcal{M}_\sigma(X)$. With $\overline{\lambda} := \left(\psi^1,\psi^2\right)_*(\lambda) \in \mathcal{M}_\sigma\left(X^2\right)$ the corresponding coupling between μ and ν, we have $\lambda(d_H) = \int d_H\left(\psi^1(z),\psi^2(z)\right)\mathrm{d}\lambda(z)$. This will prove useful in Chapter 4 to obtain couplings using some extra information outside of X^2, and ultimately obtain upper bounds on d_B. ■

Lemma 2.29. *The function d_B is a distance on $\mathcal{M}_\sigma(X)$, and $d_B(\mu,\nu)$ is always* reached *for some coupling between the measures.*

Proof. The function d_B is trivially symmetric, and $d_B(\mu,\mu) = 0$ for any measure $\mu \in \mathcal{M}_\sigma(X)$.

To prove the triangle inequality, consider three measures $\mu_1, \mu_2, \mu_3 \in \mathcal{M}_\sigma(X)$. Consider a coupling $\lambda_{1,2}$ between μ_1 and μ_2 (resp. $\lambda_{2,3}$ between μ_2 and μ_3). The measures $\lambda_{1,2}$ and $\lambda_{2,3}$ are "compatible" in the sense that they share a common projection $\pi^2_*(\lambda_{1,2}) = \pi^1_*(\lambda_{2,3}) = \mu_2$. Thence, it is known [Gla03, Chapter 6] that there exists a joining $\lambda_{1,2,3} \in \mathcal{M}_\sigma\left(X^3\right)$ between them, such that $\left(\pi^1, \pi^2\right)_*(\lambda_{1,2,3}) = \lambda_{1,2}$ and likewise $\left(\pi^2, \pi^3\right)_*(\lambda_{1,2,3}) = \lambda_{2,3}$. In particular, $\lambda_{1,2,3}$ gives us a coupling between μ_1 and μ_3:

$$\begin{aligned} d_B(\mu_1, \mu_3) &\leq \int d_H(x,z)\mathrm{d}\lambda_{1,2,3}(x,y,z) \\ &\leq \int d_H(x,y) + d_H(y,z)\mathrm{d}\lambda_{1,2,3}(x,y,z) \\ &= \int d_H(x,y)\mathrm{d}\lambda_{1,2}(x,y) + \int d_H(y,z)\mathrm{d}\lambda_{2,3}(y,z)\,. \end{aligned}$$

Now, by taking the infimum over all couplings $\lambda_{1,2}$ and $\lambda_{2,3}$, we finally obtain the triangle inequality.

Now, let two measures $\mu, \nu \in \mathcal{M}_\sigma(X)$. The set of couplings between μ and ν is a compact subset of $\mathcal{M}_\sigma\left(X^2\right)$, and the function $\lambda \mapsto \lambda\left([x_0 \neq y_0]\right)$ is continuous in the weak-$*$ topology, so its infimum $d_B(\mu,\nu)$ is reached.

Assume that $d_B(\mu,\nu) = 0$, reached for some coupling λ. Then, λ-a.s., we have $x_0 = y_0$. As λ is shift-invariant, it is more generally true for any position $k \in \mathbb{Z}^d$ that $x_k = y_k$ almost surely. By taking the countable intersection of such events, we have $x = y$ almost surely, so λ is supported on the "diagonal" of X^2. It follows that, for any measurable set $A \subseteq X$:

$$\mu(A) = \int \mathbb{1}_A(x)\mathrm{d}\lambda(x,y) = \int \mathbb{1}_A(x)\mathbb{1}_A(y)\mathrm{d}\lambda(x,y) = \int \mathbb{1}_A(y)\mathrm{d}\lambda(x,y) = \nu(A)\,,$$

so that $\mu = \nu$ (and conversely, distinct measures are distinguishable). □

Remark 2.30 (Union Bounds)**.** What d_H (and incidentally d_B) does is to measure the density of a subset of $\mathbb{Z}^d$. Naturally, we would like to be able to use union bounds to compare the density of several sets.

In this context, a countable union bound is impossible: $\mathbb{Z}^d$ has a density 1 in itself, but is a countable union of cells, each a singleton of density 0.

However, we can still use *finite* union bounds, *i.e.* the density of $A \cup B$ is less than the density of A plus the density of B (and likewise for any finite union by induction). ■

Notably, we use the shift invariance of the measures to prove that d_B distinguishes distinct points. If we try to define d_B in the general case of $\mathcal{M}(X)$, then we obtain instead a mere *pseudometric*, hence shift invariant measures represent a more natural framework to compare the global behaviour.

Lemma 2.31. *The Besicovitch topology is stronger than the weak-$*$ topology.*

Proof. Let n be a window size, and $\mu, \nu \in \mathcal{M}_\sigma(X)$. Let λ be a coupling that realises $d_B(\mu,\nu)$. Then:

$$\sum_{p \in \mathcal{A}^{I_n}} |\mu([p]) - \nu([p])| = \sum_{p \in \mathcal{A}^{I_n}} \left| \int \mathbb{1}_{[p]}(x) - \mathbb{1}_{[p]}(y)\mathrm{d}\lambda(x,y) \right| \leq \int \sum_{p \in \mathcal{A}^{I_n}} \left|\mathbb{1}_{[p]}(x) - \mathbb{1}_{[p]}(y)\right| \mathrm{d}\lambda(x,y)\,.$$

Notice now that $\sum_{p \in \mathcal{A}^{I_n}} \left|\mathbb{1}_{[p]}(x) - \mathbb{1}_{[p]}(y)\right| = 2 \times \mathbb{1}_{[x_{I_n} \neq y_{I_n}]}$, so by union bound and shift invariance of λ we obtain the upper bound $\sum_{p \in \mathcal{A}^{I_n}} |\mu([p]) - \nu([p])| \leq 2\,|I_n|\,\lambda\left([x_0 \neq y_0]\right) = 2d_B(x,y)$. At last, we have the inequality $d^*(\mu,\nu) \leq \sum_{p \in \mathcal{A}^{I_n}} |\mu([p]) - \nu([p])| + \frac{1}{2^n} \leq 2d_B(x,y) + \frac{1}{2^n}$, so the identity map $\mathrm{Id} : (X, d_B) \to (X, d^*)$ is continuous, which concludes the proof. □

Remark 2.32. This topology is actually strictly stronger. For a counter-example, consider the one-dimensional case on $\mathcal{A} = \{\mathtt{0}, \mathtt{1}\}$. Let $\omega_n = (\mathtt{0}^n\mathtt{1}^n)^\infty$ a periodic configuration and $\mu_n = \frac{1}{2n}\sum_{k=0}^{2n-1} \delta_{\sigma^k(\omega_n)}$ the corresponding ergodic measure. Naturally, $d^*\left(\mu_n, \frac{\delta_{\mathtt{0}^\infty} + \delta_{\mathtt{1}^\infty}}{2}\right) = O\left(\frac{1}{n}\right) \to 0$, but $d_B\left(\mu_n, \frac{\delta_{\mathtt{0}^\infty} + \delta_{\mathtt{1}^\infty}}{2}\right) = \frac{1}{2} \not\to 0$. ■

With the Besicovitch topology being stronger, the natural question is whether a sequence of measures converges (or admits adherence values) at all. By extension, this question of stability of SFTs in the Besicovitch topology will be studied extensively in Chapter 4. In this context, we will extend the definition of an SFT to include the concept of *obscured cells* (*i.e.* of noise in the forbidden patterns), and look at the behaviour as the amount of noise goes to 0.

In comparison, in the weak-$*$ topology, we will see that in the framework of Section 4, every SFT is stable, has a defined limit behaviour. Hence, the question shifts to *how* this convergence occurs, which will be the focal point of Chapter 5. In this context, we will look at Gibbs measures instead, which abide by the forbidden patterns more or less depending on the temperature, and look at the behaviour as the temperature of the system goes to 0.

2.7 The Robinson Tiling(s)

Let us conclude this chapter with a detailed example: that of the (aperiodic) Robinson tiling (on the lattice $G = \mathbb{Z}^2$). In this section, we will introduce three variants of the tiling, and discuss their respective properties.

Historically, the first aperiodic SFT was proposed by Berger in 1966 [Ber66], who used 20426 *Wang tiles* (*i.e.* with forbidden patterns deduced from laterally or vertically mismatching interfaces) to encode a hierarchical structure, and thus aperiodicity.

The general idea was strongly refined by Robinson [Rob71] who proposed a Wang tileset with only 56 tiles, which once again forces a hierarchical structure. Equivalently, I will here discuss the (conjugate) SFT detailed in Section 2.7.1, which also uses diagonal interactions, but brings us down to only six geometric tiles (and their rotations and symmetries, for a total of 32). I also refer the reader to lecture notes by Schwartz [Sch07] for an introduction both lengthier that the current overview and more aligned with the current notations than Robinson's seminal paper.

The Robinson tiling has many good properties, which certainly justify why it is so popular and widely studied, but it is not *minimal* (in the sense of set inclusion, which is another conjugacy invariant). To address this point, folkloric extension methods (with supplementary information on the tiles, but that still projects to the original tiling) have been used for a while in the literature [Moz89, Bar14, Gan18]. For related but different reasons that will be made clear when needed, I will use this "enhanced" structure instead of the canonical one, so I will discuss it in Section 2.7.2.

To close this whole chapter, I will lastly discuss another variant, which will be named Red-Black as it uses two alternating colours for the lines drawn on the tiles, which will form increasingly big squares avoiding each other. This structure initially served to simulate Turing machine computations in this liminal space (see Section 3.5), but we will also use it to create unstable behaviours in Chapter 4.

2.7.1 The Canonical Robinson Tiling

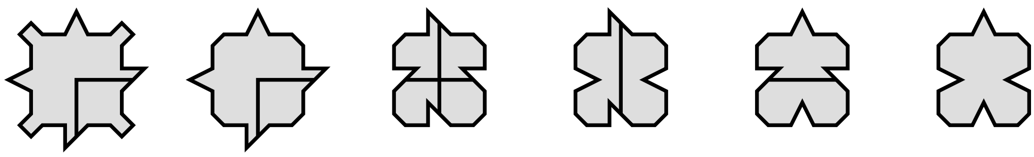

Figure 2.2: The six base Robinson tiles. The leftmost one is called a *bumpy-corner*.

For its alphabet, the canonical Robinson tiling uses the six tiles shown in Figure 2.2, as well as their rotations and symmetries, for a total of 32 tiles. The forbidden patterns naturally follow from the geometric shape of the tiles (the lines play no actual role here except to make the structure visually clearer): two laterally or vertically adjacent tiles must have matching borders, and a square of four tiles must use exactly one tile with *bumpy corners* (*i.e.* a rotation of the leftmost tile in Figure 2.2). This last "diagonal" interaction is the reason why this tileset, thus defined, isn't a set of Wang tiles.

Definition 2.33 (Macro-Tiles)**.** We define the Robinson macro-tiles inductively. First, the 1-macro-tiles are just the tiles with bumpy corners. Inductively, a $(n+1)$-macro-tile is obtained by placing four n-macro-tiles in order to draw a square around a central cross, as shown in Figure 2.3. ■

By a direct induction, an n-macro-tile is a pattern on a $(2^n - 1) \times (2^n - 1)$ square of tiles, and the big central square (formed by the crosses of $(n-1)$-macro-tiles) has a tile-length equal to $2^{n-1} + 1$.

Let us use the *orientation symbols* ◲, ◱, ◰ and ◳ to denote the four ways to fill the central cross of an n-macro-tile. For instance, the 3-macro-tile assembled in Figure 2.3 corresponds to the symbol ◲. More broadly, we will use pairs of symbols (*e.g.* ◱◳) to refer to pairs of macro-tiles *of the same scale* plus a non-specified filling for one-tile-thick interface between them (which may serve as an arm of the cross of a bigger macro-tile).

Lemma 2.34 (Non-Overlapping)**.** *Two macro-tiles cannot* partially *overlap.*

Proof. Assume two Robinson macro-tiles $p \in \mathcal{A}^I$ and $q \in \mathcal{A}^J$ overlap, *i.e.* $I \cap J \neq \emptyset$ and $p_{I \cap J} = q_{I \cap J}$. Without loss of generality, by restricting p or q to a smaller macro-tile within itself, we can assume they are both n-macro-tiles of the same scale.

Then, necessarily $I \cap J$ contains a (non-empty) rectangle grid of matching bumpy-corners, and we can then group these 1-macro-tiles into squares if $n > 1$, or else we would have incompatible alignments for the lines on the tiles (*e.g.* a ◲ tile in p but a ◰ tile in q). Then, by filling their respective central crosses, we conclude that $I \cap J$ contains a rectangle grid of matching 2-macro-tiles, and if $n > 2$ we can then group them into squares using the same orientation argument. Inductively, we repeat this process until we end up with just one whole matching n-macro-tile, *i.e.* $p = q$.

In the general case, the equality becomes of the smaller macro-tile into the bigger one. □

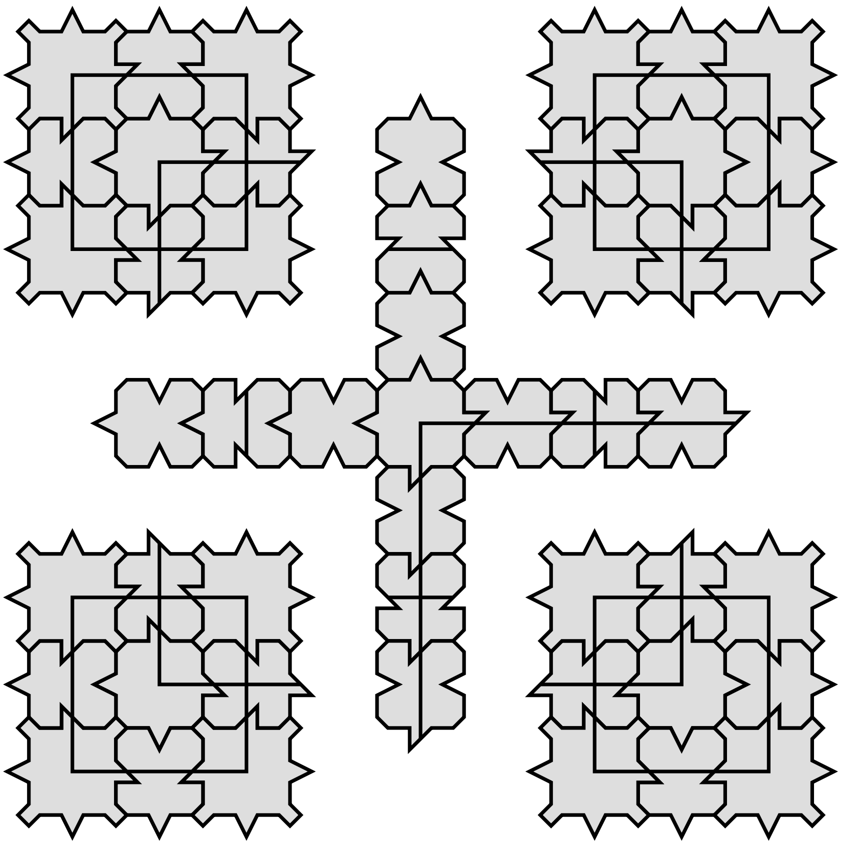

Figure 2.3: Four 2-macro-tiles around a central cross form a 3-macro-tile.

Remark 2.35 (∞-Macro-Tiles and Almost Periodicity)**.** Let ω be a globally admissible Robinson tiling.

Inductively, a given bumpy-corner (or more generally a macro-tile) must be inside arbitrarily large(r) macro-tiles. At the limit, we thus obtain an infinite area (either the whole lattice $\mathbb{Z}^2$, a half plane or a quarter plane), which we will informally call an "∞-macro-tile". Structurally, ω must be made of either just one area, four area around a central cross (in a way analogous to Figure 2.3), or two of them on both sides of an infinite line which we will call a *cut*.

In each such ∞-macro-tile, any two windows must be included within the same macro-tile after some rank. What's more, noticing that (the restriction to) n-macro-tiles behave 2^{n+1}-periodically in any bigger N-macro-tile (with $N > n$), we conclude that this result also holds in any ∞-macro-tile. In some sense, though the tilings are (strongly) aperiodic, they have an *almost* periodic structure.

However, as any ∞-macro-tile contains areas (*i.e.* the grids of n-macro-tiles) that behave with an arbitrarily high minimal period (and the overlapping property tells us that no other translation of an n-macro-tile is possible), we conclude that the ∞-macro-tile has no global periodic behaviour, and that any tiling ω is thus (strongly) aperiodic. ■

Proposition 2.36 (Unique Ergodicity)**.** *The Robinson tiling is uniquely ergodic.*

Proof. Let X be the set of Robinson tilings and $\omega \in X$. In each ∞-macro-tile of ω, n-macro-tiles behave periodically for any scale $n \in \mathbb{N}$ (with the same periods regardless of the ∞-macro-tile). It follows that the empirical distribution of any pattern $p \in \mathcal{A}^I$ converges to the same limit $\mu([p])$ regardless of the configuration ω, so $\mu_\omega = \mu$, $\mathcal{M}_\sigma(X) = \{\mu\}$ is a singleton. □

A property related to unique ergodicity is minimality of an SFT, which generalises the case where X is just the finite orbit of one single periodic configuration. The topic bears no weight on my work later on, and is well documented [FM10], so I will just give a few key ideas about it.

Definition 2.37 (Minimal SFT)**.** A subshift X is minimal (for the inclusion) if there is no strictly smaller (and non-empty) subshift $Y \subsetneq X$. ■

Among other properties, we have that if X is uniquely ergodic, with μ its ergodic measure, then the subshift $Y := \operatorname{supp}(\mu) \subseteq X$ is minimal (*and* has μ as its unique ergodic measure). This is just an inclusion, however, and the canonical Robinson tileset is a counter-example for the equality.

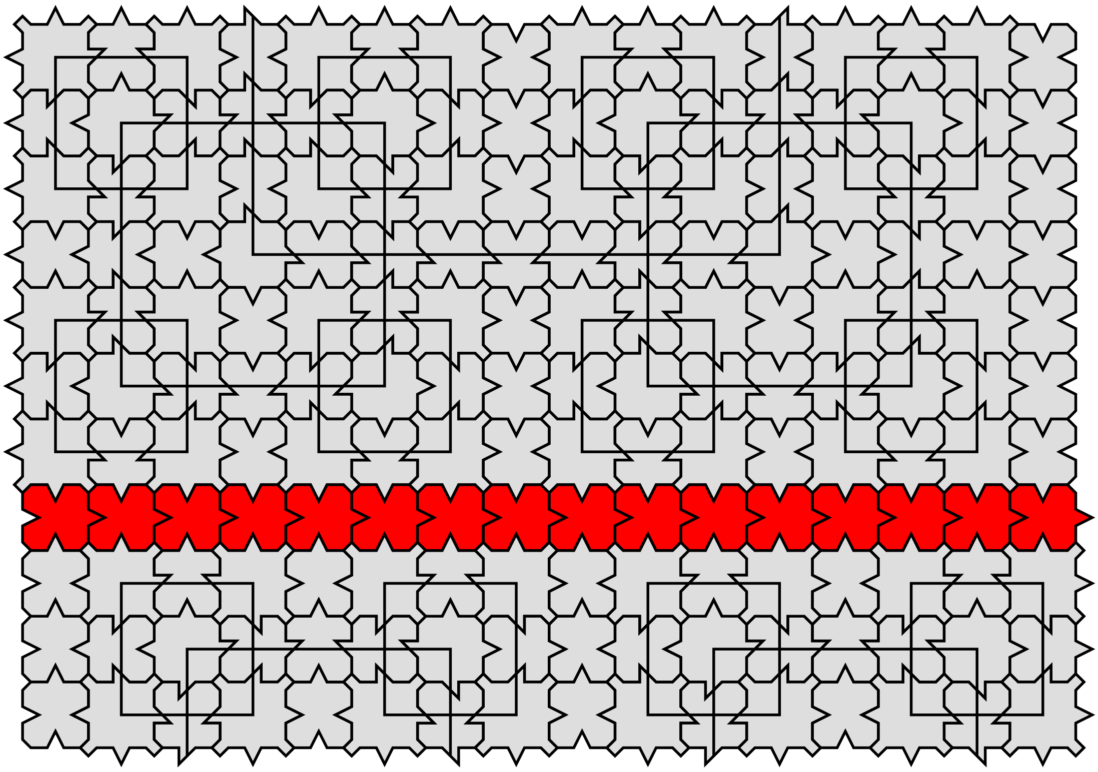

Figure 2.4: A horizontal cut, highlighted in red, between two misaligned ∞-macro-tiles.

Indeed, let $Y \subsetneq X$ be the set of Robinson tilings for which, at any scale $n \in \mathbb{N}$, n-macro-tiles behave periodically *globally* (instead of simply *within* each of the one, two or four ∞-macro-tile). Then Y is a strict subshift of X, as it doesn't contain configurations with a *misaligned cut* like in Figure 2.4 (but it's not an SFT anymore).

2.7.2 The "Enhanced" Robinson Tiling

Let's now see how we can tinker a bit with the construction of the Robinson tileset to obtain a uniquely ergodic, minimal, aperiodic but (globally) almost periodic SFT.

The issue with the Robinson tileset was that it allowed a *misaligned cut* to occur. This misalignment can of course also occur at a local scale, for locally admissible tilings, which would prove inconvenient for my stability result in Chapter 4. Conversely, to erase the misaligned cuts out of existence, we simply need to prevent them from happening at a purely local scale. We can do so by *adding* information over the canonical tiles.

More precisely, within the central cross of any macro-tile, we want to send a "signal" that communicates its orientation. This was already done in two of the four directions, as asymmetrical notch on the right side of a ◲ tile is not the same as that of a ◳ tile (they actually mirror each other), hence why a ◲◳ pattern cannot occur. However, the two other directions have the same symmetrical notch, so a ◱◳ pattern can occur for the canonical Robinson, which ultimately permits misaligned cuts.

Hence, we let the *enhanced* Robinson tileset illustrated in Figure 2.5, by adding dashed red lines (representing the top direction of a ◲ macro-tile) and dotted blue lines (representing the left direction of a ◲ macro-tile). The complete alphabet also includes the *rotations* of the two leftmost tiles (but *not* their symmetries, which would erase the orientation encoded in dotted and dashed lines), as well as the rotations and symmetries of the seven other ones (for which the top and bottom row are colour-swaps of each other, *i.e.* red dashed lines become blue dotted ones and conversely), for a total of 56 tiles.

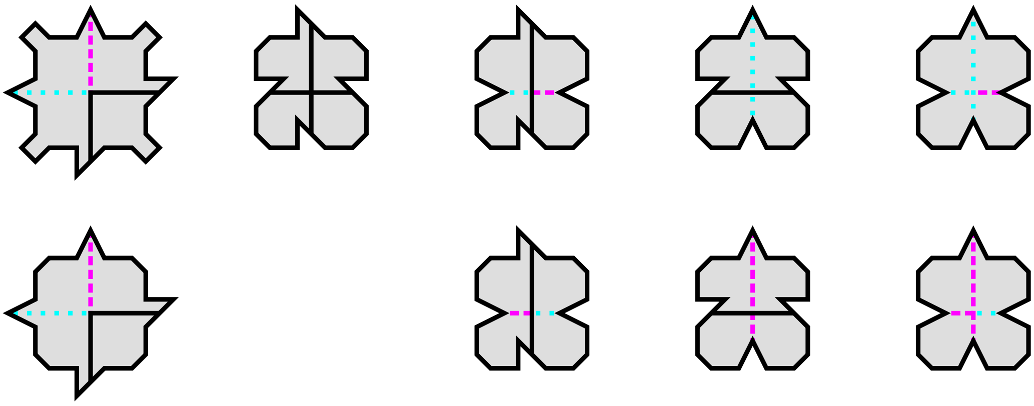

Figure 2.5: The nine base enhanced Robinson tiles.

As for the canonical Robinson tiling, the forbidden patterns here are induced by the self-evident matching rules between neighbouring tiles. Notably, the two rightmost tiles will enforce a complementary matching rule, so that the only locally admissible pairs of (macro-)tiles we can form are , , and . It follows that a misaligned cut *cannot* happen with the enhanced tileset. An illustration of how these local rules interact can be found in Figure 2.6, where the 3-macro-tile is shown.

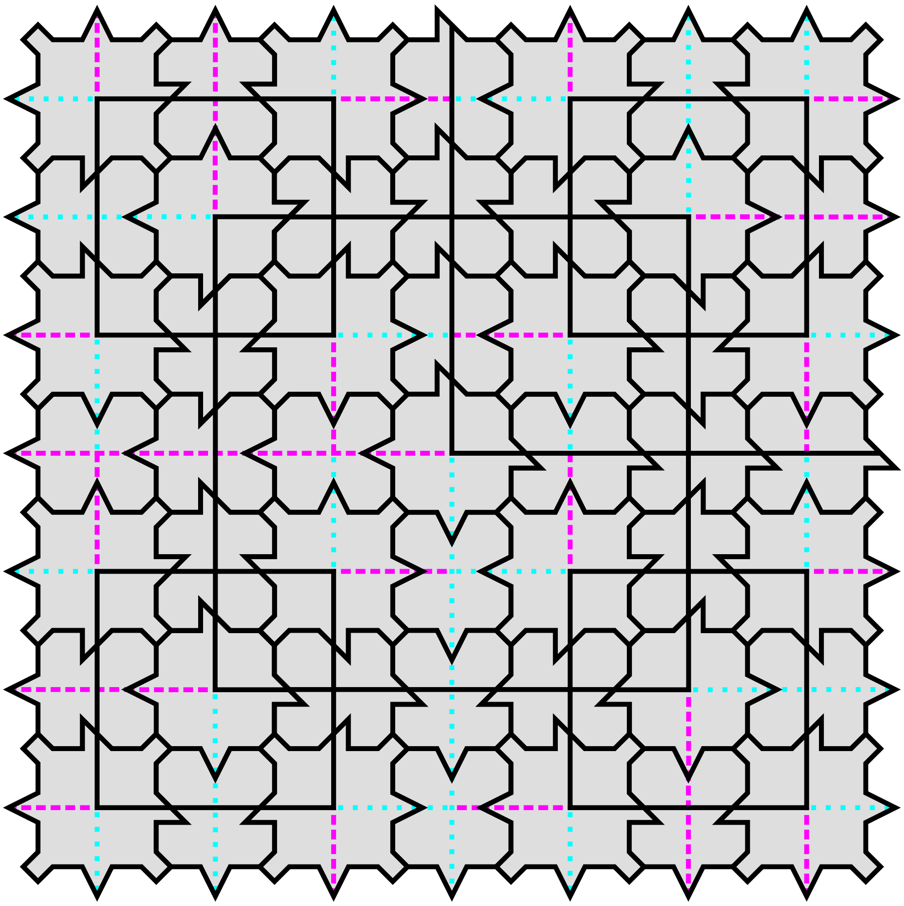

Figure 2.6: A 3-macro-tile with the enhanced tileset.

By using a mere projection that erases dotted and dashed lines, in a way compatible with the local rules on both ends, we have a morphism that sends any enhanced tiling onto a canonical one. We can thus transpose many basic structural properties from the canonical case to the enhanced one, such as non-overlapping macro-tiles and a hierarchical structure that leads to ∞-macro-tiles.

This projection can be partially reversed, in that there is a one-to-one correspondence between macro-tiles (up to and including ∞-macro-tiles). However, this morphism is neither injective nor surjective. The non-injectivity comes from the fact that, in the case of an infinite (well-aligned) cut, we map a red dashed line and a blue dotted line filling the cut to the same pattern. The non-surjectivity comes from misaligned cuts, that cannot be realised in the enhanced case.

Theorem 2.38. *The enhanced Robinson tiling is uniquely ergodic and minimal.*

Proof. Let X be the set of enhanced Robinson tilings. The unique ergodicity comes from the exact same argument on ∞-macro-tiles as for the canonical case. Now that we have eliminated misaligned cuts, however, *any* globally admissible pattern on a finite window occurs in *every* configuration $\omega \in X$, so the adherence of the orbit of ω is X itself, hence minimality. □

While the previous morphism is neither injective nor surjective, if we restrict it to configurations made of *only one* ∞-macro-tile (which happens with probability 1 for the unique invariant measure on both ends) then we obtain a bijection.

This self-aligning behaviour will prove useful in various ways later on, but I will stop the rough overview here and detail these properties in Chapters 4 and 5 as the need arises.

2.7.3 The Red-Black Robinson Tiling

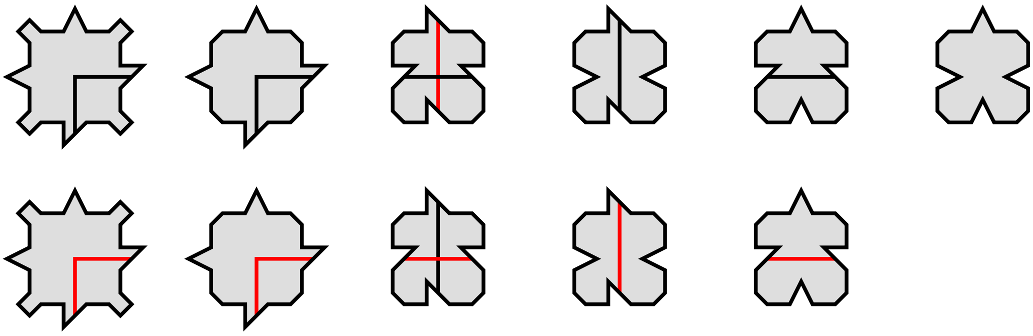

Figure 2.7: The eleven base Red-Black Robinson tiles.

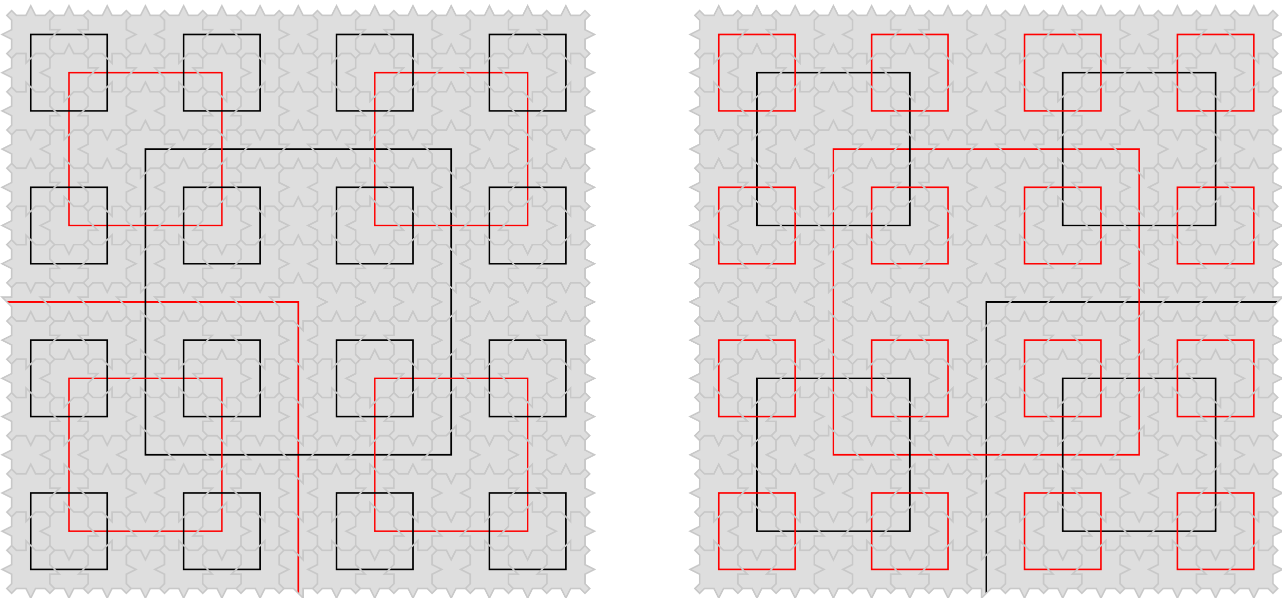

Figure 2.8: Alternating colours in Red-Black macro-tiles.

A last idea I want to introduce in this section is that we can also use the existing structure to encode more information (instead of adding supplementary structural constraints like in the enhanced case). In all generality, we can use the self-similar structure of the Robinson macro-tiles and this encoded supplementary information to simulate any substitution subshift [Moz89].

We will here describe a really basic example, with just one bit of information (which we will represent as either Red or Black lines), but more complex variants will appear in the next chapters.

Consider the tileset in Figure 2.7 (and their rotations and symmetries), and the forbidden patterns induced by the matching rules. Notably, the third column tells us that a line at a given scale (the horizontal one in the figure) must with a line encoding the other bit at the next scale of macro-tiles (the vertical one in the figure). In other words, this tileset simulates the following two-dimensional substitution scheme on $\{\mathtt{0}, \mathtt{1}\}$:

$$\mathtt{0} \mapsto \begin{smallmatrix} \mathtt{1} & \mathtt{1} \\ \mathtt{1} & \mathtt{1} \end{smallmatrix} \text{ and } \mathtt{1} \mapsto \begin{smallmatrix} \mathtt{0} & \mathtt{0} \\ \mathtt{0} & \mathtt{0} \end{smallmatrix}.$$

For this example, all the structural properties of the canonical Robinson tiling hold, including non-overlap of macro-tiles, aperiodicity and the idea of ∞-macro-tiles.

However, we now lose unique ergodicity. Instead of just having four orientations for the central cross of the n-macro-tiles that otherwise coincide everywhere, we now have two kinds of mostly different n-macro-tiles, Red or Black, depending on the colour of the bumpy-corners they use. In Figure 2.8, we have a Black ◰ 4-macro-tile on the left, and a Red ◲ 4-macro-tile on the right. Inductively, in a Red n-macro-tile, the central cross either has Red arms if n is odd, or Black ones if n is even. At the limit, we conclude that within an ∞-macro-tile, any two n-macro-tiles have the same colour pattern. Hence, a Red ∞-macro-tile induces a distribution $\mu_{\mathtt{R}}$ for which bumpy-corners are almost surely Red. Likewise, a Black ∞-macro-tile induces $\mu_{\mathtt{B}}$ for which bumpy-corners are almost surely Black, and $\mathcal{M}_\sigma(X) = [\mu_{\mathtt{R}}, \mu_{\mathtt{B}}]$ is an interval with two extremal points.

To obtain a uniquely ergodic SFT, it suffices to use only *one* of the two bumpy-corners in Figure 2.7 to define the alphabet.

References

[ABH23] Nathalie Aubrun, Nicolás Bitar, and Sacha Huriot-Tattegrain. "Strongly Aperiodic SFTs on Generalized Baumslag-Solitar groups". In: *Ergodic Theory and Dynamical Systems* (2023), pp. 1–30. 10.1017/etds.2023.44.

[AK13] Nathalie Aubrun and Jarkko Kari. "Tiling Problems on Baumslag-Solitar Groups". In: *MCU*. Vol. 128. Electronic Proceedings in Theoretical Computer Science. 2013, pp. 35–46. 10.4204/EPTCS.128.12.

[Bar14] Sebastián Barbieri. *Subshifts generados por sustituciones multidimensionales*. 2014. repositorio.uchile.cl/handle/2250/116840.

[Bar23] Sebastián Barbieri. "Aperiodic Subshifts of Finite Type on Groups Which Are Not Finitely Generated". In: *Proceedings of the American Mathematical Society* (2023). 10.1090/proc/16379.

[BC21] Jung-Chao Ban and Chih-Hung Chang. "On the Topological Entropy of Subshifts of Finite Type on Free Semigroups". In: *Taiwanese Journal of Mathematics* 25.3 (2021), pp. 579–598. 10.11650/tjm/201210.

[BDJ08] Alexis Ballier, Bruno Durand, and Emmanuel Jeandel. "Structural Aspects of Tilings". In: *STACS*. Vol. 1. 2008, pp. 61–72. 10.4230/LIPIcs.STACS.2008.1334.

[Ber17] Anton Bernshteyn. *Free Subshifts With Invariant Measures From the Lovász Local Lemma*. 2017. 10.48550/arXiv.1702.02792.

[Ber66] Robert Berger. *The Undecidability of the Domino Problem*. Memoirs of the American Mathematical Society 66. AMS, 1966. 10.1090/memo/0066.

[EM22] Solène Esnay and Etienne Moutot. "Aperiodic SFTs on Baumslag-Solitar Groups". In: *Theoretical Computer Science* 917 (2022), pp. 31–50. 10.1016/j.tcs.2022.03.010.

[FM10] Sébastien Ferenczi and Thierry Monteil. "Infinite Words With Uniform Frequencies, and Invariant Measures". In: *Combinatorics, Automata and Number Theory*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2010. Chap. 7, pp. 373–409. 10.1017/CBO9780511777653.008.

[Gan18] Silvère Gangloff. "Algorithmic Complexity of Growth-Type Invariants of SFT Under Dynamical Constraints". PhD thesis. Université d'Aix-Marseille, 2018. theses.fr/2018AIXM0231.

[Gla03] Eli Glasner. *Ergodic Theory via Joinings*. 101. AMS, 2003. 10.1090/surv/101.

[GS87] Branko Grünbaum and Geoffrey Shephard. *Tilings and Patterns*. Freeman and Co., 1987.

[Hed69] Gustav Hedlund. "Endomorphisms and Automorphisms of the Shift Dynamical System". In: *Mathematical Systems Theory* 3 (1969), pp. 320–375. 10.1007/BF01691062.

[Jea15] Emmanuel Jeandel. *Aperiodic Subshifts of Finite Type on Groups*. 2015. 10.48550/arXiv.1501.06831.

[Kar96] Jarkko Kari. "A Small Aperiodic Set of Wang Tiles". In: *Discrete Mathematics* 160 (1996), pp. 259–264. 10.1016/0012-365X(95)00120-L.

[Kel98] Gerhard Keller. *Equilibrium States in Ergodic Theory*. Vol. 42. London Mathematical Society Student Texts. Cambridge University Press, 1998. 10.1017/CBO9781107359987.

[Kre85] Ulrich Krengel. *Ergodic Theorems*. De Gruyter, 1985. 10.1515/9783110844641.

[LM21] Douglas Lind and Brian Marcus. *An Introduction to Symbolic Dynamics and Coding*. 2nd edition. Cambridge Mathematical Library. Cambridge University Press, 2021. 10.1017/9781108899727.

[Mou85] Jean Moulin Ollagnier. *Ergodic Theory and Statistical Mechanics*. Vol. 1115. Lecture Notes in Mathematics. Springer, 1985. 10.1007/BFb0101575.

[Moz89] Shahar Mozes. "Tilings, Substitution Systems and Dynamical Systems Generated by Them". In: *Journal d'Analyse Mathématique* 53 (1989), pp. 139–186. 10.1007/BF02793412.

[Rob71] Raphael Robinson. "Undecidability and Nonperiodicity for Tilings of the Plane". In: *Inventiones mathematicae* 12 (1971), pp. 177–209. 10.1007/BF01418780.

[Sch07] Richard Schwartz. *The Robinson Tiles*. 2007. math.brown.edu/reschwar/MFS/handout9.pdf.

[ST16] Brandon Seward and Robin Tucker-Drob. "Borel Structurability on the 2-Shift of a Countable Group". In: *Annals of Pure and Applied Logic* 167.1 (2016), pp. 1–21. 10.1016/j.apal.2015.07.005.

[Tem92] Arkady Tempelman. *Ergodic Theorems for Group Actions*. Vol. 78. Mathematics and its Applications. Kluwer Academic Publishers, 1992. 10.1007/978-94-017-1460-0.

Chapter 3

On Computable Analysis

A focal point of my work in the next chapters relates to *computations*, either by embedding them into the tilings or by wondering if some dynamical property can be checked by a computer, oftentimes both at once.

Hence, in this chapter, I introduce all the necessary framework to do so, as concisely as reasonable, starting from the fundamental concept of Turing machines in Section 3.1. Initially introduced by Alan Turing in 1936 [Tur36], the topic has since been discussed up and down in the literature, and I will simply direct the reader to the book by Sipser [Sip13] for an extensive introduction, among many other equally good works. Section 3.3 focuses on the specific case of unary computations, where the string $\mathtt{1}^n$ encodes the integer $n \in \mathbb{N}$.

Most of what is introduced here will relate to questions of *decidability*, *i.e.* of whether a task *can* be solved algorithmically in some sense. Notably, in Section 3.4, I introduce the arithmetical hierarchy as a way to classify the hardness of undecidable problems, where we cannot apply intuitive notions of time complexity. I refer the reader to the classical books by Rogers [Rog87] or Soare [Soa87, Soa16] for much more details on this topic, or the mathematical blog Rising Entropy [Ris20] for an informal introduction. Then, I intertwine this notion with the framework of *computable analysis* in Section 3.6, which aims at studying the computable properties of uncountable spaces such as real numbers or invariant measures. This topic is not exactly textbook material yet, but I direct the interested reader to broader introductions to the general framework [Wei00, BHW08], or its implementation for the study of Julia sets [BY09], of shift-invariant measures [GHR11, HS18], of measurable functions [Wei17] and of Cantor sets [Gan+20]. This framework will naturally appear and prove itself necessary to obtain complexity bounds, both in Chapter 4 and Chapter 5.

While the focus of my work is decidability, it's still important to spend some time discussing *complexity* (*i.e.* how *fast* can a decidable task be solved) in Section 3.2. Notably, this will play a role when simulating Turing machines within other tilings, with an unmovable time limit, as explained in Section 3.5. This notion of simulating tileset is an intrinsic part of Robinson's seminal paper [Rob71], but I refer the reader to the lecture notes by Jeandel and Vanier [JV20] for an introduction more in line with the current vocabulary.

3.1 Turing Machines

Let's begin with the basic notions regarding Turing machines and computability.

Definition 3.1 (Turing Machine)**.** A Turing machine M is a tuple $(Q, q_0, Q_A, Q_R, \mathcal{A}, \Gamma, \delta)$ where:

- Q is the finite set of internal states of the machine. In particular, q_0 is its initial state, Q_A the set of *accepting* states and Q_R the set of *rejecting* states (with $Q_A \cap Q_R = \emptyset$).
- $\mathcal{A}$ is the input alphabet of the machine, and Γ the *tape* alphabet. In particular, we have a *blank symbol* $\#$, not in the input alphabet, such that $\mathcal{A} \sqcup \{\#\} \subseteq \Gamma$.
- $\delta : Q \times \Gamma \to Q \times \Gamma \times \{\mathrm{L}, \mathrm{R}\}$ is the *transition* function.

The machine M receives a finite word $w_0 \in \mathcal{A}^*$ as its input, and then follows through a trajectory according to its transition function δ, until it possibly reaches a halting state in $Q_A \sqcup Q_R$. More precisely, initially M is in a configuration $(q_0, w_0, 0)$, and we have the transitions $(q_t, w_t, n_t) \mapsto (q_{t+1}, w_{t+1}, n_{t+1})$ defined inductively by $(q_{t+1}, b, D) := \delta(q_t, w_t[n_t])$, w_{t+1} is equal to w_t except in position n_t where $w_{t+1}[n_t] = b$, and $n_{t+1} = n_t + 1$ when $D = \mathrm{R}$ or $n_{t+1} = n_t - 1$ when $D = \mathrm{L}$ (with the convention that $0 - 1 = 0$ in $\mathbb{N}$), as illustrated in

Figure 3.1. If the machine reaches a halting state $q_t \in Q_A \sqcup Q_R$, then we simply stall the trajectory, *i.e.* $(q_{t+1}, w_{t+1}, n_{t+1}) := (q_t, w_t, n_t)$.

We can then see M as a partial function $\mathcal{A}^* \to \{\mathtt{0}, \mathtt{1}\}$ so that $M(w) = \mathtt{1}$ when the trajectory reaches an accepting state, $M(w) = \mathtt{0}$ when the trajectory reaches a rejecting state, and $w \notin \operatorname{supp}(M)$ if the trajectory doesn't halt. ■

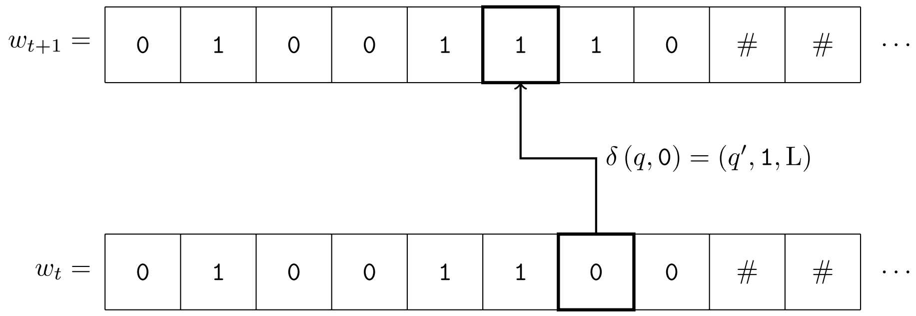

Figure 3.1: Illustration of the transition from $(q, w_t, 6)$ to $(q', w_{t+1}, 5)$ in a Turing machine.

Remark 3.2 (Multi-Tape Turing Machine)**.** In some situations, it may be helpful to picture a Turing machine with *several* tapes, each storing its own data, *e.g.* a *read-only* input tape, or a *write-only* output tape. Then, instead of giving a single direction $D \in \{\mathrm{L}, \mathrm{R}\}$, the transition function δ must return one direction per tape (we say the tapes are *asynchronous*).

If the tapes are *synchronous* (*i.e.* δ gives the same direction on each tape), then instead of having k tapes each with its own alphabet Γ_i, we can equivalently consider one single tape with the alphabet $\Gamma = \prod_{i=1}^{k} \Gamma_i$ instead. Still, picturing this as a multi-tape machine may prove helpful, in particular if *some* coordinates are read-only, which will be the case in some layers of the simulating tileset from Chapter 5. ■

Definition 3.3 (Computable Function)**.** We can adapt the previous definition to see M as a map $\mathcal{A}^* \to \mathcal{B}^*$ instead, by requiring that when M halts at time t, we also we have $n_t = 0$ and $w_t \in \mathcal{B}^*$ (so that function composition is easy to define by simply executing two machines one after another on the same tape).

Hence, a map $f : \mathcal{A}^* \to \mathcal{B}^*$ is said to be *computable* if we have a Turing machine M (with total support) that computes it, *i.e.* $M(w) = f(w)$ for any $w \in \mathcal{A}^*$. ■

Definition 3.4 (Computability on Sets)**.** Let $S \subseteq \mathcal{A}^*$. We say that S is a *decidable* (or *computable*) set if its indicator function $\mathbb{1}_S : \mathcal{A}^* \to \{\mathtt{0}, \mathtt{1}\}$ is itself computable.

We say that S is *recursively enumerable iff* there is a Turing machine $M : \mathcal{A}^* \to \{\mathtt{0}, \mathtt{1}\}$ (with partial support) such that for any $n \in S$, we have $n \in \operatorname{supp}(M)$ and $M(n) = \mathtt{1}$ (but we may have $n \notin \operatorname{supp}(M)$ if $n \notin S$). ■

Remark 3.5 (Computability on Countable Spaces)**.** Let $\mathcal{A}$ be a finite alphabet. Any countable space X, such as $\mathbb{N}$ or $\mathbb{Q}$, can be injected into the set of finite words $\mathcal{A}^*$. Such a map $\gamma : X \hookrightarrow \mathcal{A}^*$ is called an *encoding*.

Note that γ may not be a surjective map. For instance, with the canonical binary encoding $\gamma : \mathbb{N} \to \{\mathtt{0}, \mathtt{1}\}^*$, the strongest bit of $\gamma(n)$ is always equal to $\mathtt{1}$ (except for $\gamma(0) = \varepsilon$). However, we usually operate under the assumption that $\gamma(X)$ is a decidable set, that we can algorithmically identify invalid codings.

We want to perform computations on X just like we defined them on $\mathcal{A}^*$. Note that the set of bijective maps $X \to X$ is uncountable, but the set of computable maps is countable, so we may have several non-equivalent notions of what a computable subset of X is and so on. However, in practice, any two “reasonable” explicit encodings γ and γ' of X are equivalent (*i.e.* we have a computable bijective map $f : \mathcal{A}^* \to \mathcal{A}^*$ such that $\gamma' = f \circ \gamma$). Under this general philosophy, we will directly talk about computable maps $f : X \to X'$ without specifying encodings $\gamma : X \to \mathcal{A}^*$ and $\gamma' : X' \to \mathcal{B}^*$. By extension, we may also consider encodings $\gamma : X \to \mathbb{N}$ (also called *enumerations*) when needed, without ever really describing what is the input alphabet $\mathcal{A}$. ■

Hence, in this context, when we speak about a *decision problem* (such as *is this SFT stable?* in Chapter 4, or *is this potential chaotic?* in Chapter 5) we are actually studying the computational properties of the (countable) subset of encodings of elements for which the answer to the problem is *yes*.

Among other spaces, the set of Turing machines is itself countable. Hence, we can use an encoding $M \mapsto \langle M \rangle$, and use machines to do computations *on* other machines. Notably, we have the following:

Theorem 3.6 (Universal Turing Machine)**.** *There exists a* universal *Turing machine U, that simulates any machine M on any input w, so that we have an encoding $\langle M, w\rangle \in \mathrm{supp}(U)$ iff $w \in \mathrm{supp}(M)$, and then $U(\langle M, w\rangle) = M(w)$.*

Let's conclude the section on a more philosophical note, that relates Turing machines to other models.

Remark 3.7 (Church-Turing Thesis)**.** The thesis can be summarised as follows:

Any intuitively "effective" model of computation produces Turing-computable functions.

This includes strictly less expressive models, such as regular expressions or context-free grammars, but also many models that have been proven to be *equivalent* to Turing machines (*i.e.* any Turing-computable function can also be computed by the model), which we call *Turing-complete* models. We can notably think of general recursive functions, λ-calculus, random-access machines, virtually any programming language run on a real-life computer, or even *Microsoft PowerPoint* [Wil17]. ■

3.2 Time Complexity

While not the focal point of my work in the next chapters, it is necessary to understand basic notions relating to *complexity* when looking at Turing machines *simulated* in tilings (see Section 3.5), as they only have access to a finite amount of time steps to conclude their computations in this context.

Definition 3.8 (Time Complexity)**.** Let $T : \mathbb{N} \to \mathbb{N}$. We say that a Turing machine M has a time complexity of (at most) T if, for any input $w \in \mathcal{A}^*$, the computation of $M(w)$ halts in *at most* $T(|w|)$ steps. ■

Remark 3.9 (Computations by Blocks)**.** Let M be a machine on an alphabet $\mathcal{A}$ with complexity T. By using blocks of length k, we can hardcode the result of all the consecutive operations that M performs within the block, which gives us a machine M' on the input alphabet $\mathcal{A}^k$ (and likewise the tape alphabet Γ^k with the blank symbol $\#^k$) which has a complexity $\left\lceil \frac{1}{k} T \right\rceil$.

Conversely, we can always convert M to a machine using a binary alphabet. Indeed, we can encode any input letter in $\mathcal{A}$ as a binary string of length $\ell = \lceil \log_2(\mathcal{A}) \rceil$, and then hardcode a block decoder machine M' that reads the ℓ bits, writes the corresponding letter $a \in \mathcal{A}$ in its internal state, and then edits the whole block and performs the transition M would do in this situation, which takes a total of 3ℓ steps. Hence, M' has a complexity of order $n \mapsto 3\ell T\left(\frac{n}{\ell}\right)$, which is a $O(T)$ assuming for example that T is non-decreasing.

Note that, in both cases, the equivalent machine M' performs a simulation of M inside itself, either by using a bigger alphabet or supplementary internal states, which in both cases results in a *more complex* machine (*i.e.* $|\langle M' \rangle| > |\langle M \rangle|$) virtually solving the same problem or computing the same function. ■

Except for particularly fine optimality results (such as computing the minimal number of comparisons a sorting algorithm will perform), and for upper bounds in particular, we usually consider complexity $O(T)$ instead of T, *i.e.* up to a *multiplicative constant.*

However, in the case of computations simulated inside tilings, the size of the space-time diagram is fixed, not up to a constant. In this case, it may be simpler to simply require a time complexity $o(T)$, where T is the time available for inputs of a certain size. The matter of time is thus trivially avoided *after a rank*, but there is some more tinkering to be done through hardcoding to deal with computations on small inputs which may take too long.

Lemma 3.10 (Speeding-up the Machine for Small Inputs)**.** *Assume M decides a problem with complexity T* after a rank*. Then we have M' that decides the same problem in time $n \mapsto T(n) + 2n$ at any rank.*

Proof. To do so, consider X the finite set of inputs w of inputs for which M takes more than $T(|w|)$ steps to terminate. We can hardcode the finite language X into an automaton. Then, M' begins by running the automaton on the input (in linear time), and then immediately accept or reject the input w if $w \in X$ (depending on what M does), or goes back to the beginning of the tape (in linear time) and then runs M (with complexity T by assumption). □

A similar argument holds for machines that compute a function, but the supplementary term $2n$ must be replaced by $2M^*(n)$ with $M^*(n) := \max_{w \in \mathcal{A}^n}(n, |M(w)|)$ the time it takes to guarantee we can read the input or write the output.

Proposition 3.11. *Assume M computes a function in time $o(T)$, with $2M^* \leq T$ and $M^* = o(T)$. Then we have M' that computes the same function, still in time $o(T)$ asymptotically, but now terminates in* at most *T steps on any input.*

Proof. We follow the same idea as in the previous lemma. This time, let $T' = o(T)$ be the actual time complexity of M. Notably, $T' + 2n = o(T)$, so the set X of inputs of lengths such that $T'(n) + 2n > T(n)$ is finite. Inside X, we can simply hardcode M into M' as before, so that M' computes $M(w)$ in exactly $2M^*(|w|) \leq T(|w|)$ steps. Outside of X, after taking $2n$ steps to notice that $w \notin X$ and reset the machine, we simply execute M, which will take at most $T'(|w|)$ steps, hence $T'(|w|) + 2|w| \leq T(|w|)$ in total. Thus, for M', we always terminate in at most T steps, and have a complexity $T' + 2n = o(T)$ asymptotically. □

3.3 Unary Computing

For non-trivial behaviours to emerge, Turing machines need to have a tape alphabet with at least two letters. As discussed in the previous section, the input alphabet can always be brought down to $\{\mathtt{0}, \mathtt{1}\}$ without significant losses of performance (with the same time complexity up to a multiplicative constant), which then gives us a standard "unified" framework to compare the *Kolmogorov complexity* of *a priori* unrelated objects and problems.

However, in some situations, it makes sense to use a unary encoding $n \in \mathbb{N} \mapsto \mathtt{1}^n$. One reason can be to shed another light on matters related to complexity, such as pseudo-polynomial algorithms, as the *size* of the input $n \in \mathbb{N}$ exponentially explodes from $\log_2(n)$ to n (*e.g.* for the naive primality check which simply tests any divisor until it reaches $\sqrt{n}$, in polynomial time with respect to n but definitely not $\log_2(n)$). Another reason may be to have a simpler structure of the input that allows us to perform a kind of intuitive "stick-based" computing, where for instance division by n simply amounts to *removing* all-but-one stick from each bundle of size n, and then merging the leftover sticks into a bundle again. This argument is our motivation for using unary computing in the simulating tilesets in Chapter 5, as they encode *finite* but big enough space-time diagrams for our purpose (see Section 3.5 for an introduction to simulating tilesets).

Proposition 3.12 (Integer Comparison)**.** *Given two integers $i, j \in \mathbb{N}$, encoded in two synchronised tapes, testing $i = j$ (resp. $i < j$) can be done in exactly $\min(i, j) + 1$ steps.*

Proposition 3.13 (Power Checking)**.** *Let $b \geq 2$ and n be two integers. We want here to decide whether $n = b^k$ for some power $k \in \mathbb{N}$. If we assume b is fixed and hardcoded within the machine, there exists a Turing machine M_b with an asymptotic time complexity of order $n \log_b(n)$.*

Proof. To do so, we use a Turing machine that does *back-and-forth* passes on the area of length n. At each pass, the machine erases the first $b - 1$ sticks out of b, cyclically. At the end of a pass, if we have only seen one stick in total, we accept n. Else, if we have not seen a multiple of b sticks on the whole line (*i.e.* we are in the middle of a cycle of erasure), we reject it. In the remaining case, we were able to divide by b the number of sticks, and go on for another pass in the other direction.

Because each pass (which from end to end takes about n steps) divides the number of sticks by b, we can do at most $\log_b(n)$ passes before we trigger a halting condition, hence the announced complexity. □

Consider now a really similar matter, but instead of studying a decision problem we want to compute a function. In order to be able to compose functions, we suppose here that M terminates its computation only once the reading head reaches the starting position.

Proposition 3.14 (Rounded Logarithm)**.** *Let $b \geq 2$ and n be two integers. We want here to compute $\lfloor \log_b(n) \rfloor$. If we assume b is fixed and hardcoded within the machine, there exists a Turing machine M_b that computes this function with an asymptotic time complexity of order $n \log_b(n)$.*

Proof. Just like for the Power Checking, we do several passes of length (at most) n, so that the number of sticks on the tape goes from n_t to $n_{t+1} = \left\lfloor \frac{n_t}{b} \right\rfloor$, and for each pass such that $n_t > 1$, we add a stick to a second coordinate during the next pass. We then simply conclude by erasing what remains on the first coordinate. □

Proposition 3.15 (Nearest Power)**.** *Let $b \geq 2$ and n be two integers. We want here to compute $b^{\lfloor \log_b(n) \rfloor}$, the biggest power of b that is lower or equal to n. If we assume b is fixed and hardcoded within the machine, there exists a Turing machine M_b that computes this function with an asymptotic time complexity of order $2n \log_b(n)$.*

Proof. We follow here the same general idea, but we do instead a full back-and-forth for each big step. At each such step, we begin by keeping the first stick and rewriting the following $b-1$ we encounter using another symbol `$`. If at the end of the pass we found only one stick, then while going back we revert to its original value every symbol `$` we encounter. Else, if we did not encounter a multiple of b sticks during the pass, then while going back, we erase entirely every symbol `$` we encounter until we reach a stick which we also erase, and then simply go for another back-and-forth.

To put it in another way, at the t-th step we check the t-th digit expansion of n in base b, and set it to 0 if needed. □

3.4 The Arithmetical Hierarchy

So far, we discussed computable and decidable matters. However, many simple questions are *undecidable* (*i.e. not* decidable). For instance, we can only have countably many Turing machines M, which can be explicitly encoded as integers $\langle M\rangle$, so it makes sense to consider questions such as the *halting* problem (*does the machine halt on the empty input?*), the *totality* problem (*does the machine halt on each and every input?*), the *co-finality* problem (*is the complementary set of* $\mathrm{supp}(M)$ *finite?*) and so on. After giving some insight on why such problems cannot be decidable, I will introduce the arithmetical hierarchy as a way to classify their hardness.

Lemma 3.16. *The halting problem is undecidable.*

Proof. Suppose the halting problem is decided by a Turing machine H, and define the machine T as follows:

- T receives an encoding $\langle M\rangle$ as its input,
- T computes the encoding $\langle N\rangle$ of a machine N that erases its input tape, writes down $\langle M\rangle$ instead and then executes M on this input.
- T executes H on the input $\langle N\rangle$, and then simulates N on the empty input with a universal machine U. If $H(\langle N\rangle) = \mathtt{1}$ and $N(\varepsilon) = \mathtt{1}$, then $T(\langle M\rangle) = \mathtt{0}$, else $T(\langle M\rangle) = \mathtt{1}$.

As H terminates on any valid encoding, so does T. In particular, if we execute T on its own encoding, then we reach a contradiction. Indeed, if $T(\langle T\rangle) = \mathtt{0}$, then $N(\varepsilon) = \mathtt{1}$, but by definition of N this actually means that $T(\langle T\rangle) = \mathtt{1}$ (and likewise for the other way around). Thus, H cannot decide the halting problem. □

Definition 3.17 (Computable Reduction)**.** Let $S, S' \subseteq \mathbb{N}$ be two decision problems. We say that S reduces to S', and we denote $S \leq S'$, if there exists a computable map $f : \mathbb{N} \to \mathbb{N}$ such that $n \in S$ *iff* $f(n) \in S'$. ■

In particular, if S' is decidable and $S \leq S'$, then S also is decidable. This gives us a partial order on decision problems, which allows us to *compare* the relative hardness of undecidable problems, for which time complexity isn't a pertinent notion anymore. However, just like there is a *classification* of computable functions depending on their complexity (*e.g.* linear, polynomial or exponential time), we would like to classify the hardness of undecidable problems. This is where the arithmetical hierarchy comes into play.

Definition 3.18 (Arithmetical Hierarchy)**.** A problem $S \subseteq \mathbb{N}$ is Π_k-computable if there exists a computable $f : \mathbb{N}^{k+1} \to \{\mathtt{0}, \mathtt{1}\}$ such that:

$$x \in S \Leftrightarrow \underbrace{\forall y_1, \exists y_2, \forall y_3, \dots}_{k \text{ alternating quantifiers}} \; f\left(x, y_1, \dots, y_k\right) = \mathtt{1}\,.$$

Likewise, S is Σ_k-computable if the block of alternating quantifiers starts with $\exists x_1$ instead. ■

In particular, if $S' \in \Pi_k$ and $S \leq S'$, then $S \in \Pi_k$.

Definition 3.19 (Complete Problem)**.** A problem $S \subseteq \mathbb{N}$ is Π_k-complete if it is Π_k-computable and there is a computable reduction from P to S for any $P \in \Pi_k$.

Naturally, $S \in \Pi_k$ *iff* $S^c = \mathbb{N}\backslash S \in \Sigma_k$, and by extension S is Π_k-complete *iff* S^c is Σ_k-complete. ■

Proposition 3.20 (The Halting Problem in the Arithmetical Hierarchy)**.** *The halting problem is* Σ_1*-complete.*

Proof. Let T a Turing machine that receives the code of a machine M and $n \in \mathbb{N}$ as an input, simulates n steps of the computation of $M(\varepsilon)$ (using a universal machine U), returns 1 if said simulation terminates by step n. Naturally, M halts *iff* $\exists n \in \mathbb{N}, T(\langle M \rangle, n) = 1$, hence the halting problem is in Σ_1.

Now, let S a Σ_1 problem and $f : \mathbb{N}^2 \to \{0, 1\}$ the corresponding computable function. For any input $x \in \mathbb{N}$, we let T_x as the machine that, regardless of the input, uses a universal machine to simulate 1 step of computation of $f(x,1)$, then two steps of both $f(x,1)$ and $f(x,2)$, and keeps going on with n steps of $f(x,i)$ for each $1 \leq i \leq n$, until we possibly terminate a computation with $f(x,n) = 1$, in which case T_x accepts, or we just keep going forever. The map $x \mapsto \langle T_x \rangle$ is computable, and $x \in S$ *iff* T_x halts (on the empty input), so we have a computable reduction from S to the halting problem. □

Likewise, we know that the totality problem is Π_2-complete [Soa16, Theorem 4.3.2], and the co-finality problem is Σ_3-complete [Soa16, Theorem 4.3.3].

More broadly, we know that Π_k-complete problems exist for any ladder $k \in \mathbb{N}$ of the hierarchy, and that the intersection $\Delta_k = \Sigma_k \cap \Pi_k$ doesn't contain such problems [Soa16, Corollary 2.3] (which implies that $\Sigma_k \neq \Pi_k$).

Lemma 3.21. *We have* $\Sigma_k \cup \Pi_k \subsetneq \Delta_{k+1}$.

Proof. Let P be a Π_k-complete problem, S a Σ_k-complete problem, and $D = (2P) \sqcup (2S+1)$. In particular, we have the reductions $P \leq D$ and $S \leq D$. As $\Sigma_k \cup \Pi_k \subseteq \Delta_{k+1}$, we know that $D \in \Delta_{k+1}$. If we had $D \in \Pi_k$, then it would be a Π_k-complete problem, but then using $S \leq D$ we obtain the contradiction $\Sigma_k \subseteq \Pi_k$. Using the same argument, $D \notin \Sigma_k$, hence the announced result, with problems in Δ_{k+1} that are *both* Π_k-hard *and* Σ_k-hard. □

Put together, all these arguments allow us to draw a schematic representation of the arithmetical hierarchy of undecidable problems in Figure 3.2.

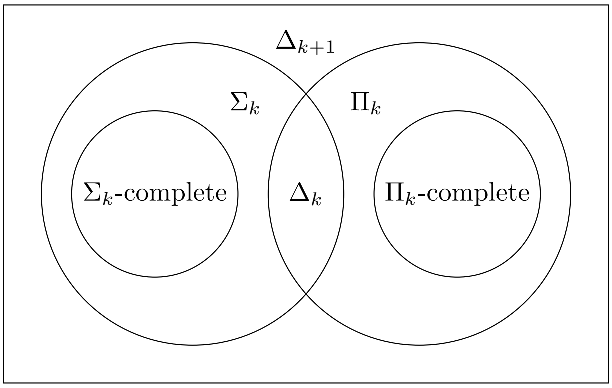

Figure 3.2: Representation of the k-th ladder of the arithmetical hierarchy.

3.5 Simulating Tilesets

In this section, I show how Turing machines can be simulated in the context of tilings with local rules introduced in the previous chapter. Historically [Rob71], these simulations have been used to prove that the *domino problem* (*i.e. does a finite family of forbidden patterns $\mathcal{F}$ induce a non-empty SFT $X_\mathcal{F}$?*) is undecidable (and actually Π_1-complete).

First, remark that $X_\mathcal{F} \neq \emptyset$ *iff* for any finite square window I_n, we have a locally admissible pattern in $\mathcal{A}^{I_n}$ ($\Rightarrow$ is direct, and we get $\Leftarrow$ by compactness). As checking the existence of a locally admissible pattern $w \in \mathcal{A}^{I_n}$ is a decidable matter, we conclude that the domino problem is a Π_1-computable problem. Now, to prove the completeness of the domino problem, we want a computable reduction from the (non) halting problem to it.

Naively, the space-time diagrams of a Turing machine M can be seen as two-dimensional tilings with local rules, using the tileset shown in Figure 3.3. Here, a time step for M corresponds to the *interface* between two consecutive rows of tiles, while each row corresponds to a *transition* from one step to the next. From left to right, we have three tiles for the case where no change occurs on the position at the transition, one for the initial internal state of the Turing machine, two corresponding to the transition and edit of the value of a given cell, and a last one for the stalled behaviour if the computation is done. Because the rules are deterministic from one given row to the next, we can see it as a kind of one-dimensional cellular automata.

Up to some more tinkering, we can use an *initial seed* (*i.e.* a finite pattern, which we will call w_ε) that will force the simulation of the machine M on the empty input. If we also include in $\mathcal{F}(M)$ the tiles that correspond to final states $q \in Q_A \sqcup Q_R$, then we obtain a tileset such that M halts *iff* w_ε is in the language of $X_{\mathcal{F}(M)}$.

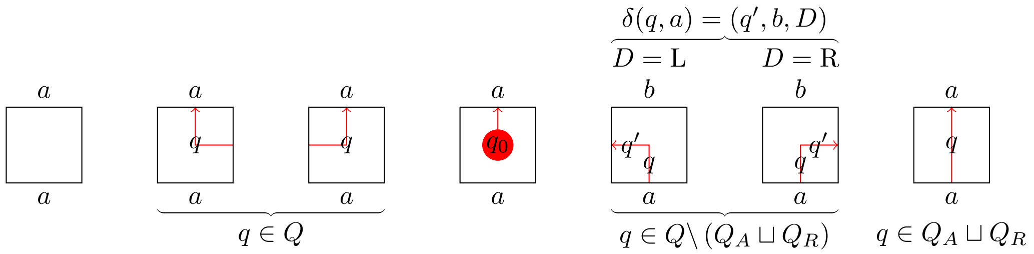

Figure 3.3: Turing space-time diagram Wang tiles for each letter $a \in \Gamma$.

In other words, the domino problem *with an initial seed* (*i.e.* whether w is in the language of $X_\mathcal{F}$ for an input $(w, \mathcal{F})$) is Π_1-complete. However, because any tileset must be able to encode arbitrarily large rectangles made of the # symbol, where no computations occur, we always have $\#^{\mathbb{Z}^2} \in X_{\mathcal{F}(M)} \neq \emptyset$, so this doesn't give us a reduction from the halting problem to the *general* domino problem. To tackle this problem, we need a way to *force* the initial seed through the local rules only. This is where the Robinson hierarchical structure comes into play.

To do so, consider the Red-Black Robinson tileset from Section 2.7.3, using the uniquely ergodic variant with only Black bumpy-corners. In particular, whenever $n = 2N + 1$ is odd, the n-macro-tile has a big Red square and a Black corner in the middle.

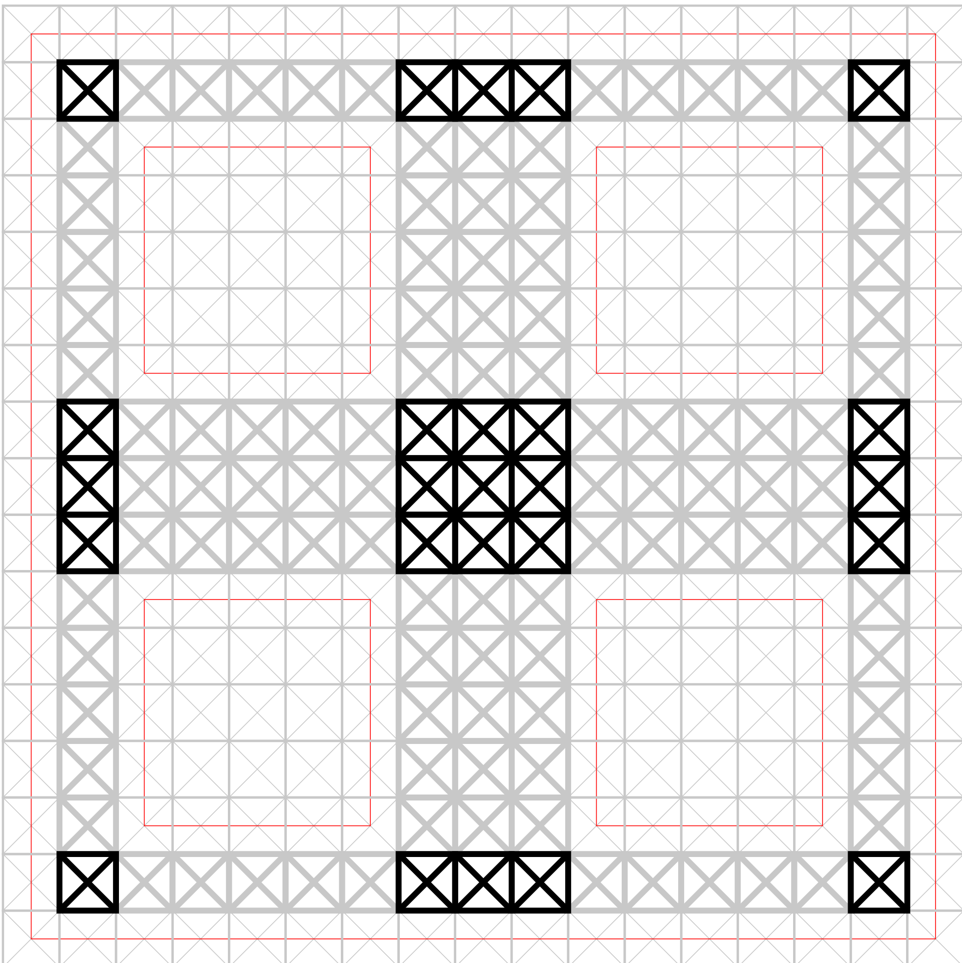

Figure 3.4: Sparse computation area inside the central Red square of a 5 macro-tile.

Notably, this central square avoids all the smaller Red squares inside it. As illustrated in Figure 3.4, once we remove the area occupied by smaller squares, we obtain a sparse square made of disconnected rectangle patches (the black thick squares) that can communicate with each other (using the rows and columns of grey thick squares). Inside a $(2N+1)$-macro-tile, this sparse area corresponds to a $\left(2^N+1\right) \times \left(2^N+1\right)$ grid of tiles (whereas the Red square itself has a tile-length of 4^N+1).

Formally, we say that a tile is in a *free column* if both the Red line above and below correspond to the *inside* faces of a Red square, and we can use local rules to check this property. The same argument holds for the columns, so this supplementary structure allows us to identify which tiles are in free columns and rows (*i.e.* the black thick squares) and which are not. Then, we can add a new layer to the Robinson tileset, implement a Wang tileset in free cells, and a synchronising behaviour in non-free cells so that when the sparse area is glued together as a single square, we obtain a locally admissible pattern for the Wang tileset. As the scale of $(2N+1)$-macro-tiles goes to infinity, so does the size of the locally admissible Wang pattern in the central square.

In particular, we can thus implement the tileset from Figure 3.3, associated to a Turing machine M. We also add local rules to force the use of an initial state tile with the empty symbol $\# \in \Gamma$ on the Wang layer in a (free) tile north-east diagonally adjacent to a Red ⊡ Robinson tile, and an a symbol $\#$ without internal state for other free tiles north adjacent to the bottom of a Red square. Hence, by using this tileset, the central square of a $(2N+1)$-macro-tile will simulate the machine M on the empty input for exactly 2^N steps of computation.

Theorem 3.22. *The domino problem is Π_1-complete.*

Proof. We have already seen that the domino problem is Π_1. We just need to prove that the previous embedding of M into the Robinson structure gives us a tileset such that M halts *iff* $X_{\mathcal{F}(M)}$ is empty.

To do so, as for the case of the domino problem with a seed, we add one last rule to forbid final states on the Wang layer, and we denote $\mathcal{F}(M)$ the corresponding set of forbidden patterns. The construction of the tileset we gave so far can be algorithmically automated, so that $M \mapsto \mathcal{F}(M)$ is a computable map. Now, as we have a projection from $X_{\mathcal{F}(M)}$ to Robinson tilings (in a way compatible with local rules, thus a morphism), we conclude that the structure on the Robinson layer still follows the arguments from Section 2.7.1, so we can decompose $\omega \in X_{\mathcal{F}(M)}$ into ∞-macro-tiles. Each such area contains arbitrarily large macro-tiles, that must simulate $M(\varepsilon)$ for an arbitrarily long time. In other words, we have $X_{\mathcal{F}(M)} \neq \emptyset$ *iff* we have admissible ∞-macro-tiles (that must contain arbitrarily large space-time diagrams) *iff* M does *not* halt on the empty-input. As the halting problem is Π_1-complete, we finally conclude on the Σ_1-completeness of the domino problem. □

3.6 Computable Analysis

Since we are interested in matters relating to SFTs $X_{\mathcal{F}}$ in the next chapters, there is always an underlying *countable* space of finite sets of forbidden patterns $\mathcal{F}$ representing the inputs. However, the questions we want to answer afterwards always relate to a subset of measures $\mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$ (or of *ground states* of Gibbs measures $\mathcal{G}_\sigma(\infty)$ in Chapter 5), in the space $\mathcal{M}_\sigma\left(\Omega_{\mathcal{A}}\right)$ which is clearly uncountable. Thence the need for a framework to do computations on uncountable sets with convenient topological properties, called *computable analysis*.

Let's introduce the idea with computable real numbers. The set $\mathbb{R}$ is uncountable, so we cannot simply encode each real number as an integer, the notion of exact computations simply doesn't make sense anymore. However, we know that the countable set $\mathbb{Q}$ is a dense family in $\mathbb{R}$, so we can identify $x \in \mathbb{R}$ to some converging sequence $(x_n) \in \mathbb{Q}^{\mathbb{N}}$, which itself can be computed by a Turing machine. The real number x is said to be *computable* there is a computable map $f : \mathbb{N} \to \mathbb{Q}$ such that such that $|f(n) - x| < 2^{-n}$. More generally, x is *limit-computable* (resp. lower, upper semi-computable) if there exists $f : \mathbb{N} \to \mathbb{Q}$ such that $\lim\limits_{n\to\infty} f(n) = x$ (resp. $\sup\limits_{n\in\mathbb{N}} f(n) = x$, $\inf\limits_{n\in\mathbb{N}} f(n) = x$). This idea extends as follows.

Definition 3.23 (Computable Metric Space)**.** A *computable metric space* is a metric space (X, d) with a countable dense family $\mathfrak{D} = (z_n)_{n\in\mathbb{N}}$, and a computable $f : \mathbb{N}^3 \to \mathbb{Q}$ such that, for any input $i, j, n \in \mathbb{N}$, we have $|d(z_i, z_j) - f(i,j,n)| \leq 2^{-n}$. An element $x \in X$ is computable when there is a computable $f : \mathbb{N} \to \mathbb{N}$ such that, for any given input precision n, $d\left(x, z_{f(n)}\right) \leq 2^{-n}$ (we say that the algorithm f *computes* x). More generally, x is *limit-computable* if there is an algorithm $f : \mathbb{N} \to \mathbb{N}$ such that $d\left(x, z_{f(n)}\right) \to 0$ but without any control on the convergence speed. ■

The notion of computation obtained depends on the order chosen to enumerate $\mathfrak{D}$ (*i.e.* the *encoding* of $\mathfrak{D}$ into $\mathbb{N}$), but we generally operate under the assumption that any two *reasonable* enumerations of the set are equivalent (in the sense that there is a computable way to transform one into the other), and consider maps directly defined on the set $\mathfrak{D}$ instead of its encoding into $\mathbb{N}$.

It follows from the definition that, for any $x, y \in \mathfrak{D}$ and $r, s \in \mathbb{Q}$, checking if $B(x,r) \cap B(y,s) \neq \emptyset$ or $B(x,r) \cap \overline{B(y,s)} \neq \emptyset$ (*i.e.* $d(x,y) < r+s$) is semi-decidable: there is an algorithm $f(x,y,r,s)$ that *always answer* the question in finite time *when the answer is yes*, but *may loop* without halting *when the answer is no* (here the infinite loop occurs when $d(x,y) = r+s$). Likewise, checking if $\overline{B(x,r)} \cap \overline{B(y,s)} \neq \emptyset$ (*i.e.* $d(x,y) \leq r+s$) is co-semi-decidable (an algorithm always answers in finite time when the answer is no).

By extension, if $(X, d, \mathfrak{D})$ is a computable space, then the space of *compact subsets* of X is computable too, using the Hausdorff distance associated to d, *i.e.* for any compact subsets $K, K' \Subset X$ we have

$$d\left(K, K'\right) := \max\left(\max_{x \in K} d\left(x, K'\right), \max_{y \in K'} d(K, y)\right),$$

and the countable dense basis of *finite subsets* of $\mathfrak{D}$ (for any compact K and $\varepsilon > 0$, there exists a finite (thus compact) subset $C \Subset \mathfrak{D}$ such that $d_H(K, C) \leq \varepsilon$).

An equivalent (and somewhat more intuitive) way of defining computability on compact subsets is to use intersecting closed rational balls (we refer the reader to an article by Cenzer and Remmel [CR02] for a discussion on the different ways to define the computability of compact sets). To do so, we identify K with the set of intersecting neighbourhoods $\mathcal{N}(K) \subseteq \mathfrak{D} \times \mathbb{Q}$, defined as:

$$\mathcal{N}(K) := \{(x,r) \in \mathfrak{D} \times \mathbb{Q} : d(x,K) \leq r\}.$$

In particular, as the space $\mathfrak{D} \times \mathbb{Q}$ is countable, we can study $\mathcal{N}(K)$ using the usual framework for computable subsets of $\mathbb{N}$, *i.e.* this set is said to be computable if the associated indicator function $\mathbb{1}_{\mathcal{N}(K)} : \mathfrak{D} \times \mathbb{Q} \to \{0,1\}$ is. By analogy with subsets of $\mathbb{R}^2$, this means that K can be "displayed" on a computer screen at any zoom level, using a algorithm to decide whether each pixel (as a "ball" of a given radius) must be lighted (which happens *iff* it intersects K).

Using this formalism, we can then naturally extend the study from computable compact sets K to the arithmetical hierarchy, by saying K is Π_k-computable when the set $\mathcal{N}(K)$ is. For the first levels in the arithmetical hierarchy, we can often obtain equivalent and seemingly more natural characterisations. For example, the compact K is Π_1-computable *iff* there is a computable sequence $(x, r) : \mathbb{N} \to \mathfrak{D} \times \mathbb{Q}$ such that $K = \left(\bigcup_{n \in \mathbb{N}} B\left(x_n, r_n\right)\right)^c$.

Definition 3.24 (Recursive Compactness)**.** Let $(X, d, \mathfrak{D})$ be a computable space. A compact set $K \Subset X$ is said to be recursively compact if there is an algorithm that semi-decides whether any given finite family of open balls $\left(B\left(x_i, r_i\right)\right)_{i \leq n}$ (with $x_i \in \mathfrak{D}$ and $r_i \in \mathbb{Q}$) covers K.

By extension, we say that the space X is locally recursively compact if any closed rational ball $\overline{B\left(x, r\right)}$ (with $x \in \mathfrak{D}$ and $r \in \mathbb{Q}$) is recursively compact. This notion naturally corresponds to the computable equivalent of locally compact spaces.

In the locally recursively compact case, any Π_1-computable compact set$K \Subset X$ is recursively compact, and in particular it is semi-decidable to know whether K is included in a rational base ball $B(x, r)$. Notably, if X is compact, it is a rational base ball (for a big radius) so it is recursively compact too. ■

A result of interest here is the fact that K is Π_2-computable *iff* $K = \mathrm{Acc}\left(x_n\right)$ is the set of all accumulation points for some computable sequence $x : \mathbb{N} \to \mathfrak{D}$. In the case of *connected* Π_2 sets, we can refine the result to also require $d\left(x_n, x_{n+1}\right) \to 0$. This was already proven specifically in the case of invariant measures on a one-dimensional subshift [HS18, Proposition 6], so we provide here a generalised statement with a more concise proof:

Proposition 3.25 (Connected Π_2 Compacts as Accumulations Sets)**.** *Let $(X, d, \mathfrak{D})$ be a locally recursively compact space.*

Then the connected Π_2*-computable compact subsets $K \Subset X$ are exactly those than can be obtained as accumulation sets $K = \mathrm{Acc}\left(x_n\right)$ with $n \in \mathbb{N} \mapsto x_n \in \mathfrak{D}$ a computable map and $d\left(x_n, x_{n+1}\right) \to 0$.*

Proof. Without loss of generality, assume X itself is compact (else we can replace X by $B := \overline{B_d\left(x_0, r\right)}$ such that $K \subseteq B$, in the rest of the proof). Since X is then *recursively* compact, it is possible to *compute* a non-decreasing sequence of finite sets $\mathfrak{D}_k \subseteq \mathfrak{D}_{k+1} \subseteq \mathfrak{D}$ such that $X = \bigcup_{x \in \mathfrak{D}_k} B_d\left(x, \frac{1}{k}\right)$ for any $k \in \mathbb{N}$.

Let K be a connected Π_2-computable compact set, and $f : \mathfrak{D} \times \mathbb{Q}^+ \times \mathbb{N}^2 \to \{0,1\}$ the associated computable function, such that $d(x,K) \leq \frac{1}{k}$ *iff* $\forall n \in \mathbb{N}, \exists t \in \mathbb{N}, f(x,k,n,t) = 1$.

For $T, k \in \mathbb{N}$, define:

$$V_k^T := \left\{ (x_i \in \mathfrak{D}_i)_{i \in [\![1,k]\!]}, \forall i,j \in [\![1,k]\!], d(x_i,x_j) < \frac{2}{\min(i,j)}, \forall n \leq k, \exists t \leq T, f(x_i,i,n,t) = 1 \right\}.$$

Thus defined, V_k^T uses the computable distance function $d : \mathfrak{D}^2 \to \mathbb{R}$. Without loss of generality, we can replace $d(x,y)$ by a computable rational upper bound $\overline{d}(x,y,k+T)$ at distance at most $\frac{1}{2^{k+T}}$ (with $\overline{d} : \mathfrak{D}^2 \times \mathbb{N} \to \mathbb{Q}$ computable, non-decreasing in its third variable), so that $(k,T) \mapsto V_k^T$ is a computable map too.

In particular, $V_k^T \subseteq V_k^{T+1} \subseteq \prod_{i=1}^k \mathfrak{D}_i$ is a non-decreasing sequence, and $\prod_{i=1}^k \mathfrak{D}_i$ is finite, so the sequence is stationary after a rank t_k. Without loss of generality, by inductively replacing t_k by $\max(t_k, t_{k-1})$, we can have $k \mapsto t_k$ non-decreasing. Likewise, if $(x_i)_{i \leq k+1} \in V_{k+1}^T$, then its prefix satisfies $(x_i)_{i \leq k} \in V_k^{T+1}$ (we must increase T by one to compensate the loss of precision for $\overline{d}$, to compute the same $\overline{d}(x_i,x_j,k+T+1)$ in both cases), so that in particular $\pi_{[\![1,k]\!]}\left(V_{k+1}^{t_{k+1}}\right) \subseteq V_k^{t_k}$ (with $\pi_{[\![1,k]\!]}$ the projection on the first k coordinates).

Let $x \in \mathfrak{D}_k$ such that $B_d\left(x, \frac{1}{k}\right) \cap K \neq \emptyset$. Then there is $(x_i) \in V_k^{t_k}$ such that $x_k = x$. Indeed, let $y \in B_d\left(x, \frac{1}{k}\right) \cap K$, and fix $i < k$. Because the $\frac{1}{i}$-neighbourhood of $\mathfrak{D}_i$ covers K, it follows that there is some $x_i \in \mathfrak{D}_i$ such that $y \in B_d\left(x_i, \frac{1}{i}\right)$. We have $B_d\left(x_i, \frac{1}{i}\right) \cap K \neq \emptyset$, so the formula

$$\forall n \leq k, \exists t \in \mathbb{N}, f(x_i,i,n,t) = 1$$

holds in particular. By triangle inequality we have $d(x_i,x_j) < \frac{1}{i} + \frac{1}{j} \leq \frac{2}{\min(i,j)}$ (and this likewise holds for $\overline{d}(x_i,x_j,k+T)$ for T big-enough, thus for t_k). Thence $(x_i)_{i \leq k} \in V_k^{t_k}$. As the $\frac{1}{k}$-neighbourhood around $\mathfrak{D}_k$ covers K and $B_d\left(x_k, \frac{1}{k}\right) \cap K = \emptyset$ when $x \in \mathfrak{D}_k \backslash \pi_k\left(V_k^{t_k}\right)$, we conclude that:

$$K \subseteq \bigcup_{x \in \pi_k\left(V_k^{t_k}\right)} B_d\left(x, \frac{1}{k}\right).$$

However, we only still have a rough covering of K, and may have outliers $x \in \pi_k\left(V_k^{t_k}\right)$ for which the intersection $B_d\left(x, \frac{1}{k}\right) \cap K = \emptyset$ is empty. Nevertheless, for a fixed scale $\ell \in \mathbb{N}$, the sequence $\left(\pi_{[\![1,\ell]\!]}\left(V_k^{t_k}\right)\right)_{k \geq \ell}$ is non-increasing, thus stationary. Denote $\varphi(\ell)$ the finite rank at which the stationary limit is reached, and $W_\ell := \pi_\ell\left(V_{\varphi(\ell)}^{t_{\varphi(\ell)}}\right)$ the corresponding projection.

Now, for $x \in \mathfrak{D}_\ell$, we have $B_d\left(x, \frac{1}{\ell}\right) \cap K \neq \emptyset$ *iff* $x \in W_\ell$. The direct implication works as before, using a sequence $(x_k \in \mathfrak{D}_k)_{k \in \mathbb{N}}$ such that $x = x_\ell$ and $\bigcap_{k \in \mathbb{N}} B_d\left(x_k, \frac{1}{k}\right) = \{y\} \subseteq K$. Conversely, if $B_d\left(x, \frac{1}{\ell}\right) \cap K = \emptyset$, then there is a rank $n \in \mathbb{N}$ such that $\forall t \in \mathbb{N}, f(x,\ell,n,t) = 0$, and thus $x \notin \pi_\ell\left(V_k^{t_k}\right)$ for $k \geq n$.

To obtain the desired computable sequence $(x_n)_{n \in \mathbb{N}}$, we first define the computable sequence of finite sets $U_{T+1} = \bigsqcup_{k \leq T+1} \pi_k\left(V_k^{T+1} \backslash V_k^T\right)$ (we use a disjoint union to underline that U_T is a *multiset*, and a single element $x \in \mathfrak{D}$ may appear once for each value of $k \leq T$).

Let's fix a scale $\ell \in \mathbb{N}$. Whenever $T > t_\ell$ and $k \leq \ell$, then $V_k^T \backslash V_k^{T-1} = \emptyset$, so U_T only uses elements for scales $\ell < k \leq T$, while the first sets of the disjoint union are empty. Thus, cumulatively, $\bigsqcup_{i \leq \ell} \pi_i\left(V_i^{t_i}\right) \subseteq \bigsqcup_{t \leq T} U_t$. As the sets $V_i^{t_i}$ are all non-empty (their $\frac{1}{i}$-neighbourhood covers $K \neq \emptyset$), it follows that $\sum_{T \in \mathbb{N}} |U_T| = \infty$, so any sequence (x_n) induced by "enumerating" the sets (U_T) will be infinite indeed. What's more, it will contain (at least) an occurrence of each element of $\left(V_k^{t_k}\right)_{k \in \mathbb{N}}$, which can approximate any element of K with precision $\frac{1}{k}$, so that $K \subseteq \mathrm{Acc}(x_n)$.

Now, let's compute a path traversing U_T. To do so, we compute the biggest scale $i \leq T$ that satisfies the following property: there exists an intermediate scale $j \in [\![i,T]\!]$ such that, using only elements of $\pi_i\left(V_j^T\right)$ in-between two elements of U_T, we can traverse U_T with distances of less than $\frac{2}{i}$ between consecutive elements (and if no such scale exists we simply enumerate U_T instead).

When $T > t_{\varphi(\ell)}$, the set $W_\ell = \pi_\ell\left(V_{\varphi(\ell)}^T\right)$ gives us such a scale i. On one hand, W_ℓ forms an optimal covering of the connected set K, with intersecting balls of radius $\frac{1}{\ell}$, so we can traverse W_ℓ itself with jumps at most $\frac{2}{\ell}$. On the other hand, any element of $x \in U_T$ is actually in a set $\pi_i\left(V_i^T\right)$ with $i \geq \varphi(\ell) \geq \ell$, thus at distance at most $\frac{2}{\ell}$ of $x_\ell \in W_\ell$ for the corresponding sequence $(x_j)_{j \leq i} \in V_i^T$, so we can visit each element of U_T while traversing W_ℓ. To summarise, when $T > t_{\varphi(\ell)}$, we can compute a path traversing U_T, using intermediate elements in $\pi_i\left(V_j^T\right)$ (with $\ell \leq i \leq j \leq T$), with distance at most $\frac{2}{\ell}$ between consecutive elements. For any such element x (either in U_T or an intermediate element), we have a corresponding finite sequence $(x_i) \in V_j^T$ (for some $\ell \leq j \leq T$), and thus in particular $x_\ell \in W_\ell$, so that $d(x,K) \leq \frac{3}{\ell}$ by triangle inequality.

Likewise, to link the last visited element $y \in U_T$ to the first element $z \in U_{T+1}$ (or more generally to the next non-empty set $U_{T'}$ instead), we may use the same algorithm, and conclude that we can link y to z with distances at most $\frac{2}{\ell}$ between elements and while staying at distance at most $\frac{3}{\ell}$ of K. Overall, by sticking the traversals of each U_T with these links, we finally obtain a well-defined computable sequence $(x_n) \in \mathfrak{D}^{\mathbb{N}}$, such that $\mathrm{Acc}\,(x_n) \subseteq K$ and $d\,(x_n, x_{n+1}) \to 0$, which concludes this implication.

For the other direction, any set $K := \mathrm{Acc}\,(x_n)$ is obviously Π_2, using the fact that $B_d\left(x, \frac{1}{k}\right) \cap K \neq \emptyset$ *iff* $\forall n \in \mathbb{N}, \exists t \geq n, d\,(x, x_t) < \frac{1}{k} + \frac{1}{n}$, with the elements x_t corresponding to a subsequence converging to an element in $B_d\left(x, \frac{1}{k}\right) \cap K$. What's more, K must be connected using roughly the same argument as in Proposition 5.3. □

Corollary 3.26. *Using the same proof minus the connectedness arguments, by simply enumerating the sets* (U_T)*, we conclude that* K *is a* Π_2*-computable set if and only* iff $K = \mathrm{Acc}\,(x_n)$ *for some computable sequence* $x : \mathbb{N} \to \mathfrak{D}$*.*

Remark 3.27 (Computations on Measures)**.** In the case of $(\mathcal{M}_\sigma(\Omega_\mathcal{A}), d^*)$, we will use $\mathfrak{P}$ the countable family of periodic points, that is to say measures $\widehat{\delta_p} := \frac{1}{|I_n|} \sum_{k \in I_n} \delta_{\sigma^k\left(p^{\mathbb{Z}^d}\right)}$ with $p \in \mathcal{A}^{I_n}$ and $n \in \mathbb{N}$. It is known that $\mathfrak{P}$ is a dense basis of $\mathcal{M}_\sigma(\Omega_\mathcal{A})$. Notice that $(x, y) \in \mathfrak{P}^2 \mapsto d^*(x, y)$ is computable, *i.e.* there is a computable $f : \mathfrak{P}^2 \times \mathbb{N} \to \mathbb{Q}$ such that $|d^*(x, y) - f(x, y, n)| \leq 2^{-n}$, hence $(\mathcal{M}_\sigma(\Omega_\mathcal{A}), d^*, \mathfrak{P})$ is a computable metric space.

A measure μ is computable, as an *element* of the computable space, if there is an algorithm $f : \mathbb{N} \to \mathfrak{P}$ such that $d^*(\mu, f(n)) \leq 2^{-n}$ (note that this implies that μ is computable as a real-valued function too). What's more, $\mathcal{M}_\sigma(\Omega_\mathcal{A})$ is locally recursively compact.

I refer the reader to the existing literature [GHR11] for an overview of these computability properties of the set of probability measures. ■

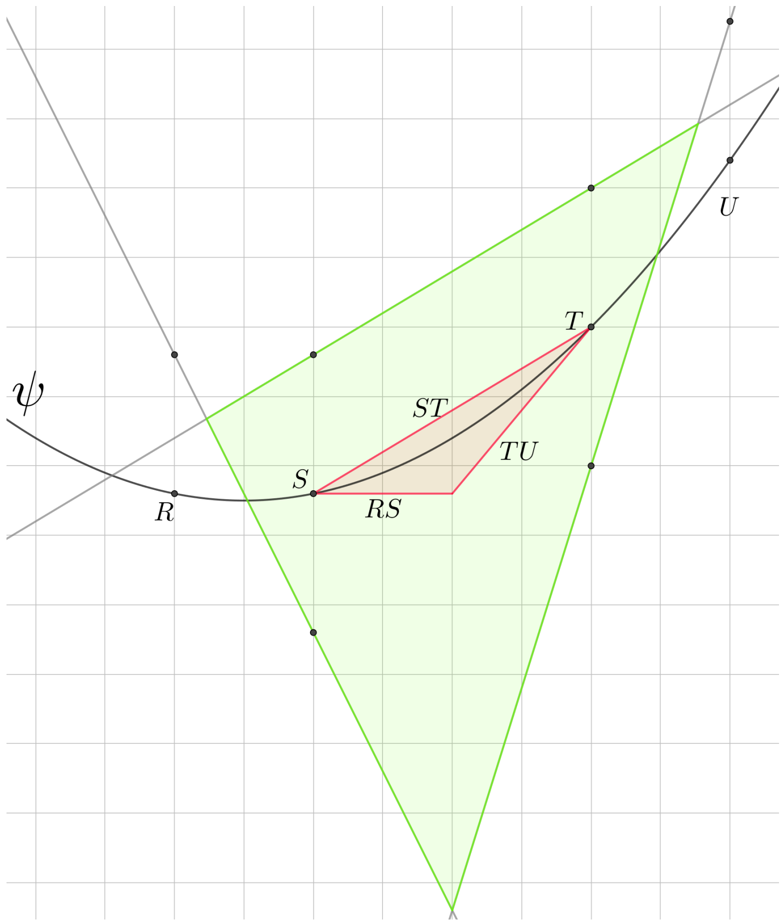

Figure 3.5: Inclusion of the graph of a convex map ψ into a computable triangle.

Definition 3.28 (Computable Maps on Computable Spaces)**.** Let $(X, d, \mathfrak{D})$ and $(X', d', \mathfrak{D}')$ be two computable spaces. The map $\psi : X \to X'$ is said to be computable if there exists a computable $f : \mathbb{N} \times \mathfrak{D}' \times \mathbb{Q} \to \mathfrak{D} \times \mathbb{Q}$ such that, for $y \in \mathfrak{D}'$ and $r \in \mathbb{Q}$:

$$\psi^{-1}\left(B_{d'}(y, r)\right) = \bigcup_{n \in \mathbb{N}} B_d(f(n, y, r))\,.$$

In particular, it follows that ψ is continuous. Likewise, a sequence $(\psi_k)_{k\in\mathbb{N}}$ of maps is (uniformly) computable if there is $f : \mathbb{N}^2 \times \mathfrak{D}' \times \mathbb{Q} \to \mathfrak{D} \times \mathbb{Q}$ such that $\psi_k^{-1}(B_{d'}(y,r)) = \bigcup_{n\in\mathbb{N}} B_d(f(k,n,y,r))$.

In the case of *locally recursively compact* spaces (*i.e.* spaces where we can semi-decide if a ball is covered by a family of balls), this notion is equivalent to the existence of a computable map $(f,s) : \mathfrak{D} \times \mathbb{Q} \to \mathfrak{D}' \times \mathbb{Q}$ such that:

$$\psi(B_d(x,r)) \subseteq B_{d'}(f(x,r), s(x,r)).$$

and $\sup_{x\in K\cap\mathfrak{D}} s(x,r) \underset{r\to 0}{\longrightarrow} 0$ for any compact $K \Subset X$. The first part tells us that we can map an approximation of $x \in X$ to an approximation of $\psi(x)$, and the second part tells us that this process converges (in a uniformly continuous way on compacts), so that whenever $x \in X$ is computable, so is $\psi(x) \in X'$. ■

Lemma 3.29. *Let $\psi : \mathbb{R} \to \mathbb{R}$ be a* convex *map. Then ψ is computable* iff *there is a computable $f : \mathbb{Q} \times \mathbb{N} \to \mathbb{Q}$ such that $|f(x,n) - \psi(x)| \leq 2^{-n}$ (*i.e. *$\psi|_{\mathbb{Q}}$ is uniformly computable).*

Proof. The direct implication holds for any computable map, by estimating $\psi(]x-\varepsilon, x+\varepsilon[)$ for increasingly smaller values of ε until we reach the desired precision on $\psi(x)$.

For the other direction, we want to estimate $\psi(]s,t[)$ using only computable estimations of $\psi(s)$, $\psi(t)$ (and two other points) given by the function $f : \mathbb{Q}^2 \to \mathbb{Q}$. More precisely, we use the idea illustrated in Figure 3.5.

Denote $\rho = \frac{|t-s|}{2} \in \mathbb{Q}$ the radius of the ball $]s,t[$, and consider two more real numbers $r = s - \rho$ and $u = t + \rho$. Denote $S = (s, \psi(s)) \in \mathbb{R}^2$ and likewise for R, T and U. On the interval $[s,t]$, we know that the graph of ψ is both below the line ST, but above the lines RS and TU, thus included in the (small red) triangle. In practice, as shown in Figure 3.5, we can replace ST by the "higher" line crossing the points $(s, f(s,\rho)+\rho)$ and $(t, f(t,\rho)+\rho)$, RS by the "lower" line crossing $(r, f(r,\rho)+\rho)$ and $(s, f(s,\rho)-\rho)$, and likewise TU by the line crossing $(t, f(t,\rho)-\rho)$ and $(u, f(u,\rho)+\rho)$. This new (big green) triangle, defined using only rational computable coordinates, contains the previous one, hence its vertical range $]m(s,t), M(s,t)[$ covers $\psi(]s,t[)$. In particular, as long as $x \in \mathbb{Q}$ stay well bounded in a compact set K, then by convexity the slopes of the previous lines behave well, so that $\sup_{x\in\mathbb{Q}\cap K} M(x-r,x+r) - m(x-r,x+r) \underset{r\to 0}{\longrightarrow} 0$, *i.e.* ψ is computable in the previously defined way. □

Lemma 3.30. *Let $\psi : (X,d,\mathfrak{D}) \to (X',d',\mathfrak{D}')$ be a computable map between locally recursively compact spaces. If $K \Subset X$ is Π_2-computable, then so is $\psi(K)$.*

Proof. Consider here the characterisation $K = \mathrm{Acc}(x_n)$ for a computable sequence $x \in \mathfrak{D}^{\mathbb{N}}$. As ψ is continuous, we have $\psi(K) = \mathrm{Acc}(\psi(x_n)) \Subset X'$.

Using the computable map $(f,s) : \mathfrak{D}\times\mathbb{Q} \to \mathfrak{D}'\times\mathbb{Q}$, we then define the sequence $y \in \mathfrak{D}'^{\mathbb{N}}$ by $y_n := f(x_n, 2^{-n})$. By compactness of (the 1-neighbourhood of) K, we know in particular that $d'(y_n, \psi(x_n)) \to 0$. It follows that $\psi(K) = \mathrm{Acc}(y_n)$ is the accumulation set of a computable sequence y, hence it is Π_2-computable too. □

We cannot hope for any kind of converse result in the general case, as we may take a constant map ψ that collapses all the complexity of the set K onto a computable singleton $\psi(K) = \{x\}$. However, under stronger assumptions on ψ (injectivity in particular), we obtain the following equivalence:

Corollary 3.31. *Consider $\psi : (X,d,\mathfrak{D}) \hookrightarrow (X',d',\mathfrak{D}')$ an injective computable map, and likewise $\psi' : X' \to X$ computable such that $\psi'\circ\psi = \mathrm{Id}_X$. Then the images of (connected) Π_2 compacts of X are exactly the (connected) Π_2 compacts of X' included $\psi(X)$.*

Proof. As ψ is a continuous injection, we have K connected *iff* $\psi(K)$ is. Using the previous lemma, we already established that if $K \Subset X$ is Π_2, then so is $K' = \psi(K)$. Conversely, any compact $K' \Subset \psi(X)$ is the image of a compact K, and if K' is Π_2, then using the previous lemma for ψ', it follows that $K = \psi'(K')$ is Π_2 too, so K' is indeed the image of a Π_2 compact set. □

References

[BHW08] Vasco Brattka, Peter Hertling, and Klaus Weihrauch. "New Computational Paradigms". In: Springer, 2008. Chap. A Tutorial on Computable Analysis, pp. 425–491. 10.1007/978-0-387-68546-5_18.

[BY09] Mark Braverman and Michael Yampolsky. *Computability of Julia Sets*. Vol. 23. Algorithms and Computation in Mathematics. Springer, 2009. 10.1007/978-3-540-68547-0.

[CR02] Douglas Cenzer and Jeffrey Remmel. “Effectively Closed Sets and Graphs of Computable Real Functions”. In: *Theoretical Computer Science* 284.2 (2002), pp. 279–318. 10.1016/S0304-3975(01)00069-X.

[Gan+20] Silvère Gangloff, Alonso Herrera, Cristóbal Rojas, and Mathieu Sablik. “Computability of Topological Entropy: from General Systems to Transformations on Cantor Sets and the Interval”. In: *Discrete and Continuous Dynamical Systems* 40.7 (2020), pp. 4259–4286. 10.3934/dcds.2020180.

[GHR11] Stefano Galatolo, Mathieu Hoyrup, and Cristóbal Rojas. “Dynamics and Abstract Computability: Computing Invariant Measures”. In: *Discrete and Continuous Dynamical Systems* 29.1 (2011), pp. 193–212. 10.3934/dcds.2011.29.193.

[HS18] Benjamin Hellouin de Menibus and Mathieu Sablik. “Characterization of Sets of Limit Measures of a Cellular Automaton Iterated on a Random Configuration”. In: *Ergodic Theory and Dynamical Systems* 38.2 (2018), pp. 601–650. 10.1017/etds.2016.46.

[JV20] Emmanuel Jeandel and Pascal Vanier. “The Undecidability of the Domino Problem”. In: *Substitution and Tiling Dynamics: Introduction to Self-Inducing Structures*. Vol. 2273. Lecture Notes in Mathematics. Springer, 2020, pp. 293–357. 10.1007/978-3-030-57666-0_6.

[Ris20] Rising Entropy. *The Arithmetic Hierarchy and Computability*. Accessed 20 July 2022. 2020. risingentropy.com/the-arithmetic-hierarchy-and-computability.

[Rob71] Raphael Robinson. “Undecidability and Nonperiodicity for Tilings of the Plane”. In: *Inventiones mathematicae* 12 (1971), pp. 177–209. 10.1007/BF01418780.

[Rog87] Hartley Rogers. *Theory of Recursive Functions and Effective Computability*. MIT press, 1987.

[Sip13] Michael Sipser. *Introduction to the Theory of Computation*. 3rd edition. Cengage, 2013.

[Soa16] Robert Soare. *Turing Computability*. Springer, 2016. 10.1007/978-3-642-31933-4.

[Soa87] Robert Soare. *Recursively Enumerable Sets and Degrees*. Springer, 1987.

[Tur36] Alan Turing. “On Computable Numbers, With an Application to the Entscheidungsproblem”. In: *Proceedings of the London Mathematical Society. Second Series* 42.3 (1936), pp. 230–265. 10.1112/plms/s2-42.1.230.

[Wei00] Klaus Weihrauch. *Computable Analysis*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2000. 10.1007/978-3-642-56999-9.

[Wei17] Klaus Weihrauch. “Computability on Measurable Functions”. In: *Computability* 6.1 (2017), pp. 79–104. 10.3233/COM-160058.

[Wil17] Tom Wildenhain. “On the Turing Completeness of MS PowerPoint”. In: *SIGBOVIK*. 2017, pp. 102–106. sigbovik.org/2017/proceedings.pdf.

Chapter 4

Besicovitch Stability in the Arithmetical Hierarchy

This chapter adapts my first two published articles [GS23a, GS23b], with the second one directly addressing questions left open in the first, in a more streamlined fashion.

In this study, the general idea of stability for Bernoulli noises came first, as *how similar are random tilings with few mistakes to globally admissible configurations? Can aperiodic structures in SFTs survive in the presence of noise?* These questions are quite natural in the context of tilings as static models of computation. Indeed, for a given model of computation, we would like to know whether the model is robust to errors, as in the case of cellular automata [Gác01], Turing machines [AC05], or some simulating tilings [DRS12].

The Besicovitch topology only came into play secondly, once it appeared that the weak-$*$ topology was ill-suited for this loosely defined notion of stability. As already discussed in Chapter 2, the main difference is that the weak-$*$ topology looks at the small picture, locally, while the Besicovitch looks at the big picture, globally. This difference allowed me to study stability through many lenses, such as percolation theory, periodicity and computability, in a way that weak-$*$ stability simply wouldn't leave room for.

The current chapter can roughly be grouped in three parts: the general framework for noisy tilings and stability, a step-by-step study of stability for (almost) periodic tilings, and lastly a study of stability as a decision problem.

First, in Section 4.1, I provide a quantitative formalism of error robustness for a given choice of local rules. Formally, we will say that an SFT is f-stable (for a given distance on measures) if any shift-invariant measure with a proportion (at most) ε of errors is at distance at most $f(\varepsilon)$ of the set of non-noisy invariant measures on the SFT. In Section 4.2, I establish that, while by definition related to the choice of local rules, this notion of stability is conjugacy invariant. Then, in Section 4.3, I exhibit a full classification of stability in the one-dimensional case, in a way that proves this specific case is decidable.

Naturally, this question is expected to be more complex in higher dimensions, as was the case for the domino problem that goes from decidable in one dimension to Π_1-complete (thus undecidable) in higher dimensions. Corollary 4.32 allows us to extend the one-dimensional (un)stable examples in higher dimensions, but without increasing their computational complexity. Thus, before studying the complexity of stability, we must see how complex structures such as simulating tilesets interact with stability.

A first step in this direction is the study of stability for periodic SFTs in Section 4.4. There, I detail a percolation argument on the clear cells (those without mistakes), that allows us to obtain a high-density connected component in which the noisy measure is globally admissible, thus similar to a non-noisy tiling. Theorem 4.40 concludes this section with the fact that periodic SFTs are linearly stable, with a $O(\varepsilon)$ speed of convergence.

This linear $O(\varepsilon)$-stability result is to put into perspective with the $O\left(1/\sqrt{\ln(1/\varepsilon)}\right)$-stability obtained in Section 4.5, using the strategy for tilesets with robust combinatorial properties described by Durand, Romashchenko and Shen [DRS12], which holds in particular for periodic tilings [BDJ10]. A key interest of their argument, however, is that it also applies to some (robust) hierarchical aperiodic tilings.

Thus, in Section 4.6, the focal point switches from periodic to hierarchical aperiodic SFTs. Notably, we see that the enhanced Robinson tiling is $O\left(\sqrt[3]{\varepsilon}\right)$-stable (Theorem 4.50). The key idea here is that Robinson configurations are almost periodic, up to a low-density grid of cells, so that we may adapt the periodic percolation argument from Section 4.4. This result is interesting for two reasons. First, the Robinson tileset is not combinatorially robust in the sense of Section 4.5, so it provides a new, perhaps simpler example of stable aperiodic tiling. Second, the speed obtained here, though not linear, is still polynomial, thus much faster than the one one from Section 4.5 for comparable self-similar aperiodic tilesets. This section concludes with a proof that the Red-Black Robinson tiling is unstable. Then, Section 4.7 generalises the scheme of proof for stability of the enhanced Robinson tileset to a class of almost periodic SFTs.

At this point, we have all the necessary tools to study the stability of simulating tilesets. By iterating upon both the stable and unstable Robinson constructions, in Section 4.8, I will step-by-step craft simulating tilings to show that deciding if an SFT is stable is Π_1-hard, Σ_1-hard and finally Π_2-hard. Notably, these hardness bounds are not only true in two dimensions but more broadly for any fixed dimension $d \geq 2$.

Lastly, in Section 4.9, I establish a complementary Π_4 upper bound on the complexity of stability in the general case. In doing so, I will use ideas typical in computable analysis, as discussed in Section 3.6, applied to the objects introduced in Section 4.1, but without using anything from the following sections except conjugacy invariance.

4.1 Noisy Framework and (Weak-$*$) Stability

In Chapter 2, I introduced the usual framework for SFTs. Notably, in this setting, a configuration either is (globally) admissible, or contains forbidden patterns, without a natural way to quantify mistakes. Hence, I will here extend the alphabet of $X_{\mathcal{F}} \subseteq \Omega_{\mathcal{A}}$ to define the noisy *clair-obscur* framework used for all the rest of the current chapter.

Before doing so, let me just discuss another general notion regarding SFTs, that of the reconstruction function, that "quantifies" the gap between locally and globally admissible patterns.

Remark 4.1 (Reconstruction Function)**.** Fix $\mathcal{F}$ a set of forbidden patterns. Consider $\varphi : \mathcal{P}_{\Subset}\left(\mathbb{Z}^d\right) \to \mathbb{N}$ the reconstruction function defined on finite windows $I \Subset \mathbb{Z}^d$ by:

$$\varphi_{\mathcal{F}}(I) = \inf\left\{k \in \mathbb{N}, w \in \mathcal{A}^{I+B_k} \text{ is locally admissible } \Rightarrow w_I \text{ is globally admissible}\right\},$$

with $B_k = [\![-k, k]\!]^d$ the ball of radius $k \in \mathbb{N}$ (for the $\|.\|_\infty$ norm), thus a $(2k+1)$-square.

A priori $\varphi_{\mathcal{F}}(I)$ could be infinite. However, as $\mathcal{A}^I$ is finite, let $L_I \subseteq \mathcal{A}^I$ the finite subset of locally admissible patterns that are *not* globally admissible. If $v \in L_I$ could be embedded into arbitrarily large admissible patterns, then by compactness it would be globally admissible. In other words, there exists a rank $k(v) \in \mathbb{N}$ after which there is no locally admissible $w \in \mathcal{A}^{I+B_k}$ such that $w_I = v$. This holds for all the patterns in L_I, so that $\varphi_{\mathcal{F}}(I) = \max_{v \in L_I} k(v) < \infty$.

As we can embed Turing machines into SFTs, as discussed in Section 3.5, the function $\varphi : (I, \mathcal{F}) \mapsto \varphi_{\mathcal{F}}(I)$ is a sort of non-computable busy beaver, that cannot be computed without deciding the domino problem. However, we will see in Section 4.4 that when $\mathcal{F}$ induces a periodic SFT, the function φ is bounded by a constant.

This reconstruction idea, in one way or another, will play a role in all the stability results of the chapter. Notably, for some aperiodic tilings, for which the traditional reconstruction arguments do not hold, we will study the "almost reconstruction" of an "almost periodic structure" instead, in Sections 4.6 and 4.7. ■

Definition 4.2 (Noisy SFT)**.** Let $X_{\mathcal{F}} \subseteq \Omega_{\mathcal{A}}$ an SFT. Consider the alphabet $\widetilde{\mathcal{A}} = \mathcal{A} \times \{\mathtt{0}, \mathtt{1}\}$. For a given cell value $(a, b) \in \widetilde{\mathcal{A}}$, whenever $b = \mathtt{0}$ we say that the cell is *clear*, and when $b = \mathtt{1}$ we say that the cell is *obscured*. We may thus identify $\mathcal{A}$ to the clear subset $\mathcal{A} \times \{\mathtt{0}\} \subsetneq \widetilde{\mathcal{A}}$. By extension, we will call patterns and configurations *clear* if they only have clear cells, as opposed to the *obscured* ones, that contain at least one *obscured* letter in $\widetilde{\mathcal{A}} \backslash \mathcal{A}$.

Using the same identification, we define the set of *clear* forbidden patterns induced by $\mathcal{F}$:

$$\widetilde{\mathcal{F}} = \left\{\left(p, \mathtt{0}^{I(p)}\right) \in \widetilde{\mathcal{A}}^{I(p)}, p \in \mathcal{F}\right\}.$$

Consequently, we have a noisy SFT $X_{\widetilde{\mathcal{F}}} \subseteq \Omega_{\widetilde{\mathcal{A}}}$, such that $X_{\mathcal{F}} \subsetneq X_{\widetilde{\mathcal{F}}}$ is the subset of clear globally admissible configurations. ■

Notably, we also have $\Omega_{\mathcal{A}} \times \{\mathtt{1}^\infty\} \subsetneq X_{\widetilde{\mathcal{F}}}$, so $\pi^1\left(X_{\widetilde{\mathcal{F}}}\right) = \Omega_{\mathcal{A}}$ (with π^1 the canonical projection from $\widetilde{\mathcal{A}}$ to $\mathcal{A}$, and likewise $\pi^2 : \widetilde{\mathcal{A}} \to \{\mathtt{0}, \mathtt{1}\}$). In other words, without any more constraints, the projection on the initial alphabet doesn't depend at all on $\mathcal{F}$, and the same remark holds for $\mathcal{M}_\sigma\left(X_{\widetilde{\mathcal{F}}}\right)$. However, the second coordinate now gives us a simple way to quantify the mistakes, and we introduce the following measure spaces.

Definition 4.3 (Noisy Probability Measures)**.** Let $\varepsilon \in [0, 1]$. An invariant measure $\nu \in \mathcal{M}_\sigma\left(\Omega_{\{\mathtt{0},\mathtt{1}\}}\right)$ is called an ε-noise if the probability of a given cell being obscured is at most ε, *i.e.* $\nu([1]) \leq \varepsilon$.

For a given class of noises $\mathcal{N} \subseteq \mathcal{M}_\sigma\left(\Omega_{\{0,1\}}\right)$, we now define the measure space:

$$\widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}}(\varepsilon) := \left\{\lambda \in \mathcal{M}_\sigma\left(X_{\widetilde{\mathcal{F}}}\right), \pi_*^2(\lambda) \in \mathcal{N} \text{ is an } \varepsilon\text{-noise}\right\}.$$

Likewise, we define the projection $\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}(\varepsilon) = \pi_*^1\left(\widetilde{M_{\mathcal{F}}^{\mathcal{N}}}(\varepsilon)\right) \subseteq \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}}\right)$. If no class is written, it is implied that $\mathcal{N} = \mathcal{M}_\sigma\left(\Omega_{\{0,1\}}\right)$, that we allow *any* noise. ■

In particular, we are interested in the limit behaviour as the amount of "permitted forbidden patterns" ε goes to 0.

Remark 4.4 (Noise vs. Impurities)**.** With the notion of *clair-obscur* noise defined above, comparing two obscured configurations is mere matter of projecting them back onto the initial alphabet $\mathcal{A}$, which results in configurations that may have some amount of forbidden patterns.

Another way to define noise would be to add a blank symbol $\square \notin \mathcal{A}$ not already in the alphabet, without changing $\mathcal{F}$. The main difference in this case is that there is no natural way to project $\square$ into $\mathcal{A}$, but we may instead directly compare them on the extended alphabet. The symbol $\square$ behaves less like noise and more like an impurity in the lattice.

From the point of view of the entropy, this changes things up. Informally, when the binary noise is maximal, we can obtain the uniform measure on $\Omega_{\mathcal{A}}$, for which the entropy is maximal. In comparison, the only measure that maximizes the amount of impurities is the Dirac measure $\delta_{\square^{\mathbb{Z}^d}}$ which has a null entropy. Studying more precisely the behaviour of the entropy in either of these settings, as a function of the amount of noise, may yield interesting further results. ■

Definition 4.5 (Classes of Dependent Noises)**.** We define $\mathcal{B} = \left\{\mathcal{B}(\varepsilon)^{\otimes \mathbb{Z}^d}, \varepsilon \in [0,1]\right\}$ the class of independent Bernoulli noises, where each cell is obscured with probability ε independently of the other cells.

More generally, we consider the class $\mathcal{D}_k$ of k-dependent noises, such that any two windows at distance at least k are independent. More formally, $\nu \in \mathcal{D}_k$ when, for any patterns $w \in \mathcal{A}^I$ and $w' \in \mathcal{A}^J$ such that $d_\infty(I,J) \geq k$, $\nu([w] \cap [w']) = \nu([w])\nu([w'])$. For any rank k, we naturally have $\mathcal{D}_k \subseteq \mathcal{D}_{k+1}$. In particular, $\mathcal{D}_1 = \mathcal{B}$. ■

A direct consequence of Definition 4.3 is that, on any class $\mathcal{N}$, for $\varepsilon < \delta$, we have the increasing inclusion $\widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}}(\varepsilon) \subseteq \widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}}(\delta)$, which naturally still holds for $\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}$ after projection. Let us notice that $\mathcal{M}_{\mathcal{F}}(0) = \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$ is non-empty, and that $\mathcal{M}_{\mathcal{F}}(1) = \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}}\right)$ is the set of shift-invariant measures on $\Omega_{\mathcal{A}}$.

We are now interested in the *stability* of the *set* of noisy measures, *i.e.* in the fact that $\mathcal{M}_{\mathcal{F}}(\varepsilon)$ gets close to $\mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$ in some sense – for some topology – as ε goes to 0. We formalise the notion as follows.

Definition 4.6 (Stability)**.** Consider here a distance d on $\mathcal{M}_\sigma\left(\Omega_{\mathcal{A}}\right)$, a noise class $\mathcal{N} \subseteq \mathcal{M}_\sigma\left(\Omega_{\{0,1\}}\right)$, and a non-decreasing function $f : [0,1] \to \mathbb{R}^+$, right-continuous at 0 with $f(0) = 0$.

The SFT induced by $\mathcal{F}$ is said to be f-stable (with respect to the distance d on the class $\mathcal{N}$) if:

$$\forall \varepsilon \in [0,1], \sup_{\mu \in \mathcal{M}_{\mathcal{F}}^{\mathcal{N}}(\varepsilon)} d\left(\mu, \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right) \leq f(\varepsilon).$$

The SFT is stable if it is f-stable for some function f. We say that $X_{\mathcal{F}}$ is linearly stable (resp. polynomially stable) if it is f-stable with $f(\varepsilon) = O(\varepsilon)$ (resp. $f(\varepsilon) = O\left(\varepsilon^\alpha\right)$ for some $0 < \alpha \leq 1$). ■

A natural topology on measures to consider first is the weak-∗ topology, but we will see here that as $\varepsilon \to 0$, any adherence point of a sequence of noisy measures (regardless of the class of noises) is in $\mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$, which implies every SFT is stable in the previous sense with this topology.

Lemma 4.7. *Let $\mu \in \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}}\right)$. If, for every pattern $p \in \mathcal{F}$, we have $\mu([p]) = 0$, then $\mu \in \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$.*

Proof. To show that $\mu \in \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$, we need to show that the measure is supported on $X_{\mathcal{F}}$ (*i.e.* $\mu\left(X_{\mathcal{F}}\right) = 1$, as $X_{\mathcal{F}}$ is closed). The complement of $X_{\mathcal{F}}$ is the set $\bigcup_{k \in \mathbb{Z}^d} \bigcup_{p \in \mathcal{F}} \sigma^k([p])$. By shift-invariance, for any $p \in \mathcal{F}$ and $k \in \mathbb{Z}^d$, we have $\mu\left(\sigma^k([p])\right) = \mu([p]) = 0$, thus $\mu\left(X_{\mathcal{F}}^c\right) = 0$ by union bound. □

Proposition 4.8. *Let $\mu_n \in \mathcal{M}_{\mathcal{F}}\left(\varepsilon_n\right)$ be a sequence of noisy measures, with $\varepsilon_n \underset{n\to\infty}{\longrightarrow} 0$. Then any adherence value of the sequence is in $\mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$.*

Proof. Consider a weakly converging subsequence $\mu_{\theta(n)} \to \mu$, with $\theta : \mathbb{N} \to \mathbb{N}$ an increasing map. Naturally, the limit μ is also an invariant measure.

Notice that, as ε-noises are defined by forcing the measure of the noise cylinder $[\mathtt{1}]$ to belong to the closed set $[0, \varepsilon]$, the set $\widetilde{\mathcal{M}_{\mathcal{F}}}(\varepsilon)$ is naturally weakly closed. Hence, by monotonous inclusion, for any $\varepsilon > 0$ we have $\mu \in \mathcal{M}_{\mathcal{F}}(\varepsilon)$.

Consider $\lambda_\varepsilon \in \widetilde{\mathcal{M}_{\mathcal{F}}}(\varepsilon)$ that projects to μ, and a forbidden pattern $p \in \mathcal{F}$. Thus, $\mu([p]) = \lambda_\varepsilon\left(\left[p, \{\mathtt{0},\mathtt{1}\}^{I(p)}\right]\right)$. In particular, because $\lambda_\varepsilon \in \mathcal{M}_\sigma\left(X_{\widetilde{\mathcal{F}}}\right)$ and $\left(p, \mathtt{0}^{I(p)}\right) \in \widetilde{\mathcal{F}}$, we have $\lambda_\varepsilon\left(\left[p, \mathtt{0}^{I(p)}\right]\right) = 0$. For the rest of the values of $b \in \{\mathtt{0},\mathtt{1}\}^{I(p)}$, we must have at least one obscure cell in $I(p)$, thus by union bound:

$$\mu([p]) \leq |I(p)| \times \lambda_\varepsilon([\mathcal{A}, \mathtt{1}]) \leq |I(p)|\varepsilon\,.$$

As ε goes to 0, we conclude that $\mu([p]) = 0$. Using the previous lemma, $\mu \in \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$. □

Proposition 4.9 (Weak-$*$ Stability)**.** *Given a metric d which induces the weak-$*$ topology, any SFT is stable with respect to the distance d.*

Proof. Assume $\mathcal{F}$ induces an unstable SFT with respect to d. Then we have a sequence $\varepsilon_n \underset{n\to\infty}{\longrightarrow} 0$ and measures $\mu_n \in \mathcal{M}_{\mathcal{F}}\left(\varepsilon_n\right)$ such that $\inf_{n\in\mathbb{N}} d\left(\mu_n, \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right) = \delta > 0$. By compactness, this sequence admits a weak-$*$ adherence value μ, and $d\left(\mu, \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right) \geq \delta$. This contradicts the previous proposition. □

Hence, all SFTs are weakly stable. As there is no *canonical* metric associated to the weak-$*$ topology, there is no real point in trying to quantify the speed of convergence in this case.

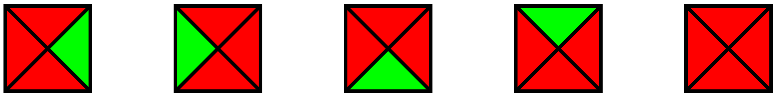

Figure 4.1: Wang domino tileset, with the rightmost tile used for empty cells.

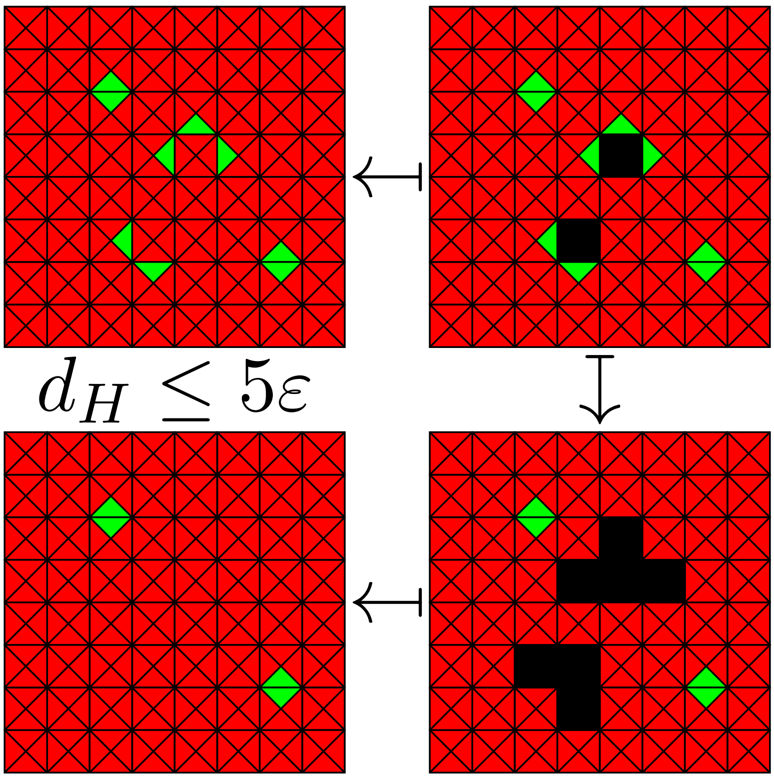

Figure 4.2: Mistakes in diluted domino tilings can be repaired locally.

The main issue with the weak-$*$ topology is that it looks at things on a local scale, on finite patterns, without really forcing any kind of behaviour on $\mathbb{Z}^d$ as a whole. In this context, this is insufficient to discriminate between SFTs. Hence, we will from now on discuss stability in the Besicovitch topology instead, looking at configurations globally, which will prove to be an appropriate scale to obtain several behaviours (both stable and unstable) with various convergence speeds.

To illustrate the notion of stability in the Besicovitch topology, let me give a stable example. Consider the diluted domino tileset in Figure 4.1, with forbidden patterns $\mathcal{F}$ on the alphabet $\mathcal{A}$. The two leftmost tiles represent the two halves of a horizontal domino, and must be paired in an admissible configuration. Likewise, the two following ones pair to form a vertical domino. The rightmost tile permits empty cells, without any domino in them, hence why the corresponding tilings are "diluted" in some sense.

In Figure 4.2, we can see how errors can be locally repaired (thus the process is overall measurable). On the top row, we have a noisy configuration on the right, admissible for $\widetilde{\mathcal{F}}$, with obscured cells in black (looking at the first coordinate $\mathcal{A}$ on the left, we can see a pattern that is not admissible for $\mathcal{F}$). We can proceed to obscure any half-domino next to an obscured cell (bottom-right), and then replace any obscured cell by a red empty cell (bottom-left). As any obscure cell has at most four such neighbours, we conclude that, starting from any configuration $(\omega, b) \in X_{\widetilde{\mathcal{F}}}$ we obtain $\omega' = \psi(\omega, b) \in X_{\mathcal{F}}$ such that $d_H(\omega, \omega') \le 5 d_H(b, 0^\infty)$, and the transformation ψ commutes with the shift action. Thus, if $(\omega, b) \sim \lambda \in \widetilde{\mathcal{M}_{\mathcal{F}}}(\varepsilon)$, it follows that:

$$d_B\left(\pi^1_*(\lambda), \mathcal{M}_\sigma(X_{\mathcal{F}})\right) \le d_B\left(\pi^1_*(\lambda), \psi_*(\lambda)\right) \le \int d_H(\omega, \psi(\omega, b)) \mathrm{d}\lambda(\omega, b) \le 5 \int d_H(b, 0^\infty) \mathrm{d}\lambda(\omega, b) \le 5\varepsilon\,.$$

As this holds for any ε-noisy measure λ, we can conclude that this diluted domino tileset is stable, with speed of convergence $f(\varepsilon) = 5\varepsilon$ at most.

This example illustrates how a typical stability proof will unravel. In due time, in Section 4.3.2, I will give an unstable example, illustrating a typical instability argument with a lower bound on d_B that does not go to 0 as ε does. This will in particular establish that not every SFT is stable, and that this notion induces indeed a non-trivial classification of SFTs.

Before doing so, however, I will discuss how stability, the definition of which intrinsically depends on the choice of $\mathcal{F}$, actually characterises the SFT $X_{\mathcal{F}}$ itself. In the next section, we will more precisely see that if two sets of patterns induce conjugate SFTs, then one is stable (for the class of all noises, or $\mathcal{B}$ the Bernoulli noises) *iff* the other is too. In other words, this notion of stability is in some ways conjugacy-invariant.

4.2 Conjugacy Invariance

As announced, from now on and for the rest of the chapter, we will always discuss stability for the Besicovitch topology defined in Section 2.6. In this topology, we quantify the frequency of differences on average.

In particular, to obtain an upper bound on the speed of convergence, what we will do in practice is start from a noisy measure $\lambda \in \mathcal{M}^{\mathcal{N}}_{\mathcal{F}}(\varepsilon)$, use an appropriate measurable map $\gamma : \Omega_{\widetilde{\mathcal{A}}} \to X_{\mathcal{F}}$ (not necessarily a morphism), and conclude that $d_B\left(\pi^1_*(\lambda), \mathcal{M}_\sigma(X_{\mathcal{F}})\right) \le \int d_H(\omega, \gamma(\omega, b)) \mathrm{d}\lambda(\omega, b)$ (*i.e.* for *any* noisy measure, we just need *one* coupling to obtain *an* upper bound).

For the lower bounds however, notably for the unstable cases, we will need to prove that there *exists* a noisy measure (for each ε) such that every configuration in its support is far from *any* element of $X_{\mathcal{F}}$. Let us begin the section by giving a universal lower bound on the convergence speed, which illustrates this idea.

Lemma 4.10 (Universal Linear Lower Bound). *Let $X_{\mathcal{F}} \subseteq \Omega_{\mathcal{A}}$ a non-empty SFT. Let $\mathcal{A}' = \mathcal{A} \sqcup \{\square\}$ and $\mathcal{F}' = \mathcal{F} \sqcup \{\square\}$, so that $X_{\mathcal{F}} = X_{\mathcal{F}'} \subseteq \Omega_{\mathcal{A}} \subseteq \Omega_{\mathcal{A}'}$. Then, $d_B(\mathcal{M}_{\mathcal{F}'}(\varepsilon), \mathcal{M}_\sigma(X_{\mathcal{F}'})) \ge \varepsilon$.*

Proof. By non-emptiness, consider an arbitrary non-noisy measure $\mu \in \mathcal{M}_\sigma(X_{\mathcal{F}'})$. Let $\psi : \mathcal{A} \times \Omega_{0,1} \to \mathcal{A}'$ be the cellular automaton such that $\psi(a, 0) = a$ and $\psi(a, 1) = \square$. In other words, $\mu_\varepsilon := \psi_*\left(\mu \otimes \mathcal{B}(\varepsilon)^{\otimes \mathbb{Z}^d}\right) \in \mathcal{M}^{\mathcal{B}}_{\mathcal{F}'}(\varepsilon)$ is the measure where any obscured cell is filled with the $\square$ symbol (that cannot appear anywhere in $X_{\mathcal{F}'}$), but the clear area is globally admissible. Hence, μ_ε-a.s. we have $\omega' \in \Omega_{\mathcal{A}'}$ such that $d_H(\omega', X_{\mathcal{F}'}) = \varepsilon$, thus at last $d_B(\mathcal{M}_{\mathcal{F}'}(\varepsilon), \mathcal{M}_\sigma(X_{\mathcal{F}'})) \ge d_B(\mu_\varepsilon, \mathcal{M}_\sigma(X_{\mathcal{F}'})) \ge \varepsilon$. ☐

Likewise, this can be adapted for specific classes of noise. Informally, this means that for any SFT, there is a conjugate SFT for which we have a linear lower bound on the convergence speed. Hence, linear stability is the best we can hope to obtain, in particular if we work up to conjugacy. As announced, we now want to see how morphisms and conjugacies affect stability.

4.2.1 Invariance for All Noises

Definition 4.11 (Thickened Noise)**.** Let $\gamma^n : \{0,1\}^{B_n} \to \{0,1\}$ be the cellular automaton locally defined by $\gamma^n(w) = \max_{k\in B_n} w_k$. We say that $\gamma^n(\omega)$ is n-thickened for $\omega \in \Omega_{\{0,1\}}$ in the sense that if the cell $c \in \mathbb{Z}^d$ is obscured in ω, then its n-neighbourhood $c + B_n$ is obscured in $\gamma^n(\omega)$. ■

These specific morphisms will allow us to entirely obscure the neighbourhood of forbidden patterns that may appear when using another morphism or a measurable application on $\Omega_{\mathcal{A}}$ later on.

Lemma 4.12. *Consider the SFTs $X_{\mathcal{F}} \subseteq \Omega_{\mathcal{A}_1}$ and $X_{\mathcal{G}} \subseteq \Omega_{\mathcal{A}_2}$, and a morphism from $X_{\mathcal{F}}$ to $X_{\mathcal{G}}$, locally defined as $\theta : \mathcal{A}_1^J \to \mathcal{A}_2$. Then for any $x, y \in \Omega_{\mathcal{A}_1}$, we have $d_H(\theta(x), \theta(y)) \le D_\theta d_H(x,y)$ with the constant $D_\theta = |J|$. Consequently, for any measures $\mu, \nu \in \mathcal{M}_\sigma(\Omega_{\mathcal{A}_1})$, we have $d_B(\theta_*(\mu), \theta_*(\nu)) \le D_\theta d_B(\mu,\nu)$.*

*There exists a radius r_θ such that $\widetilde{\theta} := (\theta, \gamma^{r_\theta}) : \Omega_{\widetilde{\mathcal{A}_1}} \to \Omega_{\widetilde{\mathcal{A}_2}}$ is a morphism from $X_{\widetilde{\mathcal{F}}}$ to $X_{\widetilde{\mathcal{G}}}$ (*i.e. *we have $\widetilde{\theta}(X_{\widetilde{\mathcal{F}}}) \subseteq X_{\widetilde{\mathcal{G}}}$). Moreover, there is a constant C_θ such that, whenever $\gamma_*^{r_\theta}(\mathcal{N}) \subseteq \mathcal{N}'$, for any $\varepsilon > 0$:*

$$\widetilde{\theta}_*\left(\widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}}(\varepsilon)\right) \subseteq \widetilde{\mathcal{M}_{\mathcal{G}}^{\mathcal{N}'}}\left(C_\theta \times \varepsilon\right) .$$

Proof. Assume that $\theta(x)_k \ne \theta(y)_k$. Then x and y *must* differ in at least one cell of the window $J + k$. Conversely, each difference between x and y can induce at most $|J|$ differences between $\theta(x)$ and $\theta(y)$, hence $d_H(\theta(x), \theta(y)) \le |J| d_H(x,y)$ by a finite union bound. Now, assuming $d_B(\mu,\nu)$ is reached for a coupling $\lambda \in \mathcal{M}_\sigma(\Omega_{\mathcal{A}_1\times\mathcal{A}_1})$, then $(\theta,\theta)_*(\lambda) \in \mathcal{M}_\sigma(\Omega_{\mathcal{A}_2\times\mathcal{A}_2})$ is a coupling between $\theta_*(\mu)$ and $\theta_*(\nu)$, and we consequently obtain the analogous bound for d_B.

Just like $\theta : \mathcal{A}_1^J \to \mathcal{A}_2$ naturally sends $\Omega_{\mathcal{A}_1}$ onto $\Omega_{\mathcal{A}_2}$, it sends any pattern $v \in \mathcal{A}_1^{J+I}$ onto $\theta(v) \in \mathcal{A}_2^I$. The "local" property that characterises $\theta(X_{\mathcal{F}}) \subseteq X_{\mathcal{G}}$ is *not* that it preserves locally admissible patterns, but that it preserves globally admissible ones, as discussed in the proof of Proposition 2.14 for example.

We cannot simply extend the morphism $\theta : \Omega_{\mathcal{A}_1} \to \Omega_{\mathcal{A}_2}$ as $\widetilde{\theta}$ by leaving the second coordinate unchanged. Indeed, if a locally admissible pattern $v \in \mathcal{A}_1^{J+I}$ is not globally admissible, nothing forbids $\theta(v) \in \mathcal{A}_2^I$ from containing forbidden patterns of $\mathcal{G}$. In such a case, let us extend v as $\omega \in \Omega_{\mathcal{A}_1}$ by filling the empty cells outside of $I + J$ with any letter $a \in \mathcal{A}_1$, and consider the noise $b = \mathbb{1}_{(I+J)^c}$ that obscures all the cells outside of $I + J$. Then naturally $(\omega, b) \in X_{\widetilde{\mathcal{F}}}$ is locally admissible but $(\theta(\omega), b) \notin X_{\widetilde{\mathcal{G}}}$ is not.

Assume that $w = \theta(v) \in \mathcal{G}$ is a forbidden pattern, with $v \in \mathcal{A}_1^{J+I(w)}$. Then v must not be *globally* admissible itself. As explained in Remark 4.1, using the reconstruction function, we have $r(w) = \varphi_{\mathcal{F}}(J + I(w)) \in \mathbb{N}$ such that, if we can extend v into a locally admissible pattern in $\mathcal{A}_1^{J+I(w)+B_{r(w)}}$, then v itself must be globally admissible.

Let us define $r_\theta = \max_{w\in\mathcal{G}} r(w) + \max_{c\in J} \|c\|_\infty$. Consider $(\omega, b) \in X_{\widetilde{\mathcal{F}}}$. If $\theta(\omega)$ contains a forbidden pattern w in the window $c + I(w)$, then it follows that the window $c + J + I(w)$ of ω is not globally admissible, so the window $c + I(w) + B_{r_\theta}$ of $\omega \in \Omega_{\mathcal{A}_1}$ must not be locally admissible (for $\mathcal{F}$). As (ω, b) is locally admissible (for $\mathcal{F}'$), this implies that at least one cell in $c + I(w) + B_{r_\theta}$ must be obscured. We proved that, if $(\omega, b) \in X_{\widetilde{\mathcal{F}}}$, then $(\theta(\omega), \gamma^{r_\theta}(b)) \in X_{\widetilde{\mathcal{G}}}$, so $\widetilde{\theta} = (\theta, \gamma^{r_\theta})$ is the morphism we wanted.

Finally, we need to exhibit the constant C_θ. Consider a noisy measure $\lambda \in \widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}}(\varepsilon)$, with an ε-noise $\nu = \pi_*^2(\lambda) \in \mathcal{N}$. Notice that $\pi^2 \circ \widetilde{\theta} = \gamma^{r_\theta}$, so the noise of $\widetilde{\theta}^*(\lambda)$ is actually $\gamma_*^{r_\theta}(\nu) \in \mathcal{N}'$. Remark that the clear configuration 0^∞ is a fixed point of γ^{r_θ}. Using a pointwise ergodic theorem, the amount of noise in ν is:

$$\nu([1]) = \int \mathbb{1}_{\{b_0\ne 0\}} \mathrm{d}\nu(b) = \int d_H\left(b, 0^\infty\right) \mathrm{d}\nu(b) = d_B\left(\nu, \delta_{0^\infty}\right) .$$

Thus, if we apply the first part of the current lemma to the morphism γ^{r_θ}, with the set $J = B_{r_\theta}$, we conclude that $d_B\left(\gamma_{r_\theta}^*(\nu), \delta_{0^\infty}\right) \le C_\theta d_B\left(\nu, \delta_{0^\infty}\right) \le C_\theta \times \varepsilon$ with the constant $C_\theta = |J| = (2r_\theta + 1)^d$. At last, $\gamma_*^{r_\theta}(\nu)$ is a $(C_\theta\varepsilon)$-noise, so $\widetilde{\theta}^*(\lambda) \in \widetilde{\mathcal{M}_{\mathcal{G}}^{\mathcal{N}'}}(C_\theta\varepsilon)$. □

Assume that the SFT $X_{\mathcal{F}}$ is f-stable, and that it is sent to $X_{\mathcal{G}}$ by θ. The lemma suggests that the *subset* $\pi_*^1\left(\widetilde{\theta}_*\left(\widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}}(\varepsilon)\right)\right) = \theta_*\left(\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}(\varepsilon)\right)$ of $\mathcal{M}_{\mathcal{G}}^{\gamma^{r_\theta}(\mathcal{N})}(C_\theta\varepsilon)$ is roughly $D_\theta \times f$-stable. However, this still does not give us enough information to obtain a full-fledged and well-defined stability property for $\mathcal{G}$. To obtain such a result, we will now assume that θ is a conjugacy between $X_{\mathcal{F}}$ and $X_{\mathcal{G}}$.

Theorem 4.13 (Conjugacy-Invariant Stability)**.** *Consider a conjugacy $\theta : X_{\mathcal{F}} \to X_{\mathcal{G}}$, and assume that $X_{\mathcal{F}}$ is f-stable with respect to d_B on a class $\gamma_*^{r_{\theta^{-1}}}(\mathcal{N})$ of noises (using the radius for the morphism θ^{-1} in the previous lemma).*

Then there exists a constant E such that $X_{\mathcal{G}}$ is g-stable on $\mathcal{N}$ with the speed

$$g : \varepsilon \mapsto D_\theta f\left(C_{\theta^{-1}}\varepsilon\right) + E\varepsilon\,.$$

Proof. We will use the results of the previous lemma for both $\theta : X_{\mathcal{F}} \to X_{\mathcal{G}}$ *and* its inverse $\theta^{-1} : X_{\mathcal{G}} \to X_{\mathcal{F}}$. Note that, on the larger domain $\Omega_{\mathcal{A}_2}$, the cellular automaton $\theta \circ \theta^{-1}$ is still well-defined, but is not necessarily the identity function outside of the domain $X_{\mathcal{G}}$. Now, if we consider two measures $\mu, \nu \in \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}_2}\right)$:

$$d_B(\mu,\nu) \le d_B\left(\mu, \left(\theta\circ\theta^{-1}\right)_*(\mu)\right) + d_B\left(\left(\theta\circ\theta^{-1}\right)_*(\mu), \left(\theta\circ\theta^{-1}\right)_*(\nu)\right) + d_B\left(\left(\theta\circ\theta^{-1}\right)_*(\nu),\nu\right)\,.$$

The idea behind this back-and-forth is that, by going from $X_{\mathcal{G}}$ to $X_{\mathcal{F}}$, we reach a stable SFT while still keeping the noise under control, and then going from $X_{\mathcal{F}}$ to $X_{\mathcal{G}}$ allows us to maintain this stability while comparing the new configuration to the old one. In particular, if $\nu \in \mathcal{M}_\sigma\left(X_{\mathcal{G}}\right)$, then it is supported on the domain $X_{\mathcal{G}}$, where $\theta \circ \theta^{-1}$ is the identity function, so that $d_B\left(\left(\theta\circ\theta^{-1}\right)_*(\nu),\nu\right) = 0$.

Consider a measure $\mu \in \mathcal{M}_{\mathcal{G}}^{\mathcal{N}}(\varepsilon)$, and $\nu_{\mathcal{F}} \in \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$ that achieves $d_B\left(\left(\theta^{-1}\right)_*(\mu), \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right)$. If we denote $\nu_{\mathcal{G}} = \theta_*\left(\nu_{\mathcal{F}}\right) \in \mathcal{M}_\sigma\left(X_{\mathcal{G}}\right)$, then:

$$d_B\left(\mu, \mathcal{M}_\sigma\left(X_{\mathcal{G}}\right)\right) \le d_B\left(\mu, \nu_{\mathcal{G}}\right) \le d_B\left(\mu, \left(\theta\circ\theta^{-1}\right)_*(\mu)\right) + d_B\left(\left(\theta\circ\theta^{-1}\right)_*(\mu), \nu_{\mathcal{G}}\right)\,.$$

In particular, using the previous lemma for θ^{-1}, we know that $\left(\theta^{-1}\right)_*(\mu) \in \mathcal{M}_{\mathcal{F}}^{\gamma_*^{r_{\theta^{-1}}}(\mathcal{N})}\left(C_{\theta^{-1}}\varepsilon\right)$. Thence, using the lemma for θ, as $X_{\mathcal{F}}$ is f-stable on $\gamma_*^{r_{\theta^{-1}}}(\mathcal{N})$, we get the bound:

$$\begin{aligned} d_B\left(\left(\theta\circ\theta^{-1}\right)_*(\mu), \nu_{\mathcal{G}}\right) &\le D_\theta d_B\left(\left(\theta^{-1}\right)_*(\mu), \nu_{\mathcal{F}}\right) \\ &= D_\theta d_B\left(\left(\theta^{-1}\right)_*(\mu), \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right) \\ &\le D_\theta f\left(C_{\theta^{-1}}\varepsilon\right)\,. \end{aligned}$$

To conclude the proof, we just need to have a linear control on $d_B\left(\mu, \left(\theta\circ\theta^{-1}\right)_*(\mu)\right)$ as $\varepsilon \to 0$. To do so, we will study $d_H\left(x, \theta\circ\theta^{-1}(x)\right)$ for any $x \in \Omega_{\mathcal{A}_2}$. More precisely, whenever $(x,b) \in X_{\widetilde{\mathcal{F}}}$, we want a bound $d_H\left(x, \theta\circ\theta^{-1}(x)\right) \le E d_H\left(b, 0^\infty\right)$. Assuming such a bound holds, consider $\lambda \in \widetilde{\mathcal{M}_{\mathcal{G}}^{\mathcal{N}}}(\varepsilon)$ that projects to μ, which naturally gives a coupling between $\mu = \pi^1_*(\lambda)$ and $\left(\theta\circ\theta^{-1}\right)_*(\mu) = \left(\pi^1 \circ \widetilde{\theta} \circ \widetilde{\theta^{-1}}\right)_*(\lambda)$. Then we obtain:

$$d_B\left(\mu, \left(\theta\circ\theta^{-1}\right)_*(\mu)\right) \le \int_{\Omega_{\widetilde{\mathcal{G}}}} d_H\left(x, \theta\circ\theta^{-1}(x)\right) \mathrm{d}\lambda(x,b) \le E \int_{\Omega_{\widetilde{\mathcal{G}}}} d_H\left(b, 0^\infty\right) \mathrm{d}\lambda(x,b) \le E\varepsilon\,.$$

The sketch of the proof from now on is pretty much the same as in the previous lemma. Let us suppose that $(x,b) \in X_{\widetilde{\mathcal{G}}}$ and that $x_k \ne \theta\circ\theta^{-1}(x)_k$ for some cell $k \in \mathbb{Z}^d$. Consider the window $J = J_{\theta^{-1}} + J_\theta$ such that the value of $\theta\circ\theta^{-1}(x)_k$ only depends on the pattern $x|_{J+k}$. Let us assume that $0 \in J$ without loss of generality. If $x|_{J+k}$ was globally admissible, then we could extend it into a globally admissible configuration $y \in X_{\mathcal{G}}$, such that $\theta\circ\theta^{-1}(x)_k = \theta\circ\theta^{-1}(y)_k = y_k = x_k$. This contradicts our hypothesis, so x_{J+k} is not globally admissible. This means that, using once again the reconstruction function φ from Remark 4.1 for the SFT $X_{\mathcal{G}}$, $x|_{J+k+B_{\varphi(J)}}$ is not locally admissible, so the same windows in b contains at least one obscured cell. Hence, $d_H\left(x, \theta\circ\theta^{-1}(x)\right) \le E d_H\left(b, 0^\infty\right)$ with the constant $E = \left|J + B_{\varphi(J)}\right|$, which concludes the proof. □

Corollary 4.14 (Conjugacy-Invariance for Stable Noise Classes)**.** *If $\mathcal{N}$ is stable under the action of any γ^n, then for any two conjugate SFTs $X_{\mathcal{F}}$ and $X_{\mathcal{G}}$, $X_{\mathcal{F}}$ is stable (resp. linearly stable, polynomially stable) on the class $\mathcal{N}$* iff *$X_{\mathcal{G}}$ is.*

In particular, this corollary holds for the class of all noises $\mathcal{N} = \mathcal{M}_\sigma\left(\Omega_{\{0,1\}}\right)$. This stability hypothesis is actually quite restrictive. For example, we naturally have the inclusion $\gamma^n(\mathcal{B}) \subseteq \mathcal{D}_{2n+1}$ but $\gamma^n(\mathcal{B}) \nsubseteq \mathcal{D}_{2n}$. Thence, $\gamma^n(\mathcal{B}) \nsubseteq \mathcal{B}$, the previous conjugacy-invariance corollary does not apply on the class $\mathcal{N} = \mathcal{B}$ of Bernoulli noises.

4.2.2 Domination and Invariance for Bernoulli Noises

We will now introduce the notion of domination between noises, which will allow us to send $\mathcal{D}_k$ back into $\mathcal{B}$, in order to obtain a conjugacy-invariant stability result for the class $\mathcal{B}$.

Definition 4.15 (Domination)**.** A Borel set $B \subseteq \{0,1\}^{\mathbb{Z}^d}$ is said to be increasing if, for any $b \in B$ and $b' \geq b$ (on each coordinate), we have $b' \in B$.

Consider $\nu_1, \nu_2 \in \mathcal{M}_\sigma\left(\Omega_{\{0,1\}}\right)$. We say that ν_2 dominates ν_1, and we denote $\nu_2 \geq \nu_1$, if $\nu_2(B) \geq \nu_1(B)$ for any increasing Borel set B. Equivalently [Lig05, Theorem 2.4], $\nu_2 \geq \nu_1$ if there exists some coupling ν_{dom} between $\nu_1 = \pi^1_*(\nu_{\text{dom}})$ and $\nu_2 = \pi^2_*(\nu_{\text{dom}})$ which is supported on $X_{\leq} := \Omega_{\{(0,0),(0,1),(1,1)\}} \subsetneq \Omega^2_{\{0,1\}}$. Note that in the reference, this equivalence is stated in a general non-shift-invariant framework, so that ν_{dom} is not *a priori* in $\mathcal{M}_\sigma\left(\Omega^2_{\{0,1\}}\right)$, but as $X_{\leq}$ is compact, we can replace ν_{dom} by any weak-$*$ adherence value of $\left(\frac{1}{|I_n|}\sum_{k\in I_n} \sigma^k_*\left(\nu_{\text{dom}}\right)\right)_{n\in\mathbb{N}}$ to obtain a likewise shift-invariant coupling.

We can extend this notion to classes of measures. Let $g : [0,1] \to [0,1]$ be a non-decreasing function, right-continuous at 0 with $g(0) = 0$. We say that the class $\mathcal{N}$ is g-dominated by $\mathcal{N}'$ if, for any $\varepsilon > 0$ and any ε-noise $\nu \in \mathcal{N}$, there exists a $g(\varepsilon)$-noise $\nu' \in \mathcal{N}'$ such that $\nu' \geq \nu$. ■

Lemma 4.16 (Disintegration Theorem [Kal02, Theorem 8.5])**.** *Let λ be any probability measure on $\Omega_{\mathcal{A}} \times \Omega_{\mathcal{A}'}$, with $\mu = \pi^1_*(\lambda)$. We can factorise $\lambda(A \times B) = \int_A \nu_x(B)\mathrm{d}\mu(x)$, such that $x \mapsto \nu_x(B)$ is measurable for any measurable set B, and that $B \mapsto \nu_x(B)$ is a probability measure for μ-a.e. $x \in \Omega_{\mathcal{A}}$.*

Using this domination property along with the disintegration theorem, we can then prove the following result, that most notably does not depend on the distance d used for the stability.

Proposition 4.17. *If the SFT $X_{\mathcal{F}}$ is f-stable on the class $\mathcal{N}'$ with respect to the distance d, and $\mathcal{N}$ is g-dominated by $\mathcal{N}'$, then $X_{\mathcal{F}}$ is $(f \circ g)$-stable on the class $\mathcal{N}$ with respect to the distance d.*

Proof. Let us assume that $X_{\mathcal{F}}$ is f-stable on the class $\mathcal{N}'$. Consider $\mu \in \mathcal{M}^{\mathcal{N}}_{\mathcal{F}}(\varepsilon)$. If $\mu \in \mathcal{M}^{\mathcal{N}'}_{\mathcal{F}}(g(\varepsilon))$, then $d_B\left(\mu, \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right) \leq f(g(\varepsilon))$ by f-stability.

In order to prove this, let us consider a measure $\lambda \in \widetilde{\mathcal{M}^{\mathcal{N}}_{\mathcal{F}}}(\varepsilon)$ such that $\pi^1_*(\lambda) = \mu$, with $\pi^2_*(\lambda) = \nu \in \mathcal{N}$ an ε-noise. By domination, there exists a $g(\varepsilon)$-noise $\nu' \in \mathcal{N}'$ such that $\nu' \geq \nu$, with a coupling ν_{dom} between them.

Then, using the disintegration theorem, for ν-a.e. $b \in \Omega_{\{0,1\}}$, there is a measure μ_b on $\Omega_{\mathcal{A}}$ such that, for any two Borel sets $A \subseteq \Omega_{\mathcal{A}}$ and $B \subseteq \Omega_{\{0,1\}}$:

$$\lambda\left(A \times B\right) = \int_B \mu_b(A)\mathrm{d}\nu(b) = \int \mu_b(A)\mathbb{1}_B(b)\mathrm{d}\nu_{\text{dom}}\left(b, b'\right) .$$

Now, we can naturally define the measure λ' on $\Omega_{\mathcal{A}} \times \Omega_{\{0,1\}}$ as:

$$\lambda'\left(A \times B\right) = \int \mu_b(A)\mathbb{1}_B\left(b'\right)\mathrm{d}\nu_{\text{dom}}\left(b, b'\right) .$$

By taking $B = \Omega_{\{0,1\}}$, it is clear that $\pi^1_*\left(\lambda'\right) = \pi^1_*(\lambda) = \mu$. Now, by taking $A = \Omega_{\mathcal{A}}$, we conclude that $\pi^2_*\left(\lambda'\right) = \pi^2_*\left(\nu_{\text{dom}}\right) = \nu'$. Moreover, consider $\widetilde{w} = \left(w, 0^{I(w)}\right) \in \widetilde{\mathcal{F}}$ a forbidden pattern. Since ν_{dom} is supported by $X_{\leq} = \{(b, b'), b \leq b'\}$:

$$\lambda'\left([\widetilde{w}]\right) = \int \mu_b([w])\mathbb{1}_{0^{I(w)}}\left(b'\right)\mathrm{d}\nu_{\text{dom}}\left(b, b'\right) \leq \int \mu_b([w])\mathbb{1}_{0^{I(w)}}(b)\mathrm{d}\nu_{\text{dom}}\left(b, b'\right) = \lambda\left([\widetilde{w}]\right) = 0 .$$

Thence, λ' is supported on $X_{\widetilde{\mathcal{F}}}$. Without loss of generality, we can replace λ' by a weak-$*$ adherence value of the averages $\left(\frac{1}{|I_n|}\sum_{k\in I_n} \sigma^k_*\left(\lambda'\right)\right)_{n\in\mathbb{N}}$, which still projects to μ and ν', but is also shift-invariant, so that $\lambda' \in \widetilde{\mathcal{M}^{\mathcal{N}'}_{\mathcal{F}}}(g(\varepsilon))$. At last, we demonstrated that $\mu = \pi^1_*\left(\lambda'\right) \in \mathcal{M}^{\mathcal{N}'}_{\mathcal{F}}(g(\varepsilon))$, which concludes the proof. □

Now, in order to use this result for Bernoulli stability, we need to dominate $\gamma_n(\mathcal{B}) \subseteq \mathcal{D}_{2n+1}$ by $\mathcal{B}$. By adapting a classical result[LSS97, Theorem 1.3], we can obtain the following bound:

Theorem 4.18 (Domination Theorem)**.** *Consider $g^d_k : \varepsilon \mapsto \min\left(2\varepsilon^{\frac{1}{(2k-1)^d}}, 1\right)$, a continuous non-decreasing function such that $g^d_k(0) = 0$.*

*The function is such that, for any k-dependent noise ν (not necessarily a shift-invariant noise $\nu \in \mathcal{D}_k$) we have the domination $\nu \leq \mathcal{B}(g(\varepsilon))^{\otimes\mathbb{Z}^d}$, where $\varepsilon = \sup_{x\in\mathbb{Z}^d} \nu\left(\{\omega_x = 1\}\right)$ (*i.e.* $\varepsilon = \nu\left([1]\right)$ in the invariant case).*

Proof. Consider the following system of inequalities:

$$\begin{cases} \beta(1-\beta)^{C_k^d} \geq \varepsilon\,, \\ \beta\xi^{C_k^d} \geq \varepsilon\,, \end{cases}$$

with $C_k^d = |B_k| - 1 = (2k-1)^d - 1$. The reason why we consider this system will be made clear at the end of the proof.

Consider here $\beta = \xi = \varepsilon^{\frac{1}{C_k^d+1}} \in]0,1[$, so that the second inequality is trivially satisfied. The first inequality holds as long as $\varepsilon \leq \frac{1}{2^{C_k^d+1}}$. If we have $g_k^d(\varepsilon) = 1$ (which implies $\varepsilon \geq \frac{1}{2^{C_k^d+1}}$), we trivially have the domination $\nu \leq \delta_{1^\infty}$. Else, the chosen couple (β,ξ) satisfies both inequalities, and we will work under this assumption for the rest of the proof. Note how, in this case, $g_k^d(\varepsilon) = \beta+\xi \geq \beta+\xi-\beta\xi = 1-(1-\beta)(1-\xi)$, which is the actual slightly tighter domination this proof will provide.

The couple (β,ξ) chosen here may not be *optimal*, but it still follows from a pretty natural heuristic. On one hand, because β and ξ play the same role in g_k^d, we want them to have the same asymptotic order of magnitude, hence $\beta = \xi$. On the other hand, if $\beta\xi^{C_k^d} > \varepsilon$, we can always *decrease* ξ until equality is reached while simultaneously decreasing $\beta+\xi$. From these two conditions it directly follows that $\beta^{C_k^d+1} = \beta\xi^{C_k^d} = \varepsilon$.

Consider the independent variables $(B_x) \sim \nu$ and $(Y_x) \sim \mathcal{B}(\xi)^{\otimes\mathbb{Z}^d}$. Then naturally $\nu \leq \nu'$ where ν' is the law of $(Z_x := \max(B_x, Y_x))_{x\in\mathbb{Z}^d}$. We now want to dominate this auxiliary measure ν'.

Let $(x_j)_{j\in\mathbb{N}}$ be a given enumeration of $\mathbb{Z}^d$ for the rest of the proof.

Consider any sequence $(z_j) \in \{\mathtt{0},\mathtt{1}\}^{\mathbb{N}}$, and assume that

$$\mathbb{P}\left(B_{x_{j+1}} = \mathtt{1} \middle| Z_1 = z_1, \ldots, Z_j = z_j\right) \leq \beta$$

whenever this quantity is well-defined. Then, as $Y_{x_{j+1}}$ is independent of all the variables in this probability:

$$\begin{aligned} &\mathbb{P}\left(Z_{x_{j+1}} = \mathtt{1} \middle| Z_1 = z_1, \ldots, Z_j = z_j\right) \\ =&1 - \mathbb{P}\left(B_{x_{j+1}} = Y_{x_{j+1}} = \mathtt{0} \middle| Z_1 = z_1, \ldots, Z_j = z_j\right) \\ =&1 - \mathbb{P}\left(B_{x_{j+1}} = \mathtt{0} \middle| Z_1 = z_1, \ldots, Z_j = z_j\right) \times \mathbb{P}\left(Y_{x_{j+1}} = \mathtt{0}\right) \\ =&1 - \mathbb{P}\left(B_{x_{j+1}} = \mathtt{0} \middle| Z_1 = z_1, \ldots, Z_j = z_j\right) \times (1-\xi) \\ \leq&1 - (1-\beta)(1-\xi) \leq g_k^d(\varepsilon)\,. \end{aligned}$$

We then define the functions $g_j(z) := \mathbb{P}\left(Z_{x_{j+1}} := \mathtt{1} \middle| Z_1 = z_1, \ldots, Z_j = z_j\right)$, with $g_j(z) = 0$ when the conditional probability is ill-defined. Starting from the sequence of uniform random variables $(U_j) \sim \mathcal{U}([0,1])^{\otimes\mathbb{N}}$, let us define $W_{x_j} = \mathbb{1}_{\left\{U_{x_j} \leq g_k^d(\varepsilon)\right\}}$, and then the sequence $V_{x_{j+1}} = \mathbb{1}_{\left\{U_{x_{j+1}} \leq g_j\left(V_{x_1},\ldots,V_{x_j}\right)\right\}}$ inductively. By construction, $V_x \leq W_x$ on any cell, V has the same law ν' as Z, and W has the law $\mathcal{B}\left(g_k^d(\varepsilon)\right)^{\otimes\mathbb{Z}^d}$, so that we have the domination we wanted on ν'.

Finally, we need to prove that $\mathbb{P}\left(B_{x_{j+1}} = \mathtt{1} \middle| Z_1 = z_1, \ldots, Z_j = z_j\right) \leq \beta$ in order to conclude. This is the critical point where we will make full use of the k-dependence, and where our initial system of inequalities will appear.

We prove this result inductively. At first, we have $\mathbb{P}\left(B_{x_1} = \mathtt{1}\right) \leq \varepsilon$ by definition. In particular, because $\varepsilon \leq \beta(1-\beta)^{C_k^d} \leq \beta$, the initialisation holds.

Assume now that the lower bound holds at any rank $i \leq j$ and for any enumeration (x_n) of $\mathbb{Z}^d$. Consider $z_1, \ldots, z_j \in \{0,1\}$ and let us partition $[\![1,j]\!]$ into three parts as follows:

- $N^{\mathtt{0}} = \left\{i \leq j, \|x_i - x_{j+1}\|_\infty < k, z_i = \mathtt{0}\right\}$,
- $N^{\mathtt{1}} = \left\{i \leq j, \|x_i - x_{j+1}\|_\infty < k, z_i = \mathtt{1}\right\}$,
- $M = \left\{i \leq j, \|x_i - x_{j+1}\|_\infty \geq k\right\}$.

Denote the events $A_b = \left\{\forall i \in N^b, Z_{x_i} = b\right\}$ with $b \in \{\mathtt{0},\mathtt{1}\}$, as well as the event $A = \{\forall i \in M, Z_{x_i} = z_i\}$ and the subset $A_1' = \{\forall i \in N^{\mathtt{0}}, Y_{x_i} = \mathtt{1}\} \subseteq A_1$. Assume $\mathbb{P}\left(A \cap A_{\mathtt{0}} \cap A_1\right) > 0$ so that we consider a well-defined conditional probability. By k-dependence, $B_{x_{j+1}}$ is independent of the event A, so:

$$\begin{aligned}\mathbb{P}\left(B_{x_{j+1}}=1\middle|Z_1=z_1,\dots,Z_j=z_j\right)&=\mathbb{P}\left(B_{x_{j+1}}=1\middle|A\cap A_0\cap A_1\right)\\&=\frac{\mathbb{P}\left(\left\{B_{x_{j+1}}=1\right\}\cap A\cap A_0\cap A_1\right)}{\mathbb{P}\left(A\cap A_0\cap A_1\right)}\\&\leq\frac{\mathbb{P}\left(\left\{B_{x_{j+1}}=1\right\}\cap A\right)}{\mathbb{P}\left(A\cap A_0\cap A_1'\right)}\\&=\frac{\mathbb{P}\left(B_{x_{j+1}}=1\right)\mathbb{P}(A)}{\mathbb{P}\left(A\cap A_0\right)\mathbb{P}\left(A_1'\right)}\\&\leq\frac{\varepsilon}{\xi^{|N_1|}}\times\frac{1}{\mathbb{P}\left(A_0|A\right)}\,.\end{aligned}$$

Let us finally use the recurrence hypothesis to find a lower bound on $\mathbb{P}\left(A_0|A\right)$. Consider an enumeration $N_0=\{y_1,\dots,y_r\}$. As the variables $(Y_{y_i})_{1\leq i\leq r}$ are independent of the rest of the terms in the conditional probabilities, we have:

$$\begin{aligned}\mathbb{P}\left(A_0|A\right)&=\prod_{i=1}^{r}\mathbb{P}\left(B_{y_i}=0\middle|A\bigcap_{n=1}^{i-1}\{B_{y_n}=0\}\right)\\&=\prod_{i=1}^{r}\mathbb{P}\left(B_{y_i}=0\middle|A\bigcap_{n=1}^{i-1}\{B_{y_n}=0\}\bigcap_{n=1}^{i-1}\{Y_{y_n}=0\}\right)\\&=\prod_{i=1}^{r}\mathbb{P}\left(B_{y_i}=0\middle|A\bigcap_{n=1}^{i-1}\{Z_{y_n}=0\}\right)\\&\geq(1-\beta)^{|N_0|}\,,\end{aligned}$$

using the fact that $A\bigcap_{n=1}^{i-1}\{Z_{y_n}=0\}$ is a condition on *at most* j coordinates of Z, so the induction hypothesis holds here. Note that the same inequality holds if $N_0=\emptyset$. Finally, by injecting this back into the previous bound:

$$\mathbb{P}\left(B_{x_{j+1}}=1\middle|Z_1=z_1,\dots,Z_j=z_j\right)\leq\frac{\varepsilon}{\xi^{|N_1|}(1-\beta)^{|N_0|}}\,.$$

Now, using the fact that (β,ξ) is a solution of the system of inequalities introduced at the beginning of the proof, $\xi\geq\left(\frac{\varepsilon}{\beta}\right)^{\frac{1}{C_k^d}}$ and likewise $(1-\beta)\geq\left(\frac{\varepsilon}{\beta}\right)^{\frac{1}{C_k^d}}$. Thence:

$$\mathbb{P}\left(B_{x_{j+1}}=1\middle|Z_1=z_1,\dots,Z_j=z_j\right)\leq\frac{\varepsilon}{\left(\frac{\varepsilon}{\beta}\right)^{\frac{|N_0|+|N_1|}{C_k^d}}}=\beta^{\frac{|N_0|+|N_1|}{C_k^d}}\varepsilon^{1-\frac{|N_0|+|N_1|}{C_k^d}}\,.$$

Finally, as $N_0\sqcup N_1\subseteq B_\infty\left(x_{j+1},k-1\right)\setminus\{x_{j+1}\}$, we have $1-\frac{|N_0|+|N_1|}{C_k^d}\geq 0$, and using one last time the fact that $\varepsilon\leq\beta$, we finally obtain:

$$\mathbb{P}\left(B_{x_{j+1}}=1\middle|Z_1=z_1,\dots,Z_j=z_j\right)\leq\beta^{\frac{|N_0|+|N_1|}{C_k^d}}\beta^{1-\frac{|N_0|+|N_1|}{C_k^d}}=\beta\,,$$

which concludes the proof. □

Remark that, because of iterative nature of the proof, the coupling that gives a domination in the proof is *not* invariant. However, once such a coupling between invariant measures is obtained, we can once again average it on its translation orbit to obtain an invariant coupling giving us the same domination.

Now, by using this domination result along with the previous conjugacy-invariant stability theorem, we obtain the following corollary:

Corollary 4.19. *If the SFT $X_{\mathcal{F}}$ is stable (resp. polynomially stable) on the Bernoulli class $\mathcal{B}$, then it is also stable (resp. polynomially stable) on any class $\mathcal{D}_k$ of dependent noises.*

Under the further assumption that there exists a conjugacy with some other SFT $\theta: X_{\mathcal{F}}\to X_{\mathcal{G}}$, then $X_{\mathcal{G}}$ is also stable (resp. polynomially stable) on the class $\mathcal{B}$.

Proof. For the first part of the result, assume that $X_{\mathcal{F}}$ is f-stable on $\mathcal{B}$. As the class $\mathcal{D}_k$ is g_k-dominated by $\mathcal{B}$ (by the domination theorem), we apply Proposition 4.17 to conclude the SFT is $(f \circ g_k)$-stable on $\mathcal{D}_k$. In particular, for the polynomial case, if f is $O(\varepsilon^{\alpha})$, then $f \circ g_k$ is $O\left(\varepsilon^{\alpha/(2k-1)^d}\right)$, still of polynomial order.

For the second part of the result, we may use the conjugacy-invariant stability theorem, as $X_{\mathcal{F}}$ is now $\left(f \circ g_{2r_{\theta^{-1}}+1}\right)$-stable on $\gamma_{r_{\theta^{-1}}}(\mathcal{B}) \subseteq \mathcal{D}_{2r_{\theta^{-1}}+1}$. In particular, if $f \circ g_{2r_{\theta^{-1}}+1}$ is a $O\left(\varepsilon^{\alpha/(4r_{\theta^{-1}}+1)^d}\right)$, then so is $D_\theta f \circ g_{2r_{\theta^{-1}}+1}(C_{\theta^{-1}}\varepsilon) + E\varepsilon$. □

Notice how, because of the noise domination, we are unable to preserve *linear* stability. Still, we have proven that stability on the class $\mathcal{B}$ is a conjugacy-invariant property, in the sense that while the *definition* of stability itself relies on a choice of $\mathcal{F}$, the *property* of stability is not dependent on this choice.

Remark 4.20 (Other Classes of Noise)**.** As $g_k(\varepsilon) \approx \varepsilon^{1/(2k-1)^d} \underset{k\to\infty}{\longrightarrow} 1$ for any fixed value of ε, we conclude that even though an SFT stable on $\mathcal{B}$ is stable on all the classes $\mathcal{D}_k$, this stability does not extend to the limit class $\mathcal{D} = \bigcup_{k\in\mathbb{N}} \mathcal{D}_k$, which would be the most natural generalisation of $\mathcal{B}$ stable under all the morphisms γ^k.

If we look at noises with infinite-range dependencies, we notably have periodic noises (*i.e.* a uniform translation of a periodic configuration $b \in \Omega_{\{0,1\}}$). In most of the interesting cases, the rigid structure of such noises allows us to explicitly construct measures that do not converge to $\mathcal{M}_\sigma(X_{\mathcal{F}})$ for d_B, as in Sections 4.3.2 and 4.4.1.

Such periodic noises have not only infinite-range interactions but also high correlations at arbitrarily long distances. The remaining in-between case would be that of infinite-range dependencies but with correlations that decrease and go to 0 as the distance goes to ∞. This case notably encompasses the Gaussian Free Field, as well as some Gibbs measures (which will be more formally introduced in the next chapter). This may be the most physically realistic case, but is also much harder to study, so we will set it aside for the rest of this exploratory chapter. ■

The class of independent noises $\mathcal{B}$ is particularly convenient to study, and will be the focal point for all the rest of this chapter.

4.3 Classification of the One-Dimensional Stability

Now that I have introduced a general framework to discuss stability, let's look at the simplest situation, the one-dimension case $G = \mathbb{Z}$.

This framework, which is the one discussed extensively by Lind and Marcus [LM21], permits many approaches that simply don't translate to the higher-dimensional case. Notably, we can see the configurations of an SFT as bi-infinite paths in a word automaton (*i.e.* a directed labelled graph). This automatic structures allows a full classification of some behaviours: for instance, the one-dimensional domino problem (the general case was discussed in Section 3.5) is decidable. Following similar ideas, we will see how stability and reconstruction can be related to mixing SFTs.

The section will be concluded by a discussion on how to transpose general SFTs from d to $d+1$ dimensions while preserving their (un)stable behaviour; this will notably imply the existence of both stable and unstable systems in any dimension d for the Besicovitch topology, which confirms the non-triviality of the question of stability seen as a decision/classification problem (as opposed to weak-$*$ stability, satisfied for any choice of $\mathcal{F}$).

4.3.1 One-Dimensional SFTs and Word Automata

In the one-dimensional case, patterns and configurations are simply *words*. Because of their linear structure, words exhibit some automatic properties not encountered in higher dimensions.

Definition 4.21 (Diameter of a Set of Forbidden Patterns)**.** Let $I \Subset \mathbb{Z}$, and denote $\mathrm{diam}(I) = \max(I) - \min(I)$ its diameter. For a word $w \in \mathcal{A}^I$, $\mathrm{diam}(w) = \mathrm{diam}(I)$. Finally, for a set of forbidden patterns $\mathcal{F}$, we denote $\mathrm{diam}(\mathcal{F}) = \max_{w\in\mathcal{F}} \mathrm{diam}(w)$ its maximal diameter. ■

Consider an automaton $G^{\ell}_{\mathcal{A}}$, a directed graph with labelled edges, where states are words in $\mathcal{A}^{\ell}$, with transitions $au \xrightarrow{b} ub$ for any $u \in \mathcal{A}^{\ell-1}$ and $a, b \in \mathcal{A}$. There is a natural correspondence between bi-infinite words $\omega \in \Omega_{\mathcal{A}}$ and bi-infinite sequences of transitions in this word automaton.

Note that this definition looks at words left-to-right (*i.e.* the next transition in a trajectory is the next letter of the word), but we could likewise look at right-to-left $ub \xrightarrow{a} au$ transitions without changing any of the following (a)periodicity properties nor the (in)stability results they imply.

Definition 4.22 (Word Automaton)**.** Let $\mathcal{F}$ a finite set of forbidden words. We define the automaton $G_{\mathcal{F}}$ induced by restricting $G_{\mathcal{A}}^{\mathrm{diam}(\mathcal{F})}$ to the states $w \in \mathcal{A}^{\mathrm{diam}(\mathcal{F})}$ that are locally admissible.

Now, bi-infinite trajectories in $G_{\mathcal{F}}$ correspond to configurations $\omega \in \Omega_{\mathcal{A}}$ where every window of diameter $\mathrm{diam}(\mathcal{F})$ is locally admissible, *i.e.* to globally admissible configurations $\omega \in X_{\mathcal{F}}$. ■

Note that likewise, for a given word automaton G (with states in $\mathcal{A}^{\ell}$), we can deduce a corresponding set of forbidden patterns $\mathcal{F} \subseteq \mathcal{A}^{\ell}$ such that $G = G_{\mathcal{F}}$, so the two characterisations are equivalent.

As the number of states is finite, an infinite path exists (*i.e.* $X_{\mathcal{F}} \neq \emptyset$) *iff* $G_{\mathcal{F}}$ contains a cycle, so the domino problem is decidable in polynomial time here.

Definition 4.23 (Weakly Irreducible Automaton)**.** Two states u and v of $G_{\mathcal{F}}$ communicate if there is a path from u to v and v to u in the directed graph.

This gives us a partial equivalence relation, whose classes are the communication classes. At least one class exists *iff* $G_{\mathcal{F}}$ contains a cycle, *iff* $X_{\mathcal{F}} \neq \emptyset$.

We say that $G_{\mathcal{F}}$ is weakly irreducible if this class is unique. Note that this does not imply that *all* the states of $G_{\mathcal{F}}$ are in the class. For example, in the directed graph represented by $a \to b \circlearrowleft$, $\{b\}$ is the only communication class, because there is no path from b to a. ■

This definition differs from the usual (stronger) notion of irreducible directed graph [LM21, Definition 2.2.13], that requires *all* the vertices to be in the unique communication class. We use here a weaker notion because "purely" transient states (not in any class) will not prevent stability in the irreducible aperiodic case in Theorem 4.27 and only affect the value of the constant in the $O(\varepsilon)$ convergence speed.

Definition 4.24 (Periodic Automaton)**.** Consider $G_{\mathcal{F}}$ a weakly irreducible automaton. We say that it is p-periodic if p is the greatest common divisor of the lengths of all the cycles found inside $G_{\mathcal{F}}$. $G_{\mathcal{F}}$ is *aperiodic* if $p = 1$.

In the p-periodic case, there exists a partition $C = \bigsqcup_{j \in \mathbb{Z}/p\mathbb{Z}} C_j$ of the communication class such that for any transition $u \to v$ of the automaton we must have $u \in C_j$ and $v \in C_{j+1}$ for some $j \in \mathbb{Z}/p\mathbb{Z}$. ■

Remark 4.25 (Aperiodic Automata and Mixing SFTs)**.** The (weakly irreducible) aperiodic case for $G_{\mathcal{F}}$ corresponds to the situation where $X_{\mathcal{F}}$ is a *mixing* subshift, *i.e.* there is a length n_0 such that for any two globally admissible words u and v, and any length $n \geq n_0$, we have a word $x \in \mathcal{A}^n$ such that uxv is globally admissible too. The constant n_0 can be computed from $G_{\mathcal{F}}$ in polynomial time. ■

Before proving that weakly irreducible aperiodic systems are stable in the general case, let us illustrate how instability can occur in the periodic case through a simple example.

4.3.2 A Uniquely Ergodic Unstable Example

For a stable SFT, as ε goes to 0, a generic noisy configuration in $X_{\widetilde{\mathcal{F}}}$ has arbitrarily few differences with a generic clear configuration in $X_{\mathcal{F}}$. In the specific case of uniquely ergodic SFTs, we expect a really simple structure for configurations of $X_{\mathcal{F}}$, which makes them natural candidates to study at first.

In the one-dimensional case, uniquely ergodic SFTs are exactly those reduced to the finite orbit of a periodic configuration. Hence, consider the simplest non-trivial uniquely ergodic one-dimensional SFT, whose only two configurations are $\omega_0 = (\mathtt{01})^{\mathbb{Z}}$ and $\omega_1 = \sigma^1(\omega_0)$ (such that $\omega_i(k) = k + i[2]$). This system is induced by the forbidden patterns $\mathcal{F} = \{\mathtt{00}, \mathtt{11}\}$, uniquely ergodic, and irreducible 2-periodic.

We define the p-periodic noise ν_p, uniform among the p translations of $\left(\mathtt{0}^{p-1}\mathtt{1}\right)^{\mathbb{Z}}$. With this noise, $\nu_p([\mathtt{1}]) = \frac{1}{p}$ goes to 0 as $p \to \infty$. Consider then $\lambda_p \in \widetilde{\mathcal{M}_{\mathcal{F}}}\left(\frac{1}{p}\right)$ such that $\pi_*^2(\lambda_p) = \nu_p$, and on each clear window of size $p-1$ (and the obscure cell on its left), we use alternatively the restriction of ω_0 or ω_1. In other words, λ_p is supported on the orbit of a $2p$-periodic configuration (ω, b). Naturally, we have $d_H(\omega, \omega_0) = d_H(\omega, \omega_1) = \frac{1}{2}$, so $d_B\left(\pi_*^1(\lambda_p), \mathcal{M}_\sigma(X_{\mathcal{F}})\right) = \frac{1}{2}$ too, *i.e.* this SFT is unstable (if we allow any noise).

We can generalise this result to all periodic SFTs without much effort, provided we use periodic noises of the form $\left(\mathtt{0}^p\mathtt{1}^d\right)^{\mathbb{Z}}$ with $\mathrm{diam}(\mathcal{F}) \leq d$ (to guarantee forbidden patterns cannot appear on two distinct clear blocks) and $p \to \infty$. Later, we will prove an analogous general periodic instability result using Bernoulli noises instead.

In the higher-dimensional case, a similar unstability result will be proven for periodic SFTs in Section 4.4.1.

4.3.3 Stability for Weakly Irreducible Aperiodic Automata

In a one-dimensional setup, as long as $X_{\mathcal{F}} \neq \emptyset$, there is always a cycle in the word automaton, thus a periodic configuration. The aperiodicity of the automaton only implies the *existence* of aperiodic configurations in $X_{\mathcal{F}}$ (thus is different from what we defined as a strongly aperiodic subshift, where *all* the configurations must be aperiodic), except if the aperiodic class is a singleton with a self-loop (*i.e.* $X = \left\{a^{\mathbb{Z}}\right\}$ for some letter $a \in \mathcal{A}$).

In this situation, the mixing behaviour will prove to be sufficient to repair faults locally around obscured cells, and ultimately obtain stability.

Denote here $L\left(X_{\mathcal{F}}\right)$ the language of $X_{\mathcal{F}}$, the set of all globally admissible words. Our goal in this subsection is to find, in an obscured configuration $(\omega, b) \in X_{\widetilde{\mathcal{F}}}$, a sequence of globally admissible words seperated by gaps of length at least n_0, which will allow us to repair the gaps to obtain a nearby clear globally admissible $\omega' \in X_{\mathcal{F}}$.

If we only exclude from ω the obscured positions in b, then we will obtain a sequence of *locally* admissible clear words instead, that may not be *globally* admissible, and the gaps between these words may be too small, of length less than n_0. By thickening the noise by a radius $\left\lceil \frac{n_0}{2} \right\rceil$, as in the previous section, we can at least make sure that any gap between two clear words is of length at least n_0, while still having a linear $O(\varepsilon)$ amount of obscured cells.

Now, the following proposition roughly tells us that in onedimensional, the reconstruction function φ from Remark 4.1 is bounded.

Proposition 4.26. *For any set $\mathcal{F}$ of one-dimensional forbidden patterns, there exists a constant $C(\mathcal{F})$ such that, for any locally admissible word $u \in \mathcal{A}^*$, by removing (at most) C letters on both ends, we obtain instead a globally admissible word $v \in L\left(X_{\mathcal{F}}\right)$.*

Proof. Note that a path of length n in $G_{\mathcal{F}}$ visits $n+1$ vertices (each a word of length $\operatorname{diam}(\mathcal{F})$), and represents a word of length $\operatorname{diam}(\mathcal{F}) + n$. Thus, we may assume that $C \geq \frac{\operatorname{diam}(\mathcal{F})}{2}$, so that we only need to consider words long enough to represent a finite path in the automaton $G_{\mathcal{F}}$.

As long as we visit vertices within the communication class of $G_{\mathcal{F}}$, we can infinitely extend the path on both directions, thence the word we encode is globally admissible.

Issues arise when we visit other states, which explicitly correspond to vertices that never occur in a bi-infinite path, *i.e.* non-globally admissible words. As there is only one communication class, no path can *cycle* through such a state. Hence, if there are k states of $G_{\mathcal{F}}$ outside of the communication class, by removing k states on each end of the path, we make sure that the path only visits the communication class, thus corresponds to a globally admissible word. Hence, $C = \max\left(k, \left\lceil \frac{\operatorname{diam}(\mathcal{F})}{2} \right\rceil\right)$ is sufficiently large.

If we want an optimal constant, we can replace k by the maximum of the length of the longest path among vertices *outside of* yet *connected to* the communication class, and half of the longest path not connected to the class. ☐

Just like n_0, this constant C can be computed from $G_{\mathcal{F}}$ in polynomial time. Now, if we remove a C-neighbourhood around each obscured cell, then we obtain a sequence of globally admissible words. Finally, by removing a D-neighbourhood with $D = \max\left(C, \left\lceil \frac{n_0}{2} \right\rceil\right)$, we make sure that we obtain a sequence of globally admissible words with fillable gaps in-between them. This is the key idea of the following theorem, whose proof mostly aims at properly explaining why the transformation we perform is a shift-invariant morphism that returns a clear globally admissible configuration.

Theorem 4.27. *Let $X_{\mathcal{F}}$ be a one-dimensional SFT, with a weakly irreducible aperiodic automaton $G_{\mathcal{F}}$. Then $X_{\mathcal{F}}$ is linearly stable, with an explicit computable constant in the $O(\varepsilon)$.*

Proof. In order to obtain linear stability, we will consider a measure $\lambda \in \widetilde{\mathcal{M}_{\mathcal{F}}}(\varepsilon)$, and build a measurable mapping $\psi : X_{\widetilde{\mathcal{F}}} \to X_{\mathcal{F}}$, so that $d_H\left(\pi^1(\omega), \psi(\omega)\right)$ is small. Let us notice that ψ does not need to be defined on all of $X_{\widetilde{\mathcal{F}}}$, but only on a high-probability space $S \subseteq X_{\widetilde{\mathcal{F}}}$. In such a case, we may add a third *independent* coordinate to λ that follows some given law in $\mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$, and project onto this third coordinate with ψ outside of the event S (to create a "if S then $\psi(\cdots)$ else" scenario), to simply use the trivial upper bound $d_H \leq 1$ outside of S. This way:

$$d_B\left(\pi^1_*(\lambda), \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right) \leq \int_S d_H\left(\pi^1(\omega), \psi(\omega)\right) \mathrm{d}\lambda(\omega) + \lambda\left(S^c\right) .$$

Let γ^D the cellular automaton on $\Omega_{\{0,1\}}$ from Definition 4.11, that thickens the noise on the D-neighbourhood of each obscure cell. This process is clearly measurable, and naturally extends as a morphism on $X_{\widetilde{\mathcal{F}}}$. We now need to map this image of $X_{\widetilde{\mathcal{F}}}$ into $X_{\mathcal{F}}$ in a measurable way.

The issue now is that, while we can manually fill each obscure gap (of length at least n_0 between two globally admissible words), issues may arise with the order of the operations. Indeed, as the SFT is mixing, we can decide on a word $w(u_1, u_2, n)$ for any globally admissible words $u_1, u_2 \in L(X_{\mathcal{F}})$ and any gap of length $n \geq n_0$, so that $|w| = n$ and $u_1 w u_2 \in L(X_{\mathcal{F}})$ too.

Naturally, if we have *three* globally admissible words u_1, u_2 and u_3 as well as two gaps of lengths i and j, then we can fill the leftmost gap first, with $v = u_1 w(u_1, u_2, i) u_2$ and then the second one with $vw(v, u_3, j) u_3$. However, doing so implies an iterative process, with a starting point, which makes it hard to ultimately obtain an map ψ that commutes with σ.

There are several ways to proceed and avoid this issue, and we chose here to be able to fill *all* those gaps simultaneously, so that the invariance of the morphism directly follows. To ensure we can fill the gaps simultaneously, we want to use only the $\operatorname{diam}(\mathcal{F})$ rightmost letters of u_1 and as many leftmost letters of u_2 (which characterise the starting point and the destination of the corresponding trajectory in $G_{\mathcal{F}}$), to chose the value of $w(u_1, u_2, n)$ accordingly. Of course this requires $\min(|u_1|, |u_2|) \geq \operatorname{diam}(\mathcal{F})$ to begin with.

Then, if we thicken the noise by obscuring the $\left\lceil \frac{\operatorname{diam}(\mathcal{F})}{2} \right\rceil$-neighbourhood of obscure cells in $\gamma^D\left(\Omega_{\{0,1\}}\right)$, and "reverse" the process by clearing the $\left\lceil \frac{\operatorname{diam}(\mathcal{F})}{2} \right\rceil$-neighbourhood of the remaining clear cells (which uses two cellular automata so is still a morphism on $X_{\widetilde{\mathcal{F}}}$), we erase any block of clear cells of length less than $\operatorname{diam}(\mathcal{F})$.

Let us name θ the cellular automaton on $\Omega_{\{0,1\}}$ obtained by applying consecutively γ^D and the two other ones, and we identify it with the morphism on $X_{\widetilde{\mathcal{F}}}$ that leaves the first coordinate unchanged. So far, any configuration $(\omega, \theta(b)) \in \theta\left(X_{\widetilde{\mathcal{F}}}\right) \subseteq X_{\widetilde{\mathcal{F}}}$ can be described as a sequence (not necessarily bi-infinite) of globally admissible words (each of length at least $\operatorname{diam}(\mathcal{F})$) separated by gaps of obscure cells (each of length at least n_0), and the density of obscure cells in $\theta(b)$ is of the same order as that of b (*i.e.* $d_H(\theta(b), 0^\infty) \leq (2D + \operatorname{diam}(\mathcal{F}) + 2)\, d_H(b, 0^\infty)$).

Now, we want to "simultaneously" fill all the obscure gaps to map $(\omega, \theta(b))$ onto $\psi(\omega, b) \in X_{\mathcal{F}}$ that coincides with ω on the clear cells of $\theta(b)$. More formally, for a given length of the gaps n, we can use a cellular automaton to rewrite the content of ω, on each obscure block of length exactly n, with $w(u_1, u_2, n)$ where u_1 is the globally admissible word of length $\operatorname{diam}(F)$ right on the left of obscure block, and likewise u_2 is the globally admissible word on the right. This automaton only needs to look at the neighbourhood of size $(n + \operatorname{diam}(\mathcal{F}))$ of each cell to compute its new value. At the limit, by applying the automaton for each length $n \geq n_0$, we obtain a measurable map ψ. The limit ψ is well-defined, as each cell will be modified at most once by the sequence of morphisms, and it commutes with σ, but is not a morphism itself as the new value in a cell may depend on what happens arbitrarily far from the cell.

There is one last issue to deal with, *i.e.* the fact that $\psi(\omega, b)$ consists of one big globally admissible clear word, but that it may have an infinite obscured window on the left or the right. Let us name S the set of configurations where this phenomenon does not happen. So far, we obtained a set S and defined a morphism $\psi : S \to X_{\mathcal{F}}$, as stated in the first paragraph of the proof, so let us now study the two terms of the bound.

First, inside of S, we only modify cells of ω such that $\theta(b)$ is obscure, so:

$$d_H(\omega, \psi(\omega)) \leq d_H(\theta(b), 0^\infty) \leq (2D + \operatorname{diam}(\mathcal{F}) + 2)\, d_H(b, 0^\infty)\ ,$$

and so we can bound the integral of this term on S by $(2D + \operatorname{diam}(\mathcal{F}) + 2)\,\varepsilon$ when $\lambda \in \widetilde{\mathcal{M}_{\mathcal{F}}}(\varepsilon)$.

Second, notice that in order to have $\theta(b) \in S^c$ (*i.e.* the configuration has a density 1 of obscured cells), we need a density at least $\frac{1}{2D+\operatorname{diam}(\mathcal{F})+2}$ of obscured cells in b to begin with. Naturally, we conclude that the frequency of obscured cells in λ satisfies:

$$\varepsilon \geq \lambda(S^c) \times \frac{1}{2D + \operatorname{diam}(\mathcal{F}) + 2} + \lambda(S) \times 0\,,$$

and so $\lambda(S^c) \leq (2D + \operatorname{diam}(\mathcal{F}) + 2)\,\varepsilon$.

At last, we obtain the explicit bound $d_B(\pi_1^*(\lambda), \mathcal{M}_\sigma(X_{\mathcal{F}})) \leq (2D + \operatorname{diam}(\mathcal{F}) + 2)\,\varepsilon$, with a constant factor computable in polynomial time. □

Remark that, when using an independent ε-Bernoulli noise, $\lambda(S) = 1$ (by Borel-Cantelli, with probability one, we have infinitely many arbitrarily large clear blocks in both directions). The case S^c is here to account for periodic noises for instance, which still will not be able to create infinite obscured windows inside $\theta(b)$ below a certain threshold.

4.3.4 Instability for Weakly Irreducible Periodic Automata

The key ingredient of the proof of the previous theorem is the fact that the SFT is mixing, that we can locally repair the mistakes that occur in obscured blocks. We may lose this repairability in two ways.

One is the loss of aperiodicity: if the system is periodic, as was the case of the simple 2-periodic example in a previous subsection, then we simply cannot repair holes of any lengths, but only those compatible with the periodicity of $G_{\mathcal{F}}$. The other is the loss of weak irreducibility. In this case, the several communication classes form a directed acyclic graph, and there are some pairs of globally admissible words for which we cannot repair a hole *at all*, regardless of the size.

Our objective is now to prove unstability with Bernoulli noises when $G_{\mathcal{F}}$ is weakly irreducible but p-periodic (with $p \geq 2$). The non-irreducible case will not be proven here, as it uses the exact same arguments.

Theorem 4.28 (Periodic Instability)**.** *Consider a one-dimensional SFT $X_{\mathcal{F}}$ such that $G_{\mathcal{F}}$ is weakly irreducible and p-periodic (with $p \geq 2$).*

Then for any $\varepsilon > 0$ there exists $\mu_\varepsilon \in \mathcal{M}^{\mathcal{B}}_{\mathcal{F}}(\varepsilon)$ such that $d_B\left(\mu_\varepsilon, \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right) \geq \frac{p-1}{p \times \mathrm{diam}(\mathcal{F})} - \varepsilon$.

Proof. Let us begin by considering the partition of $X_{\mathcal{F}}$ into p sets $(X_j)_{j \in \mathbb{Z}/p\mathbb{Z}}$ induced by the states of $G_{\mathcal{F}}$ (*i.e.* depending on the value of $\omega_{[\![0, \mathrm{diam}(\mathcal{F})-1]\!]}$, or any translation of this interval), so that if $\omega \in X_i$, then $\sigma^k(\omega) \in X_{i+k}$.

Consider also once and for all a periodic configuration $\omega_0 \in X_{\mathcal{F}}$, that corresponds to an infinite cycle of $G_{\mathcal{F}}$. Note that this cycle may not be of length p but a multiple of it – *e.g.* if $G_{\mathcal{F}}$ is made of a 6-cycle and a 10-cycle joined in a vertex, it is 2-periodic but has no 2-cycle. What matters is that ω_0 has a finite orbit under translations.

To construct $\mu_\varepsilon \in \mathcal{M}^{\mathcal{B}}_{\mathcal{F}}(\varepsilon)$, consider the noisy measure λ_ε obtained by:

1. taking the independent Bernoulli noise $\mathcal{B}(\varepsilon)^{\otimes \mathbb{Z}}$ first,
2. identifying intervals of consecutive obscured cells, of length at least $\mathrm{diam}\left(\mathcal{F}\right)$, and writing down letters of $\mathcal{A}$ uniformly at random under each such block,
3. in-between two such intervals, in a window that must have a clear cell on each end and may contain some short obscured blocks in the middle, we choose uniformly at random a translation of ω_0 to write it down on the cells, whether clear or obscured.

It is apparent that this measure has an ε-Bernoulli noise, and that it is shift-invariant by construction. The measure λ_ε is also strongly mixing, thus ergodic. Indeed, consider two finite windows $I, J \subseteq \mathbb{Z}$ such that $\min(J) - \max(I) = n > \mathrm{diam}\left(\mathcal{F}\right)$. Conditionally to the fact that the window $[\![\max(I)+1, \min(J)-1]\!]$ contains an obscured window of size $\mathrm{diam}\left(\mathcal{F}\right)$, the windows I and J behave independently from each other. As the probability of having such a window with an obscured block goes to 1 as $n \to \infty$, we deduce the mixing property on cylinders, so that λ_ε itself is strongly mixing.

Hence, if we cut $\mathbb{Z}$ into intervals of length $d := \mathrm{diam}\left(\mathcal{F}\right)$, we obtain a measure on $\left(\mathcal{A}^d \times \{\mathtt{0},\mathtt{1}\}^d\right)^{\mathbb{Z}}$. This induced measure is *also* shift-invariant and strongly mixing (thus ergodic). Hence, by Birkhoff's pointwise ergodic theorem, the frequency of a d-interval in a configuration is λ_ε-a.s. equal to its probability.

A clear d-interval has probability $(1-\varepsilon)^d$ of happening, which we bound below by $1 - d\varepsilon$. Under such an event, by construction, we can identify the state of $G_{\mathcal{F}}$ it represents, thus to which class X_i it comes from. Note that on such a clear window, if ω and ω' belong to different classes X_i and X_j, then in particular they correspond to different states of $G_{\mathcal{F}}$ thus must differ in *at least* one cell in each such d-interval.

Thus, for *any* globally admissible configuration $\omega \in X_{\mathcal{F}}$ and λ_ε-a.e. $(\omega', b) \in X_{\widetilde{\mathcal{F}}}$, we have:

$$d_H\left(\omega, \omega'\right) \geq (1 - d\varepsilon) \times \frac{p-1}{p} \times \frac{1}{d}.$$

The first factor comes from the frequency of clear windows, the second one from the probability of ω' not being in the same class as ω conditionally to some clear window, and the third one from the minimal number of differences in such a window of size $d(\mathcal{F})$ under the previous event.

It immediately follows that $d_B\left(\mu_\varepsilon, \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right) \geq \frac{p-1}{pd} - \frac{p-1}{p}\varepsilon$, which concludes the proof. □

Note that this proof can be adapted from the periodic case to the non-irreducible case where there are several communication classes, by using finite trajectories evolving inside distinct communication classes instead of words "aligned" along different periods of the system, even if all the classes are aperiodic. We thus obtain the following theorem, the proof of which we will omit for the sake of brevity, as it offers no further insight on the topic.

Theorem 4.29 (Non-Irreducible Instability)**.** *Consider an SFT $X_{\mathcal{F}}$ such that the automaton $G_{\mathcal{F}}$ is not irreducible, with p communication classes ($p \geq 2$). Then for any $\varepsilon > 0$ there exists a measure $\mu_\varepsilon \in \mathcal{M}_{\mathcal{F}}^{\mathcal{B}}(\varepsilon)$ such that $d_B\left(\mu_\varepsilon, \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right) \geq \frac{p-1}{p \times \operatorname{diam}(\mathcal{F})} - \varepsilon$.*

4.3.5 Extension to Higher Dimensions

Using the results from the previous subsection, we have the existence of both stable and unstable SFTs in the one-dimensional case. As announced, let's conclude the section with a digression on how such systems, both stable and unstable, can be extended into higher dimensions.

In doing so, I will dive deeper into the intricacies of couplings from a measure theory viewpoint, which offers a different insight on the objects I am working on, but can also be skipped without affecting the reading of the following sections.

First of all, let's quickly characterise the disintegration of shift-invariant couplings, which will be useful for the main result of this subsection.

Lemma 4.30 (Invariant Disintegration)**.** *Consider λ a probability measure on $\Omega_{\mathcal{A}} \times \Omega_{\mathcal{A}'}$, with $\mu = \pi_*^1(\lambda)$, and its disintegration $\mathrm{d}\lambda(x,y) = \mathrm{d}\nu_x(y)\mathrm{d}\mu(x)$ as in Lemma 4.16.*

Assume now that μ is shift-invariant. Then λ is also shift-invariant iff *the equality $\nu_x(B) = \nu_{\sigma^k(x)}\left(\sigma^k(B)\right)$ holds for any $k \in \mathbb{Z}^d$, any measurable cylinder $B \subseteq \Omega_{\mathcal{A}'}$, and μ-a.e. $x \in \Omega_{\mathcal{A}}$.*

Proof. Consider two cylinders A and B, as well as $k \in \mathbb{Z}^d$. As stated, we have $\lambda(A \times B) = \int_A \nu_x(B)\mathrm{d}\mu(x)$. Likewise, as μ itself is shift-invariant:

$$\lambda\left(\sigma^k(A \times B)\right) = \lambda\left(\sigma^k(A) \times \sigma^k(B)\right) = \int\limits_{\sigma^k(A)} \nu_x\left(\sigma^k(B)\right)\mathrm{d}\mu(x) = \int\limits_A \nu_{\sigma^k(y)}\left(\sigma^k(B)\right)\mathrm{d}\mu(y).$$

Now, the measure λ is invariant *iff* for any cylinder B and $k \in \mathbb{Z}^d$, we have $\lambda(A \times B) = \lambda\left(\sigma^k(A \times B)\right)$ for any cylinder A. Using the integral expressions, $\int_A \nu_x(B)\mathrm{d}\mu(x) = \int_A \nu_{\sigma^k(x)}\left(\sigma^k(B)\right)\mathrm{d}\mu(x)$. It is equivalent for this equality to hold for any A and for the functions to be μ-a.s. equal, which concludes the proof. $\square$

Note how the measures ν_x are not necessarily invariant themselves. In particular, using the Dirac measures $\nu_x = \delta_x$ – which are obviously not invariant – gives us a *diagonal* coupling between μ and itself, such that $\pi_*^2(\lambda) = \mu$ too, which is shift-invariant.

Now, given a d-dimensional SFT $X_{\mathcal{F}}$, it is possible to extend $\mathcal{F}$ into $\mathcal{F}'$ in $d+1$ dimensions, by replacing every forbidden pattern $p \in \mathcal{A}^{I(p)}$ on the window $I(p) \Subset \mathbb{Z}^d$ by $p' \in \mathcal{A}^{I(p)\times\{0\}}$ with $I(p) \times \{0\} \Subset \mathbb{Z}^{d+1}$. This way, $X_{\mathcal{F}'} = \left\{(\omega_i)_{i\in\mathbb{Z}}, \forall i \in \mathbb{Z}, \omega_i \in X_{\mathcal{F}}\right\}$. In other words, each slice (with a fixed last coordinate) represents a copy of the original SFT, with no constraints on how to align the slices. In particular, if $\mu \in \mathcal{M}_{\mathcal{F}}(\varepsilon)$, then by coupling all these layers independently, we obtain $\mu^{\otimes\mathbb{Z}} \in \mathcal{M}_{\mathcal{F}'}(\varepsilon)$.

Let us now prove that (in)stability of an SFT is in some sense preserved through this transformation. Thus, as we exhibited (un)stable 1D examples earlier in this section, this will imply the existence of (un)stable systems in *any* dimension.

Consider the projection $\zeta : b \in \{0,1\}^{\mathbb{Z}^{d+1}} \mapsto b_{\mathbb{Z}^d\times\{0\}} \in \{0,1\}^{\mathbb{Z}^d}$, that commutes with translations in $\mathbb{Z}^d$. More generally, we will use ζ as a multipurpose projector for any alphabet $\mathcal{A}$ instead of $\{0,1\}$. For a given class of $(d+1)$-dimensional noises $\mathcal{N}'$, we obtain the d-dimensional class $\mathcal{N} = \zeta^*\left(\mathcal{N}'\right)$. In particular, if $\mathcal{N}'$ is the class of $(d+1)$-dimensional Bernoulli noises, then $\mathcal{N}$ is the class of d-dimensional Bernoulli noises.

To make things easier to read, we will distinguish the Besicovitch distances d_B in d dimensions and d'_B in $d+1$ dimensions (resp. d_H and d'_H).

Proposition 4.31. *Using the projection ζ introduced in the previous paragraph, by extending the d-dimensional forbidden patterns $\mathcal{F}$ as $\mathcal{F}'$ and assuming that $\mathcal{N} = \zeta_*\left(\mathcal{N}'\right)$:*

1. *For any $\mu' \in \mathcal{M}_{\mathcal{F}'}^{\mathcal{N}'}(\varepsilon)$, we have $\mu = \zeta_*\left(\mu'\right) \in \mathcal{M}_{\mathcal{F}}^{\mathcal{N}}(\varepsilon)$.*

2. *For any $\mu \in \mathcal{M}_{\mathcal{F}}^{\mathcal{N}}(\varepsilon)$, there exists $\mu' \in \mathcal{M}_{\mathcal{F}'}^{\mathcal{N}'}(\varepsilon)$ such that $\mu = \zeta_*\left(\mu'\right)$.*

3. *In both cases,* $d_B(\mu, \mathcal{M}_{\mathcal{F}}) = d'_B(\mu', \mathcal{M}_{\mathcal{F}'})$.

Proof. In all cases, going from dimension $d+1$ to dimension d is a mere matter of projection through ζ, whereas going from dimension d to $d+1$ is a bit trickier and will require us to make use of Lemma 4.30.

The $\mathtt{n}$ in the couplings $\lambda_{\mathtt{n}}$ used afterwards, to prove the first two items, stands for the noise on its second coordinate. This is to better distinguish it from the coupling λ in the proof of the third item, which uses the same alphabet $\mathcal{A}$ for the two coordinates.

First, assume there is $\lambda'_{\mathtt{n}} \in \widetilde{\mathcal{M}^{\mathcal{N}'}_{\mathcal{F}'}}(\varepsilon)$ such that $\mu' = \pi^1_*(\lambda'_{\mathtt{n}})$. Thus, we have $\lambda_{\mathtt{n}} := \zeta_*(\lambda'_{\mathtt{n}}) \in \widetilde{\mathcal{M}^{\mathcal{N}}_{\mathcal{F}}}(\varepsilon)$, so that:

$$\mu = \zeta_*(\mu') = \zeta_*\left(\pi^1_*(\lambda'_{\mathtt{n}})\right) = \pi^1_*(\lambda_{\mathtt{n}}) \in \mathcal{M}^{\mathcal{N}}_{\mathcal{F}}(\varepsilon).$$

This proves the first item.

Conversely, consider $\lambda_{\mathtt{n}} \in \widetilde{\mathcal{M}^{\mathcal{N}}_{\mathcal{F}}}(\varepsilon)$ such that $\mu = \pi^1_*(\lambda_{\mathtt{n}})$, and let us build the desired measure μ'. Using Lemma 4.30, we have $\mathrm{d}\lambda_{\mathtt{n}}(\omega, b) = \mathrm{d}\mu_b(\omega)\mathrm{d}\nu(b)$ with $\nu = \pi^2_*(\lambda_{\mathtt{n}}) \in \mathcal{N}$ an ε-noise, and $\mathrm{d}\mu_{\sigma^k(b)}\left(\sigma^k(\omega)\right) = \mathrm{d}\mu_b(\omega)$ for ν-a.e. $b \in \Omega_{\{0,1\}}$. As $\nu \in \mathcal{N} = \zeta^*(\mathcal{N}')$, there is $\nu' \in \mathcal{N}'$ such that $\nu = \zeta_*(\nu')$. In particular, ν' is invariant and must be an ε-noise too. Now, for any families of d-dimensional layers $\omega' = (\omega_i)_{i\in\mathbb{Z}} \in \mathcal{A}^{\mathbb{Z}^{d+1}}$ and $b' = (b_i)$, we define the measures $\mathrm{d}\mu'_{b'}(\omega') = \prod_{i\in\mathbb{Z}} \mathrm{d}\mu_{b_i}(\omega_i)$ and then $\mathrm{d}\lambda'_{\mathtt{n}}(\omega', b') = \mathrm{d}\mu_{b'}(\omega')\,\mathrm{d}\nu'(b')$. Naturally, the measures $\mu'_{b'}$ are $\sigma^{e_{d+1}}$-invariant – invariant by translations on the last coordinate – by construction, and satisfy the criterion of Lemma 4.30 because the measures μ_b did. Thence, $\lambda'_{\mathtt{n}}$ is shift-inviariant, so that $\lambda'_{\mathtt{n}} \in \widetilde{\mathcal{M}^{\mathcal{N}'}_{\mathcal{F}'}}(\varepsilon)$. At last, $\mu' = \pi^1_*(\lambda'_{\mathtt{n}})$ is such that $\zeta_*(\mu') = \mu$, which proves the second item.

Lastly, consider any two invariant measures μ' and $\mu = \zeta_*(\mu')$, and let's prove the third item. We begin with the easier inequality, by considering λ' a coupling between μ' and $\nu' \in \mathcal{M}_\sigma(X_{\mathcal{F}'})$ such that:

$$d'_B(\mu', \mathcal{M}_\sigma(X_{\mathcal{F}'})) = d'_B(\mu', \nu') = \int d'_H(x', y')\,\mathrm{d}\lambda'(x', y').$$

Using Birkhoff's pointwise ergodic theorem, this is equal to $d'_B(\mu', \nu') = \int \mathbb{1}_{x'_0 \neq y'_0}(x', y')\,\mathrm{d}\lambda'(x', y')$. Likewise, $\lambda = \zeta_*(\lambda')$ is a coupling between μ and $\nu = \zeta_*(\nu') \in \mathcal{M}_\sigma(X_{\mathcal{F}})$, not necessarily such that d_B is reached, but:

$$d_B(\mu, \mathcal{M}_\sigma(X_{\mathcal{F}})) \leq d_B(\mu, \nu) \leq \int \mathbb{1}_{x_0\neq y_0}\mathrm{d}\lambda(x, y) = \int \mathbb{1}_{x'_0\neq y'_0}\mathrm{d}\lambda'(x', y') = d'_B(\mu', \mathcal{M}_\sigma(X_{\mathcal{F}'})).$$

For the reverse inequality, consider λ a coupling between the measures μ and $\nu \in \mathcal{M}_\sigma(X_{\mathcal{F}})$ such that $d_B(\mu, \mathcal{M}_\sigma(X_{\mathcal{F}})) = \int \mathbb{1}_{x_0\neq y_0}\mathrm{d}\lambda(x, y)$. As in Lemma 4.30, we can factorise $\mathrm{d}\lambda(x, y) = \mathrm{d}\nu_x(y)\mathrm{d}\mu(x)$. With configurations $x', y' \in \mathcal{A}^{\mathbb{Z}^{d+1}} = \left(\mathcal{A}^{\mathbb{Z}^d}\right)^{\mathbb{Z}}$, we define the family of measures $\mathrm{d}\nu'_{x'}(y') = \prod_{i\in\mathbb{Z}} \mathrm{d}\nu_{x_i}(y_i)$, and then $\mathrm{d}\lambda'(x', y') = \mathrm{d}\nu'_{x'}(y')\,\mathrm{d}\mu'(x')$. The measures $\nu'_{x'}$ satisfy the criterion of Lemma 4.30, so λ' is shift-invariant. This implies that it is a coupling between the measures μ' and $\nu' = \pi^2_*(\lambda') \in \mathcal{M}_\sigma(X_{\mathcal{F}'})$, once again not necessarily optimal, such that:

$$d'_B(\mu', \mathcal{M}_\sigma(X_{\mathcal{F}'})) \leq d'_B(\mu', \nu') \leq \int \mathbb{1}_{x'_0\neq y'_0}\mathrm{d}\lambda'(x', y') = \int \mathbb{1}_{x_0\neq y_0}\mathrm{d}\lambda(x, y) = d_B(\mu, \mathcal{M}_\sigma(X_{\mathcal{F}})).$$

This concludes the proof of the last item. □

Whenever moving from d to $d+1$ dimensions, the role of the disintegration theorem is here to transport the correlations between neighbouring d-dimensional layers from one measure to the next, notably to accomodate different kinds of noises $\mathcal{N}'$ (with identical layers, independent layers, or other behaviours that still project to the same d-dimensional noise class $\mathcal{N}$).

Corollary 4.32. *The d-dimensional SFT* $X_{\mathcal{F}}$ *is* f*-stable (resp. unstable) on the class* $\mathcal{N}$ *if and only if the* $(d+1)$*-dimensional SFT* $X_{\mathcal{F}'}$ *is* f*-stable (resp. unstable) on the class* $\mathcal{N}'$*.*

Proof. Going from (un)stability on $X_{\mathcal{F}'}$ to $X_{\mathcal{F}}$ is once again a simple matter of projecting measures with ζ so we won't insist further on these implications.

If $X_{\mathcal{F}}$ is f-stable on $\mathcal{N}$, and $\mu' \in \mathcal{M}^{\mathcal{N}'}_{\mathcal{F}'}(\varepsilon)$, then $\mu = \zeta^*(\mu') \in \mathcal{M}^{\mathcal{N}}_{\mathcal{F}}(\varepsilon)$ using Item 1 of the previous proposition, so that $d'_B(\mu', \mathcal{M}_\sigma(X_{\mathcal{F}'})) = d_B(\mu, \mathcal{M}_\sigma(X_{\mathcal{F}})) \leq f(\varepsilon)$ using Item 3. Thus, $X_{\mathcal{F}'}$ is f-stable.

Likewise, if $X_{\mathcal{F}}$ is unstable, we have a sequence of measures $\mu_n \in \mathcal{M}_{\mathcal{F}}(\varepsilon_n)$ with $\varepsilon_n \underset{n\to\infty}{\longrightarrow} 0$ such that $\inf_{n\in\mathbb{N}} d_B(\mu_n, \mathcal{M}_\sigma(X_{\mathcal{F}})) > 0$. Then, with the measures $\mu'_n \in \mathcal{M}^{\mathcal{N}'}_{\mathcal{F}'}(\varepsilon_n)$ given by Item 2, we conclude that $\inf_{n\in\mathbb{N}} d'_B(\mu'_n, \mathcal{M}_\sigma(X_{\mathcal{F}'})) = \inf_{n\in\mathbb{N}} d_B(\mu_n, \mathcal{M}_\sigma(X_{\mathcal{F}}))$ with Item 3, thence $X_{\mathcal{F}'}$ also is unstable. □

Using this corollary, we can in particular extend the (un)stable one-dimensional SFTs we exhibited earlier in order to obtain (un)stable SFTs in *any* dimension. Still, these examples are not entirely satisfactory and we will now strive for other higher-dimensional examples in the following sections.

Remark 4.33 (Alphabet Extension)**.** Another way to extend SFTs, not in dimension but in alphabet size, is the direct product. If we consider two d-dimensional SFTs $X_{\mathcal{F}} \subseteq \Omega_{\mathcal{A}}$ and $X_{\mathcal{F}'} \subseteq \Omega_{\mathcal{A}'}$, then we can build the SFT $X_{\mathcal{F}} \times X_{\mathcal{F}'}$ on the alphabet $\mathcal{A} \times \mathcal{A}'$.

Let us note that, with $\omega_1, \omega_2 \in \Omega_{\mathcal{A}}$ and $\omega_1', \omega_2' \in \Omega_{\mathcal{A}'}$, we have:

$$d_H(\omega_1, \omega_2) \leq d_H([\omega_1, \omega_2], [\omega_1', \omega_2']) \leq d_H(\omega_1, \omega_2) + d_H(\omega_1', \omega_2') ,$$

with the second one using a finite union bound.

Thence, if $X_{\mathcal{F}}$ and $X_{\mathcal{F}'}$ are f-stable and f'-stable on the same noise class $\mathcal{N}$, then the product $X_{\mathcal{F}} \times X_{\mathcal{F}'}$ (which also is an SFT) is $(f+f')$-stable on the class $\mathcal{N}$. If one of the SFTs is unstable, then the product is unstable with the same lower bound.

This informal alphabet extension idea, while not used verbatim later on, will be put to use to study multi-layered tilesets later on. ■

4.4 Stability for Higher-Dimensional Periodic Subshifts

In this section, we will explore the notion of stability for higher-dimensional (now denoted *2D+*) periodic SFTs. Here, by periodicity of the SFT we really mean strong periodicity, as discussed in 2.2, and not the periodificy of an associated structure like the word automaton $G_{\mathcal{F}}$ of the one-dimensional case.

First, we will use a periodic noise to illustrate how the instability argument for non-mixing one-dimensional SFTs adapts here. Then, we will focus on the Bernoulli noise and prove that, using a 2D+ percolation argument, we have linear stability for any 2D+ periodic SFT.

4.4.1 Instability for Grid Noises

As announced previously, let me generalise here the unstable example from Section 4.3.2 in the general case of periodic SFTs on $\mathbb{Z}^d$ with periodic noises.

Definition 4.34 (Grid Noise)**.** Consider $k, n \in \mathbb{N}^*$ two positive integers. We define the base pattern $b_{k,n}$ on a $(k+n)$-hypercube, such that for $x \in [\![0, k+n-1]\!]^d$, we have $b(x) = \mathtt{1}$ *iff* $\min_{1 \leq i \leq d} x_i < k$. We then identify $b_{k,n}$ to the configuration obtained by extending this base pattern periodically in all directions. In other words, $b_{k,n}$ is a $(k+n)$-periodic grid of clear n-hypercubes with obscured "walls" of thickness k. We finally define the corresponding shift-invariant noise:

$$\nu_{k,n} := \frac{1}{(k+n)^d} \sum_{x \in [\![0,k+n-1]\!]^d} \delta_{\sigma^x(b_{k,n})} .$$

The probability of an obscured cell in this noise is $1 - \left(\frac{k}{k+n}\right)^d$. ■

In higher dimensions, the previously defined diameter becomes $\operatorname{diam}(I) = \max_{i,j \in I} \|j - i\|_\infty$ for a finite window $I \Subset \mathbb{Z}^d$, $\operatorname{diam}(p) = \operatorname{diam}(I)$ for $p \in \mathcal{A}^I$ and $\operatorname{diam}(\mathcal{F}) = \max_{w \in \mathcal{F}} \operatorname{diam}(w)$ for sets of forbidden patterns. Assuming $k \geq \operatorname{diam}(\mathcal{F})$, then two distinct clear hypercubes of $\nu_{k,n}$ are "insulated" from each other by the obscured walls, and can be tiled independently from each other in a locally admissible way for $\mathcal{F}$, without any chance of creating a forbidden pattern in $\widetilde{\mathcal{F}}$.

Proposition 4.35. *For any non-trivial (*i.e.* not a singleton $\{a^\infty\}$) 2D+ N-periodic SFT $X_{\mathcal{F}}$, for any $\varepsilon > 0$, there is a measure $\mu \in \mathcal{M}_{\mathcal{F}}(\varepsilon)$ (with a grid noise) at distance at least $\frac{1}{2(N+1)^d}$ from $\mathcal{M}_\sigma(X_{\mathcal{F}})$.*

Proof. Let us assume that $X_{\mathcal{F}}$ is an N-periodic SFT (*i.e.* any translation σ^k with $k \in (N\mathbb{Z})^d$ is the identity map on $X_{\mathcal{F}}$). Then two distinct configurations $\omega \neq \omega' \in X_{\mathcal{F}}$ differ on *at least* one cell in any translation of the N-hypercube.

We will prove the result for the noises $\nu_{k,nN}$ as $n \to \infty$, for which the frequency of obscured cells is equal to $\varepsilon_n := \left(\frac{k}{k+nN}\right)^d \underset{n\to\infty}{\longrightarrow} 0$. We define the noisy measure $\lambda \in \widetilde{\mathcal{M}_{\mathcal{F}}}(\varepsilon_n)$ as follows:

- first, pick a noise grid at random following the measure $\nu_{k,nN}$,
- under any obscured cell pick a letter uniformly at random,
- then, independently from the noise, and independently on each clear nN-hypercube, pick a configuration $\omega \in X_{\mathcal{F}}$ uniformly at random, and write its restriction on the corresponding hypercube.

Fix a configuration $\omega \in X_{\mathcal{F}}$. For λ-a.e. noisy tiling (ω', b), on the first coordinate $\omega' \in \Omega_{\mathcal{A}}$, a proportion $\frac{1}{|X_{\mathcal{F}}|}$ of the clear Nn-hypercubes contain a restriction of ω. Hence, with probability $\frac{|X_{\mathcal{F}}|-1}{|X_{\mathcal{F}}|} \geq \frac{1}{2}$, the configuration chosen for this hypercube is *not* ω. For such a clear window, as an nN-hypercube, contains n^d disjoint N-hypercubes, we have at least n^d differences between ω and ω'. Finally:

$$d_H(\omega, \omega') \geq \frac{1}{2} \times \left(\frac{n}{k + nN}\right)^d .$$

This inequality holds λ-a.s. for *any* configuration $\omega \in X_{\mathcal{F}}$, so for big enough values of $n \geq k$, we obtain the lower bound $d_B\left(\pi^1_*(\lambda), \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right) \geq \frac{1}{2(N+1)^d}$. □

4.4.2 From Noisy SFTs to Percolations

In the one-dimensional case, with a Bernoulli noise, having room for aperiodicity (*i.e.* the mixing reconstruction) was what allowed us to correct defects in the noisy configurations from $X_{\widetilde{\mathcal{F}}}$ in order to couple them with configurations in $X_{\mathcal{F}}$, while intrinsic periodicity of the SFT was precisely what prevented stability. Yet, in the 2D+ case, we will see that periodicity *helps* stability, by making use of the redundancy, as long as most of the clear cells are connected to each other in an induced percolation process.

Once again, let us consider a variant of the reconstruction function from Remark 4.1. Here, $\varphi_{\mathcal{F}} : \mathbb{N} \to \mathbb{N}$ is a non-decreasing function such that, for any integer $n \in \mathbb{N}^*$, $\varphi(n) \geq n$ and whenever $\omega \in \mathcal{A}^{B_{\varphi(n)}}$ is a locally admissible pattern, its restriction ω_{B_n} is globally admissible.

For the one-dimensional case, we proved in Proposition 4.26 that this function can always be chosen as $\varphi(n) = n + c$ for some $c \in \mathbb{N}$. This property allowed us to convert a locally admissible configuration into a globally admissible clear one, up to some "peeling" around obscured cells. What we now want is to transpose this argument into the 2D+ case, using purposely the redundancy induced by the periodicity.

Lemma 4.36 (Connected Reconstruction). *Consider a 2D+ periodic SFT $X_{\mathcal{F}}$. There exists $c(\mathcal{F}) \in \mathbb{N}$ such that, for any* connected *window $I \subseteq \mathbb{Z}^d$ (not necessarily finite), if $w \in \mathcal{A}^{I+B_c}$ is locally admissible, then w_I is globally admissible.*

Proof. Let $X_{\mathcal{F}}$ an N-periodic SFT. Let us begin with the case where $I = \{e\}$ is made of a single cell. Assuming a pattern u on the window $e + B_{\lceil \frac{N}{2} \rceil}$ is *globally* admissible, then u actually is the restriction of a configuration $\omega_e \in X_{\mathcal{F}}$ that coincides with u on the window, and in particular $\omega_I = u_I$. Thus, if $c = \varphi\left(\left\lceil \frac{N}{2} \right\rceil\right)$, then whenever u is locally admissible on $e + B_c$, it is globally admissible on $e + B_{\lceil \frac{N}{2} \rceil}$ (and thus also on the subset $I = \{e\}$).

More generally, consider any connected window of cells I, and $u \in \mathcal{A}^{I+B_c}$ a locally admissible pattern. For any cell $e \in I$, we can likewise obtain a configuration $\omega^e \in X_{\mathcal{F}}$ such that, on the domain $e + B_{\lceil \frac{N}{2} \rceil}$, the pattern u and the configuration ω^e coincide.

Consider now two neighbouring cells $e, f \in I$. As we left a bit of margin to begin with, the intersection $\left(e + B_{\lceil \frac{N}{2} \rceil}\right) \cap \left(f + B_{\lceil \frac{N}{2} \rceil}\right)$ contains an N-hypercube (that contains both e and f), on which ω^e and ω^f coincide, so that $\omega^e = \omega^f$. As I is connected, by induction, the pattern u_I is actually a restriction of ω^e, hence it is globally admissible. □

To summarise, as long as we can reconstruct a globally admissible N-hypercube, we can precisely identify a corresponding element of $X_{\mathcal{F}}$. Consequently, if $X_{\mathcal{F}}$ is periodic, then for *any* configuration $(\omega, b) \in X_{\widetilde{\mathcal{F}}}$, $\gamma^c(b)$ contains several connected components I made of clear cells, and inside any such component we have ω_I globally admissible.

In all generality, with periodic noises for instance, we can obtain infinitely many finite components, without any reason for them to "synchronise" on the same $\omega \in X_{\mathcal{F}}$. However, in the specific case of Bernoulli noises, we can use a site percolation argument to conclude that there is a (unique) high-density component I.

4.4.3 Study of the Thickened Percolation

We consider here the site percolation process on $\mathbb{Z}^d$, with configurations $b \in \Omega_{\{0,1\}}$. In our framework, the *open* cells will be the clear ones, with value 0, and the *closed* ones will be the obscured ones, with value 1.

As we want specific properties on the "thickness" of the infinite component of this percolation (the connected window I such that $I+B_c$ is open), we will induce an auxiliary thickened percolation process. If $\nu \in \mathcal{M}_\sigma\left(\Omega_{\{0,1\}}\right)$ defines a random percolation, then the percolation process on $\mathbb{Z}^d$ induced by the measure $\gamma^n_*(\nu)$ will be called the n-thickened ν-percolation, with the cellular automaton γ^n from Definition 4.11.

In the article *Density and Uniqueness in Percolation* [BK89, Theorem 2], it is shown that under a condition of *finite energy* on the measure μ, defined below, the percolation process almost surely has *at most* one infinite connected component. This property holds true for any Bernoulli noise in particular.

Definition 4.37 (Finite Energy)**.** Consider $w \in \mathcal{A}^I$ a finite pattern. For a measurable set B, we define $B_w = \{\omega \in \Omega_\mathcal{A}, \exists \omega' \in B, \omega_{I^c} = \omega'_{I^c}, \omega_I = w\}$ which is also measurable. A measure μ has finite energy if, for any finite pattern w and any measurable set B, we have $\mu\left(B_w\right) > 0$ whenever $\mu(B) > 0$. ■

Note that thickened measures cannot have the finite energy property. Indeed, a consequence of finite energy is that any cylinder has a positive measure. However, for an n-thickened percolation, we cannot have three adjacent cells with the pattern 010 in a configuration $\gamma^n(b)$, as the presence of a 1 in an n-hypercube of b implies its presence in the left-translated or right-translated hypercube. The result can nonetheless be effortlessly adapted to the case of thickened measures, and we will sketch its proof here for completeness.

Lemma 4.38. *When ν has the finite energy property, any thickened ν-percolation has* at most *one infinite connected component.*

Proof. The finite energy property still holds for the measures obtained through the ergodic decomposition theorem, hence we can assume ν is ergodic. As γ_n is shift-invariant, by definition of ergodicity, if ν is ergodic, then so is the n-thickened ν-percolation.

As a shift-invariant measurable function, the number $N(b)$ of infinite components in the percolation b is $\gamma^n_*(\nu)$-a.s. constant.

If N was infinite, then for a big-enough hypercube B, the probability of encountering three different infinite components in $\gamma^n(b)$ inside of it would be positive.

In the context of site percolation processes, a trifurcation of a configuration b is an open cell that is part of an infinite component, with exactly three open neighbours such that if the cell was closed then these neighbours would each be in a different infinite component.

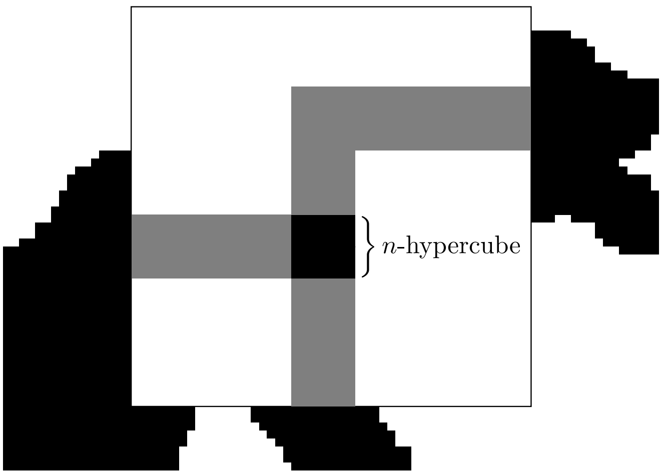

Figure 4.3: Schematic representation of a thick trifurcation in b, with open cells in black, corresponding to a standard trifurcation in $\gamma^n(b)$.

Using the finite energy property to change the configuration b inside of B when it encounters three infinite thick components, as illustrated on Figure 4.3, there is a positive probability of observing a standard trifurcation inside of B for $\gamma^n(b)$.

The rest of the proof follows as in the original theorem: if the probability that a cell is a trifurcation is positive, then so is the frequency of trifurcations by Birkhoff's ergodic theorem on $\mathbb{Z}^d$, thus it must be of order

n^d in a big hypercube. However, a theoretical $O\left(n^{d-1}\right)$ bound can be obtained on the number of trifurcations, thus a contradiction. The number N cannot be infinite.

With a similar but much simpler finite energy argument, N cannot be constant greater or equal to 2, as the probability of having at most $N-1$ components would be positive, by opening an entire hypercube encountering several components. □

Thanks to this result, we can from now on talk about *the* infinite component of the percolation process, whenever it exists. We now need to actually control the frequency of cells belonging to it. Further analyses will be done on a Bernoulli noise, but we still hope for a more general result to come from percolation theory.

Proposition 4.39 (Frequency of the Infinite Component)**.** *Consider $I(b) \subseteq \mathbb{Z}^d$ the random infinite component of the n-thickened percolation $\gamma^n(b)$, with respect to the original ε-Bernoulli percolation process $\mathbb{P} = \mathcal{B}(\varepsilon)^{\otimes \mathbb{Z}^d}$. When such an infinite component does not exist, we use the convention $I(b) = \emptyset$.*

Then the constant $C_n^d = 48(2n+1)^d$ is such that $\mathbb{P}(0 \notin I) \leq C_n^d \times \varepsilon$.

Proof. Let us describe first what the event $\{0 \notin I\}$ represents. Either the cell 0 is closed in $\gamma^n(b)$ (*i.e.* $\gamma^n(b)_0 = \mathtt{1}$) so that it belongs to no component, or it is open, but its component is finite. The first scenario happens with probability $\left(1-(1-\varepsilon)^{(2n+1)^d}\right)$.

In the second scenario, this implies that the component of 0 in the percolation process induced by $\gamma^n(x)$ on the lattice $\mathbb{Z}^2 \times \{0\}^{d-2}$ is also finite. Consider the sub-lattice $[(2n+1)\mathbb{Z}]^2 \times \{0\}^{d-2}$, where two cells are adjacent whenever one coordinate differs by $2n+1$. If two neighbouring cells e and f of this sub-lattice are open in $\gamma^n(b)$, then all the cells in $(e+B_n) \cup (f+B_n)$ must be open for b. Hence, if e and f are open, neighbours in the sub-lattice, then all the cells that lie in-between in $\mathbb{Z}^2$ are also open, so that e and f are in the same connected component of $\gamma^n(b)$. The interest of this trick is that, as those windows $e+B_n$ and $f+B_n$ are disjoint, the value of the cells e and f in $\gamma_n(b)$ are actually independent. To put it short, in this second scenario, the component of 0 in the sub-lattice $[(2n+1)\mathbb{Z}]^2$ must be finite too.

The percolation process on the sub-lattice is just a planar $\left(1-(1-\varepsilon)^{(2n+1)^d}\right)$-Bernoulli independent site percolation. In this case, if the component of 0 is finite, then the outer boundary of this component must be a cycle of closed cells, where two neighbouring cells may be diagonally adjacent, so we just need an upper bound on the probability of *this* event.

We can easily start with the upper bound $1-(1-\varepsilon)^{(2n+1)^d} \leq (2n+1)^d\varepsilon$ on the probability of a cell being closed. Now we need to count the number of cycles of a given length l. Such a cycle must necessarily intersect the half-line $\mathbb{N}^* \times \{0\}$, let's say at coordinates $(k, 0)$, and each of the columns $\{j\} \times \mathbb{Z}$ with $0 \leq j < k$ must cross the cycle at least twice, thus $l \geq 2k$ gives us an upper bound on the coordinate k. Note also that a cycle is in particular a self-avoiding path, so that, for a fixed value of k, we can upper bound the number of cycles by $9 \times 8^{l-1}$. Whenever $\varepsilon < \frac{1}{8(2n+1)^d}$, we have:

$$\begin{aligned}
\mathbb{P}(0 \notin I) &\leq (2n+1)^d\varepsilon + \sum_{l \geq 4} \frac{l}{2} \times 9 \times 8^{l-1} \times \left((2n+1)^d\varepsilon\right)^l \\
&\leq \frac{9}{16}\varepsilon \times \sum_{l \geq 1} 8(2n+1)^d \times l\left(8(2n+1)^d\varepsilon\right)^{l-1} \\
&= \frac{9}{16}\varepsilon \times \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[\sum_{l \geq 0}\left(8(2n+1)^d\varepsilon\right)^l\right] \\
&= \frac{9}{16}\varepsilon \times \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[\frac{1}{1-8(2n+1)^d\varepsilon}\right] \\
&= \frac{9}{16}\varepsilon \times \frac{8(2n+1)^d}{(1-8(2n+1)^d\varepsilon)^2}\,.
\end{aligned}$$

So far, this upper-bound is of the form $\varepsilon f(\varepsilon)$ for some function f that is positive increasing on the interval $\left[0, \frac{1}{8(2n+1)^d}\right[$ and goes to infinity on the right. If we find ε_0 in this interval such that $\varepsilon_0 f(\varepsilon_0) = 1$, then the upper bound by $f(\varepsilon_0)\,\varepsilon$ will hold on this interval as f is increasing, and the upper bound will hold for $\varepsilon_0 \leq \varepsilon \leq 1$ as $\mathbb{P}(0 \notin I) \leq 1 \leq f(\varepsilon_0)\,\varepsilon$ on this interval.

Let us denote $a = \frac{9}{16}$ and $b = 8(2n+1)^d$. Solving $\varepsilon f(\varepsilon) = 1$ equates finding the root of $b^2\varepsilon^2 - b(a+2)\varepsilon + 1$ on the interval $\left[0, \frac{1}{b}\right]$. The roots are $\varepsilon_\pm = \frac{\sqrt{a+2}}{b}\left(\frac{\sqrt{a}+\sqrt{a+2}}{2}\right)$ and only ε_- is in the desired interval. A direct computation then yields $f(\varepsilon_-) = \frac{2b}{1-a\left(\sqrt{1+\frac{2}{a}}-1\right)}$. Replacing a by its value, we obtain $1 - a\left(\sqrt{1+\frac{2}{a}}-1\right) = \frac{25-3\sqrt{41}}{16} > \frac{1}{3}$, thus finally $f(\varepsilon_-) < 6b$. At last, the constant $C_n^d = 6b = 48(2n+1)^d$ provides the desired upper bound. □

This proof depends on the specific properties of the independent percolation process, but is quite elementary in exchange. In order to adapt the following periodic stability theorem to a more general class of noises $\mathcal{N}$, one would first need to obtain a similar lower bound on the frequency of cells in the (unique) infinite connected component, so that $\sup_{\nu \in \mathcal{N}, \nu([1]) \leq \varepsilon} \nu(0 \notin I) \underset{\varepsilon \to 0}{\longrightarrow} 0$.

4.4.4 Stability Theorem

Theorem 4.40 (2D+ Periodic Stability). *Consider $X_\mathcal{F}$ a 2D+ N-periodic SFT. Then $X_\mathcal{F}$ is f-stable with respect to d_B on the class $\mathcal{B}$ of Bernoulli noises, with linear speed $f(\varepsilon) = 2C^d_{c(\mathcal{F})}\varepsilon$.*

Proof. In order to obtain linear stability, we will consider a measure $\lambda \in \widetilde{\mathcal{M}_\mathcal{F}^\mathcal{B}}(\varepsilon)$, and build a measurable mapping $\psi : X_{\widetilde{\mathcal{F}}} \to X_\mathcal{F}$, so that $d_H(\omega, \psi(\omega, b))$ is small for a λ-typical configuration $(\omega, b) \in X_{\widetilde{\mathcal{F}}}$.

Let c the constant obtained in Lemma 4.36. As $X_\mathcal{F}$ is finite, it makes sense to use it as a finite alphabet and to define the full-shift $\Omega_{X_\mathcal{F}}$. Define the morphism $\rho : X_{\widetilde{\mathcal{F}}} \to \Omega_{X_\mathcal{F}}$ such that, whenever the window B_c is clear in $\sigma^e(\omega, b) \in X_{\widetilde{\mathcal{F}}}$, then $\rho(\omega, b)_e = \omega_e$ as in Lemma 4.36, but specifically for the window B_c of $\sigma^e(\omega, b)$. This translation process is necessary to obtain a morphism ρ (or else translating (ω, b) would also "translate" the letters of $\rho(\omega, b)$ in the alphabet $X_\mathcal{F}$). If the window is partially obscured, then we may default to some fixed configuration $\omega' \in X_\mathcal{F}$. In particular, we have a local characterisation $\rho : \widetilde{\mathcal{A}}^{B_c} \to X_\mathcal{F}$ for the morphism.

Without loss of generality, assume the finite set $(X_\mathcal{F}, <)$ is strictly ordered. We may now define the adjusted majority rule cellular automaton $\theta_n : X_\mathcal{F}^{B_n} \to X_\mathcal{F}$ as follows. First, shift each configuration-letter $(\omega_e)_{e \in B_n}$ onto the configuration $\sigma^{-e}(\omega_e) \in X_\mathcal{F}$, so that we locally undo the offset introduced by ρ. Then, among them, we may apply a regular majority rule, by picking the maximal configuration for the arbitrarily introduced order in case of a tie.

Consider now the morphisms $\psi_n = \theta_n \circ \rho$ obtained by applying an adjusted majority rule after ρ. Using once again the order on $X_\mathcal{F}$, we may define the pointwise limit $\psi(\omega, b) := \overline{\lim_{n\to\infty}}\, \psi_n(\omega, b)_0$. The map $\psi : X_{\widetilde{\mathcal{F}}} \to X_\mathcal{F}$ is measurable and commutes with σ. Note that $\psi(\omega, b)$ may depend on arbitrarily far values, so ψ is *not* a morphism.

Consider the configuration $(\omega, b) \in X_{\widetilde{\mathcal{F}}}$, and let $I \subseteq \mathbb{Z}^d$ be the infinite component of the c-thickened percolation in b. As ω_{I+B_c} is locally admissible (with respect to $\mathcal{F}$, as the window is clear), ω_I is a globally admissible pattern, the restriction of some configuration $\omega_0 \in X_\mathcal{F}$. For any cell $e \in I$, we have $\rho(\omega, b)_e = \sigma^e(\omega_0)$.

Assume now that $\varepsilon < \frac{1}{2C_c^d}$, so that in the Bernoulli percolation process, I has a density greater than $\frac{1}{2}$ according to Proposition 4.39. This means that, λ-a.s., after some (random) rank n_0, strictly more than half of the cells $e \in B_n$ of (ω, b) are inside of I, thus are mapped by ρ onto translations $\sigma_f(\omega_0)$. Thence, after the very same rank n_0, $\psi_n(\omega, b)_0 = \omega_0$. Consequently, by taking the limit $n \to \infty$, λ-a.s., $\psi(\omega, b) = \omega_0$.

To sum it up, $\psi : (\omega, b) \in X_{\widetilde{\mathcal{F}}} \mapsto \omega_0 \in X_\mathcal{F}$ is a measurable map, that commutes with σ, and such that $d_H(\omega, \omega_0) \leq C_c^d\varepsilon$ whenever $\varepsilon \leq \frac{1}{2C_c^d}$. More generally, the bound $d_H(\omega, \psi(\omega, b)) \leq 2C_c^d\varepsilon$ holds λ-a.s. for any choice of ε:

$$d_B\left(\pi^1_*(\lambda), \mathcal{M}_\sigma(X_\mathcal{F})\right) \leq d_B\left(\pi^1_*(\lambda), \psi_*(\lambda)\right) \leq 2C_c^d\varepsilon\,.$$

This concludes the proof. □

This concludes our analysis of periodic SFTs in the 2D+ case. The explicit constant C_n^d could doubtlessly be improved, but such matters would require much more work without improving on the *linear* aspect of the bound.

A further track of reflection, as already mentioned earlier, may be to extend this theorem to a more general class of noises, using stronger percolation results, while leaving much of the actual proof of the theorem unchanged.

Another interesting direction here may be to consider periodic SFTs on *other* lattices than $\mathbb{Z}^d$, but where we still have results regarding the independent Bernoulli site percolations.

However, in this long quest to study the computability of the question of stability, the next milestone is to simulate Turing machines in a stable way, which in turn requires us to obtain the stability for simulating structures such as the Robinson tiling. Fortunately, the Robinson tiling is *almost* periodic, so we will be able to adapt the arguments from the periodic case, up to some tinkering.

Before doing so, however, let me digress once again to discuss another related notion, that of *robustness*.

4.5 The Case of Two-Dimensional Robust Tilesets

The aim of this section is to provide an informal analysis and restating of an already existing Besicovitch stability result in the current framework. More precisely, I will discuss the notion of stability described by Durand, Romaschenko and Shen [DRS12], which was then notably used to prove periodic stability in the two-dimensional case in a further article by Ballier, Durand and Jeandel [BDJ10]. Afterwards, we will see that this notion of robustness does *not* apply to the Robinson tileset.

Here, I will provide a rough and qualitative estimate of the convergence speed obtained with their method, starting from the definition of robustness.

4.5.1 Robust Tilesets and Sparse Sets

To obtain stability, instead of using a notion of percolation – which is best seen as one clear connected component that spans the whole otherwise obscured space – they introduce the notion of islands of errors – which is best seen as small clumps of obscured cells isolated in the whole clear space.

Definition 4.41 ((α, β)-Island of Errors)**.** Consider a noise configuration $b \in \{0, 1\}^{\mathbb{Z}^2}$ which we identify with $E \subseteq \mathbb{Z}^2$ the set of obscured cells.

A set $F \subseteq E$ is an (α, β)-island of E if F can be included in some α-square and its β-neighbourhood does not meet any other obscured cell of E, *i.e.* $(F + B_\beta) \cap (E \backslash F) = \emptyset$. ■

In this framework, the "right" way to obtain stability is to remove the islands of obscured cells, by changing the values of the tiles underneath on a small neighbourhood. This is well-encapsulated by the following notion of robustness.

Definition 4.42 ((c_1, c_2)-Robustness)**.** Let us denote by $R_{i,j} := B_{\frac{j-1}{2}} \backslash B_{\frac{i-1}{2}}$ (with $i < j$ two odd integers) the ring-shaped window obtained by removing the i-square at the centre of a j-square (with the convention that $R_{0,j}$ is just a j-square).

Let $0 < c_1 \leq c_2$ be two positive integers. A tileset $\mathcal{F}$ is (c_1, c_2)-robust if, for any $n \in \mathbb{N}$ and any locally admissible pattern $u \in \mathcal{A}^{R_{n,c_2 n}}$, there exists a locally admissible pattern $v \in \mathcal{A}^{R_{0,c_2 n}}$ such that u and v coincide on $R_{c_1 n, c_2 n}$ – which is a strict subset of the ring $R_{n,c_2 n}$ as long as $c_1 \geq 2$. ■

An explicit example of robust tileset is any one inducing a periodic SFT [BDJ10], roughly for the same reason we could obtain a globally admissible configuration by peeling a constant width of the border of any connected pattern in the previous section. However, this notion is much more general, and strongly aperiodic robust SFTs are proven to exist [DRS12].

Note that, while the constants may change in the process, this notion of robustness is invariant under conjugacy, so that we cannot prove stability of a non-robust SFT by looking for a suitable robust conjugated. Notably, we will see this framework cannot apply to the Robinson tileset.

Whenever $\beta \geq c_2\alpha$, we can "repair" an (α, β)-island of errors by changing the tiles in a $c_1\alpha$-square. In some sense, this property can be seen as the natural two-dimensional generalisation of our reconstruction on the neighbourhood of obscured blocks in the case of a one-dimensional mixing SFT. Hence, the next natural step is to obtain something analogous to our argument about being able to locally repair increasingly large obscured intervals (*i.e.* islands of errors, in this context), so we need to partition E into such "fixable" islands.

Definition 4.43 ((α, β)-Sparse Set)**.** A set $E = E_0$ is said to be sparse, given a sequence $(\alpha_k, \beta_k)_{k \in \mathbb{N}^*}$, if we can step by step remove all the (α_k, β_k)-islands from E_{k-1} to obtain a set E_k, in such a way that the decreasing limit set $E_\infty = \bigcap E_k$ is empty. ■

Up to now, the definitions introduced were formal, to underline how similar the argument can be to the one-dimensional mixing case. For the rest of this section, I will provide a qualitative and quite handwavy analysis of the convergence speed we can obtain in this framework, which already requires *much* more computations than what I did in the one-dimensional case.

4.5.2 Qualitative Analysis of the Convergence Speed

By the Borel-Cantelli theorem, any ε-Bernoulli noise will certainly contain islands for any pair (α, β), which may *a priori* be hard to correct. However, it is proven [DRS12, Lemma 3] that, assuming $8\sum_{k=1}^{n-1}\beta_k < \alpha_n \leq \beta_n$ for any $n \in \mathbb{N}^*$ and $\sum_n \frac{\ln(\beta_n)}{2^n} < \infty$, then for ε small enough the random set E is almost surely (α, β)-sparse. Unfortunately, general bounds on ε would be quite hard to obtain, but I will provide rough estimates for a choice of (α, β).

It is also proven [DRS12, Lemma 4] that in any (α, β)-sparse set E, the density of obscured cells is *at most* $\sum_n (\alpha_n/\beta_n)^2$ – the main argument is that each (α_n, β_n)-island contains *at most* α_n^2 obscured cells, among *at least* β_n^2 cells in a neighbourhood of the island disjoint of the other islands and their neighbourhoods. To properly quantify the convergence speed, we would need to take into account the density not of the islands of errors but of the $c_1\alpha$-square around them, but this approximation will suffice for the present qualitative analysis.

Consider $\alpha_n = 8^n(n-1)!n!$ and $\beta_n = 8^n(n!)^2$. It is clear that any *k-shift* of this sequence (starting at some rank $k+1$ instead of 1) will satisfy the previously stated hypotheses. For a given sparse set E, for the k-shifted sequence, the density of errors is bounded by $\sum_{n=k+1}^{\infty} \frac{1}{n^2} \leq \int_k^{\infty} \frac{1}{t^2}\mathrm{d}t = \frac{1}{k}$.

To obtain the convergence speed, we now need to estimate the maximal value of k such that E is sparse for the k-shifted sequence for a given ε. Looking at the proof of the result [DRS12, Lemma 3], it appears that the key property to obtain sparsity is that $\sum_n \frac{\ln(\beta_n)}{2^n} < \ln\left(\frac{1}{\varepsilon}\right)$. As $\ln(\beta_n) = 8\ln(n) + 2\ln(n!) \leq n^2$ after some rank, for the k-shifted sequence, we can bound the left term by $k^2 + 4k + 6$. Asymptotically, the best choice for k is thus $k(\varepsilon) \approx \sqrt{\ln(1/\varepsilon)}$, so that $f(\varepsilon) \approx \frac{1}{\sqrt{\ln(1/\varepsilon)}}$.

Note that, as $\frac{\alpha_n}{\beta_n} \to 0$, for *any* pair (c_1, c_2) (and any accordingly (c_1, c_2)-robust tileset), we will satisfy the $\beta_n \geq c_2\alpha_n$ condition after a rank n. It follows that this bound on the speed of convergence holds for any robust tileset, but on an interval $[0, \varepsilon_{\max}(c_2)]$ that depends on the constant c_2, with $k(\varepsilon_{\max})$ such that $\beta_k \geq c_2\alpha_k$. As a rule of thumb, the bigger c_2 gets, the smaller $\varepsilon_{\max}$ gets. In any case, this doesn't affect the asymptotic order of the bound outside of a multiplicative factor.

Considering all the small approximations we did on the way, what matters here is not the value of the bound but its order of magnitude. Indeed, $\frac{1}{\sqrt{\ln(1/\varepsilon)}}$ is much *much* slower than any polynomial speed, which legitimises the efforts to obtain a linear convergence speed in the periodic case.

The notion of islands and sparsity can be used as a black box to obtain percolation results [DRS12, Section 9.3], hence as a tool it is in some ways more powerful than the percolation theory we used in the previous section. However, as we have seen here, this versatility comes at the cost of the precision and simplicity of the bounds we can obtain.

4.6 Stability for the Robinson Tiling(s)

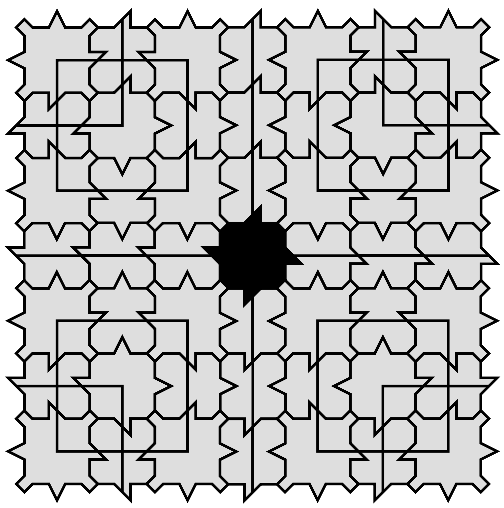

Figure 4.4: Four (arbitrarily long) lines around one obscure cell, with ∞-macro-tiles in each quarter plane, in a locally admissible way.

As mentioned in the previous section, the Robinson tileset (defined in Section 2.7) is not robust, and the counterexample in Figure 4.4 is quite simple. It is obtained by having an empty cell at the origin, an infinite arm along each axis, and an ∞-macro-tile in each quarter plane. Any attempt at a repair around the origin would require the four arms to form a cross, which is impossible.

However, the almost periodic hierarchical structure implies that, for a given scale, the corresponding macro-tiles form a periodic structure (in ∞-macro-tiles for the canonical Robinson, and globally for the enhanced Robinson), setting aside a small fraction of tiles (which corresponds to the lines used to form macro-tiles higher in the hierarchy).

The main goal of this section is to adapt the periodic scheme of proof from Section 4.4 to the enhanced Robinson tiling, now with a polynomial speed (Theorem 4.50). In the next section, a generalisation of this scheme of proof will be provided, but the key idea gets a bit drowned in the formalism, so I think there is value in proving it specifically for the enhanced Robinson tileset first. The section will be concluded by an unstable example, this time using the Red-Black Robinson tileset.

4.6.1 Local Alignment for the Enhanced Robinson Tileset

As was illustrated in Figure 2.4, using the canonical Robinson tileset, we can obtain misaligned cuts. This happens with probability 0 for the unique ergodic measure on the tiling, but is still an issue locally which prevents us from using the periodic scheme of proof, as there is no guarantee that the macro-tiles stay well-aligned within the high-density clear connected component of the thickened percolation. The enhanced Robinson tileset removes this hurdle, as it forces alignment, not just generically for the unique ergodic measure but locally for any admissible pattern.

Definition 4.44 (Well-Aligned and Well-Oriented Pairs)**.** A pair of N-macro-tiles (both having an actual position in $\mathbb{Z}^2$) is said to be well-aligned if both of their centres have one coordinate in common, and the other differs by exactly 2^N, so that there is a gap of precisely one line/column between them.

More generally, we say the two N-macro-tiles are *loosely* aligned (with $0 < k \leq 2^N - 1$ tiles in common) when one of the coordinates of their centres differs by exactly 2^N and the other by $2^N - k - 1$, *i.e.* we start with a well-aligned pair (with $2^N - 1$ tiles in common) and we translate one of them of k units in the direction of the gap in-between, as in Figure 4.5.

A pair of well-aligned macro-tiles is said to be well-oriented if they form a pattern or (or a rotation of these), which can then actually be filled by the arm of a central cross in the process of making a larger macro-tile. ■

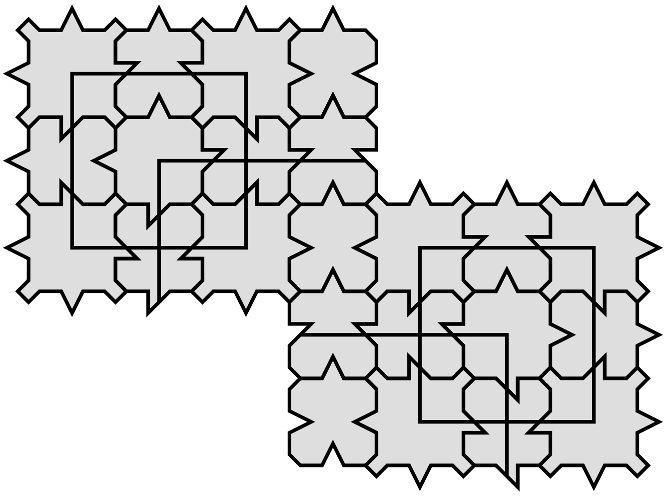

Figure 4.5: Two loosely aligned 2-macro-tiles for the canonical Robinson tileset, with a gap that can be tiled in a locally admissible way.

Notably, as illustrated in Figure 4.5, the canonical Robinson tiling allows loosely aligned pairs, with a gap filled in a locally admissible way. I will now prove that this is *not* the case for the enhanced tileset.

Definition 4.45 (Edge Words of Macro-Tiles)**.** We define the words l_N and t_N, obtained by reading the colours on the left and top edges of the N-macro-tile in a clockwise motion, with blue dotted lines encoded as a `0` and red dashed lines as a `1`. ■

For a binary word, we define $\bar{b} = 1 - b$ the binary complement of a letter, extended to binary words by a direct induction (letter by letter). We also define the mirror map such that $\text{mirror}(uv) = \text{mirror}(v)\text{mirror}(u)$, that returns the same word but backwards. Both of these mappings are involutions and they commute with each other.

Lemma 4.46. *Let l_N and t_N be the previously defined edge words of an enhanced N-macro-tile. For any $N \in \mathbb{N}^*$, we have $t_N = \overline{\text{mirror}(l_N)}$. What is more, $|l_N| = |t_N| = 2^N - 1$ is odd, and these words actually differ of exactly one letter in their middle.*

Proof. For $N = 1$, we simply have $l_1 = 0$ and $t_1 = 1$. When building an $(N+1)$-macro-tile, on the left edge from bottom to top, we first have a N-macro-tile that reads as t_N, then we read the $\mathtt{0}$ given by the blue dotted arm of the central cross, and finally l_N on the N-macro-tile, so that $l_{N+1} = t_N \mathtt{0} l_N$. Likewise, $t_{N+1} = t_N \mathtt{1} l_N$. Hence:

$$\overline{\text{mirror}(l_{N+1})} = \overline{\text{mirror}(t_N \mathtt{0} l_N)} = \overline{\text{mirror}(l_N)}\mathtt{1}\overline{\text{mirror}(t_N)} = t_N \mathtt{1} l_N = t_{N+1}\,,$$

which concludes the proof by induction. □

For instance, in Figure 2.6, we can observe that $l_3 = \mathtt{1100100}$ and $t_3 = \mathtt{1101100}$.

Proposition 4.47 (Local Alignment of Macro-Tiles)**.** *Consider the enhanced Robinson tileset, introduced in Section 2.7.2. For any scale $N \in \mathbb{N}^*$, a pair of loosely aligned N-macro-tiles with a tileable gap in-between (that can be filled in a* locally admissible *way as in Figure 4.5)* must *be well-aligned and well-oriented.*

Proof. Assuming the pair is well-aligned, the local matching rules where their two central arms join guarantee that the ill-oriented pairs cannot match: tries to join a black arm with a non-black one, tries to join two dashed red lines, etc. The only compatible pairs are thus the well-oriented ones, where two black arms join to form a half-square, or a red dashed arm joins a blue dotted one.

Now, at the scale $N = 1$, if two tiles are loosely aligned they are actually well-aligned, thus if the gap is tileable they are well-oriented. This allows us to initialise the induction.

Assume the result holds up to scale $N \in \mathbb{N}^*$ and consider a pair of $(N+1)$-macro-tiles, once again loosely aligned with a tileable gap. The macro-tiles cannot have exactly one tile in common, which would imply that we have two 1-macro-tiles well-aligned with a tileable gap but ill-oriented, hence $k \geq 2$.

What is more, k cannot be even. Assuming k is even, this pair of $(N+1)$-macro-tiles contains a pair of 2-macro-tiles with a tileable gap and 2 tiles in common. It is clear that this cannot happen, by an exhaustion of cases. For example, looking at a well-aligned , if we move the right macro-tile of one unit upwards, then the right arm of the left macro-tile will face the bottom-left corner of the right macro-tile, both with a red dashed line, which is impossible.

This concludes the case $N + 1 = 2$, as $k \geq 3$ must then be equal to 3, maximal, so that the 2-macro-tiles are well-aligned. Likewise, when $N + 1 > 2$, the N-macro-tiles must be well-aligned with k odd, so either the $(N+1)$-tiles are well-aligned, or only half of the N-macro-tiles actually face the gap and are well-aligned. In the second scenario, we are once again in a tileable ill-oriented case, impossible. Finally, the $(N+1)$-macro-tiles must be well-aligned thus well-oriented, which concludes the induction. □

With this local tool, we can now adapt the local reconstruction that was used for the periodic case. However, this will now give us a reconstruction *specifically* for the macro-tiles of a certain scale (that behave periodically) without giving us control on what happens in the grid around them.

Proposition 4.48 (Almost Reconstruction)**.** *For any scale $N \geq 2$, let $C_N = 2^N - 1$. This constant is such that for any $n \in \mathbb{N}$ and any clear locally admissible pattern ω on B_{n+C_N}, its restriction ω_{B_n} is almost globally admissible, in the sense that up to a one-tile thick ignored grid, ω_{B_n} is the restriction of an enhanced Robinson tiling, a grid of well-aligned and well-oriented N-macro-tiles, as in Figure 4.7.*

Proof. We will demonstrate a slightly stronger result here, *i.e.* that by removing *at most* C_N layers of tiles on the top, bottom, left and right sides of any locally admissible *square*, and not necessarily the same amount of layers on each side, we obtain an actual grid of N-macro-tiles, well-aligned and well-oriented with their neighbours. Thence, by actually peeling C_N layers on each side, we obtain the stated result.

To do so, we need to proceed inductively, as before. We cannot really initialise the result at $N = 1$ without proving the case $N = 2$ at the same time, which is why we only give C_N for $N \geq 2$ in this proposition (though C_2 will work for the case $N = 1$ too).

Hence, let us now prove the case $N = 2$, by peeling 3 layers of any admissible k-square B to obtain the announced restriction (remember that B_n is a *ball* of *radius* n, hence a $(2n+1)$-square). First, it is known that with the canonical Robinson tileset, if a locally admissible 3-square has a ◳ tile in the bottom-left corner, then it is a 2-macro-tile. This property still holds for the enhanced tileset, and can be easily checked by enumerating all the cases.

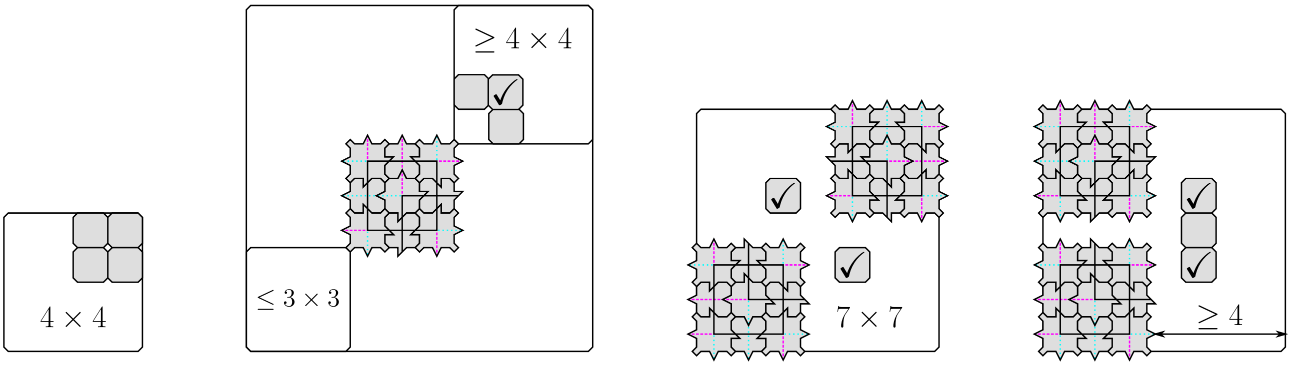

Figure 4.6: From left to right, key steps (a), (b), (c) and (d) of the case $N = 2$.

We will inductively build a rectangle of well-aligned 2-macro-tiles in the k-square B, assuming that $k \geq 10$ for now. If we look at the 4-square in the bottom-left corner, one of the four cells highlighted in the step (a) of Figure 4.6 must contain a 1-macro-tile, with bumpy corners. Then this bumpy corner is actually part of a 2-macro-tile in B. So far, we have a 1×1 rectangle of 2-macro-tiles.

As illustrated in step (b), considering where our first bumpy corner was, there is at *most* a 3×3 rectangle in the bottom-left corner (diagonally adjacent to the 2-macro-tile), and $k \geq 10$, so the top-right corner is at *least* a 4×4 rectangle of tiles. One of the three highlighted tiles in step (b) must be a bumpy corner too. If there was a corner in one of the unchecked tiles, it would be part of a 2-macro-tile, that should either intersect the one drawn on Figure 4.6 – which is impossible even for the canonical Robinson tileset – or be loosely aligned with it – which is impossible for the enhanced Robinson tileset according to Proposition 4.47. Hence the checked cell must contain a tile with bumpy corners, and more precisely a ◳ tile for the same reasons. This tile can then be completed into a 2-macro-tile, which brings us to step (c). There, the two checked cells must contain a 1-macro-tile too, and each can be completed into its own 2-macro-tile, so that we obtain at last a 2×2 rectangle of 2-macro-tiles.

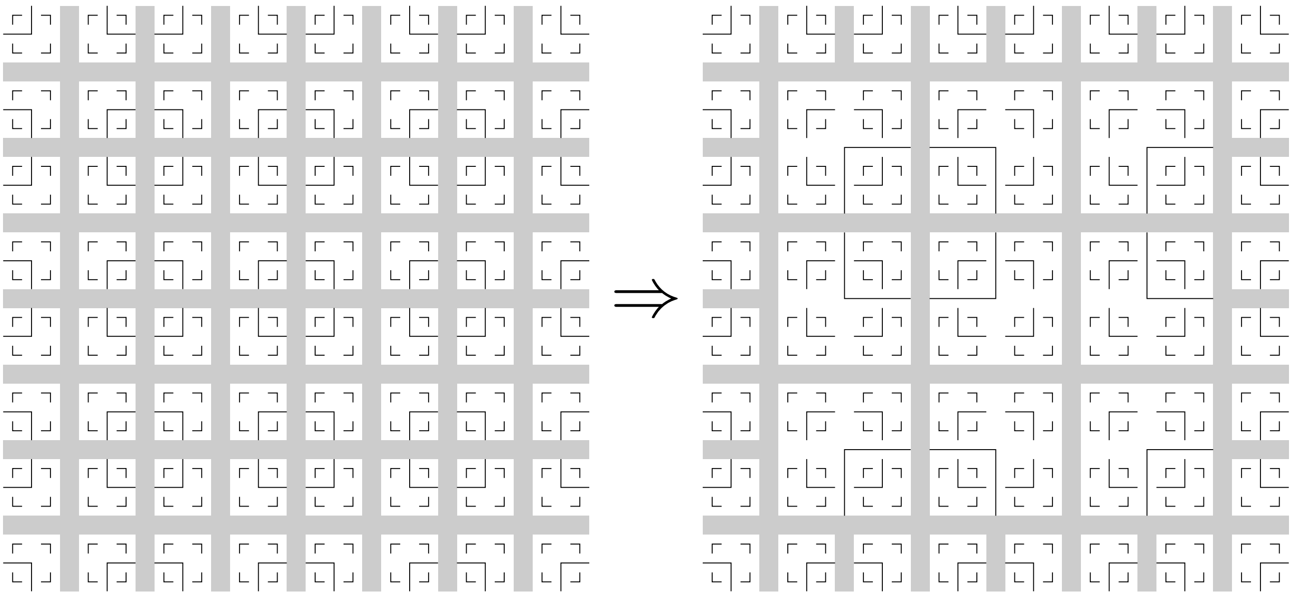

Figure 4.7: By filling the tileable grid around N-macro-tiles, we obtain $(N + 1)$-macro-tiles, up to one outer layer of N-macro-tiles.

Just like the two diagonally adjacent 2-macro-tiles present in step (c) imply a square of 2-macro-tiles, the presence of two laterally adjacent 2-macro-tiles in step (d) implies a square of 2-macro-tiles. Thus, now that we have a rectangle with *at least* 2 macro-tiles on each side, we can repeat step (d) in each direction as long as 4 tiles or more remain. Hence, as long as $k \geq 10$, $C_2 = 3$ works well.

More generally, we can trivially peel a 9-square into one single tile of grid if we remove 4 layers on each side, so that $C_2 = 4$ works in this case. However, a more careful study of the cases $k \in \{7, 8, 9\}$ allows us to conclude that $C_2 = 3$ works for these cases and is optimal (to do so consider a 9-square *centered* on a 2-macro-tile, so that all the adjacent ones will be missing a layer). When $k \leq 6$, $C_2 = 3$ trivially works too, which concludes our study of $N = 2$.

Assume now that the result holds at rank N with the constant C_N and let us prove it at rank $N+1$. We can start by peeling away at most C_N tiles, using our induction hypothesis, to obtain a grid of well-aligned and well-oriented N-macro-tiles. A square of well-aligned N-macro-tiles can either form one $(N+1)$-macro-tile, represent the lateral interface between two $(N+1)$-macro-tiles or represent the central corner between four $(N+1)$-macro-tiles. Thus, by peeling at most one layer of N-macro-tiles on each border – an N-macro-tile not part of an $(N+1)$-macro-tile and the following grid, so 2^N tiles in total – we remove the incomplete interfaces and corners to obtain a grid of well-aligned $(N+1)$-macro-tiles (hence well-oriented by the previous proposition) as seen in Figure 4.7. In conclusion, the result holds at rank $N+1$ with the constant $C_{N+1} = C_N + 2^N$, hence $C_N = 2^N - 1$ by a direct induction. □

4.6.2 Almost Stability at a Fixed Scale and Stability

The previous proposition can now be used to obtain an analogous to Lemma 4.36, by reconstructing a family of well-oriented and well-aligned N-macro-tiles in the infinite connected component of the thickened percolation, with a thickness that depends on C_N. I will do so directly while building the coupling, in the following proposition.

Proposition 4.49 (Almost Stability)**.** *Let $X_{\mathcal{F}}$ be the enhanced Robinson tiling (see Section 2.7.2). For any choice of $\varepsilon > 0$, any scale $N \geq 2$ of macro-tiles, and any measure $\mu \in \mathcal{M}_{\mathcal{F}}^{\mathcal{B}}(\varepsilon)$, we have a coupling s.t.:*

$$d_B\left(\mu, \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right) \leq 96\left(2^{N+2}+1\right)^2 \varepsilon + \frac{1}{2^{N-1}} .$$

Proof. For a given scale N, we want to apply the percolation argument as if we had a $\left(2 \times 2^N\right)$-periodic SFT. The factor 2 specifically comes from the fact that N-macro-tiles (of size 2^N when counting the ignored grid) alternate between two orientations on each given row or column (for instance ◳ and ◰ on a row, and ◲ and ◱ on the next one) in any globally admissible tiling $\omega \in X_{\mathcal{F}}$, hence a 2^{N+1} periodic behaviour.

By looking at a *globally admissible* 2^N-square, we can always identify one, two or four partial N-macro-tiles. Thus, we can actually identify to which translation of the 2^{N+1}-periodic pattern this window corresponds. Note that unlike in the general k-periodic case, where we needed to look at k-squares to identify the translation, we only need to look at a window of size $\frac{k}{2}$ here because the Robinson tiling has a lot of intrinsic redundancy already. Just like in the periodic case, we can then look at the c-thickened percolation, with

$$c = \left\lceil \frac{2^{N+1}+1}{2} \right\rceil + C_N = 2^N + 1 + 2^N - 1 = 2^{N+1} ,$$

as explained in Lemma 4.36 but using the almost reconstruction property from Proposition 4.48. As stated in Proposition 4.39, the infinite component of the c-thickened percolation has density at least $1 - 48\left(2c+1\right)^2 \varepsilon$.

Let's add a blank symbol $\square \notin \mathcal{A}$ to the original alphabet. Then, following the proof of Theorem 4.40, we can measurably map a noisy configuration (ω, b) onto a globally admissible configuration $\psi(\omega, b) \in X_{\mathcal{F}}$ but on the extended alphabet $\mathcal{A} \sqcup \{\square\}$, by writing the $\square$ on the thin grid around the periodic N-macro-tiles, such that almost surely, by finite union bound:

$$d_H\left(\omega, \psi(\omega, b)\right) \leq 96\left(2^{N+2}+1\right)^2 \varepsilon + \frac{2^{N+1}-1}{2^{2N}} .$$

The first term of the sum comes from the density of the thickened percolation, and the second one from the density of the thin grid made of $\square$ symbols $\psi(\omega, b)$, of the one-tile thick grid itself, which is equal to $1 - \left(\frac{2^N-1}{2^N}\right)^2$.

In order to conclude, we need to explain how to measurably project $\psi(\omega, b)$ back onto the original alphabet $\mathcal{A}$, how to fill-in the grid, so that we obtain an actual globally admissible enhanced Robinson tiling. To do so, we can simply consider some measure $\widetilde{\mu} \in \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$, take a configuration $y \in X_{\mathcal{F}}$ at random independently of the rest following $\widetilde{\mu}$, and then replace $\psi(\omega, b)$ by $\psi'(\omega, b, y)$ which is the unique translation of y by a vector

$k \in [\![0, 2^{N+1} - 1]\!]^2$ such that the N-macro-tiles of $\psi'(\omega, b, y)$ and $\psi(\omega, b)$ are aligned. This whole process is measurable, shift-invariant, and only changes the values of $\psi(\omega, b)$ on the $\square$ tiles which were already taken into account in the upper bound, so that the same bound holds for $d_H(\omega, \psi'(\omega, b, y))$.

Thence, we have a coupling such that $d_B(\mu, \mathcal{M}_\mathcal{F}) \leq 96\left(2^{N+2} + 1\right)^2 \varepsilon + \frac{1}{2^{N-1}}$, which proves the bound. $\square$

By taking N arbitrarily large, and then $\varepsilon \to 0$, we directly deduce the *stability* of our enhanced Robinson tiling for the Besicovitch distance. By optimising over N for a given value of ε, we will now conclude this analysis with an explicit non-linear upper bound on this speed.

Theorem 4.50 (Enhanced Robinson Stability)**.** *Let $X_\mathcal{F}$ be the enhanced Robinson tiling. Then $X_\mathcal{F}$ is f-stable with respect to d_B on the class of Bernoulli noises $\mathcal{B}$, with $f(\varepsilon) = 48\sqrt[3]{6\varepsilon}$ for small-enough values of ε. In particular, $X_\mathcal{F}$ is polynomially stable.*

Proof. To simplify things, we start by bounding $\left(2^{N+2}+1\right)^2 \leq 2^{2N+5}$, so that we are now trying to minimise $2\left(4^N \times 2^9 3\varepsilon + \frac{1}{2^N}\right)$ instead. If we denote $c(\varepsilon) = \sqrt[3]{2^9 3\varepsilon} = 8\sqrt[3]{3\varepsilon}$, then the upper-bound can now be rewritten as $2c\left(\left(2^N c\right)^2 + \frac{1}{2^N c}\right)$.

If we treat $x = 2^N c$ as a real-valued parameter, then $x^2 + \frac{1}{x}$ is minimal at $x = \sqrt[3]{\frac{1}{2}}$, and equal to $\frac{3}{\sqrt[3]{4}}$. This gives us a $24\sqrt[3]{6\varepsilon}$ bound. As N must be integer, we cannot have $N = \log_2\left(\frac{x}{c}\right)$, but by replacing it with the nearest integer (at distance at most $\frac{1}{2}$), we obtain the previous bound up to a factor $4^{\frac{1}{2}} = 2$, thus the announced bound. In order for this rounding argument to give a valid scale $N \geq 2$ in the previous proposition, we need $N(\varepsilon) \geq \frac{3}{2}$ to begin with, hence $\varepsilon \leq \frac{1}{49152\sqrt{2}}$. $\square$

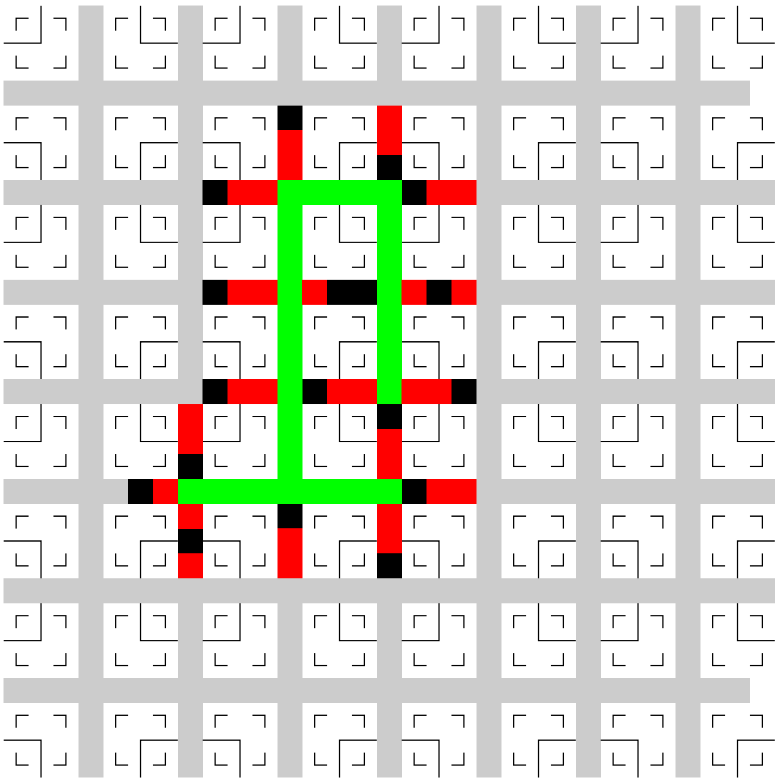

Figure 4.8: Bond percolation induced in the grid by obscured cells, with a finite connected component in green.

Remark 4.51 (Limits of the Notion of Stability)**.** We now have polynomial stability for the (aperiodic) enhanced Robinson tiling. One of the initial motivations for studying stability in the Besicovitch topology is that d_B accounts for the *global* structure of configurations, thus may be allow us to know whether the aperiodic structure of quasicrystals is preserved.

However, for the Robinson tilings here, this aperiodic structure is *not* preserved, regardless of how little there is, and I will now give a rogh idea of why it is.

Fix here a globally admissible grid of N-macro-tiles, with a one-tile thick empty grid around them, as was the case in Figure 4.7, and then add the ε-Bernoulli noise. We can see this empty grid as a bond percolation process on $\mathbb{Z}^2$, with the sites being the intersection points of the grid, and the bonds being the $\left(2^N-1\right)$-long lines and columns between neighbouring sites. Here, a bond is open *iff* all the cells inside are clear, with probability $(1-\varepsilon)^{2^N-1} \underset{N\to\infty}{\longrightarrow} 0$ (for any fixed choice of ε). When a bond between two sites is closed (*i.e.* there is at least one obscured tile in the line) there is no constraint forcing the corners to be well-oriented. This process is illustrated in Figure 4.8, with a grid of 2-macro-tiles: the obscured cells are represented in black, the corresponding broken bonds in red, which leaves the green connected component insulated from the outside. In particular, for a big-enough scale N, this percolation process has no infinite connected component. Each connected component can then be independently filled in a locally admissible way for the Robinson tileset. We thus obtain random noisy tilings arbitrarily close to $\mathcal{M}_\sigma(X_\mathcal{F})$, for any choice of $\varepsilon > 0$, but always such that the hierarchical structure is broken at big-enough scales of macro-tiles.

Hence, d_B fails to quantify how well the aperiodic structure of $X_\mathcal{F}$ is preserved, at least when this structure can be stored in areas with arbitrarily small density, which is the case of the Robinson tiling. ■

4.6.3 An Unstable Example

Let's conclude the section by studying the Red-Black Robinson tileset, which will provide us an interesting unstable example.

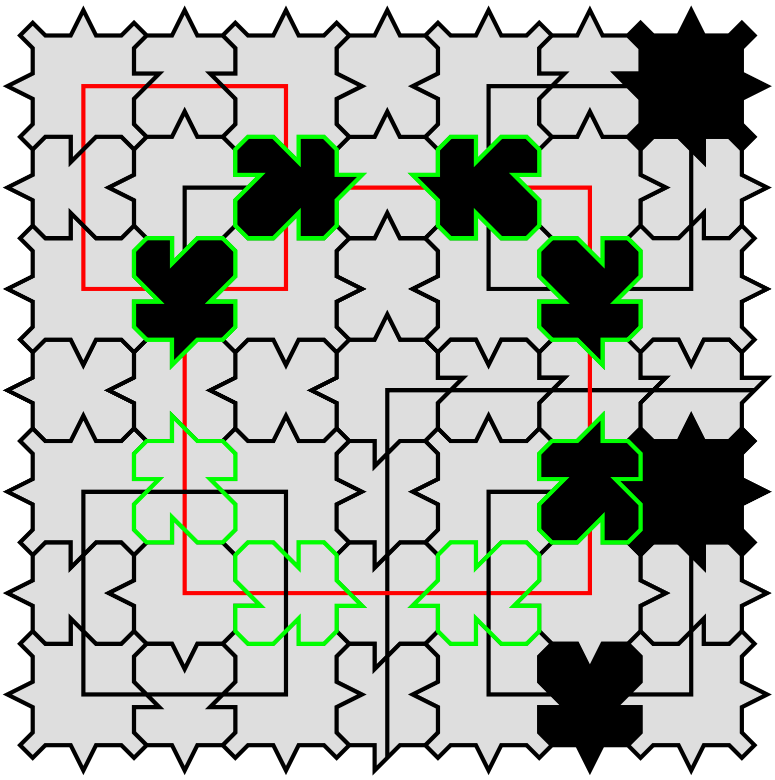

Figure 4.9: A locally admissible 3-macro-tile with obscured cells

Proposition 4.52. *Let $X_\mathcal{F}$ be the Red-Black Robinson tiling. For any $\varepsilon > 0$, there is $\mu \in \mathcal{M}^\mathcal{B}_\mathcal{F}(\varepsilon)$ such that $d_B\left(\mu, \mathcal{M}_\sigma\left(X_\mathcal{F}\right)\right) \geq \frac{1}{8}$. Thus, the SFT is unstable.*

Proof. The goal of this proof is to convert a generic tiling $\omega \in X_\mathcal{F}$ into a random noisy tiling $\lambda_{\omega,b}$ on $X_\mathcal{F}$, with b a random variable on $\Omega_{\{0,1\}}$. Using a generic Bernoulli noise b in the input, we will obtain a noisy tiling for which its bumpy-corners are now half Red and half Black, which will yield the announced result since bumpy-corners have frequency $\frac{1}{4}$ in the Robinson tiling.

I will build this measure λ iteratively, as a limit of locally-defined random transformations. At each step of the construction, the actual structure of the Robinson tiling will be left unchanged, and only the colours will be mismatched, so we may still talk about n-macro-tiles in this structural sense, even though they are not admissible. We initialise $\lambda_1 = \delta_{(\omega,b)}$ as a constant Dirac measure.

Let me now explain how we obtain λ_2 out of λ_1. This transformation will be done independently on *each* of the 2-macro-tiles of ω. We distinguish two cases, both illustrated in Figure 4.9, where the black cells c represent *obscured* tiles (where $b_c = 1$). A macro-tile is said to be *flippable* if both of its bi-coloured crosses, highlighted with green borders in the figure, are obscured tiles. In such a situation, we will flip its colours (Black lines become Red and conversely) with probability $\frac{1}{2}$, independently of the rest, which still preserves the local rules inside the macro-tile. In the figure, the top-left 2-macro-tile is flipped, the top-right 2-macro-tile is flippable but not flipped, and the two bottom 2-macro-tiles are not flippable.

Likewise, we go from λ_{n-1} to λ_n by flipping independently at random any flippable n-macro-tile (*except* the two ends of the central arm that are "after" the bi-coloured crossed tiles, which must match the colour of the yet-unflipped corresponding $(n+1)$-macro-tile). This process guarantees that, if we denote $\omega' \sim \lambda_n$ the new colouring, then $(\omega', b) \in X_{\widetilde{\mathcal{F}}}$ necessarily.

Notice how the highlighted cells that decide whether a given macro-tile is flippable are disjoint for each macro-tile and at each scale. Hence, assuming that $B \sim \mathcal{B}(\varepsilon)^{\mathbb{Z}^2}$ is a Bernoulli noise, each macro-tile at each scale is flippable with probability ε^2, independently of the rest. For any initial choice of $\omega \in X_{\mathcal{F}}$, the weak-$*$ limit measure $\lambda_{\omega,B} = \lim_n \lambda_n$ is well-defined, supported by $X_{\widetilde{\mathcal{F}}}$.

Notice how $\lambda_{\omega,B}$ is not shift-invariant (as the monochromatic Robinson structure of ω is preserved), but it "commutes" with σ in that $\sigma_*^k\left(\lambda_{\sigma^k(\omega),B}\right) = \lambda_{\omega,B}$. It follows that, if we average $\lambda_{\omega,B}$ over $\omega \sim \mu_0 \in \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$ (chosen independently from the noise B), we obtain a shift-invariant measure $\mu \in \mathcal{M}_{\mathcal{F}}^{\mathcal{B}}(\varepsilon)$.

Now, consider $G(\omega) \subseteq \mathbb{Z}^2$ the set of cells containing a bumpy corner in a generic configuration $\omega \in \Omega_{\mathcal{F}}$. For a given cell $c \in G(\omega)$, we denote by $\mathrm{flip}_{c,n}$ the random variable equal to 1 when the n-macro-tile containing c is flippable for the pair (ω, B). Hence the variables $\mathrm{flip}_{c,n} \sim \mathcal{B}\left(\varepsilon^2\right)$ are *iid*. Conditionally to the event $\mathrm{flip}_{c,n} = 1$, the colour of the cell c is uniformly distributed in λ_n after rank n. Thence, by Borel-Cantelli lemma, the colour of c is uniformly distributed in $\lambda_{\omega,B}$ (and thus μ).

Likewise, consider two distant cells $c, d \in G(\omega)$. As $d_\infty(c,d) \to \infty$, the smallest rank $n_0(c,d)$ such that c and d belong to the same n-macro-tile of ω goes to infinity. The families $\left(\mathrm{flip}_{c,n}\right)_{n<n_0}$ and $\left(\mathrm{flip}_{d,n}\right)_{n<n_0}$ are independent, and conditionally to the fact that both of these sequences contain at least a 1, the colours of cells c and d are independently uniform (in the measures λ_n after rank n_0, hence for $\lambda_{\omega,B}$ and μ).

Now, let $\omega \sim \mu$. Without loss of generality, assume $0 \in G(\omega)$, so that $G = (2\mathbb{Z})^2$. Then the family (colour of the cell $2c)_{c\in\mathbb{Z}^2}$ describes a shift-invariant ergodic dynamical system, so that we may apply a pointwise ergodic theorem. This implies that the *frequency* of both Black and Red bumpy-corners is generically equal to $\frac{1}{2}$ in μ. As G has density $\frac{1}{4}$ in $\mathbb{Z}^2$, we conclude that for almost-any $\omega' \sim \mu$ and any generic $\omega_0 \in \Omega_{\mathcal{F}}$ (with monochromatic bumpy-corners), we have the bound $d_H\left(\omega_0, \omega\right) \geq \frac{1}{2} \times \frac{1}{4} = \frac{1}{8}$ assuming bumpy-corners overlap between the two configurations, and even a $\frac{1}{2}$ bound if they are misaligned. In conclusion, μ satisfies $d_B\left(\mu, \mathcal{M}_\sigma\left(\Omega_{\mathcal{F}}\right)\right) \geq \frac{1}{8}$. □

Remark that this scheme of proof still applies almost verbatim as long as we can do an admissible colour flip process. Notably, if we implement the Red-Black structure in an (otherwise stable) enhanced Robinson tileset, then once again we obtain an unstable tiling. In later sections, we will implement this unstable flippable structure inside of simulating tilesets, by conditioning its emergence to the halting of the simulated Turing machine for example. This will allow us to study the computational complexity of the question of stability.

Before doing so, however, I will generalise the almost periodic scheme of proof into a black box theorem which will then be used in the following section.

4.7 Generalised Almost Periodic Stability

In order to prove the stable cases later on, I will here state a direct generalisation of Theorem 4.50 to a class of *almost periodic* SFTs. This framework will notably apply both to the already studied periodic SFTs from Section 4.4 and the enhanced Robinson tileset from Section 4.6, but also to the yet-to-be-introduced classes of simulating tilesets in the next section. This section is here mostly for the sake of technical completeness, and doesn't add any new idea to the proof of stability for the enhanced Robinson tileset.

Definition 4.53 (Almost Periodic SFT)**.** Let $X_{\mathcal{F}} \subseteq \Omega_{\mathcal{A}}$ an SFT, $p \in \mathbb{N}^*$ a period and $\rho > 0$ a density.

We say that $X_{\mathcal{F}}$ is ρ-almost p-periodic if there is a p-periodic set $G \subseteq \mathbb{Z}^d$ (invariant under translations in $(p\mathbb{Z})^d$) acting as a "grid" of density at most ρ, such that any configuration restricted to a translation of G^c is made periodic. By this, we mean that for any $\omega \in X_{\mathcal{F}}$, there is a unique translation of G (given by a non-necessarily unique $k \in [\![0, p-1]\!]^d$) such that ω_{G^c+k} is p-periodic.

In this case, assuming $\square \notin \mathcal{A}$, we can define $\omega^\square$ by overwriting ω_{G+k} by the blank symbol $\square$. Thence, $X_{\mathcal{F}}^\square = \left\{\omega^\square, \omega \in X_{\mathcal{F}}\right\}$ is a finite p-periodic SFT. ■

The non-uniqueness of k comes from the fact that we may have G a q-periodic grid instead (with q a factor of p), with some more redundancy in its structure, *e.g.* for the enhanced Robinson tiling, with 2^{N+1}-periodicity at the scale of N-macro-tiles once a 2^N-periodic grid is ignored.

By analogy with the enhanced Robinson tileset, we will want to obtain tilesets with a *sequence* of scales at which we have ρ_n-almost p_n-periodicity, with $\rho_n \to 0$. Notably, for periodic SFTs, we can directly have $\rho = 0$ without the need of a multi-scale argument. The next logical step is to translate the idea of reconstruction in this framework.

Definition 4.54 (C-Reconstruction Function)**.** Let $X_{\mathcal{F}}$ a ρ-almost p-periodic SFT, and G the associated grid.

The SFT has the C-reconstruction property if, for any locally admissible tiling ω of $B_{\lceil \frac{p}{2} \rceil + C}$ there is a unique translation of G such that $\left[\omega_{B_{\lceil \frac{p}{2} \rceil} \cap (G^c+k)}\right]^{\square}$ (obtained by filling $B_{\lceil \frac{p}{2} \rceil} \cap (G+k)$ with $\square$ symbols) is globally admissible in $X_{\mathcal{F}}^{\square}$ (thence $\omega_{B_{\lceil \frac{p}{2} \rceil} \cap (G^c+k)}$ is globally admissible in $X_{\mathcal{F}}$). What's more, the translation of G depends only on what appears in any p-square included in the central window $B_{\lceil \frac{p}{2} \rceil}$ (which is either a $(p+1)$ or $(p+2)$-square depending on the parity). ■

As $X_{\mathcal{F}}^{\square}$ is p-periodic, there is a unique choice $\omega^{\square} \in X_{\mathcal{F}}^{\square}$ of configuration that will match the pattern $\left[\omega_{B_{\lceil \frac{p}{2} \rceil} \cap (G^c+k)}\right]^{\square}$. Now, using this idea, Proposition 4.49 can be restated as:

Proposition 4.55 (Besicovitch Bound)**.** *Let $X_{\mathcal{F}}$ a ρ-almost p-periodic SFT with C-reconstruction. Then, for any $\varepsilon > 0$ and $\mu \in \mathcal{M}_{\mathcal{F}}^{\mathcal{B}}(\varepsilon)$, we have the bound $d_B(\mu, \mathcal{M}_\sigma(X_{\mathcal{F}})) \leq 48\left(2\left(C + \lceil \frac{p}{2} \rceil\right) + 1\right)^d \varepsilon + \rho$.*

Proof. The proof is naturally really similar to Proposition 4.49, so I will just underline the general ideas in the current formalism.

Consider $\lambda \in \widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{B}}}(\varepsilon)$ and $(\omega, b) \in X_{\widetilde{\mathcal{F}}}$ a λ-generic noisy configuration. The percolation argument from Proposition 4.39 tells us that, almost-surely, we can forget about the $\left(C + \lceil \frac{p}{2} \rceil\right)$-neighbourhood of obscured cells and still have a unique connected component of clear cells with density of at least $1 - 48\left(2\left(C + \lceil \frac{p}{2} \rceil\right) + 1\right)^d \varepsilon$.

Each clear cell c of this connected component is the center of a clear window I_c of diameter $2\lceil \frac{p}{2} \rceil + 1$, the C-neighbourhood of which is clear and locally admissible, so by the C-reconstruction property, there is a unique translation of G and a unique periodic configuration $\omega_c^{\square} \in X_{\mathcal{F}}^{\square}$ that matches ω on $I_c \cap (G+k)^c$ (for the right translation). We can do likewise for any other cell.

Now, two neighbouring cells $c, c' \in \mathbb{Z}^d$ share a common p-square window which fixes the same choice of translation for G. Hence, $\omega_c^{\square}$ and $\omega_{c'}^{\square}$ overlap on this p-square, and $X_{\mathcal{F}}^{\square}$ is p-periodic, so they are equal. Thus, all the cells of the infinite connected component $I(b)$ must agree on the same $\omega^{\square}$. The map $\varphi : (\omega, b) \mapsto \omega^{\square}$ is measurable (for ε small-enough, so that I has density greater than $\frac{1}{2}$).

In particular, ω and $\omega^{\square}$ can only differ outside of the component $I(b)$, or on the translation of the grid G, so that $d_H\left(\omega, \omega^{\square}\right) \leq \text{density}\left(I^c\right) + \text{density}(G) \leq 48\left(2\left(C + \lceil \frac{p}{2} \rceil\right) + 1\right)^d \varepsilon + \rho$, and the same bound holds for $d_B\left(\pi_*^1(\lambda), \varphi_*(\lambda)\right)$. At last, we can fill-in the $\square$ symbols of G in an appropriate random way, in order to send $\varphi^*(\lambda) \in \mathcal{M}_\sigma\left(X_{\mathcal{F}}^{\square}\right)$ into $\mathcal{M}_\sigma(X_{\mathcal{F}})$ without changing the bound on d_B. □

In particular, this proposition gives us a linear $O(\varepsilon)$ bound for f the speed of stability of any periodic SFT (which will be 0-almost periodic, with $G = \emptyset$ and C-reconstruction for some C).

Expectedly, the unstable Red-Black Robinson tiling doesn't fit in this framework, as we can have two ∞-macro-tiles with respectively Red and Black bumpy corners on both sides of a cut, which breaks both the almost periodicity globally and the reconstruction locally. To "stabilise" the Red-Black structure, we can use the dashed and dotted lines of the enhanced Robinson tileset to synchronise the Red-Black bit in *all four directions*. Naturally, this cancels out the flip argument, which would now create many forbidden patterns on the edges of the flipped macro-tile.

Corollary 4.56 (Stability)**.** *Assume there is a* sequence *of triplets (p_n, ρ_n, C_n) for which the previous proposition applies to $X_{\mathcal{F}}$, with $\rho_n \to 0$. Then $X_{\mathcal{F}}$ is a stable SFT.*

This will be sufficient to obtain stability when necessary. However, as I did for the enhanced Robinson tiling in Theorem 4.50, let me now provide another black box to transform the previous family of bounds for triplets (p_n, ρ_n, C_n) into a convergence speed $f(\varepsilon)$ for the stability.

Lemma 4.57 (Meta Multi-Scale-to-Polynomial Bound)**.** *Let $D_k = \varepsilon\alpha^k + \beta^k$ with $k \in \mathbb{Z}$ and $0 < \beta < 1 < \alpha$. Denote $\theta = \log_\alpha\left(\frac{1}{\beta}\right) = \frac{-\ln(\beta)}{\ln(\alpha)} > 0$. Then, for any choice $K \in \mathbb{Z}$, the following bound holds as long as $\varepsilon \leq \frac{\theta}{\alpha^{K(1+\theta)}}$:*

$$\min_{k \geq K} D_k \leq \max\left(\sqrt{\alpha}, \sqrt{\frac{1}{\beta}}\right) \times \left(\theta^{\frac{1}{1+\theta}} + (1/\theta)^{\frac{1}{1+1/\theta}}\right) \times \varepsilon^{\frac{\theta}{1+\theta}}.$$

Proof. Assume we have an optimal bound, but for a parameter $k \in \mathbb{R}$. Then, by replacing k with the nearest integer we will either increase the power of α by $\frac{1}{2}$ or decrease that of β by $\frac{1}{2}$, which gives us the leftmost factor in the announced upper bound. Note that this rounding is more reasonable under the assumption that $\alpha \approx \frac{1}{\beta}$. If one is much bigger than the other, we may simply decide on which side we always round k, with an added factor α or $\frac{1}{\beta}$ instead. In any case, this choice doesn't affect the order of magnitude of the speed.

Now, consider the parameter $x := \alpha^k \in \mathbb{R}^{+*}$. Thus, $k = \log_\alpha(x)$ so:

$$\beta^k = \exp\left(\frac{\ln(x)}{\ln(\alpha)} \times \ln(\beta)\right) = \exp\left(-\theta \ln(x)\right) = x^{-\theta}.$$

With this rewriting, $D(x) := \varepsilon x + x^{-\theta}$ is much easier to minimise. Indeed, D can be seen as a positive convex function that goes to $+\infty$ on 0^+ and $+\infty$, hence is minimised when $D'(x) := \varepsilon - \theta x^{-\theta-1} = 0$, thus at $x = \left(\frac{\theta}{\varepsilon}\right)^{\frac{1}{\theta+1}}$. Using this value of x in D directly gives us the rest of the expected bound.

Now, for the domain of validity, for us to be ably to round k properly, we simply require $k = \log_\alpha(x) \geq K$, which translates as $\varepsilon \leq \frac{\theta}{\alpha^{K(1+\theta)}}$. When the bound doesn't hold, when K is greater that the optimal value, the optimal choice is simply D_K. ☐

Corollary 4.58 (Polynomial Stability)**.** *Assume there is a* sequence *of triplets (p_n, ρ_n, C_n) for which the previous proposition applies to $X_\mathcal{F}$. If $C_n + p_n = O\left(\alpha^{\frac{n}{d}}\right)$ and $\rho_n = O\left(\beta^n\right)$, then using the previous lemma gives us a polynomial bound $O\left(\varepsilon^{\frac{\theta}{1+\theta}}\right)$ on the speed of convergence, with $\theta = \frac{-\ln(\beta)}{\ln(\alpha)}$.*

As I said, the enhanced Robinson tileset is $\frac{1}{2^N}$-almost 2^{N+1}-periodic at the scale of N-macro-tiles for any scale N. What's more, using Proposition 4.48, we also obtain 2^N-reconstruction at that same scale. Hence, this corollary applies with $(\alpha, \beta) = \left(4, \frac{1}{2}\right)$, so $\theta = \frac{1}{2}$ and $\frac{\theta}{1+\theta} = \frac{1}{3}$. In other words, we obtain a $O\left(\sqrt[3]{\varepsilon}\right)$ bound comparable to that of 4.50 (with a comparable multiplicative constant).

With this general stability result, we will now finally move onto the matter of decidability of the question of stability.

4.8 Undecidability of the Stability

I will now make full use both of the stable structure of the enhanced Robinson tileset and the unstable colour flipping argument of the Red-Black Robinson tileset to produce several classes of tilesets, where a computable reduction from a known decision problem (see Section 3.4) to stability occur.

In all the rest of this chapter, I will treat stability (with respect to the distance d_B and the class $\mathcal{B}$ of Bernoulli noises) as a *decision* problem, as the question *does $\mathcal{F}$ induce a (non-empty) stable SFT?* Notice that, as already discussed in Section 3.5, the domino problem (*i.e.* deciding whether $X_\mathcal{F} = \emptyset$) is itself undecidable, Π_1-complete. However, this (undecidable) non-emptiness assumption will not play a significant role when studying stability, which as we will see is a strictly harder problem. In any case, all the tilesets $\mathcal{F}$ in this section will be such that $X_\mathcal{F} \neq \emptyset$.

Matter-of-factly, proving the Π_2-hardness of stability would directly imply the weaker bounds I introduce first. However, the Π_2-hard construction relies on the Σ_1-hard one, and I believe the Π_1-hard one uses a complementary and more intuitive idea that will help get the point across.

4.8.1 Π_1-hard Construction

First, we will make use of the halting problem. As this problem is Σ_1-complete, we need to reduce it to the question of instability to prove that stability is Π_1-hard.

In other words, we want to simulate Turing machines in tilesets, as in Section 3.5, in a stable way, and make it so that if and when a machine terminates, an unstable structure emerges.

As already seen in Section 4.6, the enhanced tileset gives us a basic stable structure, such that N-macro-tiles locally align with their neighbours, and are mostly identical with each other (up to an arbitrarily low-density set as $N \to \infty$). Likewise, the Red-Black tileset uses a simple-to-adapt colour flipping argument to prove instability, a key part being that we end up with two (or more) groups of macro-tiles with a lower bound (that doesn't go to 0) on how similar the groups are (*e.g.* by grouping macro-tiles depending on the colour of their bumpy-corners, either Red or Black, with a frequency of differences of at least $\frac{1}{4}$ between the groups).

Description of the Tileset

Let me first describe the tileset used in this section. I will use an alphabet with two layers $\mathcal{A} \subseteq \mathcal{A}_R \times \mathcal{A}_M$, where $\mathcal{A}_R$ stands for the common Robinson layer, and $\mathcal{A}_M$ for the simulating layer specific to a given Turing machine M. Consequently, I will denote $X_{P_1(M)}$ the corresponding SFT.

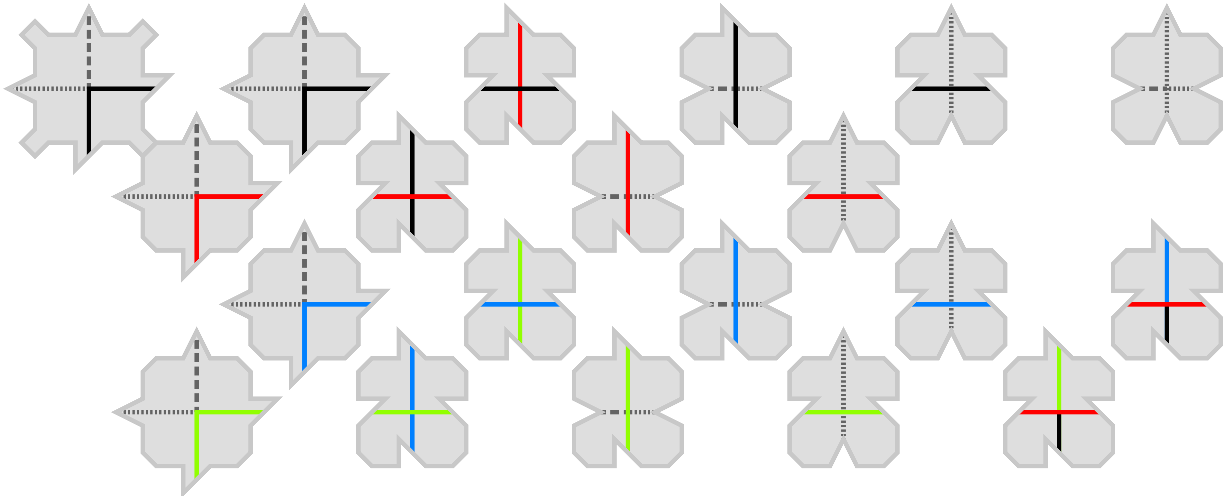

Figure 4.10: Main tiles of the alphabet $\mathcal{A}_R$.

Let's first describe the common layer $\mathcal{A}_R$. As we can see in Figure 4.10, the tileset uses four main colours, as well as grey dotted and dashed lines that enforce the locally enhanced structure from Section 2.7.2. Notice how bumpy-corners must be Black, after which we alternate between Black and Red, so this part of the structure is still stable. At some point, to-be-decided by the layer $\mathcal{A}_M$, we may transition from the Red-Black (stable) regime to the Blue-Green (unstable) regime using one of the two transition tiles on the bottom-right of Figure 4.10. The whole set $\mathcal{A}_R$ is given by all the rotations of the first three columns containing the tiles with a monochromatic V-shaped arm (but no symmetry, so that we may preserve the "chirality" of the macro-tiles encoded in their crosses) and rotations and symmetries of the rest, which brings us to a total of $|\mathcal{A}_R| = 172$ tiles.

The initial Red-Black structure, with Black bumpy-corners, allows us to simulate a Turing machine M on the layer $\mathcal{A}_M$ as explained in Section 3.5. More precisely, each Red square (of size 4^n+1 in a $(2n+1)$-macro-tile) will encode a space-time diagram (of size 2^n+1 at the n-th scale of simulations). Here, we initialise the machine M on the empty input ε.

If the computation $M(\varepsilon)$ terminates in the alloted time at a certain scale, then on the top border of the Red square it will "notice" the machine has reached a final state $q \in Q_A \sqcup Q_R$, and accordingly force a transition from the Red-Black to the Blue-Green regime. Notably, the next colour can be freely chosen as Blue or Green, without constraints except that both "arms" of the central cross use the same colour out of these two (which can be locally enforced by encoding a supplementary bit on Red lines in the layer $\mathcal{A}_M$ for example).

Note that a description of the machine M can be algorithmically converted into a description of the tileset $P_1(M)$ (along with its corresponding alphabet) in finite time, hence a computable reduction.

Theorem 4.59 (Stability is Π_1-hard)**.** *Consider a Turing machine M. Then the SFT $X_{P_1(M)}$ is stable* iff *M does* not *halt on the empty input.*

As the halting problem is Σ_1-complete, we deduce that the question of stability is Π_1-hard.

The following subsubsections will each focus on one of the implications, which put together directly give the previous result.

The Stable Case

For the stable case, assume that M doesn't halt on the empty input. Because of this, at any scale of admissible macro-tiles, the previously described transition from the Red-Black to the Blue-Green regime cannot occur, and the two last lines of tiles in Figure 4.10 may as well not exist in $\mathcal{A}_R$. My goal is to prove that the framework of Section 4.7 applies here.

Notice how we can project the current tileset to the enhanced Robinson tileset (directly at the scale of the alphabet) in a way compatible with the local rules. Hence, structurally, Propositions 4.47 and 4.48 both apply.

If we were only interested in the first coordinate $\mathcal{A}_R$, then we may stop here, erase the information *around* the $(2n+1)$-macro-tiles and obtain a 4^{n+1}-periodic behaviour, with (2×4^n)-reconstruction. However, here, we cannot overlook what happens on the layer $\mathcal{A}_M$. Of course, what happens *within* a given Red square is purely deterministic, but a sparse area of any N-macro-tile must stay available for the computations of higher macro-tiles.

If we extend the periodic grid from the one surrounding the $(2n+1)$-macro-tiles in a globally admissible configuration $\omega \in X_{P_1(M)}$ to G_n which now also includes any area of a $(2n+1)$-macro-tile outside of the Red squares it contains, and doesn't depend on the choice of ω up to translation. We then have the following lemma.

Lemma 4.60. *The SFT $X_{P_1(M)}$ is ρ_n-almost p_n-periodic with C_n-reconstruction, with $p_n = 4^{n+1}$, $C_n = 2 \times 4^n$ and ρ_n the density of the p_n-periodic set G_n defined above.*

Proof. As announced, the set G_n is 4^n-periodic (its period is half of that of macro-tiles, which must alternate between ◲ and ◰ on a given row and so on). As G_n includes the grid *around* $(2n+1)$-macro-tiles, Proposition 4.48 applies for the layer $\mathcal{A}_R$.

Regarding the alignment of G_n, it must correspond to that of the grid around $(2n+1)$-macro-tiles, hence it is fixed by the C_n-reconstruction property of the enhanced tileset.

Remark that, on the layer $\mathcal{A}_M$, because the Turing machine is deterministic, everything that happens on the *inside* of a given admissible Red square is fixed, insulated from outside interference. Hence, on this layer (and using of course the alignment of G_n given by the layer $\mathcal{A}_R$) we obtain a p_n-periodic behaviour outside of G_n on the layer $\mathcal{A}_M$ too. □

In order to conclude, we need to compute ρ_n the density of G_n.

Lemma 4.61. *In a $(2n+1)$-macro-tile, we have $O\left(12^n\right)$ tiles outside of the Red squares.*

Proof. The general idea of the proof is that Red squares form a kind of Sierpiński carpet inside macro-tiles.

Denote r_n the number of tiles *inside* the Red squares in a $(2n+1)$-macro-tile. As we can see on Figure 2.8, in the process of forming a $(2n+3)$-macro-tile, we will create a big central square around four $(2n+1)$-macro-tiles, surrounded by twelve $(2n+1)$-macro-tiles. As we already know the size of this big square, we obtain the following recurrence:

$$r_{n+1} = 12r_n + \left(4^{n+1}+1\right)^2 \geq 12r_n + 16^{n+1}\,.$$

As $r_1 = 25 \geq 16$, we obtain by induction $r_n \geq 4^{n+1}\left(4^n - 3^n\right)$. At the same time, a $(2n+1)$-macro-tile has $\left(2^{2n+1}-1\right)^2 \leq 4^{2n+1}$ tiles, so *at most* $4^{n+1}3^n = 4 \times 12^n$ tiles in total outside the Red squares. □

Hence, as $(2n+1)$-macro-tiles use $\Theta\left(16^n\right)$ tiles in total, we conclude that G_n has density $\rho_n = O\left(\left(\frac{3}{4}\right)^n\right)$.

Proposition 4.62. *If M doesn't halt on the empty input, then $X_{P_1(M)}$ is polynomially stable, with convergence speed $O\left(\varepsilon^r\right)$ at rate $r = \frac{2-\log_2(3)}{6-\log_2(3)} \approx 0.094$.*

Proof. We apply Corollary 4.58 with constants $\alpha = 4$ and $\beta = \frac{3}{4}$, so $r = \frac{\theta}{1+\theta}$ gives the announced rate. □

The Unstable Case

Proposition 4.63. *Assume M halts on the empty input. Then for any $\varepsilon > 0$ we have a measure $\mu \in \mathcal{M}^{\mathcal{B}}_{P_1(M)}(\varepsilon)$ such that $d_B\left(\mu, \mathcal{M}_\sigma\left(X_{P_1(M)}\right)\right) \geq \frac{1}{4^{n+1}}$, where n is the last scale of simulation, at which M halts.*

Proof. Consider $N = 2(n+1)$ the first scale at which N-macro-tiles have a big Blue or Green square in the middle. Assuming two aligned N-macro-tiles don't use the same colour for the square (of diameter $d = 2^{N-1}+1$ tiles), then we obtain at least $p = 4 \times (d-1) = 2^{N+1}$ differences.

By following the very same colour-flipping process as in Proposition 4.52, but on the Blue-Green bit starting at the scale of N-macro-tiles, we obtain a generic colour-flipped configuration ω (with monochromatic Blue or Green squares in the N-macro-tiles).

Thus, for any generic $\omega \in X_{P_1(M)}$ that aligns with ω' up to the scale of N-macro-tiles, we obtain a lower bound $d_H(\omega, \omega') \geq \frac{1}{2} \times \frac{p}{4^N} = \frac{1}{2^N} = \frac{1}{4^{n+1}}$, with the factor $\frac{1}{2}$ coming from the frequency of Blue and Green big squares in ω', whereas *all* such squares of ω must be of the same colour.

Now, assume that N-macro-tiles in ω and ω' don't align well. By choosing the best pairing of N-macro-tiles between ω and ω', we still have a rectangle with both sides of length at least $2^{N-1} - 1$ (*i.e.* the size of a $(N-1)$-macro-tile) where the N-macro-tiles of both tilings overlap. In this area, both macro-tiles have a Blue or Green corner of their big square, made of at least $2 \times \left(2^{N-1} + 1\right)$ tiles. As these two corners intersect in at most 2 tiles, and the rest of the area is guaranteed to use only Black or Red communication channels, we have at least 2^N differences between ω and ω' in this window. As this process repeats 2^N periodically in both directions, without even having to take the colour-flipping into account, we still obtain a lower bound $d_H(\omega, \omega') \geq \frac{1}{2^N} = \frac{1}{4^{n+1}}$. □

Remark 4.64. More generally, as long as we can guarantee *one* difference between the two kinds of macro-tiles which we colour-flip, we obtain a lower bound on d_B of order $\frac{1}{\text{macro-tile area}}$. We will directly invoke this "obvious" lower bound for further unstable cases, to avoid tedious computations that don't bring any more understanding.

Still, the order of magnitude $\frac{1}{\text{macro-tile diameter}}$ obtained in the previous proposition is the best one can reasonably hope for in general, as a signal that transits through a macro-tile will typically only cross a number of tiles proportional to the diameter, normalised by the tile area. ■

4.8.2 Σ_1-hard Construction

We can "flip around" the previous construction, by adding an *unstable* information atop of the structure simulating the Turing machine, in such a way that the information gets frozen and becomes *stable* if the machine halts. We will first describe the construction of $S_1(M)$ out of a machine M, and then state the corresponding undecidability result.

In the previous tileset $P_1(M)$, the Robinson layer $\mathcal{A}_R$ used one communication channel with four different colours. Here, for $S_1(M)$, we use *two* communication channels in the lines of the Robinson structure, each one having two possible values. First, the Red-Black channel must be initialised as Black in bumpy corners, and then alternate, in order to have the right structure to simulate the machine M. Second, the Blue-Green channel can be freely initialised. However, if M halts at a given scale of simulation, then the border of the Red square *must* be Blue on the other channel, which we call a *freeze*. Note that here, we can keep simulating M at higher scales after it halts for the first time, as subsequent freezes will just occur at scales of macro-tiles where the Blue-Green channel would be frozen into Blue anyway.

Proposition 4.65. *The SFT $X_{S_1(M)}$ is stable* iff *M halts on the empty input. Thus, stability is Σ_1-hard.*

Proof. First, assume that M doesn't halt. Then we can freely do a colour-flipping process starting from any $\mu \in \mathcal{M}_\sigma\left(X_{S_1(M)}\right)$, just like in Proposition 4.52. We can start flipping the Blue-Green channel at the scale of bumpy corners, hence instability with a $\frac{1}{8}$ lower bound on d_B.

Now, assume M halts. Then, in any tiling $\omega \in X_{S_1(M)}$, the Blue-Green channel is retroactively frozen all the way down to the Green bumpy-corners. By using the same grid G_n as in Lemma 4.60, we can likewise ignore everything that happens outside of Red squares, and control everything inside, hence a ρ_n-almost p_n-periodic tiling with $p_n = O(4^n)$ and $\rho_n = O\left(\left(\frac{3}{4}\right)^n\right)$.

Finally, denote n_{halt} the first scale of simulation at which M halts in $S_1(M)$. If we try to reconstruct things locally at scales lower than n_{halt}, then we will reach a family of well-aligned and well-oriented $(2n+1)$-macro-tiles, but without any freezing happening in the tiles, hence this Blue-Green channel that may not behave in a globally admissible way, all the way down to the high-density set of bumpy-corners. Still, as long as $n \geq n_{\text{halt}}$, the freezing prevents this from happening, and using the same $C_n = O(4^n)$ as in Lemma 4.60, we conclude that this scale of the tiling has indeed C_n-reconstruction.

Lastly, starting at high-enough scales, for low-enough values of ε, the proof of Proposition 4.62 applies *verbatim*, so we have stability with a polynomial $O(\varepsilon^r)$ convergence rate. □

4.8.3 Π_2-hard Construction

In the construction for $S_1(M)$, we obtained stability *iff* there exists a time step such that M halts *on the empty input*. Consequently, if we manage to twist the construction to include *any possible input*, then we may equate stability with the Π_2-complete totality problem.

There are several ways to proceed, but I choose here to use the method of Toeplitz encoding of the input, because it is quite versatile, and may more generally be able to convert a (structurally close to) uniquely ergodic SFT encoding a Σ_k-hard problem into a (definitely not uniquely ergodic anymore) SFT encoding a Π_{k+1}-hard problem.

Toeplitz Input

The Toeplitz encoding of an infinite sequence $u \in \Gamma^{\mathbb{N}^*}$ on an alphabet Γ consists of inductively filling with u_n half of the holes still free after the previous iterations, which gives a sequence $u_1 * u_1 * u_1 * u_1 * \dots$, then $u_1u_2u_1 * u_1u_2u_1 * \dots$, and so on. Toeplitz sequences have been studied as dynamical systems for a long while [JK69], and have since been encoded in higher-dimensional SFTs [CV21].

The idea of the method is to sequentially write the wanted input u into the consecutive scales hierarchical structure, which will appear as a Toeplitz encoding $u_1u_2u_1\dots$ from the point of view of the simulated Turing machine, and then adapt the machine to decode it back into its original form u at first. This method was already used by Barbieri and Sablik [BS19] in particular.

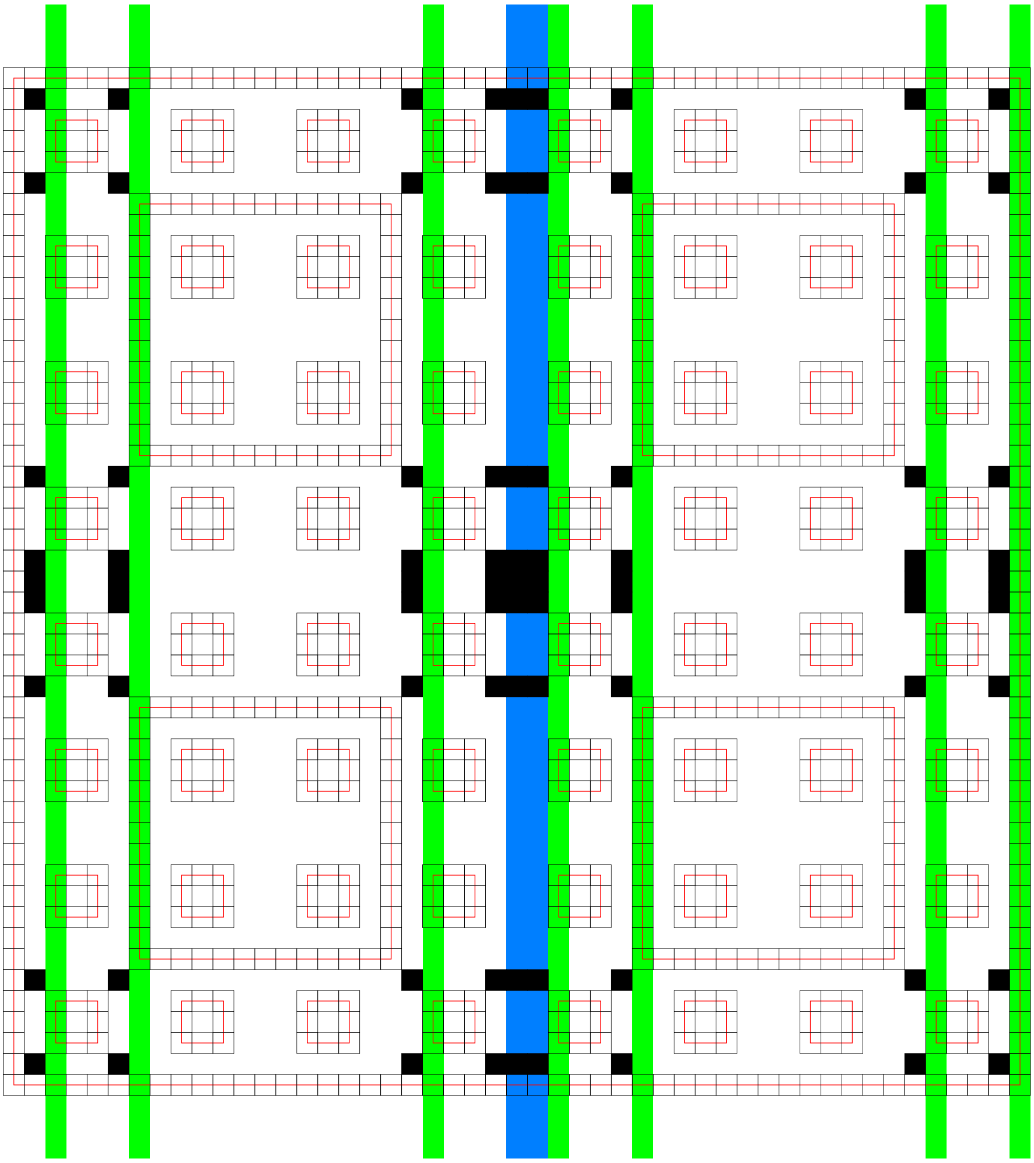

Figure 4.11: Structure of the read-only input tape, with read-only values stored in the highlighted columns, whereas the machine operates within the black patches.

More precisely, I build the tileset $P_2(M)$ as follows. For the Robinson structure, use the same parallel Red-Black and Blue-Green bits as for $S_1(M)$. Then, add another channel that can take values in the alphabet $\Sigma \sqcup \{\#, \$^\Sigma, \$^\#\}$ where Σ is the input alphabet of the machine M, $\#$ the blank tape symbol, and the $\$^*$ symbols two supplementary letters (not in the tape alphabet of M). On Black channels, we can freely use any symbol $\$^*$ following a letter from Σ on the previous Red scale, but we *must* use $\$^\#$ following $\#$. On Red channels, we *must* use a letter from Σ following a $\$^\Sigma$ symbol on the previous Black scale, and use $\#$ following $\$^\#$. If we look only at the Red channels, this gives an infinite word $u \in \Sigma^*\#^\mathbb{N} \sqcup \Sigma^\mathbb{N}$. When $u \in \Sigma^*\#^\mathbb{N}$, we will identify it with its prefix in Σ^*, followed by $\#^\mathbb{N}$.

Quite importantly, the choice of a letter is not only communicated along the regular Red-Black channels in two directions (following the arms in the central cross), but also along the other two dashed and dotted alignment channels of the enhanced Robinson self-aligning structure. Thus, any two (well-aligned) neighbouring N-macro-tiles must encode the same sequence.

On the simulation layer, the Turing machine is able to read which symbol is written down in the column directly on the right of its current position. Hence, from the point of view of the Turing machine simulated in a Red square, this represents a *read-only* second tape. In order to adequately use u as an input, we first need to explain what the machine sees.

Lemma 4.66 (Toeplitz Encoding of the Input)**.** *Let $u \in \Sigma^*\#^\mathbb{N} \sqcup \Sigma^\mathbb{N}$. Define the words $w_n = w_{n-1}u_n w_{n-1}$ by induction, initialised with the empty word w_0. The word w_n is a prefix of the Toeplitz encoding of the whole sequence u.*

At the n-th scale of simulation, from the point of view of the Turing machine, the read-only tape reads as $w_{n-1}\$^*\$^* w_{n-1}u_n$.

Proof. The last letter of the read-only tape obviously corresponds to the right border of the n-th Red square, hence reads as u_n. The central $\$^*$ symbols come from the fact, as highlighted by the blue columns in Figure 4.11, they correspond to the $(n+1)$-th scale for Black squares followed by the first scale of bumpy corners.

The rest of the word, that reads as w_{n-1} on both sides of these central columns, can be explained by the inductive construction of macro-tiles. Indeed, each quarter of the n-th Red square is actually a whole $(2n-1)$-macro-tile with a central Red square, and the Red squares are themselves stacked in a Toeplitz way within the macro-tile, with a gap in-between each that allows to read the letter on them. □

From Decoding the Input to Computations

Let's now explain what Turing machine is encoded into $P_2(M)$, and how it affects the Blue-Green channel.

First, the machine will have to *decode* the Toeplitz input, while keeping the Blue-Green channel *stable* (*e.g.* by using a third non-alternating colour). More precisely, the machine will step by step read the letters at positions 2^k on the read-only tape and write them one after another at the beginning of its working tape. This process will decode the Toeplitz encoding w_n back into the sequence $u_1 \dots u_n$. Using a unary counter, which we multiply by two after reading each letter, reaching the k-th letter will require about $\Theta\left(4^k\right)$ steps of computation.

Now, this process can halt in two ways. First, we read a $\$^*$ symbol, meaning that we reached halfway through the read-only tape. In this case, the machine simply idles for the rest of its finite runtime, without unfreezing the Blue-Green bit when it reaches the top border of the Red square. This won't happen at big-enough scales of simulation, considering it would take about $\Theta\left(4^n\right)$ steps but the n-th machine only has a finite horizon of 2^n steps, but it can occur at the initial scales of simulation and in particular at the very first one where the first symbol is $\$^*$. Second, we read a $\#$ symbol before reaching the central $\$^*$, in which case the decoding of the word $u \in \Sigma^*$ is complete. Without waiting, the machine starts simulating M on u (this will occur roughly at the $2|u|$-th scale of simulation). This will signal the Red square to *ignite* the unstable Blue-Green bit (if it was not already done at a lower scale), as was the case for $P_1(M)$ in Subsection 4.8.1. Then, if M halts on u, this will signal the Red square to *freeze* the Blue-Green bit, as was the case for $S_1(M)$ in Subsection 4.8.2.

Undecidability of the Stability

Lemma 4.67. *Assume M doesn't have a total support, and in particular doesn't halt on the input $u \in \Sigma^*$. Consider $\mu_u \in \mathcal{M}_\mathcal{F}$ an invariant measure with $u\#^\mathbb{N}$ written in the Red scales of any generic configuration. Then, by colour-flipping the Blue-Green channel after the initial decoding scales, we obtain the measure $\mu_u^\varepsilon \in \mathcal{M}_\mathcal{F}^\mathcal{B}(\varepsilon)$, such that $\inf_{\varepsilon>0} d_B\left(\mu_u^\varepsilon, \mathcal{M}_\sigma\left(X_\mathcal{F}\right)\right) > 0$.*

Proof. As in Proposition 4.63, if we compare two macro-tiles with a Blue or Green square, corresponding to the same input $u\#^{\mathbb{N}}$, we can obtain a lower bound on their density of mismatched tiles purely through the Blue-Green square.

Likewise, if we compare such a macro-tile with a macro-tile corresponding to *another* input, then they must differ in one of the first $|u|+1$ letters in a Red channel. In this case, we can also obtain a lower bound independent of ε, even if they are perfectly aligned, using this mismatching letter in the input. □

Proposition 4.68. *Denote $\varphi(n)$ the first* scale of simulation *at which, for any input $u \in \Sigma^{\leq n}$, both the* decoding *phase and the* computation *phase are over. Assume M has total support, so that $\varphi(n) < \infty$. Then, using the notations of Lemma 4.60, the SFT $X_{P_2(M)}$ is ρ_n-almost p_n-periodic with $C_{\varphi(n)}$-reconstruction.*

Proof. If we follow the same scheme of proof as in the lemma, then we almost obtain a ρ_n-almost p_n-periodic SFT with C_n-reconstruction up to *one* detail.

Here, the Red-Black channel and all the computations in Red squares behave deterministically, so they are fixed for a given input (which synchronises between neighbouring tiles), but the Blue-Green channel is not. However, if we exploit the $C_{\varphi(n)}$-reconstruction of the Robinson structure, then either:

- a given Red square isn't done decoding its input, so the Blue-Green bit is still frozen, uniquely determined,
- a given Red square has decoded its input u, at a scale of simulation lower than n, which implies $|u| < n$, but the Red square actually fits into a bigger $(2\varphi(n)+1)$-macro-tile, which will terminate its simulation of M on u, thus freeze the Blue-Green bit of this Red square.

In both cases, we indeed guaranteed that the area inside Red squares is globally admissible, hence the n-th scale of simulation admits $C_{\varphi(n)}$-reconstruction. □

In particular, $\rho_n \to 0$ so $X_{P_2(M)}$ is stable, according to Corollary 4.56. However, because φ can be as big as any computable function (literally, by computing $n = |u|$, computing $f(n)$ in unary for some computable function f, which necessarily takes at least $f(n)$ steps, and only then doing the initially planned computation $M(u)$), we can't expect any explicit bound on $C_{\varphi(n)}$ to apply Corollary 4.58, and will not obtain polynomial stability this time. The next theorem directly follows:

Theorem 4.69. *Let M a Turing machine. The machine M has total support* iff *$X_{P_2(M)}$ is stable. As the totality problem is Π_2-complete, we deduce that stability is Π_2-hard.*

In some sense, the Toeplitz encoding of inputs allowed us to transform the initial Σ_1-hard construction into a Π_2-hard one. However, this process doesn't adapt to translate the Π_1-hard construction into a Σ_2-hard one. In order to do this, we would need an added universal quantifier, which cannot work if we encode only *one* input at a time in a ground configuration. Hence, in this case we would need to enumerate the inputs inside the tiling in any case.

Remark 4.70 (Alternate Construction for Π_2-Hardness)**.** Let me conclude this section by briefly describing another computable reduction from the totality problem to stability, this time without having to encode any input.

The main idea is here to stack ignition-freezing blocks onto each other. In the tiling $P_2'(M)$, we enumerate the words of Σ^*, *e.g.* following a lexicographical order biased by increasing lengths. After enumerating a new word u we simulate M on it. Once this simulation ends, we both freeze the lower scales of the Blue-Green bit *and* ignite an independent Blue-Green bit for higher scales. Then, we enumerate the next word, rinse and repeat.

If M doesn't have a total support, then $P_2'(M)$ will have finitely many ignition-freezing blocks, and the last (infinite) one never freezes its Blue-Green bit, which we will be able to colour-flip. Conversely, if M has total support, then at any given scale of simulation, there exists a higher scale of simulation at which M terminates on *some* word, which will guarantee the current Blue-Green bit is frozen. ■

Remark 4.71 (Extension to Higher Dimensions)**.** In Section 4.3, we established that in the one-dimensional case, stability is decidable. In this section, we just established that in the two-dimensional case, stability is undecidable, and Π_2-hard. In higher dimensions $d > 2$, using Corollary 4.32, we can simply use an independent copy of $X_{P_2(M)}$ for each coordinate of $\mathbb{Z}^{d-2}$, which gives a d-dimensional SFT that is stable *iff* $X_{P_2(M)}$ is. Thus, we have a computable reduction from totality to stability for *any* fixed dimension $d \geq 2$. ■

This concludes this section. Now that we have a *lower* bound on the complexity of stability, let's move onto an altogether different approach, using arguments of computable analysis, to obtain an *upper* bound on the complexity.

4.9 Upper Bound on the Complexity of Stability

As announced, I will now need to dig deeper into the framework of computable analysis on invariant measures to describe how much computation power is actually needed to decide our notion of stability.

For the rest of this section, I will consider an alphabet $\mathcal{A}$ and a set of forbidden patterns $\mathcal{F}$, without assuming anything anything more (notably, we still don't know if $X_{\mathcal{F}} \neq \emptyset$). From there, my goal in this section is to explain a process to conclude on whether $X_{\mathcal{F}}$ is stable or not.

Even though the goal is to study convergence for the Besicovitch distance, as discussed in Section 3.6, and Remark 3.27 in particular, the natural framework to describe (invariant) measures to a computer is the weak-$*$ topology. Notably, the "periodic measures" (*i.e.* the uniform measures on the finite orbits of the periodic configurations) are a natural choice of countable dense basis in this topology.

Before doing anything substantial, I will need to introduce some supplementary framework, and in particular the (family of) computable weak-$*$ distances that will allow us to study distances in various spaces of measures (and notably couplings) in a way compatible with projections.

Once this preliminary work is done, we will see how the measure sets $\mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$, $\mathcal{M}_{\mathcal{F}}^{\mathcal{B}}(\varepsilon)$ and $J\left(\mu, \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right)$ (*i.e.* the set of couplings between μ and $\mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$) can be described in the framework of computable analysis.

At last, I will use these sets to prove a Π_4 upper bound on the stability problem.

4.9.1 Complements of Computable Analysis on Invariant Measures

Definition 4.72. Let $\mathcal{A}$ a finite alphabet. The weak-$*$ topology on $\mathcal{M}_\sigma\left(\Omega_{\mathcal{A}}\right)$ can be induced by the following distances:

$$d_r^+(\mu,\nu) := \sum_{n\in\mathbb{N}} \frac{1}{2^n} \times \frac{1}{r^{|I_n|}} \times \frac{1}{|\mathcal{A}^{I_n}|} \sum_{w\in\mathcal{A}^{I_n}} |\mu([w]) - \nu([w])| \ ,$$

with $r \geq 1$ a normalisation factor (with the convention $d^+ = d_1^+$). ■

Note how, if $V_1 \subseteq \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}_1}\right)$ and $V_2 \subseteq \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}_2}\right)$ are both closed, then so is the set of their couplings $J\left(V_1, V_2\right) \subseteq \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}_1\times\mathcal{A}_2}\right)$ (in their respective weak-$*$ topologies).

The interest of having a whole family of distances (instead of just d^* using the same formula regardless of the space) is that we can use them to relate distances between measures and couplings:

Lemma 4.73 (Projection Lemma)**.** *Consider two measures* $\lambda, \lambda' \in \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}_1\times\mathcal{A}_2}\right)$. *Then:*

$$d_{|\mathcal{A}_2|}^+\left(\pi_*^1(\lambda), \pi_*^1\left(\lambda'\right)\right) \leq d^+\left(\lambda,\lambda'\right) \ .$$

Proof. We have:

$$\begin{aligned}
& d_{|A_2|}^+\left(\pi_*^1(\lambda), \pi_*^1\left(\lambda'\right)\right) \\
=& \sum_{n\in\mathbb{N}} \frac{1}{2^n} \times \frac{1}{\left|\mathcal{A}_2^{I_n}\right|} \times \frac{1}{\left|\mathcal{A}_1^{I_n}\right|} \sum_{w_1\in\mathcal{A}_1^{I_n}} \left|\pi_*^1(\lambda)\left([w_1]\right) - \pi_*^1\left(\lambda'\right)\left([w_1]\right)\right| \\
=& \sum_{n\in\mathbb{N}} \frac{1}{2^n} \times \frac{1}{\left|\mathcal{A}_2^{I_n}\right|} \times \frac{1}{\left|\mathcal{A}_1^{I_n}\right|} \sum_{w_1\in\mathcal{A}_1^{I_n}} \left| \sum_{w_2\in\mathcal{A}_2^{I_n}} \lambda\left([w_1,w_2]\right) - \lambda'\left([w_1,w_2]\right) \right| \\
\leq& \sum_{n\in\mathbb{N}} \frac{1}{2^n} \times \frac{1}{\left|\left(\mathcal{A}_1\times\mathcal{A}_2\right)^{I_n}\right|} \sum_{(w_1,w_2)\in(\mathcal{A}_1\times\mathcal{A}_2)^{I_n}} \left|\lambda\left([w_1,w_2]\right) - \lambda'\left([w_1,w_2]\right)\right| \\
=& d^+\left(\lambda,\lambda'\right) \ ,
\end{aligned}$$

i.e. the announced bound. □

We define $\overline{\mathrm{B}}(\mu,\varepsilon) = \left\{\nu \in \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}}\right), d^+(\mu,\nu) \leq \varepsilon\right\}$ the closed ball around $\mu \in \mathcal{M}_{\mathcal{A}}$ of radius $\varepsilon > 0$, and extend this notation to denote the ε-neighbourhood of any closed set of measures.

Let's now try to compute the compact sets we will need in order to study stability (*i.e.* $\mathcal{M}_\sigma(X_\mathcal{F})$ etc.), using the countable dense basis of finite sets of periodic points. The first set I will thus compute is the space $\mathcal{M}_\sigma(\Omega_\mathcal{A})$ itself.

Lemma 4.74 (Covering Lemma for $\mathcal{M}_\sigma(\Omega_\mathcal{A})$)**.** *There is $\psi : \mathbb{Q}^{+*} \times \mathbb{N}^2 \to \mathbb{N}$ a computable map such that, for any finite alphabet $\mathcal{A}$, any dimension d and any rational $\delta > 0$:*

$$\mathcal{M}_\sigma(\Omega_\mathcal{A}) = \bigcup_{w \in \mathcal{A}^{I_{\psi(\delta,|\mathcal{A}|,d)}}} \overline{\mathrm{B}}\left(\widehat{\delta_w}, \delta\right).$$

Proof. Denote s_n the partial sum up to rank n associated to the distance d^+. We can bound $r_n = d^+ - s_n \leq \frac{1}{2^{n-1}}$ independently of the dimension d, of $\mathcal{A}$ and of any pair of measures. Hence, for a given value of δ, we first compute $n(\delta) = 2 + \left\lceil \log_2\left(\frac{1}{\delta}\right)\right\rceil$, such that $r_n \leq \frac{\delta}{2}$. We now want to cover $\mathcal{M}_\sigma(\Omega_\mathcal{A})$ with balls of radius $\frac{\delta}{2}$ for the pseudo-distance s_n.

Notice that for any $k \leq n$ and any pattern $w \in \mathcal{A}^{I_k}$, we have $\mu([w]) = \sum_{v \in \mathcal{A}^{I_n}, v_{I_k} = w} \mu([v])$. It follows that:

$$\begin{aligned} s_n(\mu,\nu) &= \sum_{k \leq n} \frac{1}{2^k |\mathcal{A}^{I_k}|} \sum_{w \in \mathcal{A}^{I_k}} |\mu([w]) - \nu([w])| \\ &\leq \sum_{k \leq n} \frac{1}{2^k} \sum_{w \in \mathcal{A}^{I_k}} |\mu([w]) - \nu([w])| \\ &\leq \sum_{k \leq n} \frac{1}{2^k} \sum_{w \in \mathcal{A}^{I_n}} |\mu([w]) - \nu([w])| \\ &\leq 2 \sum_{w \in \mathcal{A}^{I_n}} |\mu([w]) - \nu([w])| \\ &\leq 2 \left|\mathcal{A}^{I_n}\right| \sup_{w \in \mathcal{A}^{I_n}} |\mu([w]) - \nu([w])|. \end{aligned}$$

Hence, we now need to uniformly approximate any $\mu \in \mathcal{M}_\sigma(\Omega_\mathcal{A})$ on the window I_n by a periodic measure to conclude.

To do so, consider μ_m the restriction of μ to I_m. We identify μ_m, a measure on $\mathcal{A}^{I_m}$, with the measure on $\Omega_\mathcal{A}$ that charges a periodic configuration $w^{\mathbb{Z}^d}$ (with $w \in \mathcal{A}^{I_m}$) with probability $\mu([w]) = \mu_m(\{w\})$. Remark that μ_m is not shift-invariant, but is $(m\mathbb{Z})^d$-periodic under translation, so we can define the corresponding averaged measure $\widehat{\mu_m} := \sum_{w \in \mathcal{A}^{I_m}} \mu_m(\{w\}) \times \widehat{\delta_w}$ which is shift-invariant.

In particular for any $w \in \mathcal{A}^{I_n}$, as long as $k + I_n \subseteq I_m$ (*i.e.* $k \in I_{m-n}$), then $[w]_k := \sigma^k([w])$ is still a cylinder defined inside I_m. Hence, for any such translation we have $\mu_m([w]_k) = \mu([w]_k) = \mu([w])$. Now, we have:

$$\widehat{\mu_m}([w]) := \frac{1}{|I_m|} \sum_{k \in I_m} \mu_m([w]_k) = \frac{|I_{m-n}|}{|I_m|} \mu([k]) + \frac{1}{|I_m|} \sum_{k \in I_m \setminus I_{m-n}} \mu_m([w]_k),$$

hence $\widehat{\mu_m}([w]) = \mu([w]) + O\left(\frac{n}{m}\right)$, where the computable domination constant depends on d. Now, if we use instead μ_m^k a dyadic approximation of μ_m on $\mathcal{A}^{I_m}$, with precision $\frac{1}{2^k}$, we obtain a measure $\widehat{\mu_m^k}$ for which:

$$\widehat{\mu_m^k}([w]) = \mu([w]) + O\left(\frac{n}{m}\right) + O\left(\frac{|\mathcal{A}^{I_m}|}{2^k}\right).$$

This new term simply uses the domination constant 1. Remark in particular that there is only a finite amount of such dyadic measures on the window I_m with precision $\frac{1}{2^k}$. We just need to be able to approximate these by periodic measures to conclude.

We can decompose any such dyadic measure as $\widehat{\mu_m^k} = \frac{1}{2^k} \sum_{w \in \mathcal{A}^{I_m}} p(w) \widehat{\delta_w}$ with weights $p(w) \in \mathbb{N}$ that sum to 2^k. Consider now $M = (m+1) \times 2^k - 1$. On the corresponding window I_M, we can fit a total of 2^k slices, each made of windows I_m stacked in all directions but one. In $p(w)$ such consecutive slices, we write w in each box I_m. This gives us a configuration $\overline{w} \in \mathcal{A}^{I_M}$ such that, for any $w \in \mathcal{A}^{I_n}$, we have $\widehat{\delta_{\overline{w}}}([w]) = \widehat{\mu_m^k}([w]) + O\left(\frac{n}{m}\right)$, once again with a computable domination constant that depends on d, so that:

$$\left|\widehat{\delta_{\overline{w}}}([w]) - \mu([w])\right| = O\left(\frac{n}{m}\right) + O\left(\frac{|\mathcal{A}^{I_m}|}{2^k}\right).$$

Thus, we can actually compute integers $m(\delta, |\mathcal{A}|, d)$ and $k(\delta, |\mathcal{A}|, d)$ such that $\left|\widehat{\delta_{\overline{w}}}([w]) - \mu([w])\right| \leq \frac{\delta}{2} \times \frac{1}{2|\mathcal{A}^{I_n}|}$, which we can replace in the supremum bound for s_n. At last, we proved that there exists a pattern $w \in \mathcal{A}^{I_{\psi(\delta,|\mathcal{A}|,d)}}$ such that $\mu \in \overline{\mathrm{B}}\left(\widehat{\delta_w}, \delta\right)$, with $\psi(\delta, |\mathcal{A}|, d) := (m(\delta, |\mathcal{A}|, d) + 1)2^{k(\delta,|\mathcal{A}|,d)} - 1$ a function that can be computed by a Turing machine. □

Such a lemma isn't very surprising, given that the family of periodic points is dense (the density of which can be obtained by a simpler non-computable compactness argument), and we have a very explicit formula for the distance. These coverings are merely a stepping stone that will allow us to see that other measure sets, such as $\mathcal{M}_\sigma(X_\mathcal{F})$, can be obtained as the weak-$*$ limits of uniformly computable sequences of likewise (finite) compact sets.

Note how we always have $d_r^+ \leq d_1^+$ when $r \geq 1$, so this covering lemma more generally applies for all these distances.

4.9.2 Computable Descriptions of Measure Sets

Now, from the point of view of Turing machines, the main obstruction to discuss the notion of stability is that it is not obvious how we should proceed to compute $d_B(\mu, \mathcal{M}_\sigma(X_\mathcal{F}))$. To do so, I will step-by-step reach a computable characterisation of the sets $\mathcal{M}_\sigma(X_\mathcal{F})$, $\widetilde{\mathcal{M}_\mathcal{F}^\mathcal{B}}(\varepsilon)$, and at last $J(\mu, \mathcal{M}_\sigma(X_\mathcal{F}))$, following the tracks of the previous covering lemma for $\mathcal{M}_\sigma(\Omega_\mathcal{A})$.

Lemma 4.75 (Covering Lemma for $\mathcal{M}_\sigma(X_\mathcal{F})$)**.** *Consider ψ the radius-to-scale function obtained in the Covering Lemma 4.74 for $\mathcal{M}_\sigma(\Omega_\mathcal{A})$. Assume $\mathcal{F} \subseteq \mathcal{A}^{I_k}$ for some $k \in \mathbb{N}$. Then:*

$$\mathcal{M}_\sigma(X_\mathcal{F}) \subseteq \bigcup_{w \in \mathcal{W}_\mathcal{F}(\rho)} \overline{\mathrm{B}}\left(\widehat{\delta_w}, \rho\right)$$

where $\mathcal{W}_\mathcal{F}(\rho) \subseteq \mathcal{A}^{I_{\psi(\rho,|\mathcal{A}|,d)}}$ is the set of patterns on the window $I_{\psi(\rho,|\mathcal{A}|,d)}$ that contain at most $\varphi(\rho, k, |\mathcal{A}|, d)$ forbidden patterns from $\mathcal{F}$, with the computable map $\varphi(\rho, k, |\mathcal{A}|, d) := \left\lfloor 2^k \times \left|\mathcal{A}^{I_k}\right| \times \rho \times \left|I_{\psi(\rho,\dots)}\right|\right\rfloor$.

Note that the set $\mathcal{W}_\mathcal{F}(\rho)$ depends on k,$|\mathcal{A}|$ and d, but we hide them from the notation as they are either directly "written" in the computer representation of $\mathcal{F}$ or can be deduced from it.

Proof. The first covering lemma gives us $\mathcal{M}_\sigma(\Omega_\mathcal{A}) = \bigcup_{w \in \mathcal{A}^{I_{\psi(\rho,|\mathcal{A}|,d)}}} \overline{\mathrm{B}}\left(\widehat{\delta_w}, \rho\right)$. If we prove that we have $\mathcal{M}_\sigma(X_\mathcal{F}) \cap \overline{\mathrm{B}}\left(\widehat{\delta_w}, \rho\right) = \emptyset$ whenever w has more than $\varphi(\rho, \dots)$ forbidden patterns, this will conclude the proof.

Consider such a pattern $w \notin \mathcal{W}_\mathcal{F}(\rho)$ and $\mu \in \mathcal{M}_\sigma(X_\mathcal{F})$. For any forbidden pattern $u \in \mathcal{F} \subseteq \mathcal{A}^{I_k}$, we have $\mu([u]) = 0$, thence:

$$d^+\left(\mu, \widehat{\delta_w}\right) \geq \frac{1}{2^k} \times \frac{1}{|\mathcal{A}^{I_k}|} \sum_{u \in \mathcal{F}} \widehat{\delta_w}([u]) .$$

Now, by summing the number of occurrences of each forbidden pattern in w, we obtain a total of at least $\varphi(\rho, \dots) + 1 > 2^k \left|\mathcal{A}^{I_k}\right| \times \rho \times \left|I_{\psi(\rho,\dots)}\right|$. The rightmost factor is precisely the normalisation constant used to define $\widehat{\delta_w}$, so that $d^+\left(\mu, \widehat{\delta_w}\right) > \rho$, which concludes the proof. □

Remark that we can always assume $\mathcal{F} \subseteq \mathcal{A}^{I_k}$ without loss of generality. Indeed, in the general case, for any set of forbidden patterns $\mathcal{F}'$, let k big enough so that $I(p) \subseteq I_k$ for any $p \in \mathcal{F}$. Let $\mathcal{F}' \subseteq \mathcal{A}^{I_k}$ be the set of non locally admissible patterns (with respect to $\mathcal{F}$). Then we directly have $X_{\mathcal{F}'} = X_\mathcal{F}$, the two SFTs are conjugate so one is stable *iff* the other one is, and $\mathcal{F} \mapsto \mathcal{F}'$ is a computable reduction.

Corollary 4.76. *Using the previous notations, we have:*

$$\mathcal{M}_\sigma(X_\mathcal{F}) = \bigcap_{\rho > 0} \bigcup_{w \in \mathcal{W}_\mathcal{F}(\rho)} \overline{\mathrm{B}}\left(\widehat{\delta_w}, \rho\right) .$$

Proof. The covering lemma for $\mathcal{M}_\sigma(X_\mathcal{F})$ holds for any $\rho > 0$, hence by taking the intersection we directly obtain the inclusion $(\subseteq)$.

Conversely, consider $\mu \in \bigcap \bigcup (\cdots)$. There exists a sequence of radii $\rho_n \to 0$ and corresponding patterns w_n such that $d^+\left(\mu, \widehat{\delta_{w_n}}\right) \leq \rho_n$. For any given forbidden pattern $u \in \mathcal{F}$, we have:

$$\widehat{\delta_{w_n}}([u]) \leq \frac{\varphi(\rho_n, \dots) + O\left(\psi(\rho_n, \dots)^{d-1}\right)}{\left|I_{\psi(\rho_n,\dots)}\right|} = O(\rho_n) + O\left(\frac{1}{\psi(\rho_n, \dots)}\right) \to 0 ,$$

with the first part of the bound relating to the occurrences of u within w_n and the second to the occurrences on the interface between two square blocks of $w_n^{\mathbb{Z}^d}$. Thus, for the weak-$*$ limit μ we have $\mu([u]) = 0$, so $\mu \in \mathcal{M}_\sigma(X_\mathcal{F})$ by Lemma 4.7. □

On one hand, the covering lemma tells us that any measure in $\mathcal{M}_\sigma(X_\mathcal{F})$ can be approximated by some periodic measures with an explicit bound on the speed of convergence. More precisely, all measures $\mu \in \mathcal{M}_\mathcal{F}$ are ρ-close to some $\widehat{\delta_w}$ with $w \in \mathcal{W}_\mathcal{F}(\rho)$, but not *all* such $\widehat{\delta_w}$ are necessarily ρ-close to $\mathcal{M}_\sigma(X_\mathcal{F})$. This is roughly correlated to the nuance between locally admissible and globally admissible tilings. On the other hand, the corollary tells us that, while we have no computable bound on the speed of convergence, we still necessarily converge to measures $\mathcal{M}_\sigma(X_\mathcal{F})$, *i.e.* the ρ-coverings converge to the set $\mathcal{M}_\sigma(X_\mathcal{F})$ in the corresponding Hausdorff topology as $\rho \to 0$. If we identify $\mathcal{W}_\mathcal{F}(\rho)$ with the finite set of induced periodic measures, then $(\mathcal{W}_\mathcal{F}(\rho))_\rho$ is a uniformly computable family of base points, such that $\mathcal{M}_\sigma(X_\mathcal{F})$ is limit-computable by taking $\rho \to 0$.

The next landmark is the case of the noisy framework, in order to obtain a limit-computability result for $\widetilde{\mathcal{M}^\mathcal{B}_\mathcal{F}}(\varepsilon)$, as a subset of $\mathcal{M}_\sigma\left(X_{\widetilde{\mathcal{F}}}\right)$.

In the following proposition, we denote s_n the partial sum for $d^+_{|\mathcal{A}|}$ on the space of noises $\Omega_{\{0,1\}}$ (with $\mathcal{A}$ the alphabet of $\mathcal{F}$). In particular, if we use the computable rank $n(\rho)$ introduced in the proof of the covering lemma for $\mathcal{M}_\sigma(\Omega_\mathcal{A})$, then we can guarantee $s_{n(\rho)} \leq d^+_{|\mathcal{A}|} \leq s_{n(\rho)} + \rho$.

Lemma 4.77 (Covering Lemma for $\widetilde{\mathcal{M}^\mathcal{B}_\mathcal{F}}(\varepsilon)$)**.** *As in the Covering Lemma 4.75 for $\mathcal{M}_\sigma(X_\mathcal{F})$, we assume here that $\mathcal{F} \subseteq \mathcal{A}^{I_k}$. Note that $\mathcal{M}_{\widetilde{\mathcal{F}}}$ uses the alphabet $\mathcal{A} \times \{0,1\}$, of cardinality $2|\mathcal{A}|$. For any $\varepsilon \in [0,1]$, we can refine the covering of $\mathcal{M}_\sigma\left(X_{\widetilde{\mathcal{F}}}\right)$ into a covering of $\widetilde{\mathcal{M}^\mathcal{B}_\mathcal{F}}(\varepsilon)$:*

$$\widetilde{\mathcal{M}^\mathcal{B}_\mathcal{F}}(\varepsilon) = \bigcap_{\rho > 0} \bigcup_{(w,b) \in \widetilde{\mathcal{W}^\varepsilon_\mathcal{F}}(\rho)} \overline{\mathrm{B}}\left(\widehat{\delta_{(w,b)}}, \rho\right),$$

where $\widetilde{\mathcal{W}^\varepsilon_\mathcal{F}}(\rho) \subseteq \mathcal{W}_{\widetilde{\mathcal{F}}}(\rho)$ is the subset for which $s_{n(\rho)}\left(\widehat{\delta_b}, \mathcal{B}(\varepsilon)^{\otimes \mathbb{Z}^d}\right) \leq \rho$ holds on the second coordinate.

Proof. As before, we need to prove both inclusions.

Consider first ($\subseteq$). As $\widetilde{\mathcal{M}^\mathcal{B}_\mathcal{F}}(\varepsilon) \subseteq \mathcal{M}_{\widetilde{\mathcal{F}}}$, we already have the inclusion for any $\rho > 0$ *if* we forget about the new condition on $s_{n(\rho)}$, using the Covering Lemma 4.75 for $\mathcal{M}_\sigma\left(X_{\widetilde{\mathcal{F}}}\right)$. Hence, it suffices to prove that if $(w,b) \in \mathcal{W}_{\widetilde{\mathcal{F}}}(\rho)$ does not satisfy the new condition, then $d^+\left(\widehat{\delta_{(w,b)}}, \widetilde{\mathcal{M}^\mathcal{B}_\mathcal{F}}(\varepsilon)\right) > \rho$. This directly follows from the fact that for any $\lambda \in \widetilde{\mathcal{M}^\mathcal{B}_\mathcal{F}}(\varepsilon)$, using the Projection Lemma 4.73 we have:

$$d^+\left(\widehat{\delta_{(w,b)}}, \lambda\right) \geq d^+_{|\mathcal{A}|}\left(\widehat{\delta_b}, \mathcal{B}(\varepsilon)^{\otimes \mathbb{Z}^d}\right) \geq s_{n(\rho)}\left(\widehat{\delta_b}, \mathcal{B}(\varepsilon)^{\otimes \mathbb{Z}^d}\right) > \rho.$$

Conversely, consider ($\supseteq$). For any measure $\lambda \in \bigcap\bigcup(\cdots) \subseteq \mathcal{M}_\sigma\left(X_{\widetilde{\mathcal{F}}}\right)$, we have a sequence $\rho_n \to 0$ and patterns $(w_n, b_n) \in \widetilde{\mathcal{W}^\varepsilon_\mathcal{F}}(\rho_n)$ such that $d^+\left(\widehat{\delta_{(w_n,b_n)}}, \lambda\right) \leq \rho_n$. We just need to prove that $\pi^2_*(\lambda) = \mathcal{B}(\varepsilon)^{\otimes \mathbb{Z}^d}$. This comes from the fact that:

$$\begin{aligned} d^+_{|\mathcal{A}|}\left(\pi^2_*(\lambda), \mathcal{B}(\varepsilon)^{\otimes \mathbb{Z}^d}\right) &\leq d^+_{|\mathcal{A}|}\left(\pi^2_*(\lambda), \widehat{\delta_{b_k}}\right) + d^+_{|\mathcal{A}|}\left(\widehat{\delta_{b_k}}, \mathcal{B}(\varepsilon)^{\otimes \mathbb{Z}^d}\right) \\ &\leq d^+\left(\lambda, \widehat{\delta_{(w_k,b_k)}}\right) + d^+_{|\mathcal{A}|}\left(\widehat{\delta_{b_k}}, \mathcal{B}(\varepsilon)^{\otimes \mathbb{Z}^d}\right) \\ &\leq d^+\left(\lambda, \widehat{\delta_{(w_k,b_k)}}\right) + s_{n(\rho_k)}\left(\widehat{\delta_{b_k}}, \mathcal{B}(\varepsilon)^{\otimes \mathbb{Z}^d}\right) + \rho_k \\ &\leq 3\rho_k\,. \end{aligned}$$

As $k \to \infty$, we have $d^+_{|\mathcal{A}|}\left(\pi^2_*(\lambda), \mathcal{B}(\varepsilon)^{\otimes \mathbb{Z}^d}\right) = 0$, which concludes the proof. □

Using rational parameters ε and ρ, the function $(b, \varepsilon, \rho) \mapsto s_{n(\rho)}\left(\widehat{\delta_b}, \mathcal{B}(\varepsilon)^{\otimes \mathbb{Z}^d}\right)$ is computable. Hence, as for the case of $\mathcal{M}_\sigma(X_\mathcal{F})$, this covering lemma tells us *both* that we *can* mathematically approximate any $\lambda \in \widetilde{\mathcal{M}^\mathcal{B}_\mathcal{F}}(\varepsilon)$ with an explicit bound ρ, and that we have a way of describing $\widetilde{\mathcal{M}^\mathcal{B}_\mathcal{F}}(\varepsilon)$ as a limit-computable set (uniformly in ε), without control on the convergence speed.

Remark 4.78 (Approximating $J(\mu, \mathcal{M}_\sigma(X_\mathcal{F}))$)**.** Later on, in order to approximate $\lambda \in J(\mu, \mathcal{M}_\sigma(X_\mathcal{F}))$, we will consider a periodic measure $\widehat{\delta_{(w_1,w_2)}}$ on $\Omega_{\mathcal{A}\times\mathcal{A}}$ such that $\widehat{\delta_{w_1}}$ is close to $\widehat{\delta_w}$ an approximation of $\mu \in \widetilde{\mathcal{M}^\mathcal{B}_\mathcal{F}}(\varepsilon)$ (*i.e.* we have $\widehat{\delta_{(w,b)}} \in \widetilde{\mathcal{W}^\varepsilon_\mathcal{F}}(\rho)$) and that $\widehat{\delta_{w_2}}$ is close to $\widehat{\delta_{w'}} \in \mathcal{W}_\mathcal{F}(\rho)$ an approximation of some measure in $\mathcal{M}_\sigma(X_\mathcal{F})$. ■

4.9.3 Equivalent Characterisations of Stability

In the current context, if $X_{\mathcal{F}} = \emptyset$, then there is some window I_k that doesn't admit locally admissible tilings. Consequently, as long as $\varepsilon < 1$, by Borel-Cantelli lemma, with probability 1, a measure $\lambda \in \widetilde{\mathcal{M}^{\mathcal{B}}_{\mathcal{F}}}(\varepsilon)$ should have infinitely many clear windows the size of I_k, which cannot be tiled, thus $\widetilde{\mathcal{M}^{\mathcal{B}}_{\mathcal{F}}}(\varepsilon) = \emptyset$. Likewise, it follows that under some rank ρ_0, the computable sets $\widetilde{\mathcal{W}^{\varepsilon}_{\mathcal{F}}}(\rho)$ must also be empty. The idea to take out of this paragraph is that, while I won't insist on it in the conclusion of the chapter to keep things more readable, the fact that $X_{\mathcal{F}} \neq \emptyset$ is directly implied by the existence of the objects written in the following characterisations.

Let the measurable event $\Delta := \bigcup_{a\neq b\in\mathcal{A}}[(a,b)] \subseteq \Omega_{\mathcal{A}^2}$. Because we consider shift-invariant measures, by an ergodic theorem, $d_B(\mu,\nu) = \inf_{\lambda\in J(\mu,\nu)} \lambda(\Delta)$. In particular, as $\lambda \mapsto \lambda(\Delta)$ depends only on a finite window in $\Omega_{\mathcal{A}^2}$, it is a continuous map for the weak-$*$ topology.

Let me remind what it means for $\mathcal{F}$ to induce a stable SFT:

$$\forall\delta > 0, \exists\varepsilon > 0, \forall\mu \in \mathcal{M}^{\mathcal{B}}_{\mathcal{F}}(\varepsilon), d_B\left(\mu, \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right) \leq \delta\,.$$

Notice how, by monotonicity of the definition, we can restrict this formula by quantifying ε and δ over the countable set $\mathbb{Q}^{+*}$ of positive rationals instead. What's more, using the previous rewriting of d_B through joinings, we obtain the following characterisation:

$$\forall\delta \in \mathbb{Q}^{+*}, \exists\varepsilon \in \mathbb{Q}^{+*}, \forall\mu \in \mathcal{M}^{\mathcal{B}}_{\mathcal{F}}(\varepsilon), \exists\lambda \in J\left(\mu, \mathcal{M}_{\mathcal{F}}\right), \lambda(\Delta) \leq \delta. \tag{4.1}$$

As in the previous subsection, without loss of generality (up to a computable reduction), assume that $\mathcal{F} \subseteq \mathcal{A}^{I_k}$ for some integer k.

Proposition 4.79. *The SFT $X_{\mathcal{F}}$ is stable* iff *it satisfies the following formula:*

$$\begin{gathered}\forall\delta \in \mathbb{Q}^{+*}, \exists\varepsilon \in \mathbb{Q}^{+*}, \forall\rho \in \mathbb{Q}^{+*},\\ \forall\mu \in \mathcal{M}^{\mathcal{B}}_{\mathcal{F}}(\varepsilon), \exists\left(w_1,w_2\right) \in \left(\mathcal{A}^2\right)^{I_{\psi\left(\rho,|\mathcal{A}^2|,d\right)}},\\ \left[d^+_{|\mathcal{A}|}\left(\widehat{\delta_{w_1}},\mu\right) \leq \rho\right] \wedge \left[d^+_{|\mathcal{A}|}\left(\widehat{\delta_{w_2}},\mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right) \leq \rho\right] \wedge \left[\widehat{\delta_{(w_1,w_2)}}(\Delta) \leq \delta + |\mathcal{A}|^2\rho\right].\end{gathered} \tag{4.2}$$

Proof. Consider first the implication (4.1 $\Rightarrow$ 4.2). Assuming Equation 4.1 is satisfied, let us fix δ, ε and μ such that there exists a joining $\lambda \in J\left(\mu, \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right)$ for which $\lambda(\Delta) \leq \delta$. Using the Covering Lemma 4.74 for $\mathcal{M}_\sigma\left(\Omega_{\mathcal{A}^2}\right)$, we know that for any rational $\rho \in \mathbb{Q}^{+*}$, we have some couple $(w_1,w_2) \in \left(\mathcal{A}^2\right)^{I_{\psi\left(\rho,|\mathcal{A}^2|,d\right)}}$ such that $d^+\left(\widehat{\delta_{(w_1,w_2)}},\lambda\right) \leq \rho$. The first two inequalities in Equation 4.2 follow directly from the Projection Lemma 4.73. For the third one:

$$\begin{aligned}\widehat{\delta_{(w_1,w_2)}}(\Delta) &\leq \lambda(\Delta) + \sum_{a,b\in\mathcal{A}}\left|\lambda([(a,b)]) - \widehat{\delta_{(w_1,w_2)}}([(a,b)])\right|\\ &\leq \lambda(\Delta) + |\mathcal{A}|^2 d^+\left(\lambda, \widehat{\delta_{(w_1,w_2)}}\right)\\ &\leq \delta + |\mathcal{A}|^2\rho\,.\end{aligned}$$

Remark that the consecutive universal blocks $\forall\mu \in \mathcal{M}^{\mathcal{B}}_{\mathcal{F}}(\varepsilon)$ and $\forall\rho \in \mathbb{Q}^{+*}$ do not depend on each other, so we can freely reorder them as in Equation 4.2, which concludes this implication.

Conversely, suppose Equation 4.2 holds and let us prove (4.2 $\Rightarrow$ 4.1). Likewise, fix δ, ε and μ for which the rest of the formula (*i.e.* $\forall\rho\ldots$) is satisfied. Consider a sequence $\rho_n \to 0$ and the consequent patterns $\widehat{\delta_{w_1^n,w_2^n}} \in \left(\mathcal{A}^2\right)^{I_{\psi\left(\rho_n,|\mathcal{A}^2|,d\right)}}$. Up to extraction of a subsequence, we can assume that the sequence weakly converges to a measure $\lambda \in \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}^2}\right)$. The first inequality of Equation 4.2 gives us $\pi^1_*(\lambda) = \mu$ at the limit. The second inequality gives us $\pi^2_*(\lambda) \in \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$ (as $\mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$ is closed), so $\lambda \in J\left(\mu, \mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right)$. At last, the third inequality becomes $\lambda(\Delta) \leq \delta$ at the limit by continuity of $\lambda \mapsto \lambda(\Delta)$, which concludes the proof. □

This brings us one step closer to the desired characterisation, but we still need to let go of $\mu \in \mathcal{M}^{\mathcal{B}}_{\mathcal{F}}(\varepsilon)$ and $\mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)$ to obtain a formula using only "computable" objects.

Proposition 4.80. *The SFT $X_{\mathcal{F}}$ is stable* iff *it satisfies the following formula:*

$$\begin{gathered}\forall\delta\in\mathbb{Q}^{+*},\exists\varepsilon\in\mathbb{Q}^{+*},\forall\rho\in\mathbb{Q}^{+*},\exists\gamma\in\mathbb{Q}^{+*},\gamma\leq\rho,\\ \forall(w,b)\in\widetilde{\mathcal{W}^{\varepsilon}_{\mathcal{F}}}(\gamma),\exists w_0\in\mathcal{W}_{\mathcal{F}}(\rho),\exists\,(w_1,w_2)\in\left(\mathcal{A}^2\right)^{I_{\psi(\rho,|\mathcal{A}^2|,d)}},\\ \left[d^+_{|\mathcal{A}|}\left(\widehat{\delta_{w_1}},\widehat{\delta_w}\right)\leq 2\rho\right]\wedge\left[d^+_{|\mathcal{A}|}\left(\widehat{\delta_{w_2}},\widehat{\delta_{w_0}}\right)\leq 2\rho\right]\\ \wedge\left[\widehat{\delta_{(w_1,w_2)}}(\Delta)\leq\delta+|\mathcal{A}|^2\rho\right].\end{gathered}\tag{4.3}$$

Proof. Let us prove that (4.2 ⇒ 4.3). Assume Equation 4.2 holds true and fix δ, ε and ρ so that the rest of the formula holds true. The Covering Lemma 4.77 for $\widetilde{\mathcal{M}^{\mathcal{B}}_{\mathcal{F}}}(\varepsilon)$ implies that:

$$\max_{(w,b)\in\widetilde{W^{\varepsilon}_{\mathcal{F}}}(\gamma)} d^+\left(\widehat{\delta_{(w,b)}},\widetilde{\mathcal{M}^{\mathcal{B}}_{\mathcal{F}}}(\varepsilon)\right)\underset{\gamma\to 0}{\longrightarrow}0\,.$$

In particular, there is a rational $\gamma\leq\rho$ such that $\max d^+(\cdots)\leq\rho$. This will allow us to "merge" the universal block that should replace $\forall\mu$ directly into the already existing $\forall\rho\in\mathbb{Q}^{+*}$.

Now, for such a choice of γ, and any $(w,b)\in\widetilde{W^{\varepsilon}_{\mathcal{F}}}(\gamma)$, there always exists some $\mu\in\widetilde{\mathcal{M}^{\mathcal{B}}_{\mathcal{F}}}(\varepsilon)$ such that $d^+\left(\widehat{\delta_{(w,b)}},\mu\right)\leq\rho$. As $\pi^1_*(\mu)\in\mathcal{M}^{\mathcal{B}}_{\mathcal{F}}(\varepsilon)$, Equation 4.2 applies to it, and we can chose a corresponding pair $(w_1,w_2)\in\left(\mathcal{A}^2\right)^{I_{\psi(\rho,|\mathcal{A}^2|,d)}}$. Hence:

$$\begin{aligned}d^+_{|\mathcal{A}|}\left(\widehat{\delta_{w_1}},\widehat{\delta_w}\right)&\leq d^+_{|\mathcal{A}|}\left(\widehat{\delta_{w_1}},\pi^*_1(\mu)\right)+d^+_{|\mathcal{A}|}\left(\pi^*_1(\mu),\widehat{\delta_w}\right)\\ &\leq d^+_{|\mathcal{A}|}\left(\widehat{\delta_{w_1}},\pi^*_1(\mu)\right)+d^+\left(\mu,\widehat{\delta_{(w,b)}}\right)\\ &\leq 2\rho\,.\end{aligned}$$

Likewise, we have $\nu\in\mathcal{M}_\sigma(X_{\mathcal{F}})$ such that $d^+_{|\mathcal{A}|}\left(\widehat{\delta_{w_2}},\nu\right)\leq\rho$ in Equation 4.2, thus by the Covering Lemma 4.75 for $\mathcal{M}_\sigma(X_{\mathcal{F}})$ we have a pattern $w_0\in W_{\mathcal{F}}(\rho)$ such that $d^+_{|\mathcal{A}|}\left(\nu,\widehat{\delta_{w_0}}\right)\leq\rho$, hence $d^+_{|\mathcal{A}|}\left(\widehat{\delta_{w_2}},\widehat{\delta_{w_0}}\right)\leq 2\rho$. The third inequality does not change, which concludes the implication (4.2 ⇒ 4.3).

Now, suppose Equation 4.3 is true and let us prove (4.3 ⇒ 4.1). Fix δ, ε in the formula. Consider any sequence $\rho_n\to 0$, and the corresponding γ_n in the formula.

Let $\mu\in\mathcal{M}^{\mathcal{B}}_{\mathcal{F}}(\varepsilon)$. Using the Covering Lemma 4.77, we know there exists a sequence $(w^n,b^n)\in\widetilde{\mathcal{W}^{\varepsilon}_{\mathcal{F}}}(\gamma_n)$ such that $d^+_{|\mathcal{A}|}\left(\mu,\widehat{\delta_{w^n}}\right)\leq\gamma_n\leq\rho_n\to 0$. At any rank, we may chose $w^n_0\in\mathcal{W}_{\mathcal{F}}(\rho_n)$ and $(w^n_1,w^n_2)\in\left(\mathcal{A}^2\right)^{I_{\psi(\rho_n,|\mathcal{A}^2|,d)}}$ accordingly in Equation 4.3. Up to extraction, $\widehat{\delta_{w^n_0}}$ converges to $\nu\in\mathcal{M}_\sigma(X_{\mathcal{F}})$ and $\widehat{\delta_{(w^n_1,w^n_2)}}$ to $\lambda\in\mathcal{M}_\sigma(\Omega_{\mathcal{A}^2})$.

At the limit $\rho_n\to 0$, the first inequality of Equation 4.3 tells us that $\pi^1_*(\lambda)=\mu$. Likewise, the second one tells us that $\pi^2_*(\lambda)=\nu\in\mathcal{M}_\sigma(X_{\mathcal{F}})$, thence $\lambda\in J(\mu,\mathcal{M}_\sigma(X_{\mathcal{F}}))$. The third inequality naturally becomes $\lambda(\Delta)\leq\delta$, hence Equation 4.1 holds true, the SFT is stable. □

Theorem 4.81 (Upper Bound for Stability)**.** *The problem of stability is at most* Π_4*.*

Proof. We just proved that $\mathcal{F}$ induces a (non-empty) stable SFT *iff* it satisfies Equation 4.3. For the two blocks $d^+_{|\mathcal{A}|}\left(\widehat{\delta_{w_i}},\widehat{\delta_w}\right)\leq 2\rho$ in Equation 4.3, we can replace 2ρ by 3ρ to have a strict inequality instead (while still having an equivalent characterisation, as the proof of (4.3 ⇒ 4.1) only relies on the fact that $\rho\to 0$).

The interest of this variant point of view is that, as $d^+_{|\mathcal{A}|}\left(\widehat{\delta_{w_i}},\widehat{\delta_w}\right)$ is a computable real number, checking $d^+_{|\mathcal{A}|}\left(\widehat{\delta_{w_i}},\widehat{\delta_w}\right)<3\rho$ becomes a *semi-decidable* problem, adding a countable existential quantifier that can be merged into the $\exists\gamma$ block.

This variant formula starts with $[\forall\delta\in\mathbb{Q}^{+*},\exists\varepsilon\in\mathbb{Q}^{+*},\forall\rho\in\mathbb{Q}^{+*},\exists\gamma\in\mathbb{Q}^{+*}]$, *i.e.* four alternating layers of countable quantifiers. The following quantifiers are over finite (uniformly) computable sets, and then the three inequalities are decidable. Hence, this whole block can be decided in finite time. □

This concludes the current chapter, and my study of the Besicovitch stability. In the next chapter, as announced, we will dive into other aspects related to stability, but in the weak-∗ topology.

References

[AC05] Eugène Asarin and Pieter Collins. “Noisy Turing Machines”. In: *ICALP*. Vol. 3580. Lecture Notes in Computer Science. Springer, 2005, pp. 1031–1042. 10.1007/11523468_83.

[BDJ10] Alexis Ballier, Bruno Durand, and Emmanuel Jeandel. “Tilings Robust to Errors”. In: *LATIN 2010: Theoretical Informatics*. Springer Berlin Heidelberg, 2010, pp. 480–491. 10.1007/978-3-642-12200-2_42.

[BK89] Richard Max Burton and Michael Keane. “Density and Uniqueness in Percolation”. In: *Communications in Mathematical Physics* 121 (1989), pp. 501–505.

[BS19] Sebastián Barbieri and Mathieu Sablik. “A Generalization of the Simulation Theorem for Semidirect Products”. In: *Ergodic Theory and Dynamical Systems* 39.12 (2019), pp. 3185–3206. 10.1017/etds.2018.21.

[CV21] Antonin Callard and Pascal Vanier. “Computational Characterization of Surface Entropies for $\mathbb{Z}^2$ Subshifts of Finite Type”. In: *ICALP*. Vol. 198. LIPIcs. 2021, 122:1–122:20. 10.4230/LIPIcs.ICALP.2021.122.

[DRS12] Bruno Durand, Andrei Romashchenko, and Alexander Shen. “Fixed-Point Tile Sets and Their Applications”. In: *Journal of Computer and System Sciences* 78.3 (2012), pp. 731–764. 10.1016/j.jcss.2011.11.001.

[Gác01] Peter Gács. “Reliable Cellular Automata With Self-Organization”. In: *Journal of Statistical Physics* 103.1-2 (2001), pp. 45–267. 10.1023/A:1004823720305.

[GS23a] Léo Gayral and Mathieu Sablik. “On the Besicovitch-Stability of Noisy Random Tilings”. In: *Electronic Journal of Probability* 28 (2023), pp. 1–38. 10.1214/23-EJP917.

[GS23b] Léo Gayral and Mathieu Sablik. “Arithmetical Hierarchy of the Besicovitch-Stability of Noisy Tilings”. In: *Theory of Computing Systems* (2023). 10.1007/s00224-023-10142-y.

[JK69] Konrad Jacobs and Michael Keane. “0–1-Sequences of Toeplitz Type”. In: *Zeitschrift für Wahrscheinlichkeitstheorie und Verw. Gebiete* 13.2 (1969), pp. 123–131. 10.1007/BF00537017.

[Kal02] Olav Kallenberg. *Foundations of Modern Probability*. 3rd edition. Springer, 2002. 10.1007/978-3-030-61871-1.

[Lig05] Thomas Liggett. *Interacting Particle Systems*. Classics in mathematics. Springer, 2005. 10.1007/b138374.

[LM21] Douglas Lind and Brian Marcus. *An Introduction to Symbolic Dynamics and Coding*. 2nd edition. Cambridge Mathematical Library. Cambridge University Press, 2021. 10.1017/9781108899727.

[LSS97] Thomas Liggett, Roberto Schonmann, and A. M. Stacey. “Domination by Product Measures”. In: *Annals of Probability* 25.1 (1997), pp. 71–95. 10.1214/aop/1024404279.

CHAPTER **5**

Uniformly Chaotic Models of Gibbs Measures

When using the weak-$*$ topology, every non-empty SFT is stable in the sense of the previous chapter. However, this doesn't mean there isn't anything to say with this topology. This chapter adapts my third article [GST23], that looks at what happens in this case. In doing so, I switch to a more granular notion of convergence, of individual trajectories (*i.e.* families of measures indexed by a decreasing noise parameter) instead of the whole sets of noisy measures. In this context, the weak-$*$ topology becomes a much better fit, as we are not trying to discriminate *whether* a model converges, but rather *how* the convergence happens and what is the complexity of the limit behaviour.

More precisely, I let go of the noisy framework previously introduced to focus instead on the thermodynamic framework. In this context, the Gibbs measures we consider are equilibrium states, that maximise the pressure function $h(\mu) - \beta\mu(\varphi)$, where $\beta := 1/T$ is the *inverse temperature* parameter (here, the temperature T represents a role analogous to ε in the previous chapter), and φ a potential function (that will later be related to forbidden patterns in the finite-range case). We will denote $\mathcal{G}_\sigma(\beta)$ the set of such measures, and study the behaviour of trajectories $(\mu_\beta)_{\beta\in\mathbb{R}^+}$ as $\beta \to \infty$ (with $\mu_\beta \in \mathcal{G}_\sigma(\beta)$). This formalism will be properly introduced in Section 5.1.

The model is said to have *chaotic* temperature dependence near zero temperature if every trajectory diverges. Note that, due to compactness, every trajectory has convergent subsequences. The limit of any such subsequence is called a *ground state* (the term is used inconsistently in the literature, and I redirect the reader to a comparison of the various notions [EFS93, Appendix B.2]). In other words, the ground states of the model are the accumulation points of $\mathcal{G}_\sigma(\beta)$ as $\beta \to \infty$. The set of ground states will be denoted by $\mathcal{G}_\sigma(\infty)$.

Characterising the ground states and their stability at positive temperature is one of the most fundamental problems in equilibrium statistical mechanics. In the one-dimensional case, it was proven [Bré03, Lep05, CGU11] that when using finite-range potentials, there is uniqueness of the Gibbs measure for any β, and the unique trajectory (μ_β) converges to a limit, so chaotic dynamics cannot occur. However, for infinite-range potentials, such chaotic models exist, and these were “simulated” using finite-range potentials in three dimensions [CH10] first, and more recently in two dimensions [CS20, Bar+22]. The thermodynamic argument they used is generalised in Section 5.3, so that it may be used as black box for my purposes.

The goal of the current chapter is to expand these ideas, and to add computability arguments to obtain a precise characterisation of the sets $\mathcal{G}_\sigma(\infty)$ that can be obtained. Computability theory was already used in the previous constructions as an elegant tool to construct complex examples, but the results of the current chapter show that the notion of computability is essential in understanding the phenomenon of chaotic temperature dependence.

The main results of the chapter, Theorem 5.55 in particular, concern a stronger form of chaos, termed *uniform chaos*, in which all trajectories accumulate to the same set of (all) ground states $\mathcal{G}_\sigma(\infty)$. For such uniform systems, $\mathcal{G}_\sigma(\infty)$ must not only be connected (which is a general obstruction) but also Π_2-computable (the general complexity bound being Π_3) assuming the potential φ is computable. These obstructions are detailed in Section 5.2.

The rest of the chapter is dedicated to step-by-step proving that this upper bound is optimal in some sense, that any connected Π_2-computable set can be obtained as a limit set $\mathcal{G}_\sigma(\infty)$ (up to a computable affine homeomorphism) using a two-dimensional finite-range potential. Given that Π_2-computable sets can realised as accumulation points of computable sequences (Proposition 3.25), the analogy with the accumulation set $\mathcal{G}_\sigma(\infty)$ came naturally.

First, in Section 5.4, I construct a class of tilesets (with the associated potentials) that will satisfy the assumptions of the black box thermodynamic argument. In doing so, I once again start from the Robinson tiling to encode Turing machines performing computations, this time so that we obtain a fine control on the structure and the entropy of admissible patterns. Then, in Section 5.5, I apply the black box framework to relate the Gibbs measures to measures on words "simulated" by the Turing machines embedded in the admissible tilings. Lastly, in Section 5.6, I tie some loose ends regarding which measures on words can be simulated, notably regarding the required time complexity (which absent from the general definition of Π_2-computable objets, but cannot be ignored when using a simulating tileset with a finite space-time horizon).

5.1 Thermodynamic Framework

My primary interest lies in two-dimensional lattice models with finite "spin" alphabets as well as finite-range interactions. Nevertheless, the results of Section 5.2 work in any dimension $d \geq 1$ and for a very general class of interactions, and the construction presented in Sections 5.4–5.6 can be easily adapted to higher dimensions. Likewise, the results of Section 5.3 apply to finite-range models in any dimension.

As initially announced, we are now considering $\mathcal{M}_\sigma(\Omega_\mathcal{A})$ endowed with the weak-$*$ topology, and we will use the distance d^* defined in Section 2.4.

5.1.1 Interactions and Potentials

An *interaction* is a family $\Phi = (\Phi_I)_{I \Subset \mathbb{Z}^d}$ of maps $\Phi_I : \Omega_\mathcal{A} \to \mathbb{R}$ such that $\Phi_I(x)$ depends only on the restriction x_I. The value $\Phi_I(x)$ is interpreted as the energy contribution of the pattern x_I. The total *energy content* of a window $J \Subset \mathbb{Z}^d$ is the sum $E_J(x) = \sum_{I \subseteq J} \Phi_I(x)$. We restrict ourselves to interactions that are *shift-invariant*, meaning $\Phi_{I-k}(x) = \Phi_I(\sigma^k(x))$ for every $I \Subset \mathbb{Z}^d$ and $k \in \mathbb{Z}^d$. For brevity, we drop the adjective *shift-invariant* when talking about interactions.

An interaction Φ is said to have *finite range* if there exists an $r \in \mathbb{N}$ such that $\Phi_I = 0$ for every I with a diameter larger than r. Such an r is referred to as the *range* of the interaction. A broader family of interactions are those that are *absolutely summable*, that is:

$$\sum_{I \ni 0} \|\Phi_I\| < \infty ,$$

where $\|\cdot\|$ denotes the uniform norm. The latter condition covers (more or less) all "physically relevant" interactions [EFS93].

By a *potential* we shall mean a map $\varphi : \Omega_\mathcal{A} \to \mathbb{R}$. Every absolutely summable interaction Φ (more generally, every interaction satisfying a weaker summability condition) defines a continuous potential φ where

$$\varphi(x) := \sum_{I \ni 0} \frac{1}{|I|} \Phi_I(x) .$$

The value of φ is interpreted as the energy contribution of the site at the origin. The potential associated to a finite-range interaction is a *local* function, that is, there exists a finite window $D \Subset \mathbb{Z}^d$ such that the restriction x_D uniquely determines $\varphi(x)$.

Given any finite set of forbidden patterns $\mathcal{F}$, one can define a corresponding non-negative finite-range interaction $\Phi = (\Phi_I)_{I \Subset \mathbb{Z}^d}$ that assigns energy 1 to each *fault*, that is,

$$\Phi_I(x) = \begin{cases} 1 & \text{if } x_I \text{ is a translation of an element of } \mathcal{F}, \\ 0 & \text{otherwise.} \end{cases}$$

We call this the *fault interaction* corresponding to $\mathcal{F}$. The potential associated to Φ will be denoted by $\varphi_\mathcal{F}$ and is called the *fault potential* associated to $\mathcal{F}$. Fault interactions have earlier been exploited in models of quasicrystals [Rad86, Mię90, Mię97]. The interaction we will use in the following construction is of this type.

5.1.2 Pressure and Equilibrium States

Let $H(q) = -\sum_{s \in S} q(s) \log_2(q(s))$ be the *entropy* of a probability distribution q on a finite set S. We will measure entropies in bits, hence the base-2 logarithm in the definition of $H(q)$. The *entropy per site* (*i.e.* the Kolmogorov–Sinai entropy) of a shift-invariant measure $\mu \in \mathcal{M}_\sigma(\Omega_\mathcal{A})$ is:

$$h(\mu) = \lim_{n \to \infty} \frac{1}{|I_n|} H(\mu_{I_n}) = \inf_{n \in \mathbb{N}} \frac{1}{|I_n|} H(\mu_{I_n}) ,$$

where μ_{I_n} denotes the projection of μ on the patterns on I_n, such that $\mu_{I_n}(w) = \mu([w])$ for $w \in \mathcal{A}^{I_n}$.

Consider a model with a continuous potential φ. According to the variational principle of equilibrium statistical mechanics, the *equilibrium states* of the model at *inverse temperature* β are precisely the shift-invariant measures $\mu \in \mathcal{M}_\sigma(\Omega_\mathcal{A})$ that maximise the *pressure*

$$p_\mu(\beta) := h(\mu) - \beta\mu(\varphi)\,.$$

The maximal value, denoted $p(\beta)$, is called the *topological pressure*.

The Dobrushin–Lanford–Ruelle theorem [Rue04, Theorem 4.2] tells us that, if φ is the potential associated to an absolutely summable interaction Φ, then the equilibrium states of φ coincide with the shift-invariant Gibbs measures for Φ. We will denote the set of all equilibrium states at inverse temperature β by $\mathcal{G}_\sigma(\beta)$. This set is not necessarily a singleton, but is never empty. We will use this notation even when φ is not associated to an absolutely summable potential.

I refer the reader to classical books [Wal82, DGS76, Rue04] for a broader background on ergodic theory and the thermodynamic formalism. While the version of the Dobrushin–Lanford–Ruelle theorem used here has been known long-enough to be textbook material, finding new generalisations of the result for more general groups and/or interactions is still an active research topic [BM23].

5.1.3 Ground States, Stability and Chaos

An accumulation point of equilibrium states as $\beta \to \infty$ is called a *ground state*. In other words, the measure $\nu \in \mathcal{M}_\sigma(\Omega_\mathcal{A})$ is a ground state if there exists a sequence of $\beta_n \to \infty$ and equilibrium states $\mu_n \in \mathcal{G}_\sigma(\beta_n)$ such that $\mu_n \to \nu$. We denote the set of all ground states by $\mathcal{G}_\sigma(\infty)$. By compactness of $\mathcal{M}(\Omega_\mathcal{A})$, the set $\mathcal{G}_\sigma(\infty)$ is non-empty.

We say that a ground state ν is *stable* (in the weak-$*$ topology) if one can choose an equilibrium state μ_β, for each $\beta > 0$, in such a way that $\mu_\beta \to \nu$ as $\beta \to \infty$. A family (μ_β) identifies a *trajectory* of the system as the temperature goes to zero.

A model is said to have *chaotic* temperature dependence near zero temperature if none of its ground states is stable. A *uniform* model is one in which every trajectory (μ_β) has the entire set $\mathcal{G}_\sigma(\infty)$ of ground states as its accumulation points. A uniform model with at least two ground states is clearly chaotic. We call such a model a *uniformly chaotic* model.

Remark 5.1. A stable model is necessarily uniform, and a uniform model that is not stable is necessarily chaotic. However, not all chaotic models are uniformly chaotic.

Indeed, consider a potential φ on the symbolic space $\Omega_\mathcal{A}$, φ' on $\Omega_{\mathcal{A}'}$ (with the same dimension), and $\psi : (x, y) \in \Omega_{\mathcal{A}\times\mathcal{A}'} \mapsto \varphi(x) + \varphi'(y)$. The model induced by ψ is thus made of two layers, on the alphabets $\mathcal{A}$ and $\mathcal{A}'$, that behave independently. If φ induces a chaotic model, then so does ψ as no trajectory converges (on its first coordinate). Likewise, if φ' induces a *non-uniform* model, then so does ψ.

On one hand, for φ, we can use a two-dimensional potential inducing a chaotic model [CS20, Bar+22]. On the other hand, for φ', we can use the two-dimensional Ising model which is non-uniform, as the trajectories corresponding to the *all-plus* and *all-minus* boundary conditions both converge to their own limit. Thence, two-dimensional non-uniformly chaotic models exist. ■

Given the common occurrence of accumulation points in this chapter, I will use the following notation. The set of accumulation points of a sequence (x_n) of points in a metric space will be denoted by $\mathrm{Acc}_{n\to\infty}(x_n)$. Likewise, we write $\mathrm{Acc}_{n\to\infty}(M_n)$ for the set of the accumulation points of a sequence (M_n) of subsets of a metric space. We use a similar notation for families indexed by real numbers rather than integers:

$$\begin{aligned}
\operatorname*{Acc}_{\beta\to\infty}(x_\beta) &= \{y, \forall\varepsilon > 0, \forall\beta_0 > 0, \exists\beta \geq \beta_0, d(x_\beta, y) < \varepsilon\}\,,\\
\operatorname*{Acc}_{\beta\to\infty}(M_\beta) &= \{y, \forall\varepsilon > 0, \forall\beta_0 > 0, \exists\beta \geq \beta_0, \exists x \in M_\beta, d(x, y) < \varepsilon\}\,.
\end{aligned}$$

With this notation, $\mathcal{G}_\sigma(\infty) = \mathrm{Acc}_{\beta\to\infty}(\mathcal{G}_\sigma(\beta))$, and the model is uniform if and only if, for every cooling trajectory (μ_β), we have $\mathcal{G}_\sigma(\infty) = \mathrm{Acc}_{\beta\to\infty}(\mu_\beta)$.

5.2 Topology and Computational Complexity of $\mathcal{G}_\sigma(\infty)$

The goal of this section is two-fold. First, to describe the topology of $\mathcal{G}_\sigma(\infty)$, we will see that $\mathcal{G}_\sigma(\infty)$ is connected in general, and that it is equal to $\mathrm{Acc}\left(\mathcal{G}_\sigma\left(\beta_n\right)\right)$ for some sequence $\beta_n \to \infty$ in the case of a uniform model. Second, we want to obtain a computational upper bound on $\mathcal{G}_\sigma(\infty)$. To do so, I will prove that, assuming the potential $\varphi : \Omega_{\mathcal{A}} \to \mathbb{R}$ is a computable function, the set $\mathcal{G}_\sigma(\infty)$ is at most Π_3 in general and Π_2 for a uniform model.

5.2.1 Topological Properties of $\mathcal{G}_\sigma(\infty)$

Lemma 5.2 (Weak Continuity of Gibbs Measures)**.** *Let $\beta_n \to \beta \in \mathbb{R}^+$ be a converging sequence of inverse temperatures, along with a converging sequence of Gibbs measures $\mu_n \to \mu$ with $\mu_n \in \mathcal{G}_\sigma\left(\beta_n\right)$. Then we have $\mu \in \mathcal{G}_\sigma(\beta)$. It directly follows that, for any $\beta > 0$:*

$$\bigcap_{\varepsilon>0} \overline{\bigcup_{\substack{|\beta'-\beta|\leq\varepsilon \\ \beta'\neq\beta}} \mathcal{G}_\sigma\left(\beta'\right)} \subseteq \mathcal{G}_\sigma(\beta)\,.$$

In other words, $\beta \mapsto \mathcal{G}_\sigma(\beta)$ is upper semi-continuous using the Hausdorff distance on closed subsets of $\mathcal{M}_\sigma\left(\Omega_{\mathcal{A}}\right)$ and the partial order given by inclusion.

Proof. We know that the pressure function $(\mu,\beta) \in \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}}\right) \times \mathbb{R}^+ \mapsto p_\mu(\beta)$ is upper semi-continuous, and continuous w.r.t. β for any fixed measure. Here, $p_\mu(\beta) \geq \overline{\lim}\, p_{\mu_n}\left(\beta_n\right) = \overline{\lim}\, p\left(\beta_n\right)$. Consider $\mu_\beta \in \mathcal{G}_\sigma(\beta)$, so that $p_{\mu_\beta}(\beta) = p(\beta)$. Then quite naturally:

$$p\left(\beta_n\right) \geq p_{\mu_\beta}\left(\beta_n\right) \underset{n\to\infty}{\longrightarrow} p_{\mu_\beta}(\beta) = p(\beta)\,.$$

As $p_\mu(\beta) \geq p(\beta)$, μ reaches the topological pressure at β, so $\mu \in \mathcal{G}_\sigma(\beta)$. ☐

Proposition 5.3. *The set $\mathcal{G}_\sigma(\infty)$ is compact and connected.*

Proof. The set $\mathcal{G}_\sigma(\infty)$ is closed by definition, and a subset of the compact set $\mathcal{M}_\sigma\left(\Omega_{\mathcal{A}}\right)$, so it is compact too. Assume now that $\mathcal{G}_\sigma(\infty)$ is not connected. Then there exists two disjoint open sets $A, B \subseteq \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}}\right)$ such that both intersect $\mathcal{G}_\sigma(\infty)$ and $\mathcal{G}_\sigma(\infty) \subseteq A \sqcup B$. By compactness, there is a threshold β_0 such that for any $\beta \geq \beta_0$, we have $\mathcal{G}_\sigma(\beta) \subseteq A \sqcup B$ – or else we would have a sequence of Gibbs measures $\left(\mu_{\beta_n}\right)_{n\in\mathbb{N}}$ (with $\beta_n \to \infty$) in the compact $(A \sqcup B)^c$, thus some adherence value outside of $\mathcal{G}_\sigma(\infty)$.

For any such $\beta \geq \beta_0$, since $\mathcal{G}_\sigma(\beta)$ itself is convex (thus connected), one has either $\mathcal{G}_\sigma(\beta) \subseteq A$ or $\mathcal{G}_\sigma(\beta) \subseteq B$. If all the sets were included in A (resp. B) after a rank then we would have $B \cap \mathcal{G}_\sigma(\infty) = \emptyset$, a contradiction. Thus, we have two inverse temperatures $\beta_A, \beta_B \geq \beta_0$ such that $\mathcal{G}_\sigma\left(\beta_A\right) \subseteq A$ and $\mathcal{G}_\sigma\left(\beta_B\right) \subseteq B$ (and $\beta_A < \beta_B$ without loss of generality).

Consider now the partition $\left[\beta_A, \beta_B\right] = I_A \sqcup I_B$ such that $G(\beta) \subseteq A$ *iff* $\beta \in I_A$. As both sets are non-empty, one of them must have an accumulation point on the other. Without loss of generality, assume we have $\left(\beta_n\right) \in I_A^{\mathbb{N}}$ that converges to $\beta \in I_B$. Up to extraction, we also have a sequence of Gibbs measures $\left(\mu_{\beta_n}\right)$ all in A, within the compact B^c, that converges to $\mu \notin B$. However, using the previous lemma, $\mu \in \mathcal{G}_\sigma(\beta) \subseteq B$. Thus, we have a contradiction that concludes the proof, $\mathcal{G}_\sigma(\infty)$ is indeed connected. ☐

This is as far as we will go for the general case of a weak-$*$ continuous potential φ with summable variation. Let's now focus on the uniform case, to obtain a preliminary result that will be useful for the incoming computational bound.

Given a compact $K \subseteq \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}}\right)$, we will denote $\mathrm{diam}\,(K) = \max_{\mu,\mu'\in K} d^*\left(\mu,\mu'\right)$ its diameter. Then, let $A_\varepsilon := \left\{\beta > 0, \mathrm{diam}\left(\mathcal{G}_\sigma(\beta)\right) < \varepsilon\right\}$ for any $\varepsilon > 0$. The weak continuity lemma implies that $\beta \mapsto \mathrm{diam}\left(\mathcal{G}_\sigma(\beta)\right)$ is upper semi-continuous, so the sets A_ε are all open.

Proposition 5.4. *If the model is uniform, then* $\sup A_\varepsilon = \infty$ *for any* $\varepsilon > 0$, i.e. *the set contains arbitrarily big values for β. Moreover, $\mathcal{G}_\sigma(\infty) = \mathrm{Acc}_{\beta\to\infty}\left(\left(\mathcal{G}_\sigma(\beta)\right)_{\beta\in A_\varepsilon}\right)$.*

Proof. First, assume that A_ε is bounded by β_0, *i.e.* $\mathrm{diam}\left(\mathcal{G}_\sigma(\beta)\right) \geq \varepsilon$ for any $\beta > \beta_0$. Consider a measure $\mu \in \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}}\right)$. For any $\beta > \beta_0$, we can pick some $\mu_\beta \in \mathcal{G}_\sigma(\beta)$ such that $d^*\left(\mu, \mu_\beta\right) \geq \frac{\varepsilon}{2}$. Thus, $\mu \notin \mathrm{Acc}\left(\mu_\beta\right)$. This holds in particular for any $\mu \in \mathcal{G}_\sigma(\infty)$, which contradicts the uniformity of the model, hence A_ε is not bounded.

It follows that the set of accumulation points $\mathrm{Acc}_{\beta\to\infty}\left((\mathcal{G}_\sigma(\beta))_{\beta\in A_\varepsilon}\right)$ is well-defined, and the inclusion $\supset$ trivially holds. Suppose that the other inclusion doesn't hold, and let a ground state $\mu \notin \mathrm{Acc}\left((\mathcal{G}_\sigma(\beta))_{\beta\in A_\varepsilon}\right)$. If $\beta \in A_\varepsilon$, we pick any $\mu_\beta \in \mathcal{G}_\sigma(\beta)$, and if $\beta \notin A_\varepsilon$, we can pick μ_β such that $d^*(\mu, \mu_\beta) \geq \frac{\varepsilon}{2}$ as before. Then $\mu \notin \mathrm{Acc}(\mu_\beta)$, which once again contradicts uniformity of the model. □

5.2.2 Computability Obstructions on $\mathcal{G}_\sigma(\infty)$

In order to obtain a general upper bound on the complexity of $\mathcal{G}_\sigma(\infty)$, a first step will be to obtain a computable description of $\mathcal{G}_\sigma(\beta)$ (*e.g.* for computable values of $\beta > 0$). More generally, we will obtain a description of the sets $\mathcal{G}_\sigma([\beta^-, \beta^+]) := \bigcup_{\beta\in[\beta^-,\beta^+]} \mathcal{G}_\sigma(\beta)$.

Lemma 5.5. *For any $0 \leq \beta^- \leq \beta^+$, the set $\mathcal{G}_\sigma([\beta^-, \beta^+])$ is compact.*

Proof. Let a sequence of measures $(\mu_n) \in \mathcal{G}_\sigma([\beta^-, \beta^+])^{\mathbb{N}}$ converging to μ. It can be associated to a sequence of inverse temperatures $(\beta_n) \in [\beta^-, \beta^+]^{\mathbb{N}}$, converging to $\beta \in [\beta^-, \beta^+]$ (up to extraction). It follows from Lemma 5.2 that $\mu \in \mathcal{G}_\sigma(\beta) \subseteq \mathcal{G}_\sigma([\beta^-, \beta^+])$, hence the set is closed (in the compact space $\mathcal{M}_\sigma(\Omega_\mathcal{A})$). □

As $\mathcal{G}_\sigma(\beta)$ is by definition the set of measures that realise the topological pressure $p(\beta)$, a first step in studying these sets is to study the pressure itself. In the one-dimensional case, when the phase space is an SFT $X_\mathcal{F}$ (*i.e.* we maximise the pressure $\mu \mapsto p_\mu(\beta)$ within $\mathcal{M}_\sigma(X_\mathcal{F})$, instead of among all the measures $\mu \in \mathcal{M}_\sigma(\Omega_\mathcal{A})$), it is known that the pressure $\beta \mapsto p(\beta)$ is computable [Bur+22]. In particular, $p(0)$ is the *topological entropy* of the subshift $X_\mathcal{F}$. Consequently, their scheme of proof doesn't hold in the higher-dimensional case, as it is known that the topological entropy of a 2D SFT is not always computable [HM10]. However, we obtain here the higher-dimensional result when the phase space is the full-shift $\Omega_\mathcal{A}$, in which case the topological entropy $p(0)$ is simply equal to $\log_2(|\mathcal{A}|)$.

Remark 5.6 (Computable Potentials)**.** Let's take some time to properly convey what it means for a potential $\varphi : \Omega_\mathcal{A} \to \mathbb{R}$ to be *computable*. Let's translate the general framework of Definition 3.28 into something more natural.

As already done before implicitly, we endow $\mathbb{R}$ with the countable dense basis $\mathbb{Q}$. We can use the convenient countable topological basis of cylinder sets $([w])_{w\in\mathfrak{W}}$ (for the weak-$*$ topology). with $\mathfrak{W}$ the set of finite patterns $w \in \mathcal{A}^{[\![-r,r]\!]^d}$ with $r \in \mathbb{N}$. A potential φ is computable when we have a pair of computable functions $f, \rho : \mathfrak{W} \to \mathbb{Q}$, such that $\varphi([w]) \subseteq]f(w) - \rho(w), f(w) + \rho(w)[$, and $\max_{w\in\mathcal{A}^{[\![-r,r]\!]^d}} \rho(w) \underset{r\to\infty}{\longrightarrow} 0$. In particular, a computable potential is necessarily continuous.

If Φ is a family of local interactions, that only takes (finitely many) computable real values, then it follows that the associated potential φ_Φ is computable (as soon as r goes beyond the range of interactions of Φ, then $\varphi_\Phi([w])$ is just a singleton containing a computable real number). If Φ only takes rational values, then we can compute exact rational values for the potential $\varphi_\Phi([w])$ for such big-enough words. All of this holds in particular for any SFT-like potential $\varphi_\mathcal{F}$. ■

Proposition 5.7. *Let $\varphi : \Omega_\mathcal{A} \to \mathbb{R}$ be a computable potential. Then the pressure $\beta \mapsto p(\beta)$ is computable.*

Proof. Let $h_n : \mu \in \mathcal{M}_\sigma(\Omega_\mathcal{A}) \mapsto -\frac{1}{|I_n|}\sum_{w\in\mathcal{A}^{I_n}} \mu([w])\log_2(\mu([w]))$ be the approximate entropy on the finite window I_n, converging to the metric entropy h. Unlike the entropy h, h_n is continuous on $\mathcal{M}_\sigma(\Omega_\mathcal{A})$, and the family (h_n) is even uniformly computable (so $h = \inf h_n$ is limit-computable by above).

Let μ be a measure on $\mathcal{A}^{I_n}$, and denote $\mathcal{M}_n$ the set of such measures (which we may represent as a finite-dimensional simplex). We define the measure $\widehat{\mu}$ by dividing $\mathbb{Z}^d$ into a grid of windows I_n, picking independently a labelling of each window with distribution μ, and then averaging over the finite translation orbit so that $\widehat{\mu} \in \mathcal{M}_\sigma(\Omega_\mathcal{A})$. In particular, $\widehat{\mathcal{M}_n} := \{\widehat{\mu}, \mu \in \mathcal{M}_n\}$ contains the periodic points $\widehat{\delta_w}$ with $w \in \mathcal{A}^{I_n}$, so that $\mathfrak{P} \subseteq \bigcup_{n\in\mathbb{N}} \widehat{\mathcal{M}_n}$.

As h_n only depends on the restriction of a measure μ to the window I_n, it still makes sense to study h_n for measures $\mu \in \mathcal{M}_n$. We can define the approximate pressure function as $p_n : (\mu, \beta) \in \mathcal{M}_n \times \mathbb{R} \mapsto h_n(\mu) - \beta\widehat{\mu}(\varphi)$. For any measure $\mu \in \mathcal{M}_n$, we actually have $h_n(\mu) = h(\widehat{\mu})$, thus the bound $p_n(\mu, \beta) = p_{\widehat{\mu}}(\beta) \leq p(\beta)$.

At last, we introduce the approximate maximal pressure $p_n^* : \beta \mapsto \max_{\mu\in\mathcal{M}_n} p_n(\mu, \beta)$. Because p_n is computable on $\mathcal{M}_n \times \mathbb{R}$, which is a locally recursively compact space, the function p_n^* is computable too. Actually, because computing p_n^* involves roughly the same algorithm regardless of the dimension of the simplex (which itself is a computable function of n), the functions (p_n^*) are uniformly computable.

We already established that, for any $n \in \mathbb{N}$, we have $p_n^* \leq p$. Consider an actual Gibbs measure $\mu \in \mathcal{G}_\sigma(\beta)$, such that $p_\mu(\beta) = p(\beta)$, and denote $\mu_n = \mu|_{I_n} \in \mathcal{M}_n$. We have:

$$\begin{aligned} p(\beta) &\leq h_n(\mu_n) - \beta\mu(\varphi) \\ &= p_n(\mu_n, \beta) + \beta\left(\widehat{\mu_n}(\varphi) - \mu(\varphi)\right) \\ &\leq p_n^*(\beta) + \beta\|\varphi\|_{d^*} \sup_{\mu \in \mathcal{M}_\sigma(\Omega_\mathcal{A})} d^*\left(\mu, \widehat{\mu_n}\right). \end{aligned}$$

where $\|\varphi\|_{d^*}$ is the operator norm of $\mu \mapsto \mu(\varphi)$ (on the vector space of finite shift-invariant measures on $\Omega_\mathcal{A}$, in which d^* corresponds to a well-defined norm for the weak-$*$ topology). In particular, the only part where a measure $\mu \in \mathcal{M}_\sigma(\Omega_\mathcal{A})$ and $\widehat{\mu_n}$ may significantly differ is on the boundary of the window I_n, hence we can obtain an explicit bound $O\left(\frac{1}{n}\right)$ on this term. It follows that we have a computable bound $|p - p_n^*| = O\left(\frac{\beta}{n}\right)$, hence p is uniformly computable on $\mathbb{Q}$.

At last, $p = \sup p_\mu$ is a supremum of affine (thus convex) functions, so it is convex too. We conclude that p is not simply computable on $\mathbb{Q}$ but indeed a computable function on $\mathbb{R}$ by Lemma 3.29. □

Remark that we can use the computable potential φ as a countable parameter too. Doing so tells us that $\beta \mapsto p(\beta)$ is uniformly computable in φ, or equivalently that $(\beta, \varphi) \mapsto \sup_\mu h(\mu) - \beta\mu(\varphi)$ is also a computable map.

Proposition 5.8. *Let φ be a computable potential. It is semi-decidable to know whether the quadruplet $(x, r, \beta^-, \beta^+) \in \mathfrak{P} \times \mathbb{Q}^3$ is such that $\mathcal{G}_\sigma\left([\beta^-, \beta^+]\right) \cap \overline{B(x,r)} = \emptyset$.*

Proof. Let $\psi_n(\mu, \beta) := p(\beta) - (h_n(\mu) - \beta\mu(\varphi))$. The family (ψ_n) is uniformly computable (*i.e.* the function $(n, \mu, \beta) \mapsto \psi_n(\mu, \beta)$ is computable). What's more, there is a large-enough value of $A \in \mathbb{Q}$ such that all the functions ψ_n take bounded values, in the interval $[-A, A]$.

Notice that $\sup_n \psi_n(\mu, \beta) = p(\beta) - p_\mu(\beta) \geq 0$, with equality to 0 *iff* $\mu \in \mathcal{G}_\sigma(\beta)$. Thus:

$$\bigcup_{\beta \geq 0} \mathcal{G}_\sigma(\beta) \times \{\beta\} = \bigcap_{n \in \mathbb{N}} \psi_n^{-1}([-A, 0]).$$

Since $(\psi_n)_n$ is uniformly computable, the sets $\psi_n^{-1}([-A, 0])$ are uniformly Π_1-computable, *i.e.* there is a computable $f : \mathbb{N}^2 \times \mathbb{Q}^2 \times \mathfrak{P} \times \mathbb{Q}$ such that:

$$\left(\overline{B(x,r)} \times \left[\beta^-, \beta^+\right]\right) \cap \psi_n^{-1}([-A, 0]) \neq \emptyset \Leftrightarrow \forall k \in \mathbb{N}, f(n, k, \beta^-, \beta^+, x, r) = 1.$$

It follows that:

$$\overline{B(x,r)} \cap \mathcal{G}_\sigma(\left[\beta^-, \beta^+\right]) \neq \emptyset \Leftrightarrow \forall k, n \in \mathbb{N}, f(n, k, \beta^-, \beta^+, x, r) = 1.$$

Hence, the complementary question is semi-decidable. □

In other words, the compact sets $\mathcal{G}_\sigma\left([\beta^-, \beta^+]\right)$ are uniformly Π_1-computable. This holds in particular for the sets $\mathcal{G}_\sigma(\beta)$ with $\beta \in \mathbb{Q}^+$. Consequently, according to Remark 3.27, it is semi-decidable to know if $\mathcal{G}_\sigma(\beta) \subseteq B(x,r)$ for an input $(x, r, \beta) \in \mathfrak{P} \times \mathbb{Q}^2$.

Proposition 5.9. *Let φ be a computable potential. Then the compact set $\mathcal{G}_\sigma(\infty)$ is Π_3-computable.*

Proof. Let $(x, r) \in \mathfrak{P} \times \mathbb{Q}$. We have:

$$\begin{aligned} & \mathcal{G}_\sigma(\infty) \cap \overline{B(x,r)} \neq \emptyset \\ \Leftrightarrow\ & \forall \beta_0 \in \mathbb{N}, \forall \delta \in \mathbb{Q}^{+*}, \exists \beta \in \mathbb{R}^+_{\geq \beta_0}, \mathcal{G}_\sigma(\beta) \cap \overline{B(x, r+\delta)} \neq \emptyset \\ \Leftrightarrow\ & \forall \beta_0 \in \mathbb{N}, \forall \delta \in \mathbb{Q}^{+*}, \exists b \in \mathbb{N}_{\geq \beta_0}, \mathcal{G}_\sigma([b, b+1]) \cap \overline{B(x, r+\delta)} \neq \emptyset. \end{aligned}$$

The first equivalence comes from the fact that any ground state in $\mathcal{G}_\sigma(\infty) \cap \overline{B(x,r)}$ will be approached by Gibbs measures in $\mathcal{G}_\sigma(\beta)$ at distance δ (after any rank β_0 and with any precision $\delta > 0$, which we can both assume to be rational by monotonicity of the properties considered). For the second one, simply notice that such a real number $\beta \geq \beta_0$ *must* be included in an integer interval $[b, b+1]$.

Using Proposition 5.8, we deduce that checking if the set $\mathcal{G}_\sigma([b, b+1]) \cap \overline{B(x, r+\delta)}$ is (non-)empty is a (co-)recursively-enumerable problem. We can thus reduce the rightmost part to a formula $\forall k, f(k, b, x, r+\delta)$ with a computable f. Thence, $\mathcal{G}_\sigma(\infty)$ is a Π_3-computable compact set. □

Theorem 5.10. *Let φ a computable potential, and assume it induces a uniform model. Then the set of ground states $\mathcal{G}_\sigma(\infty)$ is Π_2-computable.*

Proof. Let $(x,r) \in \mathfrak{P} \times \mathbb{Q}$. We have:

$$\begin{aligned}
& \mathcal{G}_\sigma(\infty) \cap \overline{B(x,r)} = \emptyset \\
\Leftrightarrow \quad & \forall \beta_0 \in \mathbb{N}, \forall \delta \in \mathbb{Q}^{+*}, \exists \beta \in \mathbb{R}^+_{\geq \beta_0}, \mathcal{G}_\sigma(\beta) \cap \overline{B(x, r+\delta)} \neq \emptyset \\
\Leftrightarrow \quad & \forall \varepsilon \in \mathbb{Q}^{+*}, \forall \beta_0 \in \mathbb{N}, \forall \delta \in \mathbb{Q}^{+*}, \exists \beta \in \mathbb{R}_{\geq \beta_0}, \exists y \in \mathfrak{P}, \\
& \mathcal{G}_\sigma(\beta) \subseteq B(y,\varepsilon) \text{ and } B(y,\varepsilon) \cap \overline{B(x, r+\delta)} \neq \emptyset \\
\Leftrightarrow \quad & \forall \varepsilon \in \mathbb{Q}^{+*}, \forall \beta_0 \in \mathbb{N}, \forall \delta \in \mathbb{Q}^{+*}, \exists \beta \in \mathbb{Q}_{\geq \beta_0}, \exists y \in \mathfrak{P}, \\
& \mathcal{G}_\sigma(\beta) \subseteq B(y,\varepsilon) \text{ and } B(y,\varepsilon) \cap \overline{B(x, r+\delta)} \neq \emptyset\,.
\end{aligned}$$

As in the previous proposition, the first equivalence comes from the definition of $\mathcal{G}_\sigma(\infty)$ as an accumulation set. The second equivalence comes from the fact that the model is uniform, thus using Proposition 5.4, we have $\mathcal{G}_\sigma(\infty) = \mathrm{Acc}\left((\mathcal{G}_\sigma(\beta))_{\beta \in A_\varepsilon}\right)$, the accumulation points can be reached using only scales $\beta \in A_\varepsilon$, for which the diameter of $\mathcal{G}_\sigma(\beta)$ is less than ε. In particular, as A_ε is an open set, by weak continuity of Gibbs measures (Lemma 5.2) we can instead consider $A_\varepsilon \cap \mathbb{Q}$ and still reach all the accumulation points.

Since the sets $\mathcal{G}_\sigma(\beta)$ are uniformly Π_1-computable for $\beta \in \mathbb{Q}$, and $\mathcal{M}_\sigma(\Omega_\mathcal{A})$ is recursively compact, we deduce that it is semi decidable to say if $\mathcal{G}_\sigma(\beta) \subseteq B(y,\varepsilon)$. Likewise, in any computable space, checking whether $B(y,\varepsilon) \cap \overline{B(x,r)} \neq \emptyset$ is semi-decidable. In both cases, this only adds more existential quantifiers on the right of the formula, hence we can conclude that $\mathcal{G}_\sigma(\infty)$ is Π_2-computable. □

These two upper bounds suggest a profound link between the computational complexity of the set $\mathcal{G}_\sigma(\infty)$ and the dynamic properties of the model (*i.e.* whether it is uniform or not). Ultimately, proving the "optimality" of both bounds would allow for a more affirmative conclusion on this matter. The rest of this chapter will be dedicated to a first step in this direction: the optimality of the Π_2 bound for uniform models.

5.3 Uniform Marker Distribution in Gibbs States

In this section, we present a general framework for understanding the Gibbs measures of a certain class of finite-range models at moderately low temperatures. The models in question have the property that their ground patterns consist mostly of distinguishable, non-overlapping blocks we call markers. The main result of this section (Theorem 5.20) states that, under some entropy assumption on markers and ground patterns, every shift-invariant Gibbs measure within a certain temperature interval induces, with high probability, a large grid of markers, each chosen uniformly at random. This result will later be applied, in the context of the current construction, to a hierarchy of marker sets at larger and larger scales.

The key ideas in this section are borrowed from the infinite-range, one-dimensional construction of Chazottes and Hochman [CH10], but here we adapt them to the framework of finite-range models in higher dimensions, and present them in a more streamlined fashion.

5.3.1 Ground Configurations and Ground Patterns

In this section, we let Φ be a finite-range interaction with range r that takes only non-negative values (*i.e.* $\Phi_I \geq 0$ for every $I \Subset \mathbb{Z}^d$), and let φ be the associated potential. For such an interaction, the *total energy*

$$E(x) = \sum_{I \Subset \mathbb{Z}^d} \Phi_I(x) = \sum_{k \in \mathbb{Z}^d} \varphi\left(\sigma^k(x)\right)$$

of a configuration x is well-defined, though it can possibly take the value $+\infty$. We let Z_Φ denote the set of all configurations x with $E(x) = 0$. We shall assume that $Z_\Phi \neq \emptyset$. An interaction of this type is called a *non-frustrated* interaction [Mię98]. The elements of Z_Φ are called non-frustrated ground configurations, or (more concisely) *null-energy* configurations.

Definition 5.11 (Ground Patterns)**.** A *ground pattern* is a pattern with null energy content, *i.e.* a pattern $w \in \mathcal{A}^I$ with $E_I(w) = 0$. The set of all ground patterns on a window I is denoted by G_I. The elements of Z_Φ are precisely the configurations whose restrictions to any finite window is a ground pattern. ■

The interaction appearing in the construction will be a fault interaction associated to a tileset. Observe that for such an interaction, the non-frustrated ground configurations are precisely the valid tilings, and the ground patterns are precisely the locally admissible tilings on finite windows.

5.3.2 Ubiquity of Ground Markers at Low Temperatures

For the rest of this section, we let

$$\alpha := \min\left\{\Phi_I(x), x \in \Omega_{\mathcal{A}}, I \Subset \mathbb{Z}^d, \Phi_I(x) > 0\right\} .$$

We shall assume that Φ is not constantly null, thus the latter set is non-empty and we have $\alpha > 0$. In the case of a fault interaction, $\alpha = 1$. Note that, whenever $w \in \mathcal{A}^I \setminus G_I$, we have $E_I(w) \geq \alpha$.

Lemma 5.12 (High Probability of Ground Patterns)**.** *Let $I \Subset \mathbb{Z}^d$ a finite window and $\varepsilon > 0$, and define the threshold $\beta_0 := \frac{\log_2(|\mathcal{A}|)}{\alpha\varepsilon}|I|$. Then, for every $\beta \geq \beta_0$ and $\mu \in \mathcal{G}_\sigma(\beta)$, we have $\mu([G_I]) \geq 1 - \varepsilon$.*

Proof. Let $\beta \geq \beta_0$ and $\mu \in \mathcal{G}_\sigma(\beta)$. Since Z_Φ is a non-empty subshift, it supports at least one shift-invariant measure. Let $\nu \in \mathcal{M}_\sigma(Z_\Phi)$ be such a measure. Then, $\nu(\varphi) = 0$, hence:

$$0 \leq h(\nu) = h(\nu) - \beta\nu(\varphi) = p_\nu(\beta) \leq p_\mu(\beta) = h(\mu) - \beta\mu(\varphi) \leq \log_2(|\mathcal{A}|) - \beta\mu(\varphi) .$$

It follows $\mu(\varphi) \leq \frac{\log_2(|\mathcal{A}|)}{\beta}$. Using the Markov inequality, we then have:

$$1 - \mu([G_I]) = \mu(\{x \in \Omega_{\mathcal{A}}, E_I(x) \geq \alpha\}) \leq \frac{1}{\alpha}\mu(E_I) \leq \frac{1}{\alpha}|I|\mu(\varphi) ,$$

which concludes the proof. □

I will now introduce the notion of markers. Let's recall $I_n = [\![0, n-1]\!]^d$ is a hypercube of size n.

Definition 5.13 (Ground Markers)**.** Let $\ell, m \in \mathbb{N}$. An (ℓ, m)*-marker set* is a set $Q \subseteq G_{I_\ell}$ of ground patterns on I_ℓ that satisfies the following conditions:

- *Non-overlapping*: For every $u, v \in Q$ and $k \in \mathbb{Z}^d$, if $u_{I_\ell \cap (k+I_\ell)} = (\sigma^k(v))_{I_\ell \cap (k+I_\ell)}$, then either $k = 0$ or $I_\ell \cap (k + I_\ell) = \emptyset$.
- *Covering*: For every $w \in G_{I_m}$, there is a translation k such that $(\sigma^k(w))_{I_\ell} \in Q$.

If $m = (2 + \tau)\ell - 1$ for some $\tau \geq 0$, we say that Q has *margin factor* τ. ■

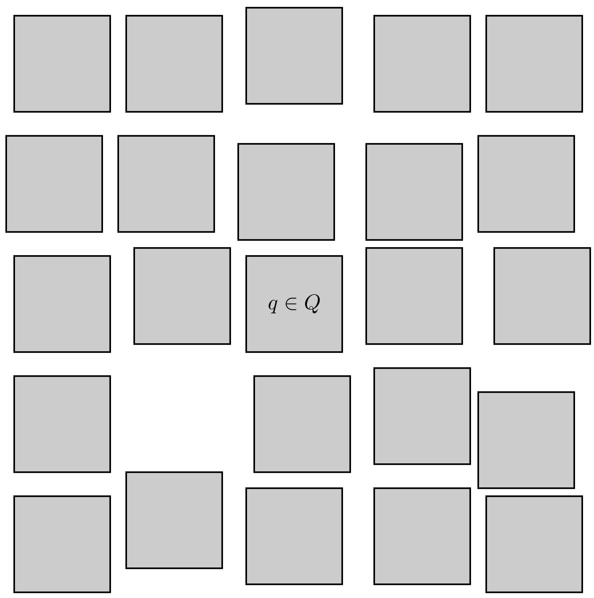

Figure 5.1: A family of markers tiling a high-density portion of a window.

To motivate the latter definition, observe that if $Q \subseteq G_{I_\ell}$ is a marker set with margin factor $\tau = 0$, then every large ground pattern is essentially covered (except for a margin of width $\ell - 1$ along its boundary) by non-overlapping translations of markers from Q. A non-zero margin factor allows for a non-zero space and some wiggle room between the markers (see Figure 5.1).

Lemma 5.14 (Covering Lemma)**.** *Let Q be a marker set with margin factor τ on I_ℓ. For $x \in \Omega_{\mathcal{A}}$, let $f_n(x)$ denote the proportion of x_{I_n} covered by (disjoint) markers from Q. For every $\varepsilon, \delta > 0$, there exists an inverse temperature β_0 (given by the previous lemma, corresponding to I_m and $\frac{\varepsilon}{2}$) such that, for every $\beta \geq \beta_0$ and every ergodic $\mu \in \mathcal{G}_\sigma(\beta)$, there exists an $n_0 \in \mathbb{N}$ such that, for every $n \geq n_0$, we have:*

$$\mu\left(\{x \in \Omega_{\mathcal{A}}, f_n(x) \geq 1 - \varepsilon'\}\right) \geq 1 - \delta\,,$$

with $\varepsilon' = 1 - \frac{1-\varepsilon}{(1+\tau)^d}$.

Proof. Let $\beta_0 = \frac{2\log_2(|\mathcal{A}|)m^d}{\alpha\varepsilon}$. Let $I = I_\ell$ and $J = I_m$. The previous lemma gives $\mu([G_J]) \geq 1 - \frac{\varepsilon}{2}$ for any $\beta \geq \beta_0$ and $\mu \in \mathcal{G}_\sigma(\beta)$.

By the convergence in probability version of the ergodic theorem:

$$\mu\left(\left\{x \in \Omega_{\mathcal{A}}, \left|\mu([G_J]) - \frac{1}{|I_n|}\sum_{k \in I_{n-m+1}} \mathbb{1}_{[G_J]} \circ \sigma^{-k}(x)\right| > \frac{\varepsilon}{2}\right\}\right) \underset{n\to\infty}{\longrightarrow} 0\,.$$

It follows that, after some rank n_0, with μ-probability of at least $1-\delta$:

$$\frac{1}{|I_n|}\sum_{j \in I_{n-m+1}} \mathbb{1}_{[G_J]} \circ \sigma^{-j}(x) \geq 1 - \varepsilon\,.$$

Thus, letting $\mathcal{J}_n(x) := \left\{j \in I_{n-m+1}, \sigma^{-j}(x) \in [G_J]\right\}$, we have $\frac{|\mathcal{J}_n(x)|}{|I_n|} \geq 1 - \varepsilon$ with μ-probability at least $1-\delta$.

Consider a configuration x satisfying $\frac{|\mathcal{J}_n(x)|}{|I_n|} \geq 1 - \varepsilon$. For every $j \in \mathcal{J}_n(x)$, x contains a translation of a pattern from Q within $J + j$. That is, there is an $i \in \mathbb{Z}^d$ such that $I + i \subseteq J + j$ and $\sigma^{-i}(x)_I \in Q$. Such a window $I + i$ can fit in at most

$$[J : I] := \left|\{j, I \subseteq J + j\}\right| = |I_{m-\ell+1}| \leq (1+\tau)^d \ell^d$$

positions inside J.

Let $\mathcal{P}_n(x)$ denote the set of pairs (j,k) with $k \in I + i \subseteq J + j \subseteq I_n$ for some i, such that x_{I+i} is a marker from Q and x_{J+j} is a ground pattern. Let $\mathcal{K}_n(x)$ denote the set of positions $k \in I_n$ that are covered in x by some marker from Q, so that $f_n(x) = \frac{|\mathcal{K}_n(x)|}{|I_n|}$. By double counting the elements of $\mathcal{P}_n(x)$, we get:

$$|I| \times |\mathcal{J}_n(x)| \leq |\mathcal{P}_n(x)| \leq |\mathcal{K}_n(x)| \times [J : I]\,.$$

The left inequality comes from the fact that for each position $j \in \mathcal{J}_n(x)$, x contains a marker from Q on $J + j$, hence there are at least $|I|$ possible values k for which $(j,k) \in \mathcal{P}_n(x)$. To see why the inequality on the right holds, observe that every $k \in \mathcal{K}_n(x)$ is covered by a unique marker in x (recall that markers do not overlap!), hence a unique i such that x_{I+i} is a marker and $k \in I + i$. In turn, for every such i, there exists at most $[J : I]$ values for j such that $I + i \subseteq J + j \subseteq I_n$ and x_{J+j} is a ground pattern.

We conclude that:

$$f_n(x) = \frac{|\mathcal{K}_n(x)|}{|I_n|} \geq \frac{|I|}{[J : I]}\frac{|\mathcal{J}_n(x)|}{|I_n|} \geq \frac{1}{(1+\tau)^d} \times (1 - \varepsilon) = 1 - \varepsilon'\,,$$

for x in a set with μ-probability at least $1 - \delta$. □

Intuitively, for a Gibbs measure μ at low temperature, most of any μ-typical configuration is covered by markers. The lower the temperature, the higher the density regions covered by markers. The next lemma formulates the fact that, when μ is ergodic, among the occurrences of markers in a μ-typical configuration, the relative frequency of each marker follows the law of μ.

Lemma 5.15 (Relative Frequencies)**.** *Let Q be a set of patterns on a window $I \Subset \mathbb{Z}^d$, and $\mu \in \mathcal{M}_\sigma(\Omega_{\mathcal{A}})$ any ergodic measure such that $\mu([Q]) > 0$.*

Then, for every $\delta, \kappa > 0$, there exists an n_0 such that, for every $n \geq n_0$, with μ-probability at least $1 - \delta$, the frequency of a pattern $q \in Q$ among all occurrences of Q (completely) inside I_n is in the interval $[\mu([q]|[Q]) - \kappa, \mu([q]|[Q]) + \kappa]$.

Proof. By the ergodic theorem, $\frac{1}{|I_n|} \sum_{i, I+i \subseteq I_n} \mathbb{1}_{[Q]} \circ \sigma^{-i} \overset{\mu\text{-a.s.}}{\longrightarrow} \mu([Q]) > 0$. This convergence also holds for the cylinder $[q]$, so by taking their ratio we conclude that the frequency of $q \in Q$ satisfies:

$$\frac{\sum_{I+i \subseteq I_n} \mathbb{1}_{[q]} \circ \sigma^{-i}}{\sum_{I+i \subseteq I_n} \mathbb{1}_{[Q]} \circ \sigma^{-i}} \overset{\mu\text{-a.s.}}{\longrightarrow} \frac{\mu([q])}{\mu([Q])} = \mu([q] | [Q]) \,.$$

Almost sure convergence implies convergence in probability, which happens uniformly in the finite set Q, hence the claimed result. □

Lemma 5.15 in particular applies under the hypotheses of Lemma 5.14. The goal of the next two subsections is to provide technical tools to obtain explicit information on the conditional distribution $\mu_Q : q \mapsto \mu([q] | [Q])$.

5.3.3 Tighter Bounds on the Pressure

The variational characterisation of the Gibbs measures offers the following approach to deducing information about them: In order to show that every element of $\mathcal{G}_\sigma(\beta)$ satisfies a property $\star$, one can by contraposition show that if μ does not satisfy $\star$, then the pressure of μ is not maximal, that is, $p_\mu(\beta) < p(\beta)$. Following this general idea, we will later study the property that the empirical distribution of markers is almost uniform. Since the exact value of the pressure is hardly ever explicitly known, using bounds is a reasonable way to proceed.

In Lemma 5.12, we used the trivial pressure bounds $p(\beta) \geq 0$ (which is independent of the parameter β) and $p_\mu(\beta) \leq \log_2(|\mathcal{A}|) - \beta\mu(\varphi)$. The following lemmas provide tighter bounds on $p(\beta)$ and $p_\mu(\beta)$.

Lemma 5.16 (Lower Bound on the Topological Pressure)**.** *Let $n \geq 2r$, and let $S \subseteq \mathcal{A}^{I_n}$ be a non-empty set of patterns. Let $\hat{E} \in \mathbb{R}$ be such that $\max_{q \in S} E_{I_n}(q) \leq \hat{E}$. Then, for every $\beta > 0$:*

$$p(\beta) \geq \frac{\log_2(|S|) - \beta \hat{E}}{|I_n|} - \beta \frac{|I_n| - |I_{n-2r}|}{|I_n|} \|\varphi\| \,,$$

and the term on the right is of order $O\left(\frac{\beta}{n}\right)$.

Proof. We use the fact that $p(\beta) \geq p_\mu(\beta)$ for every measure $\mu \in \mathcal{M}_\sigma(\Omega_\mathcal{A})$. Consider the measure μ defined as follows. First, partition the lattice $\mathbb{Z}^d$ into blocks $(I_n + nk)_{k \in \mathbb{Z}^d}$. Then, for each block, pick a pattern from S uniformly at random and independently of the other blocks. This defines a non-invariant measure $\nu \in \mathcal{M}(\Omega_\mathcal{A})$, but periodic for the shift action. Finally, average this measure ν over its finite translation orbit to obtain $\mu \in \mathcal{M}_\sigma(\Omega_\mathcal{A})$.

The entropy per site of μ is $h(\mu) = \frac{1}{|I_n|} \log_2(|S|)$. The expected energy per site under μ can be bounded as follows:

$$\begin{aligned}
\mu(\varphi) &= \frac{1}{|I_n|} \sum_{i \in I_n} \nu\left(\varphi \circ \sigma^{-i}\right) \\
&= \frac{1}{|I_n|} \sum_{i \in I_n} \sum_{C \ni i} \frac{1}{|C|} \nu\left(\Phi_C\right) \\
&= \frac{1}{|I_n|} \left(\sum_{C \subseteq I_n} \nu\left(\Phi_C\right) + \sum_{C \not\subseteq I_n} \frac{|C \cap I_n|}{|C|} \nu\left(\Phi_C\right) \right) \\
&\leq \frac{1}{|I_n|} \left(\nu\left(E_{I_n}\right) + \sum_{i \in I_n \setminus [\![r, n-r-1]\!]^d} \nu\left(\varphi \circ \sigma^{-i}\right) \right) \\
&\leq \frac{1}{|I_n|} \left(\hat{E} + \left(|I_n| - |I_{n-2r}|\right) \|\varphi\| \right) .
\end{aligned}$$

The claimed lower bound for $p(\beta)$ follows. □

Lemma 5.17 (Upper Bound on the Pressure)**.** *Let $\mu \in \mathcal{M}_\sigma(\Omega_\mathcal{A})$ be an ergodic measure. Let $n_1, n_2, \dots$ be a sequence with $n_i \nearrow \infty$, and for each i, consider a set of patterns $R_i \subseteq \mathcal{A}^{I_{n_i}}$. Let $\gamma > 0$ and $\check{E}_i$ be such that $\mu([R_i]) \geq \gamma$ and $\min_{q \in R_i} E_{I_{n_i}}(q) \geq \check{E}_i$ for every i. Then, for every $\beta > 0$ we have:*

$$p_\mu(\beta) \leq \underline{\lim}_{i \to \infty} \frac{\log_2\left(|R_i|\right) - \beta \check{E}_i}{|I_{n_i}|} \,.$$

Proof. By the Shannon-McMillan-Breiman theorem (with convergence in probability), for every $\alpha_1 > 0$, we have:

$$\mu\left(\left\{x \in \Omega_{\mathcal{A}}, \left|\frac{-\log_2\left(\mu\left([x_{I_n}]\right)\right)}{|I_n|} - h(\mu)\right| > \alpha_1\right\}\right) \underset{n\to\infty}{\longrightarrow} 0\,.$$

Likewise, using the ergodic theorem (with convergence in probability), for every $\alpha_2 > 0$, we have:

$$\mu\left(\left\{x \in \Omega_{\mathcal{A}}, \left|\frac{E_{I_n}(x)}{|I_n|} - \mu\left(\varphi\right)\right| > \alpha_2\right\}\right) \underset{n\to\infty}{\longrightarrow} 0\,.$$

Combining the two, for every $\alpha > 0$, we obtain:

$$\mu\left(\left\{x \in \Omega_{\mathcal{A}}, \left|\frac{-\log_2\left(\mu\left([x_{I_n}]\right)\right) - \beta E_{I_n}(x)}{|I_n|} - p_\mu(\beta)\right| > \alpha\right\}\right) \underset{n\to\infty}{\longrightarrow} 0\,.$$

Now, let $0 < \varepsilon < \gamma$. For n large enough, the latter probability is at most ε, thus, with probability at least $1-\varepsilon$, we have $\mu\left([x_{I_n}]\right) = 2^{(-p_\mu(\beta)\pm\alpha)|I_n| - \beta E_{I_n}(x)}$. Let $R_i' \subseteq R_i$ denote the subset of patterns for which the latter estimate holds. Then, for i big enough, we have:

$$0 < \gamma - \varepsilon \leq \mu\left([R_i']\right) = \sum_{q \in R_i'} \mu([q]) \leq |R_i| \times 2^{-(p_\mu(\beta)-\alpha)|I_{n_i}| - \beta \check{E}_i}\,.$$

This can be rewritten as:

$$p_\mu(\beta) \leq \frac{\log_2\left(|R_i|\right) - \beta \check{E}_i}{|I_{n_i}|} - \frac{\log_2(\gamma-\varepsilon)}{|I_{n_i}|} + \alpha\,.$$

We conclude by taking the $\underline{\lim}$ as $i \to \infty$ and noticing that $\alpha > 0$ is arbitrary. □

5.3.4 Counting Templates

We shall refer to the sets R_i used in Lemma 5.17 as *templates*. In order to apply the lemma, we will need appropriate choices for templates. Furthermore, we need a lower bound $\check{E}_i$ for the energy contents of the elements of R_i, and an upper bound for the entropy of R_i. In this section, we introduce an appropriate family of templates in terms of markers, and give a sharp upper bound for their entropies. As it turns out, the lower bound $\check{E}_i = 0$ for the energy contents will be sufficient for us.

Definition 5.18 (Marker Templates)**.** Let Q be a marker set on I_ℓ, $\mathcal{W}$ a set of probability measures on Q, $\varepsilon > 0$ and $n \geq \ell$. Let us define $R_n\left[Q, \mathcal{W}, \varepsilon\right]$ as the set of patterns $p \in \mathcal{A}^{I_n}$ such that:

- p is ε-covered by markers, that is, $\left|\left\{k \in I_n, \exists i \text{ s.t. } k \in I_\ell + i \subseteq I_n, p_{I_\ell+i} \in Q\right\}\right| \geq (1-\varepsilon)|I_n|$,
- the empirical distribution of markers from Q covering p belongs to $\mathcal{W}$.

We call $R_n\left[Q, \mathcal{W}, \varepsilon\right]$ the $(Q, \mathcal{W}, \varepsilon)$-*template* on I_n. ■

Lemma 5.19 (Entropy of Marker Templates)**.** *Let Q a marker set on I_ℓ, and $\mathcal{W}$ a closed set of probability measures on Q. Then, for every $0 < \varepsilon < \frac{1}{2}$ and n sufficiently large:*

$$\frac{1}{|I_n|}\log_2\left(|R_n[Q,\mathcal{W},\varepsilon]|\right) \leq \varepsilon\left(1 + \log_2(|\mathcal{A}|)\right) + H(\varepsilon) + \frac{H^*}{|I_\ell|}\,,$$

where $H(\varepsilon) := -\varepsilon\log_2(\varepsilon) - (1-\varepsilon)\log_2(1-\varepsilon)$ is the binary entropy function, and $H^* = \max_{w\in\mathcal{W}} H(w)$ stands for the maximal entropy among the elements of $\mathcal{W}$.

Proof. Note that the positions of each of the markers can be directly deduced from the set of coordinates not covered by markers. An element of $R_n[Q,\mathcal{W},\varepsilon]$ can thus be described as a set of r non-covered sites (with $r \leq \varepsilon|I_n|$), a symbol from $\mathcal{A}$ for each such site, and a family of $M_r := \frac{|I_n|-r}{|I_\ell|}$ markers from Q whose empirical distribution belongs to $\mathcal{W}$. Hence:

$$|R_n[Q,\mathcal{W},\varepsilon]| \leq \sum_{r \leq \varepsilon|I_n|} \binom{|I_n|}{r} |\mathcal{A}|^r \sum_{\rho\in\mathcal{W}} |T_{M_r}(\rho)|\,,$$

with $T_m(\rho)$ the set of families of elements in Q of length m with empirical distribution ρ. In order for $T_m(\rho)$ to be non-empty, ρ must be a rational measure with common denominator m, of which at most $(m+1)^{|Q|}$ exist.

For such probability measures, a known upper bound [CT06, Theorem 11.1.3] gives $T_m(\rho) \leq 2^{mH(\rho)}$. It follows that:

$$\begin{aligned}|R_n[Q,\mathcal{W},\varepsilon]| &\leq \sum_{r\leq\varepsilon|I_n|}\binom{|I_n|}{r}|\mathcal{A}|^r\,(M_r+1)^{|Q|}\,2^{M_rH^*}\\ &\leq \sum_{r\leq\varepsilon|I_n|}\binom{|I_n|}{r}|\mathcal{A}|^{\varepsilon|I_n|}\,(M_0+1)^{|Q|}\,2^{M_0H^*}\\ &\leq 2^{H(\varepsilon)|I_n|}|\mathcal{A}|^{\varepsilon|I_n|}\,(M_0+1)^{|Q|}\,2^{M_0H^*}\,,\end{aligned}$$

where, in the last step, we have used the general inequality $\sum_{r\leq\varepsilon m}\binom{m}{r}\leq 2^{H(\varepsilon)m}$, which holds whenever $\varepsilon<\frac{1}{2}$ (see [Gal14, Theorem 3.1]). By taking the logarithm, we obtain:

$$\frac{1}{|I_n|}\log_2\left(|R_n[Q,\mathcal{W},\varepsilon]|\right)\leq H(\varepsilon)+\varepsilon\log_2(|\mathcal{A}|)+|Q|\frac{\log_2\left(|I_n|+|I_\ell|\right)-\log_2\left(|I_\ell|\right)}{|I_n|}+\frac{H^*}{|I_\ell|}\,.$$

In particular, as $n\to\infty$, we have $|Q|\frac{\log_2(|I_n|+|I_\ell|)-\log_2(|I_\ell|)}{|I_n|}\to 0$, hence the $\varepsilon\times 1$ term in the announced bound for n large enough. □

5.3.5 Equidistribution for Relative Frequencies of Markers

Exploiting the results of Section 5.3.2 and the bounds in Section 5.3.3, we now want to show that, at moderately low temperatures, markers cover almost all of the lattice, and with an almost uniform empirical distribution among themselves.

More specifically, we already know (from Lemmas 5.14 and 5.15) that below a certain temperature, the markers cover most of the lattice and their empirical distribution is close to the conditional measure μ_Q (recall that $\mu_Q:=\mu([q]|[Q])$). We now identify a lower temperature above which the entropy of μ_Q is close to $\log_2(|Q|)$, the entropy of the uniform distribution λ_Q. Hence, within the designated temperature interval, markers appear in a "disordered" fashion (*i.e.* with more or less uniform frequencies).

Let us emphasize that at much lower temperatures, markers group themselves into "ordered" blocks, hence their frequencies will be far from uniform.

Theorem 5.20 (Equidistribution)**.** *Let Φ be a non-negative, finite-range interaction (of range r) with a non-empty set Z_Φ of null-energy configurations. Suppose that Φ is not identically null, and define the positive constant $\alpha=\min\{\Phi_I(x),\Phi_I(x)>0\}$.*

Let Q be a marker set on $I:=I_\ell$ with margin factor τ (see Definition 5.13), and let $J:=I_m$ with $m=(2+\tau)\ell-1$. Let $\kappa,\varepsilon>0$, $n\geq 2r$, and $\varepsilon':=1-\frac{1-\varepsilon}{(1+\tau)^d}$, and assume $\kappa,\varepsilon'<\frac{1}{2}$. Suppose that the following two criteria are satisfied:

- *Entropy:* $\frac{\log_2(|G_{I_n}|)}{|I_n|}\geq(1-\kappa)\frac{\log_2(|Q|)}{|I|}$,
- *Temperature:* $\frac{2\log_2(|\mathcal{A}|)}{\alpha\varepsilon}|J|\leq\beta\leq\frac{\varepsilon}{\|\varphi\|}\frac{|I_n|}{|I_n|-|I_{n-2r}|}$.

Then, for every $\mu\in\mathcal{G}_\sigma(\beta)$, we have the following properties:

1. *Covering:* $\mu(\{x\in\Omega_\mathcal{A},\exists i\in\mathbb{Z}^d,0\in I+i,x_{I+i}\in Q\})\geq 1-\varepsilon'$,
2. *Uniformity:* $H\left(\mu_Q\right)\geq(1-2\kappa)\log_2(|Q|)-\left[\varepsilon+\varepsilon'\left(1+\log_2(|\mathcal{A}|)\right)+H\left(\varepsilon'\right)\right]\times|I|-H(\kappa)$.

Proof. The covering claim follows directly from Lemma 5.14 for ergodic measures, as long as the lower bound of the temperature criterion is satisfied. The extension to all of $\mathcal{G}_\sigma(\beta)$ then follows by ergodic decomposition, once we recall that $\mathcal{G}_\sigma(\beta)$ is a Choquet simplex with the ergodic Gibbs measures as its extremal points [Rue04, Corollary 3.14].

To prove the second claim, consider $\mathcal{W}:=\{w,\|w-\mu_Q\|_{\mathrm{TV}}\leq\kappa\}$, the κ-neighbourhood of μ_Q with respect to the total variation distance $\|w-w'\|_{\mathrm{TV}}:=\frac{1}{2}\sum_{q\in Q}|w'(q)-w(q)|$. Observe that, if $|w(q)-\mu_Q(q)|\leq\frac{2\kappa}{|Q|}$ for every $q\in Q$, then $w\in\mathcal{W}$.

Let $0<\delta<1$. Using the temperature criterion again, we can apply Lemma 5.14 and 5.15 (for $\frac{2\kappa}{|Q|}$ instead of κ). Reformulating the result in the vocabulary of Section 5.3.4, for each n, we have $\mu(R_n)\geq 1-\delta$ where

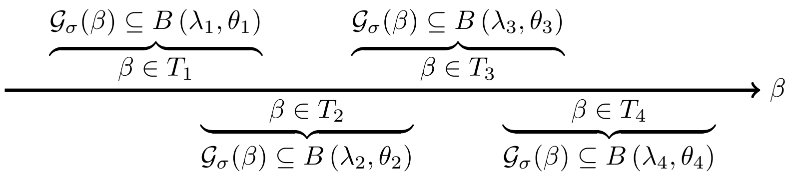

Figure 5.2: Overlapping temperature intervals for Theorem 5.20.

$R_n := R_n[Q, \mathcal{W}, \varepsilon']$ is the $(Q, \mathcal{W}, \varepsilon')$-template. Using these templates in Lemma 5.17, with $\gamma = 1 - \delta > 0$ and $\breve{E}_i = 0$, we have $p(\beta) = p_\mu(\beta) \leq \underline{\lim} \frac{\log_2(|R_i|)}{|I_i|}$. Now, applying Lemma 5.19, we obtain the bound

$$p(\beta) \leq \frac{H^*}{|I|} + \varepsilon' \left(1 + \log_2(|\mathcal{A}|)\right) + H\left(\varepsilon'\right) ,$$

where $H^* = \max_{w \in \mathcal{W}} H(w)$.

On the other hand, using Lemma 5.16, we have:

$$\begin{aligned} p(\beta) &\geq \frac{\log_2\left(|G_{I_n}|\right)}{|I_n|} & - \beta \frac{|I_n| - |I_{n-2r}|}{|I_n|} \|\varphi\| \\ &\geq (1-\kappa) \frac{\log_2(|Q|)}{|I|} - \varepsilon \,, \end{aligned}$$

with the bound on the left term coming from the entropy criterion, and the bound on the right term following from temperature criterion.

Comparing the two bounds on $p(\beta)$ we find that:

$$H^* \geq (1-\kappa) \log_2(|Q|) - \left[\varepsilon + \varepsilon' \left(1 + \log_2(|\mathcal{A}|)\right) + H\left(\varepsilon'\right)\right] \times |I| \,.$$

The uniformity claim now follows from the general bound $|H(w) - H(w')| \leq H(\kappa) + \kappa \log_2(|Q|)$ which holds whenever w and w' are two measures on Q with $\|w - w'\|_{\mathrm{TV}} \leq \kappa \leq \frac{1}{2}$ [Zha07]. □

The goal is to construct a model for which the above theorem can be applied to a sequence of marker sets Q_k at larger and larger scales, and with $\kappa_k, \varepsilon_k \to 0$. The model will be tuned in such a way that the temperature intervals corresponding to consecutive scales overlap, as illustrated in Figure 5.2. This will provide us with control over the Gibbs measures at *all* temperatures, and in turn, allow us to engineer the set of ground states.

In the next section, I construct a class of interactions that achieve the above objective.

5.4 Constructing Well-Behaved Simulating Tilesets

In this section, the goal is to build step by step (and layer by layer) a structure satisfying the requirements of Theorem 5.20, and in which a class of *well-behaved* Turing machines can be encoded. In the following sections, I will use these machines to force specific distributions on words.

A general sketch of the construction will be given in Section 5.4.1, after which each following subsection will focus on one of the layers of the structure.

5.4.1 General Ideas of the Construction

The final tiles have 6 different interdependent layers. At each step of the construction we obtain a non-empty SFT $X_{\mathcal{F}_j}$ on an alphabet $\mathcal{A}_j$, and each layer will be responsible for one aspect of the construction.

- First, we start with a $X_{\mathcal{F}_1}$ that merges the enhanced and uniquely ergodic Red-Black structure on the alphabet $\mathcal{A}_1$, as was the case in the previous chapter. This tileset provides us with self-avoiding Red squares in a hierarchical arrangement, which will allow us to encode increasingly large space-time diagrams of Turing machines, in the sparse area of any Red square that avoids smaller squares. Some particular scales of macro-tiles will later be chosen for the marker sets Q_k.
- I add a three-phased layer $\mathcal{A}_{\text{phase}} := \{\mathtt{F}, \mathtt{B}, \mathtt{H}\}$ (*i.e.* Frozen, Blocking and Hot) and define the tileset $\mathcal{A}_2 \subseteq \mathcal{A}_1 \times \mathcal{A}_{\text{phase}}$ as follows. Tiles with a Red line on the layer $\mathcal{A}_1$ can be anything, other tiles must be either $\mathtt{F}$ or $\mathtt{H}$. For the corresponding local rules $\mathcal{F}_2$, neighbouring tiles linked by a Red line on their

interface, or without any Red line on them, must have the same phase value. For tiles with a Red straight line (setting aside Red corners), as we have a clearly identifiable "inside" and "outside" faces, from outside to inside we only allow the following phase transitions: FFF, HHH and HBF. This last transition represents cases where, in a globally admissible tiling, the Red square is Blocking, so that all of its inside is Frozen while the outside is Hot. In practice, in a typical (locally admissible) Robinson macro-tile, we will have a connected area made of Hot tiles that blur communication, with several Blocking squares, each encoding some signal synchronised among all of its Frozen interior. This behaviour contrasts with generic globally admissible tilings, which will be entirely Frozen.

- The entropy criterion of Theorem 5.20 is easier to satisfy (at increasingly large scales of markers Q_k while maintaining $\kappa_k \to 0$) for positive-entropy tilings (or null-entropy tilings for which $\frac{1}{|I_n|} H\left(G_{I_n}\right) \to h(X) = 0$ as slowly as possible). However, the most naive embeddings of Turing machines into the Robinson tiling give a really tight structure with a fast-decreasing entropy. Thus, to preserve the desired $\kappa_k \to 0$ behaviour, we need to make room for more entropy to spawn. As stated right before, the goal of Frozen squares in this construction is to encode a signal within, and to synchronise it with neighbouring Frozen squares. Because such tiles will only encode very little information (about n bits of internal information and 2^n bits of outside communication with neighbouring squares for 16^n tiles in total), they will be the biggest source of entropy decrease here. To counterbalance this phenomenon, we add a layer $\mathcal{A}_{\text{scale}}$ to obtain $\mathcal{A}_3$ with forbidden patterns $\mathcal{F}_3$ that will limit which scales of Robinson macro-tiles are allowed be Blocking. More precisely, a square *is allowed* to be Blocking (it is *Blockable*) *iff* the scale at which it occurs (N for a $(2N+1)$-macro-tile) is a power of 3 ($N = 3^k$). This will give us the k-th scale of markers:

$$Q_k := \left\{\omega \in G_{I_{l_{n_k}}}, \pi_{\mathcal{A}_1}(\omega) \text{ is a Robinson } \left(2 \times 3^k + 1\right)\text{-macro-tile}\right\}.$$

- Then, a new problem arises in that we eventually want to control the uniform distribution λ_{Q_k} on k-markers. In particular, for the temperature intervals of Theorem 5.20 to be able to overlap from one scale to the next, we need to keep the distributions λ_{Q_k} and $\lambda_{Q_{k+1}}$ as close as reasonably possible. To do so, in the rules $\mathcal{F}_4$ for the alphabet $\mathcal{A}_4$ including a new layer $\mathcal{A}_{\text{odometer}}$, we encode a 2D odometer between Red squares, so that only a low density $\frac{1}{t_k}$ of k-markers within a bigger Hot $(k+1)$-marker are indeed Blocking (and thus perform computations), while the rest stay Hot. We want t_k to be big-enough (to have $\lambda_{Q_{k+1}}$ as close to λ_{Q_k}), but at the same time not too large (so that configurations in $X_{\mathcal{F}_4}$ have a density 1 of Frozen tiles). Choosing $t_k \approx \log_2(k)$ will allow us to satisfy both wanted properties.

- With this structure in mind, we add the last general-purpose layer $\mathcal{A}_{\text{signal}} = \{\texttt{-1}, \texttt{0}, \texttt{+1}\}$ to obtain $\mathcal{A}_5$, which encodes a symbol in each Red line, synchronised in all directions, so that two neighbouring Frozen tiles encode the same bit ±1 in their central cross, but Blocking and Hot lines transmit a 0 (and a Blocking square insulates the ±1 signals on its inside from synchronising with the 0 on its outside). As a globally admissible tiling is generically Frozen everywhere, without an infinite cut, it encodes a single infinite word in $\{\pm 1\}^{\mathbb{N}}$.

  Locally, this synchronising property of $\mathcal{F}_5$ will make it so that the central Blocking Red square of a k-marker will encode a well-defined binary word w of length $3^k - 1$, which can be seen as a (read-only) Toeplitz encoding (meaning that $123\ldots$ reads as $1213121\ldots$, each following element with a period twice as long) for Turing machines operating inside the sparse space-time diagram.

- At last, given a well-behaved Turing machine M (see Definition 5.32), we explain how the tileset $\mathcal{A}_{6(M)}$ is obtained by adding a layer $\mathcal{A}_{\text{simul}(M)}$ to $\mathcal{A}_5$, so that inside each Blocking Red square, computations occur according to M (and the rules $\mathcal{F}_{6(M)}$), in order to force a given value $M(s)$ on the first bits of w conditionally to each possible value of the non-deterministic "random" seed s given to the machine M as an input.

The sum of the results presented in the rest of the section may be summarised as follows:

Theorem 5.21. *Let M be a* well-behaved *Turing machine with two synchronous tapes. We assume that, given a binary seed s on the read-only input tape, M computes a binary word $M(s)$ (on the alphabet $\{\pm 1\}$) of length $\lfloor \log_2(|s|) \rfloor$ with a time complexity $o\left(2^{3^{|s|}}\right)$.*

Then there exists a simulating tileset $X_{\mathcal{F}_{6(M)}}$ on the alphabet $\mathcal{A}_{6(M)}$, such that after some scale $k_0(M)$, a Blocking k-marker checks that the prefix of the word it encodes is $M(s)$, where s is the free seed of the Blocking square.

For this tileset (and the potential associated to it), Theorem 5.20 applies to the marker sets (Q_k), with overlapping intervals for the temperature criterion, and such that $\varepsilon_k, \kappa_k \to 0$.

What's more, there is a common alphabet $\mathcal{A}_0 \subseteq \mathcal{A}_{6(M)}$ independent of M (corresponding to Frozen tiles in which no Turing computations occurs), such that the restriction $\mathcal{F}_0$ of $\mathcal{F}_{6(M)}$ to $\mathcal{A}_0$ doesn't depend on M either, and that $\mathcal{M}_\sigma\left(X_{\mathcal{F}_{6(M)}}\right) = \mathcal{M}_\sigma\left(X_{\mathcal{F}_0}\right)$. This common subshift is "universal", in the sense that we have an affine bijection $\gamma : \mathcal{M}\left(\{\pm 1\}^{\mathbb{N}}\right) \to \mathcal{M}_\sigma\left(X_{\mathcal{F}_0}\right)$.

5.4.2 Layer 1: Robinson Tiles

I already discussed all the basic notions about the Robinson tiling in Section 2.7, and the specific topic of Turing machine simulations in Section 3.5. As announced, in the current construction, I will use the enhanced structure along with the Red-Black structure with Black bumpy-corners, which results in the tileset illustrated in Figure 5.3. Using the same argument as in Proposition 2.36, this SFT $X_{\mathcal{F}_1}$ is uniquely ergodic.

Remark 5.22 (Reminder for Robinson)**.** Considering the highly intricate nature of the current construction, I will remind some basic properties here.

An n-macro-tile tiles a square I_{l_n}, with $l_n := 2^n - 1$ (hence a 2×2^n-periodic behaviour for n-macro-tiles in a globally admissible tiling). Its central square naturally has half the size, $2^{n-1} \pm 1$ (depending on whether we include the border tiles in the square or not).

Each n-macro-tile has a big square in its center, of one colour, and a central cross with a V-shaped "corner" of the other colour. The colours alternate with the parity of n. In particular, whenever $n = 2N + 1$ is odd, then the n-macro-tile has a big Red square and a Black corner in the middle.

At such scales, the Red square contains a $\left(2^N + 1\right) \times \left(2^N + 1\right)$ sparse area, made of rectangle patches avoiding smaller Red squares, connected to each other by communication channels. This sparse area will allow us to encode arbitrarily big spacetime diagrams, which will be of use in one way or another in most of the next layers. ■

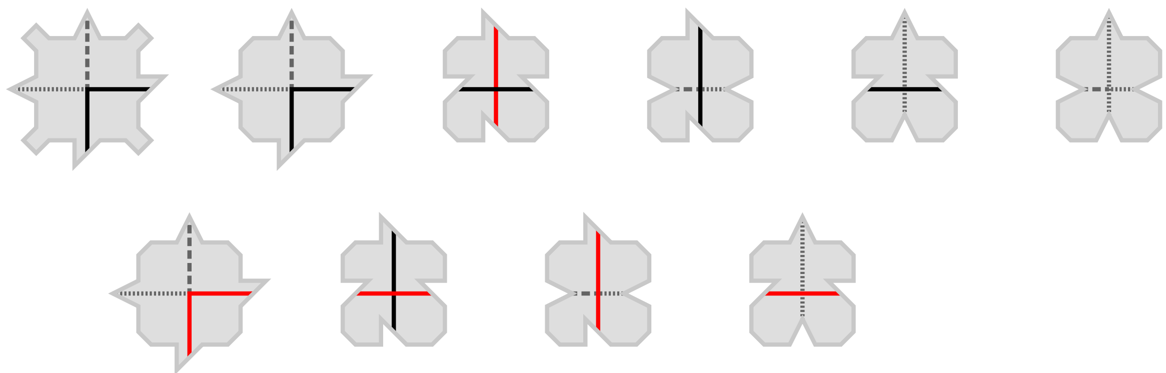

Figure 5.3: Variant of the Robinson tileset,
using two colours for the full lines to obtain non-intersecting squares.
The dotted and dashed lines help to strengthen the structure.

The canonical Robinson tileset is sufficient to obtain the non-overlapping property (of marker sets) for macro-tiles. If we try to use the canonical Robinson tileset to define marker sets, however, then the covering property may be hard to obtain, in comparison to the enhanced case where it will follow from Proposition 4.48 (the proof of which applies verbatim for the current tileset).

With this proposition, to obtain a marker set of macro-tiles, we so far have a covering property that guarantees that an n-macro-tile (defined on a l_n-square) can be encountered in a $3l_n$-square (by peeling at most l_n tiles on each border). This bound is optimal if we want to end up with a grid of markers inside any square, but if we want to obtain just *one* marker in a given square it can be slightly improved:

Lemma 5.23 (Covering Lemma)**.** *Let $n \geq 2$. For any locally admissible tiling $\omega \in \mathcal{A}_1^{I_{2l_n+5}}$ there is a translation $J \subseteq I_{2l_n+5}$ of I_{l_n} such that ω_J is an n-macro-tile.*

The result translates more broadly for admissible tilings on the alphabets $\mathcal{A}_j$ (with $1 \le j \le 6$ and forbidden patterns $\mathcal{F}_j$) if we project $\omega \in \mathcal{A}_j^{I_{2l_n+5}}$ onto its first coordinate $\mathcal{A}_1$.

Proof. Let me remind that, once we have made a square with four n-macro-tiles, the only locally admissible way to fill the central cross is to produce a $(n+1)$-macro-tile. Hence, if we have a 3×3 grid of (well-aligned and well-oriented) n-macro-tiles, we are certain to find among them a square pattern that will actually form a $(n+1)$-macro-tile. More generally, if we start with a grid of $(2u+1) \times (2u+1)$ n-macro-tiles, then we can group them into squares so that we extract a $u \times u$ grid of $(n+1)$-macro-tiles. By a direct induction, starting from a $\left(2^{n-1}-1\right) \times \left(2^{n-1}-1\right)$ grid of (well-aligned and well-oriented) 2-macro-tiles, we can find in it an n-macro-tile (in the form of a 1×1 grid).

Now, according to Proposition 4.48, we can always observe such a grid of 2-macro-tiles in a locally admissible tiling of I_m with

$$m = 4 \times \left(2^{n-1}-1\right) + 7 = 2^{n+1} + 3 = 2l_n + 5\,,$$

which concludes the proof. □

It follows that, if we consider Q the set of n-macro-tiles, then Q is a marker set with margin factor $\tau = \frac{6}{l_n}$. In particular, for reasons that will be explained later on, we define the following marker sets:

$$Q_k := \left\{\omega \in G_{I_{l_{n_k}}}, \pi_{\mathcal{A}_1}(\omega) \text{ is a Robinson } n_k\text{-macro-tile}\right\}\,,$$

with $N_k := 3^k$ and $n_k := 2N_k + 1$, and with margin factor $\tau_k := \frac{6}{l_{n_k}}$. In other words, Q_k is the set of locally admissible tilings of a square $I_{l_{n_k}}$, on the alphabet $\mathcal{A}_j$ (with $j \in [\![1,6]\!]$ depending on the context), for which the layer $\mathcal{A}_1$ contains a well-formed n_k-macro-tile. Hence, the k-th scale of markers correspond the the 3^k-th scale of macro-tiles with a Red square.

5.4.3 Layer 2: Frozen, Blocking, or Hot

As said earlier, for the second layer, we consider $\mathcal{A}_{\text{phase}} = \{\mathtt{F}, \mathtt{B}, \mathtt{H}\}$ a three-phased alphabet, or rather two phases, the Frozen tiles F and Hot tiles H, with Blocking tiles B playing the role of a thin interface between the other two.

We define the tileset $\mathcal{A}_2 \subseteq \mathcal{A}_1 \times \mathcal{A}_{\text{phase}}$ as follows. Tiles with a Red line on the layer $\mathcal{A}_1$ can be anything, but other tiles must be either F or H and cannot be of type B. The B tiles are schematically represented in Figure 5.4, with a red H area on their "outer" side and a blue F area on their "inner" side.

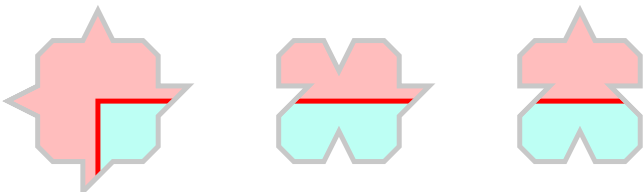

Figure 5.4: Schematic representation of Blocking tiles with a Red line, with a Frozen area on the inside and a Hot area on the outside.

For the local rules, we want neighbouring tiles to have "matching" phases. Hence, neighbouring tiles linked by a Red line on their interface, or without any Red line on them, must have the same phase value. For tiles with a Red straight line (setting aside Red corners), as we have a clearly identifiable "inside" and an "outside" faces, from outside to inside we only allow the following phase transitions: FFF, HHH and HBF. This last HBF case corresponds to the background colours in Figure 5.4.

Consequently, in a macro-tile (or a globally admissible tiling), we will typically have several B Red squares, with all their inside being F, floating around in a connected H puddle, as seen in Figure 5.5. Alternatively, we may have an entirely F tiling. If we halt the construction at this layer, then just by assuming all the scales are Hot except the very first one, we obtain a periodic grid of of squares in 3-macro-tiles, that can each be independently B of H, hence a positive topological entropy. The goal of the next two layers is to control which scales and which tiles can be B, resulting in a return to the zero-entropy case (which makes the entropy criterion in Theorem 5.20 harder to satisfy while still having $\kappa_k \to 0$) but gives us a much more rigid structure as a result (which will allow us to quantify the distribution of the yet-to-be-encoded words).

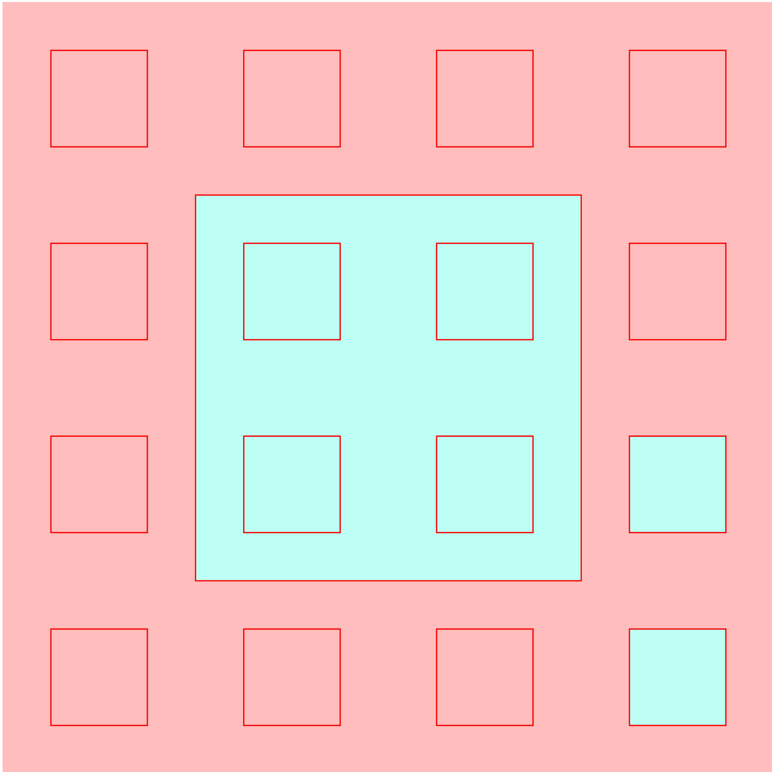

Figure 5.5: Example of admissible arrangement of Blocking squares in a macro-tile.

5.4.4 Layer 3: Forcing the Blockable Scales

As stated in the introductory section, in order for the entropy to decrease to 0 slowly enough, we want to control what scales of macro-tiles are *allowed* to have a B square, which we will call *Blockable* scales. This process is done in two steps, which we will ultimately identify as the new layer $\mathcal{A}_{\text{scale}}$. First, we will use a *sequential* cellular automaton [Wol02] to "obtain" N the step of computation (of a $(2N+1)$-macro-tile). Second, we will use N as an input for a Turing machine to check whether $N = 3^k$ for some $k \in \mathbb{N}$.

Computing the Scale

For this subsection, forget about the previous tiling, and consider the sequential cellular automaton with the transition rules from Figure 5.6.

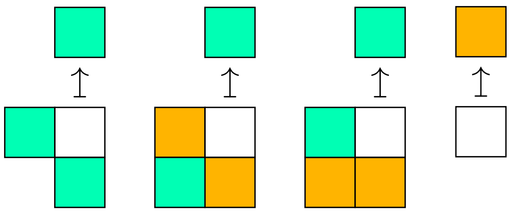

Figure 5.6: Rules of the sequential cellular automaton,
applied from left to right, in a *if/then/else* fashion.

A one-dimensional cellular automaton C can be simulated in a two-dimensional SFT on $\mathcal{A}$, with local rules that determine how to compute $x_{t+1} = C(x_t)$ with $x_t \in \mathcal{A}^{\mathbb{Z}}$. A space-time diagram of the automaton thus corresponds to a tiling of the subshift. Unlike a parallel cellular automaton where each cell's value at time $t+1$ depends on its neighbourhood at time t, for sequential automata we also need to take into account the left-neighbourhood at time $t+1$ to update sequentially all the cells, by doing a kind of left-to-right sweep.

In this specific case, illustrated in Figure 5.6, the value $x_{t+1}(k)$ of the position $k \in \mathbb{Z}$ at time $t+1$ depends on $x_t(k-1)$ and $x_t(k)$ at the previous time, as well as its updated leftmost neighbour $x_{t+1}(k-1)$. Informally, the automaton can be explained as follows. With the leftmost rule, black cells represent static walls, and will stay unchanged from time t to $t+1$ regardless of the left-neighbourhood. The rest of the rules translate the fact that, *if* at time t, the colour of two neighbouring cells goes from green to yellow, *then* at time $t+1$ we force a color switch at the corresponding position (either green-to-yellow or yellow-to-green), *else* the colour stays unchanged. Note that, for a transition $x_{t+1} = C(x_t)$ to be well-defined, one must force an "initialisation" on the left of each cell, which is done through black walls which force their right neighbour to be initialised as yellow at the next step. Hence, if we have a finite window bounded by black cells on both ends, then f_t the number of colour-flips at time t satisfies the recurrence relation $f_{t+1} = \left\lfloor \frac{f_t}{2} \right\rfloor$, as can be seen in Figure 5.7. The following proposition directly follows.

Proposition 5.24. *Consider a bounded initial configuration with a yellow initial block, such that $f_0 = 2^N - 1$, with $N \geq 1$. Then N is written in unary in the rightmost column (which is green up to and including step $N-1$, after which the whole configuration is a stable yellow block).*

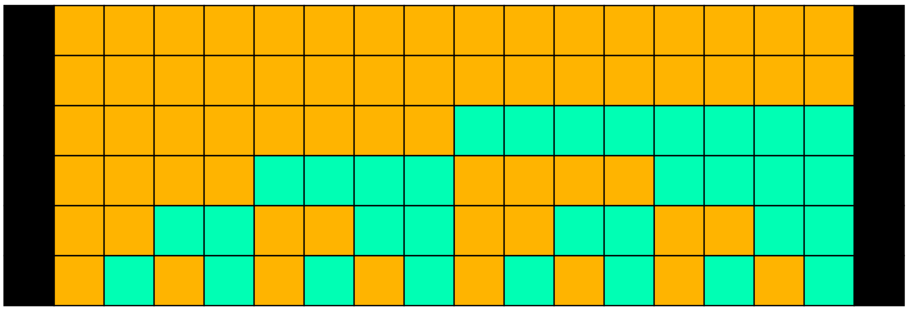

Figure 5.7: Example of spatially bounded space-time diagram, with $f_0 = 15$.

In particular, if we rotate the whole construction clockwise (with time transitions from left to right), adapt the transition rules to work within the sparse computation area of a simulating Red square, initialise the first cell as black and then alternate from yellow to green and forth, we can obtain a unary coding of N on the bottom row of the sparse area of the N-th scale of Red squares, ready to be used by a Turing machine.

Checking the Scale

We have now a $\left(2^N+1\right) \times \left(2^N+1\right)$ sparse computing area (in the N-th Red square), with a unary input N. We want to check whether N is a power of three, in 2^N+1 steps at most. According to Section 3.3, we can do so with time complexity $O(N \ln(N))$. In particular, $N \ln(N) = o\left(2^N+1\right)$, so we guarantee that for big enough values of N the Turing machine will always terminate (and either accept or reject N) fast-enough. Once the machine terminates, we make it go idle so that the decision can then reach the upper border of the bounded spacetime diagram.

Thus, we set-up the tileset so that a Red square is Blockable *iff* the Turing machine *accepts* the input (and forbid it to be Blockable if the machine rejets or simply has not enough time to terminate its computation).

As stated earlier, the set of k-markers Q_k corresponds to the N_k-th scale of Red squares, with an n_k-macro-tile on the Robinson layer and any locally admissible behaviour on the rest. We will say that Q_k corresponds to the k-th scale of Blockable Red squares.

This new layer has two effects on the *number* of k-markers. On one hand, this number is decreased by the mere fact that most Red square are *not* blockable, hence we cut out a source of combinatorial explosion. On the other hand, while this layer acts in a deterministic way inside a (locally admissible) Red square, a k-marker still has a lot of area outside of Red squares, that may be used by bigger markers for their own sparse computation areas, and this adds a new source of entropy to take into account. These effects will be taken into account later on when we estimate the size of Q_k.

Partitioning the Marker Set

From now on, we will in particular distinguish $Q_k = Q_k^{\mathtt{H}} \sqcup Q_k^{\mathtt{F}} \sqcup Q_k^{\mathtt{B}}$, depending on whether the Red square of a k-marker is `H` in the inside (thus on the sparse area outside), `F` (everywhere), or a `B` interface (Hot outside, Frozen inside). We can already prove an upper bound on the cardinality of $Q_k^{\mathtt{F}}$:

Lemma 5.25. *Let $Q_k^{\mathtt{F}}$ be the set of Frozen k-markers, on the alphabet $\mathcal{A}_j$ (with $j \geq 2$ layers). Then we have $|Q_k^{\mathtt{F}}| \leq |\mathcal{A}_j|^{\mathrm{per}_k}$, with $\mathrm{per}_k := 4 \times (l_{n_k} - 1) = 8\left(4^{3^k} - 1\right)$ the perimeter of a k-marker.*

Proof. The hypothesis $j \geq 2$ means that $\mathcal{A}_j$ contains in particular the second layer $\mathcal{A}_{\text{phase}}$, so that the subset $Q_k^{\mathtt{F}} \subseteq Q_k$ is well-defined. By definition of $Q_k^{\mathtt{F}}$, the layer $\mathcal{A}_{\text{phase}}$ is here fixed, entirely equal to F.

For the first layer $\mathcal{A}_1$, the only degree of freedom is the orientation of its central cross, which is fixed by outline of the pattern.

In the case $j \geq 3$, on the layer $\mathcal{A}_{\text{scale}}$, we distinguish two areas. If we are looking within a given complete Red square, in which case the odometer performs a complete computation in a deterministic fashion regardless of the marker $q \in Q_k^{\mathtt{F}}$ we consider. If we are looking at the sparse area outside of the Red squares, then it may be used to perform odometer computations for Red squares at higher scales $k' \geq k$, but then the computations are deterministic conditionally to a labelling of the outline of the marker.

This argument of *fixed conditionally to the outline* (either for the sparse computation area, or for some information encoded in Red squares in a deterministic way or synchronised with neighbouring Red squares at the same scale) will likewise hold for the layers $\mathcal{A}_{\text{odometer}}$, $\mathcal{A}_{\text{signal}}$ and $\mathcal{A}_{\text{simul}(M)}$ introduced in the following subsections, hence the upper bound $|A_j|^{\text{per}_k}$ with per_k the number of tiles in the outline of a k-marker. □

We now want to obtain similar upper bounds for $Q_k^{\mathtt{B}}$ and $Q_k^{\mathtt{H}}$, thus on Q_k. Likewise, we want a *lower* bound for $Q_k^{\mathtt{H}}$, which will allow us to conclude that $\frac{|Q_k^{\mathtt{H}}|}{|Q_k|}$ goes to 1 fast-enough, that $\lambda_{Q_k} \approx \lambda_{Q_k^{\mathtt{H}}}$. Computing these will require a much more intricate analysis (and an accordingly intricate structure), which is why we need to wait until all the layers are properly defined (*i.e.* the case $j = 6$) until we can make sense of it.

5.4.5 Layer 4: Controlling the Density of Blocking Squares

On this new layer, we want to set-up a communication at the scale of k-markers between *neighbouring* Red squares, in order to control precisely which Blockable tiles are in H or B, in particular within a bigger marker. To do so, we will implement an odometer with period $t_k := 2^{\lfloor \log_2(\lfloor \log_2(k) \rfloor) \rfloor} - 1$ at the scale of k-markers, using the sparse communication channels *in-between* neighbouring squares of a given scale, in all four directions, in such a way that:

- A Red Blockable non-F square receives on the bottom a unary coding of an integer $i \in \mathbb{N}$ (within the sparse area not already used by smaller markers).
- This square computes t_k using the unary input N_k from the previous layer (also sparse), and then compares it to the input i so that:
  - if $i < t_k$, then the square is forced to be H, and the square outputs $i + 1$ above,
  - if $i = t_k$, then the square is forced to be B, and outputs 0 above,
  - if $i > t_k$, then the square *rejects* the input, which results in an invalid tiling.
- This unary integer i is "mirrorred" on the left and right sides of the square, and must match the input of the left and right neighbouring non-Frozen tiles.
- The Blockable square of a non-F k-marker forcefully initialises at 0 the input of any smaller l-marker (for scales $l < k$) directly next to its top border.
- Nothing can be transmitted *through* F markers, except in their sparse area that may be used by higher-scale B markers.

All these properties are illustrated in Figure 5.8 where we see, schematically, a B $(k+1)$-marker on the bottom, which both transmits the integer 0 to the next $(k+1)$-marker above, forcing it to be H, as well as initialising the k-markers next to its top side to 0 too. This signal then cycles periodically modulo $(t_k + 1)$ within each column of k-markers. The signal also synchronises laterally, so that the columns on the sides behave correctly even though they don't have an initialisation to 0 within them.

In order for this process to work correctly, we thus need to be able to subdivide a $(k+1)$-marker into a grid of $(t_k + 1) \times (t_k + 1)$ blocks of k-markers, that align well with the Red square. As the Red square itself is a grid of k-markers of length $\frac{2^{2\times 3^{k+1}}}{2^{3^k+1}} = 2^{5\times 3^k - 1}$, it suffices to have a smaller power of 2 for t_k in order to obtain the wanted periodic behaviour. Hence, the choice $t_k = 2^{\lfloor \log_2(\lfloor \log_2(k) \rfloor) \rfloor} - 1 \approx \log_2(k)$ is compatible with these constraints. Assume for now that the tiles can compute t_k and compare it to i fast-enough. The following lemma directly follows.

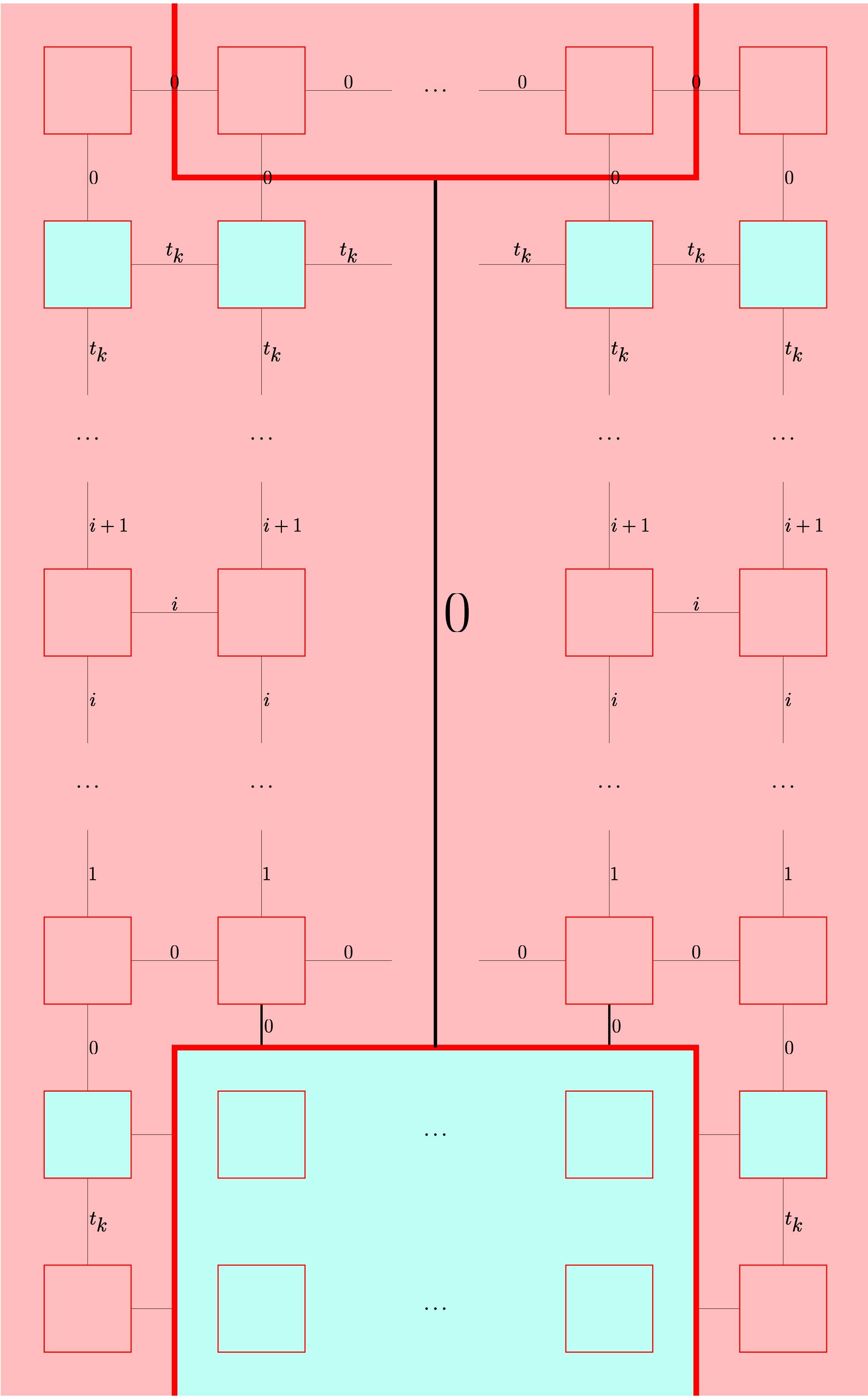

Figure 5.8: Structure of the odometer at the scale of k-markers.

Lemma 5.26. *Consider $\omega \in Q^{\mathtt{H}}_{k+1}$ a* H *$(k+1)$-marker on $\mathcal{A}_j$ $(j \geq 4)$. It can be seen as a grid of k-markers, with a proportion exactly $\frac{1}{t_k+1}$ of* B *k-markers and all the rest being* H *ones.*

Proposition 5.27. *Denote $\mathrm{freq}^{\mathtt{B},k}_{\mathtt{F},l}$ (resp. $\mathrm{freq}^{\mathtt{H},k}_{\mathtt{F},l}$) the proportion of l-markers in a* B *(resp.* H*) k-marker that are* F*. We naturally have the initialisation $\mathrm{freq}^{\mathtt{H},k+1}_{\mathtt{F},k} = 0$, as a* H *$(k+1)$-marker is a grid of* H *and* B *k-markers. Then, for $k \geq l$:*

$$\begin{cases} \mathrm{freq}^{\mathtt{H},k+1}_{\mathtt{F},l} = \dfrac{1}{t_k+1}\mathrm{freq}^{\mathtt{B},k}_{\mathtt{F},l} + \dfrac{t_k}{t_k+1}\mathrm{freq}^{\mathtt{H},k}_{\mathtt{F},l}\,, \\ \mathrm{freq}^{\mathtt{B},k}_{\mathtt{F},l} = \dfrac{1}{4} + \dfrac{3}{4}\mathrm{freq}^{\mathtt{H},k}_{\mathtt{F},l}\,. \end{cases}$$

It follows that $1 - \mathrm{freq}^{\mathtt{H},k+1}_{\mathtt{F},l} = \left(1 - \frac{1}{4(t_k+1)}\right)\left(1 - \mathrm{freq}^{\mathtt{H},k}_{\mathtt{F},l}\right)$. Thus, by induction,

$$\mathrm{freq}^{\mathtt{H},l+j+1}_{\mathtt{F},l} = 1 - \prod_{i=1}^{j}\left(1 - \frac{1}{4\left(t_{l+i}+1\right)}\right)\,.$$

What's more:

$$\prod_{i=j}^{k-1}\left(1 - \frac{1}{4\left(t_i+1\right)}\right) \leq \exp\left(-\frac{1}{4}\sum_{i=j}^{k-1}\frac{1}{t_i+1}\right) \leq \exp\left(-\frac{1}{4}\int_j^k \frac{1}{\log_2(x)}\mathrm{d}x\right)$$
$$\leq \exp\left(-\frac{1}{4}\left(\frac{k}{\log_2(k)} - j\right)\right) \to 0\,.$$

Hence $\mathrm{freq}^{\mathtt{H},k}_{\mathtt{F},l} = 1 - o(1)$ when $k \to \infty$ for a fixed l (and even when $l \to \infty$ as long as we stay in the asymptotic regime $l = o\left(\frac{k}{\log_2(k)}\right)$).

Proposition 5.28. *Generically, a (globally admissible) tiling of $X_{\mathcal{F}_j}$ (with $j \geq 4$) is entirely Frozen.*

Proof. Consider $\omega \in X_{\mathcal{F}_j}$ and fix a scale of small l-markers. Structurally, for any choice of k, we can cut ω into an infinite grid of k-markers. It follows from the previous proposition that the proportion of F l-markers in their own infinite grid is $\lim_{k\to\infty} \mathrm{freq}^{*,k}_{\mathtt{F},l} = 1$. Then, notice that the density of this grid of l-markers in $\mathbb{Z}^d$ goes to 1 as $l \to \infty$, so that overall, ω has a density 1 of tiles equal to F on the layer $\mathcal{A}_{\text{phase}}$. □

Corollary 5.29. *The tileset $X_{\mathcal{F}_4}$ is uniquely ergodic.*

Proof. The previous proposition tells us that the layer $\mathcal{A}_{\text{phase}}$ is generically constant equal to F. Likewise, we know that the layer $\mathcal{A}_1$ is uniquely ergodic (in a way that's compatible with the constant behaviour of $\mathcal{A}_{\text{phase}}$). Because any part of the plane is generically inside the Red square of a F marker for big-enough scales, we conclude that likewise nothing happens on the layer $\mathcal{A}_{\text{odometer}}$ generically. The argument for the remaining (third) layer $\mathcal{A}_{\text{scale}}$ is a bit more intricate, but the general idea is that computations within each scale of Red squares are identical regardless of the square we consider, hence we end up with the same well-defined frequencies regardless of the tiling $\omega \in X_{\mathcal{F}_4}$ we consider, which concludes the argument. □

Remark 5.30. The notable consequence of this proof is that, generically, we only observe a strict subset $\mathcal{A}' \subsetneq \mathcal{A}_4$ in configurations, made of F tiles without any odometer structure. Hence, if we restrict $\mathcal{F}_4$ to $\mathcal{F}'$ which uses only forbidden patterns with tiles in $\mathcal{A}'$, then we can conclude that $\mathcal{M}_\sigma\left(X_{\mathcal{F}'}\right) = \mathcal{M}_\sigma\left(X_{\mathcal{F}_4}\right)$ is the same singleton. This remark will later be improved to hold for a whole family of non-uniquely-ergodic tilesets. ■

This generic behaviour is in stark contrast to what we expect from a typical k-marker, to the highly-intricate arrangement of B squares of many scales guided by all the odometers initialised by the big Red square in the center of the marker. This mismatch between the "typical" (uniform) behaviour for k-markers and their behaviour within globally admissible tilings is precisely what will allow us to force highly complex signals in the presence of noise, through computations that will disappear once the amount of noise goes to 0, leaving only the signals at the limit.

We now need to actually justify that non-F k-markers have enough computation time to compute t_k from $N_k = 3^k$ and then compare it to the odometer input i. According to the time complexities in Section 3.3, we can step by step compute $N_k \mapsto k = \lfloor \log_3(N_k) \rfloor$, then $k \mapsto k' := \lfloor \log_2(k) \rfloor$, and finally $k' \mapsto t_k = 2^{\lfloor \log_2(p) \rfloor} - 1$. If $k \leq 1$, we cannot actually have a well defined $t_k \geq 0$ in this case, so we just manually deal with it by setting

$t_0 = t_1 = t_2 = 0$. Each step of this process adds a new layer of logarithm on the complexity, so that $N = 3^k \mapsto t_k$ still has an overall complexity of order $N \log_3(N)$. As long as $4N \leq 2^N + 1$ (*i.e.* $N \geq 3$, thus $k \geq 1$), we can guarantee that the Turing machine has enough time not only to compute t_k, but also compare it with i (in at most N steps) and then update the value of i to either $i+1$ or 0 (in at most N steps). We can simply manually deal with the case $k = 1$ by making this scale non-Blockable instead. Another notable detail is that we will back-propagate the computed value of k within the sparse area, so that the last layer $\mathcal{A}_{\text{simul}(M)}$ can in turn use it as an input for another Turing machine.

With all this being said and done, we are back to a uniquely ergodic zero-entropy situation, but now with some structure allowing computation at carefully chosen places (the `B` squares) in the typical (`H`) k-markers. We now need to encode words in the tiles, so that there is *something* to be computed. This will also serve as a source of "local" entropy (still resulting in zero-entropy tilings at the limit but with a slower decrease rate, with some amount of combinatorial explosion).

5.4.6 Layer 5: Encoding Words

On this new layer, we want a `F` marker to encode a binary word, and to synchronise it with neighbouring `F` markers. This way, we will have many "independent" short words in `B` squares in a big `H` marker, but only *one* infinite word in generic `F` configurations.

To do so, we add a signal `±1` or `0` onto each Red line of the Robinson, with the following rules. Red lines on `F` tiles encode a bit `±1`, while Red lines on non-`F` tiles encode a `0`. The signals are synchronised in all four directions (so that two neighbouring Red `F` squares encode the same bit), but when a Red line crosses a `B` interface, the `H` side will naturally be equal to the neutral symbol `0` while the `F` side will take a binary `±1` value.

Consequently, a `B` k-marker in $Q_k^{\texttt{B}}$ encodes a well-defined binary word (on the alphabet $\{\texttt{±1}\}$) of length $b_{\text{total}}(k) = 3^k - 1$ inside its central Red square. This word can even be accessed by a Turing machine within the sparse area as a Toeplitz encoding (*i.e.* the word *abcde*... reads as *abacabadabac*..., each letter with a doubled period of occurrence) according to Lemma 4.66.

Of course, this encoding spans the whole length of the finite input tape of the Turing machine here, so we cannot even reasonably decode all of it (and write the result at the beginning of the input tape) without running out of time. Hence, in the next section, for the last layer, we will consider Turing machines that only read the first $b_{\text{read}}(k)$ bits of the signal in k-markers. For now, the only thing we require is $b_{\text{read}}(k) \to \infty$ (so that in the thermodynamic framework, asymptotically when $\beta \to \infty$, as the typical scale of k-markers goes to ∞, we control more and more bits of an infinite binary string). However, we want $b_{\text{read}}(k)$ to stay small-enough, so that computing $b_{\text{read}}(k)$ and then decoding these first bits of the Toeplitz input takes only a negligible proportion of the $2^{3^k} + 1$ space-time horizon of the `B` square of a k-marker).

The construction up to the previous layer gave us a uniquely ergodic structure. Now, with this layer, the encoded words add a degree of freedom that breaks unique ergodicity. However, we still have a well-understood structure. This time, we don't have a uniquely ergodic system, but we may identify ergodic measures with infinite binary strings. This can be formalised as follows:

Proposition 5.31. *Let $\mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right)$ be the set of probability measures on infinite binary words. We have an affine injective map $\gamma : \mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right) \to \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}_5}\right)$ such that $\operatorname{Im}(\gamma) = \mathcal{M}_\sigma\left(X_{\mathcal{F}_5}\right)$.*

*In particular, we have a formal correspondence between (ergodic) extremal measures in $\mathcal{M}_\sigma\left(X_{\mathcal{F}_5}\right)$ and extremal measures $\mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right)$ (*i.e.* Dirac measures δ_w with $w \in \{\texttt{±1}\}^{\mathbb{N}}$).*

Proof. Let $\omega \in X_{\mathcal{F}_5}$. Following the tracks of Remark 5.30, generically, ω is all `F`, and induces an ergodic measure when we project it to $\mathcal{A}_4$.

Now, as neighbouring `F` k-markers synchronise their $b_{\text{total}}(k)$ bits `±1`, it follows that the all-`F` generic tiling ω encodes a well-defined infinite word $w \in \{\texttt{±1}\}^{\mathbb{N}}$. Thus, ω has well-defined frequencies and induces a unique shift-invariant measure μ_ω. What's more, μ_ω actually only depends on the *word* w encoded in ω, so we can denote it μ_w.

This gives us a bijective correspondence $\gamma : \delta_w \mapsto \mu_w$ between extremal points of $\mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right)$ and $\mathcal{M}_\sigma\left(X_{\mathcal{F}_5}\right)$. What's more, γ is continuous, using the weak-* topology on both ends. As $\mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right)$ is a Bauer simplex, it follows [Alf71, Theorem II.4.1] that we can extend γ as $\gamma : \mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right) \to \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}_5}\right)$ the announced convex map. ☐

At this point, we are done introducing the structural general-purpose layers, shared by all the class of simulating tilesets we are about to define. At last, we need to properly define the simulating tileset corresponding to a Turing machine M, and quantify its effect on the entropy.

5.4.7 Layer 6: Controlling Words Through Entropy

For a given Turing machine M, we want to add a new layer to $\mathcal{A}_5$, to obtain the complete alphabet $\mathcal{A}_{6(M)}$. We will define the behaviour on the new layer $\mathcal{A}_{\text{simul}(M)}$ as follows. Suppose that M, being given a binary string s as an input, returns a binary string $M(s)$ of length $b_{\text{read}}(|s|)$.

Then, in the sparse area of a $\mathtt{B}$ Red square, we perform the following computations, using four synchronised tapes, three of them being read-only. On the first tape, we can access a unary integer k (given by the layer $\mathcal{A}_{\text{odometer}}$ in the tiling). On the second one, we have access to a binary "random" seed s (random in the sense that it is not fixed by the local rules of the simulating tileset, and be uniformly distributed for the measure λ_{Q_k}). On the third one, M can access the Toeplitz encoding of some binary string (given by the layer $\mathcal{A}_{\text{signal}}$ in the tiling). We then use the fourth tape to perform the following tasks:

- Check that the seed s is of length k (break local rules if not),
- Decode the Toeplitz input into a prefix u of length $b_{\text{read}}(k)$,
- Compute a word $M(s)$ (of length $b_{\text{read}}(k)$),
- Check if $M(s)$ is equal to u (break local rules if not).

Within $\mathtt{F}$ squares, we setup an "empty" computation, and likewise for $\mathtt{H}$ tiles, so that we see this blank symbol on most of the layer $\mathcal{A}_{\text{simul}(M)}$ *except* the sparse computation area of $\mathtt{B}$ squares.

Remember that at the k-th scale of computations, we have a space-time horizon of size $\left(2^{3^k}+1\right)$. Following the ideas of Section 3.3, the first task is just a back-and-forth, and takes more or less $2k = o\left(2^{3^k}\right)$ steps. Likewise, by a similar back-and-forth, the last task takes about $b_{\text{read}}(k) = o\left(2^{3^k}\right)$ steps. By using a copy of the unary string k, we may decode the first $b_{\text{read}}(k) = \lfloor \log_2(k) \rfloor$ bits of the Toeplitz input in $2k\log_2(k) = o\left(2^{3^k}\right)$ steps, by eliminating one digit of k out of two at each step. Hence, as long as $M(s)$ can be computed in $o\left(2^{3^{|s|}}\right)$ steps, then up to hardcoding the behaviour of the computations in the simulating layer $\mathcal{A}_{\text{simul}(M)}$ up to a finite scale k_0, we can guarantee that any $\mathtt{B}$ k-marker can perform all the previous tasks before running out of time. We will explain why this time complexity constraint is not an actual obstruction in Section 5.6.

Definition 5.32 (Well-Behaved Machine)**.** In the context of this paper, we will say that a Turing machine M is *well-behaved* if it satisfies the previous assumptions, *i.e.* it computes a binary signal $M(s)$ of length $b_{\text{read}}(|s|)$ on a read-only input binary seed s (written on a synchronous tape), with time complexity $o\left(2^{3^{|s|}}\right)$. ■

To summarise, this new layer serves to make sure that, in a marker $q \in Q_k^{\mathtt{B}}$ on the alphabet $\mathcal{A}_{6(M)}$, the central Red square encodes a (non-fixed) binary seed s of length k and forces $M(s)$ as the prefix (of length $b_{\text{read}}(k)$) of the word encoded on the layer $\mathcal{A}_{\text{signal}}$. Assuming s follows the uniform distribution U_k on $\{\pm 1\}^k$, we obtain a prefix $M(s)$ distributed according to $M^*(U_k)$. In particular, setting aside the matter of computation time for now, we can realise dyadic measures on $\{\pm 1\}^{b_{\text{read}}(k)}$ with precision $\frac{1}{2^k}$ with this process. Again, the study of which measures can actually be obtained will be done in Section 5.6.

Theorem 5.33. *Regardless of the machine M we consider, there is a common alphabet $\mathcal{A}_0 \subseteq \mathcal{A}_{6(M)}$, such that $\mathcal{F}_{6(M)}$ always restricts to the same set of forbidden patterns $\mathcal{F}_0$ on the alphabet $\mathcal{A}_0$. What's more, we have $\mathcal{M}_\sigma\left(X_{\mathcal{F}_{6(M)}}\right) = \mathcal{M}_\sigma\left(X_{\mathcal{F}_0}\right)$.*

Then, we have an affine injection $\gamma : \mathcal{M}\left(\{\pm 1\}^{\mathbb{N}}\right) \to \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}_0}\right)$ with $\operatorname{Im}(\gamma) = \mathcal{M}_\sigma\left(X_{\mathcal{F}_0}\right)$.

Proof. Let $\mathcal{A}_0$ be the common subset of $\mathtt{F}$ tiles that contain an empty symbol on the layers $\mathcal{A}_{\text{odometer}}$ and $\mathcal{A}_{\text{simul}(M)}$, and $\mathcal{F}_0$ the restriction of $\mathcal{F}_{6(M)}$. Using once again the ergodicity argument from Remark 5.30, generically, a tiling $\omega \in X_{\mathcal{F}_{6(M)}}$ is entirely $\mathtt{F}$, so the new layer $\mathcal{A}_{\text{simul}(M)}$ only performs empty computations. In other words, generically, $\omega \in X_{\mathcal{F}_0}$. The rest of the proof follows *exactly* that of Proposition 5.31. □

Notably, as was the case for the odometer structure, the new Turing machine computations can only occur *locally*, in finite markers, but vanish as the scale goes to infinity until we obtain globally admissible tilings without any computation.

At long last, let's bring up again the question of the cardinality of $Q_k^{\mathtt{H}}$ and $Q_k^{\mathtt{B}}$, thus of the set Q_k.

Definition 5.34. Consider a well-behaved machine M and the corresponding tileset $\mathcal{A}_{6(M)}$. Let $\widetilde{Q_k^{\mathtt{H}}}$ (resp. $\widetilde{Q_k^{\mathtt{B}}}$) be the set of tuples of seeds s (of length l in a B square of scale l) and compatible words u (of length $b_{\text{total}}(l)$ in the same square, such that $M(s)$ is a prefix of u of length $b_{\text{read}}(l)$), for *all* the B squares in a H (resp. B) k-marker. ■

Lemma 5.35. *Let* $* \in \{\mathtt{H}, \mathtt{B}\}$. *If* $t_k > 0$, *we have the following bounds:*

$$|\widetilde{Q_k^*}| \leq |Q_k^*| \leq |\widetilde{Q_k^*}| \times |\mathcal{A}_{6(M)}|^{\text{per}_k}.$$

Proof. Structurally, the behaviour within each B square of a given marker is independent, *i.e.* we can freely change the value of the compatible seed-signal (s, u) pair it encodes in a locally admissible way without affecting the rest of the marker. Complementarily, we can change the information flowing *outside* of B squares (*e.g.* the value of the odometer for the central square of a H marker), regardless of the seed-signal pairs encoded in each B square. If we denote $Q_k^{*,\text{sparse}}$ this complementary information, then we have the product decomposition $Q_k^* = \widetilde{Q_k^*} \times Q_k^{*,\text{sparse}}$.

The assumption $t_k > 0$ guarantees that both H and B k-markers exist, hence $1 \leq |Q_k^{*,\text{sparse}}|$. The upper bound $|Q_k^{*,\text{sparse}}| \leq |\mathcal{A}|^{\text{per}_k}$ comes from the same argument as in Lemma 5.25, from the fact that any non-fixed information of the k-marker (except seed-signal pairs) represents a computation spanning the whole tile, fixed by the boundary condition on the perimeter of the k-marker. □

Henceforth, we can shift the focus from the markers Q_k^* to the encoded words $\widetilde{Q_k^*}$, which has the added interest of encoding *only* highly-structured hierarchical behaviour, enforced by the odometers, without any dependence on the alphabet $\mathcal{A}_{6(M)}$. In cases where $t_k = 0$, by construction, there cannot be H k-markers, so the previous lower bound on $Q_k^{\mathtt{H}}$ cannot hold, as we want the hierarchical structure of $\widetilde{Q_k^{\mathtt{H}}}$ to be defined nonetheless in the following upper bound (*i.e.* have $\widetilde{Q_k^{\mathtt{H}}} \neq \emptyset$ in any case, by encoding the tuples H markers *would* encode were they to actually exist). Contrary to what was previously said, for the sake of having simpler computations, we will assume that there is a one bit seed at the scale of 0-markers (instead of none), so that $|\widetilde{Q_0^{\mathtt{B}}}| = 2$. However, the overall order of the following bound still holds with the initial definition for $\mathcal{A}_{6(M)}$ (and more broadly as long $t_k \approx \log_2(k)$ asymptotically, regardless of what happens on the first scales).

Proposition 5.36. *We have the following inductive behaviour:*

- $|\widetilde{Q_{k+1}^{\mathtt{H}}}| = |\widetilde{Q_k^{\mathtt{H}}}|^{256^{3^k} \frac{t_k}{t_k+1}} \times |\widetilde{Q_k^{\mathtt{B}}}|^{256^{3^k} \times \frac{1}{t_k+1}}$,
- $|\widetilde{Q_k^{\mathtt{B}}}| = |\widetilde{Q_k^{\mathtt{H}}}|^{\frac{3}{4}} \times 2^{\rho(k)}$, *with* $\rho(k) = k + b_{\text{total}}(k) - b_{\text{read}}(k)$ *the total number of bits used by the seed-signal pair in the* B *square of a* k*-marker.*

Proof. For the first point, notice that a given H $(k+1)$-marker is just a $16^{3^k} \times 16^{3^k}$ grid of smaller k-markers, either H or B, each one with its own family of B squares encoding independent seed-signal pairs. Hence, we have a total of 256^{3^k} independent k-markers. By construction, the odometer forces a proportion $\frac{1}{t_k+1}$ of them to be B, and the rest H, from which the formula follows.

Likewise, for the second equality, notice that the area outside of the central B square can be subdivided in well-behaved grids of smaller B and H markers, in the same proportions as in a H marker. This outside area occupies $\frac{3}{4}$ of the marker, hence the $|\widetilde{Q_k^{\mathtt{H}}}|^{\frac{3}{4}}$ factor. The second factor follows directly from the definition of $\rho(k)$, of the fact that an input seed s in the B square (of length k) forces the value of the prefix $M(s)$ of length $b_{\text{read}}(k)$ out of the $b_{\text{total}}(k)$ bits of the signal, hence $2^{\rho(k)}$ seed-signal pairs in total. □

Asymptotically, most of the entropy comes from the unchecked bits of the signal, so that $\rho(k) \approx 3^k$. Later on, we will use the brutal bound $\rho(k) \leq l_{n_k}$, which will still prove good-enough for our purposes, and also holds regardless of the length of the seed (both the seed and the signal must fit on the border of the B square).

Proposition 5.37. *We have* $|\widetilde{Q_k^{\mathtt{H}}}| = C_k^{16^{3^k}}$, *with* $2^{4^{-k}} \leq C_k \leq 2$.

Proof. Denote $h_k = \log_2\left(|\widetilde{Q_k^{\mathtt{H}}}|\right)$ and $b_k = \log_2\left(|\widetilde{Q_k^{\mathtt{B}}}|\right)$. The sets are non-empty by definition so that the logarithms are always well-defined, non-negative. We have the system:

$$\begin{cases} h_{k+1} = 256^{3^k}\left(\frac{t_k}{t_k+1}h_k + \frac{1}{t_k+1}b_k\right), \\ b_k = \frac{3}{4}h_k + \rho(k). \end{cases}$$

By replacing b_k in the first line, we obtain a first-order recursion scheme for h_k. In particular, if we normalise $u_k := \frac{h_k}{16^{3^k}}$, then we obtain the induction:

$$u_{k+1} = u_k\left(1 - \frac{1}{4\,(t_k+1)}\right) + \frac{\rho(k)}{(t_k+1)\,16^{3^k}}.$$

In particular, using $C_k := 2^{u_k}$, we will obtain the desired expression for $|\widetilde{Q_k^{\mathtt{H}}}|$.

In order to get the upper bound on C_k, we just notice that $\frac{4t_k+3}{4(t_k+1)} \leq 1$, and $u_0 = 0$, so u_k can be bounded by $u = \sum_{n=1}^{\infty} \frac{\rho(n)}{(t_n+1)16^{3^n}} \leq \sum_{n=1}^{\infty} \frac{4^{3^n}}{1\times 16^{3^n}} \leq 1 < \infty$.

For the lower bound, we use the fact that $u_{k+1} \geq \left(1 - \frac{1}{4(t_k+1)}\right)u_k$, so $u_k \geq u_1 \times \prod_{n=1}^{k-1}\left(1 - \frac{1}{4(t_n+1)}\right)$ by a direct induction. Notice that $b_0 = \rho(0) = 1$ and $t_0 = 0$ so that $u_1 = 16b_0 \geq 1$. Now, using the lower bound $\ln(1-\varepsilon) \geq -4\ln(4)\varepsilon$ for $\varepsilon \in \left[0, \frac{1}{4}\right]$, we can obtain:

$$\begin{aligned} \prod_{n=1}^{k-1}\left(1 - \frac{1}{4\,(t_n+1)}\right) &= \exp\left(\sum_{n=1}^{k-1}\ln\left(1 - \frac{1}{4\,(t_n+1)}\right)\right) \\ &\geq \exp\left(\left(-4\ln(4)\sum_{n=1}^{k-1}\frac{1}{4\,(t_n+1)}\right)\right) \\ &= 4^{-\sum\limits_{n=1}^{k-1}\frac{1}{t_n+1}} \\ &\geq 4^{-k}, \end{aligned}$$

hence $2^{4^{-k}}$ works for the lower bound. □

Note that for an optimal lower bound, we would have to use $\sum_{n=2}^{k-1}\frac{1}{\ln(n)} \sim \mathrm{li}(k) \sim \frac{k}{\ln(k)}$, which would complicate both the proof and the expression of the bound, without improving further results.

This is where we close this section. We will continue studying the size of Q_k in the next section, but this time by relating it explicitly to the hypotheses of Theorem 5.20 from the previous section.

5.5 Uniform Marker Distribution for Simulating Tilesets

The goal of this section is to intertwine together the results on the tileset $\mathcal{A}_{6(M)}$, from Section 5.4, with the results on Gibbs measures from Section 5.3, using the potential φ_M associated to $\mathcal{F}_{6(M)}$. As in the previous section, we suppose here that the machine M is well-behaved (see Definition 5.32).

5.5.1 More Marker Bounds for the Entropy Criterion

In this subsection, we state all the counting arguments we will need so that Theorem 5.20 from Section 5.3 applies to the k-markers from Section 5.4 in a useful way.

In the previous section, we computed estimations for $|Q_k^{\mathtt{H}}|$, $|Q_k^{\mathtt{B}}|$ and $|Q_k^{\mathtt{F}}|$. These estimates will allow us to derive some important bounds relating to markers.

First, let us prove that, for the uniform distribution λ_{Q_k}, the typical k-marker is $\mathtt{H}$. This will then allow us to rewrite the entropy criterion of Theorem 5.20 in a more usable way.

Lemma 5.38. *Denote $p_k^{\mathtt{H}} := \lambda_{Q_k}(Q_k^{\mathtt{H}})$. Assume $t_k > 0$, so that Lemma 5.35 applies. Then:*

$$p_k^{\mathtt{H}} \geq 1 - 2^{-\frac{16^{3^k}}{4^{k+1}} + O\left(4^{3^k}\right)}.$$

Proof. We have:

$$p_k^{\mathtt{H}} = \frac{|Q_k^{\mathtt{H}}|}{|Q_k^{\mathtt{H}}| + |Q_k^{\mathtt{B}}| + |Q_k^{\mathtt{F}}|} \geq 1 - \frac{|Q_k^{\mathtt{B}}| + |Q_k^{\mathtt{F}}|}{|Q_k^{\mathtt{H}}|} \geq 1 - \frac{|\mathcal{A}_{6(M)}|^{l_{n_k}} |\widetilde{Q_k^{\mathtt{B}}}| + |\mathcal{A}_{6(M)}|^{4l_{n_k}}}{|\widetilde{Q_k^{\mathtt{H}}}|} .$$

The first inequality comes from $|Q_k| = |Q_k^{\mathtt{H}}| + |Q_k^{\mathtt{B}}| + |Q_k^{\mathtt{F}}| \geq |Q_k^{\mathtt{H}}|$, and for the second one we use Lemma 5.35 to bound $|Q_k^{\mathtt{B}}|$ and Lemma 5.25 to bound $|Q_k^{\mathtt{F}}|$. Then, we use Propositions 5.36 and 5.37:

$$p_k^{\mathtt{H}} \geq 1 - |\mathcal{A}_{6(M)}|^{l_{n_k}} \frac{|\widetilde{Q_k^{\mathtt{H}}}|^{\frac{3}{4}} \times 2^{\rho(k)}}{|\widetilde{Q_k^{\mathtt{H}}}|} - \frac{|\mathcal{A}_{6(M)}|^{4l_{n_k}}}{|\widetilde{Q_k^{\mathtt{H}}}|} \geq 1 - \frac{2|\mathcal{A}_{6(M)}|^{4l_{n_k}} 2^{l_{n_k}}}{2^{4^{-(k+1)} \times 16^{3^k}}} ,$$

hence the announced bound. □

Let's now look at the entropy criterion of Theorem 5.20.

Lemma 5.39. *Assume that $t_k, t_{k+1}, t_{k+2} > 0$. Let $B_k := I_{l_{n_k}}$ be the window where k-markers are defined, $J_k := I_{l_{n_{k+2}}}$ the window for $(k+2)$-markers, and G_k the set of locally admissible tilings of J_k. We have:*

$$\frac{\log_2(|G_k|)}{|J_k|} \geq (1 - \kappa_k) \frac{\log_2(|Q_k|)}{|B_k|} ,$$

with $\kappa_k = \frac{2+o(1)}{\log_2(k)}$.

Proof. Let's first obtain an upper bound for the term on the right, using in particular Lemma 5.35 and Proposition 5.37:

$$\begin{aligned} \log_2(|Q_k|) &= \log_2\left(|Q_k^{\mathtt{H}}|\right) && + \log_2\left(1 + \frac{|Q_k^{\mathtt{B}}| + |Q_k^{\mathtt{F}}|}{|Q_k^{\mathtt{H}}|}\right) \\ &= \log_2\left(|Q_k^{\mathtt{H}}|\right) && + \log_2\left(1 + \frac{1 - p_k^{\mathtt{H}}}{p_k^{\mathtt{H}}}\right) \\ &\leq h_k + \mathrm{per}_k \log_2\left(|\mathcal{A}_{6(M)}|\right) + \frac{1 - p_k^{\mathtt{H}}}{\ln(2) p_k^{\mathtt{H}}} . \end{aligned}$$

This window J_k can notably contain an admissible $\mathtt{H}$ $(k+2)$-marker (as $t_{k+2} > 0$). This marker is a $16^{3^{k+1}} \times 16^{3^{k+1}}$ grid of $(k+1)$-markers, a proportion $\frac{t_{k+1}}{t_{k+1}+1}$ of which are $\mathtt{H}$. Likewise, each such $\mathtt{H}$ $(k+1)$-marker contains $256^{3^k} \times \frac{t_k}{t_k+1}$ $\mathtt{H}$ k-markers, each of them encoding independent seed-signal pairs on their coordinate $\widetilde{Q_k^{\mathtt{H}}}$. It follows that:

$$\begin{aligned} \log_2(|G_k|) &\geq 256^{3^{k+1}} \frac{t_{k+1}}{t_{k+1}+1} \times 256^{3^k} \frac{t_k}{t_k+1} \times h_k \\ &= 256^{4\times 3^k} \times \left(1 - \frac{1}{t_{k+1}+1}\right)\left(1 - \frac{1}{t_k+1}\right) h_k \\ &\geq 256^{4\times 3^k} \times \left(1 - \frac{1}{t_k+1}\right)^2 h_k . \end{aligned}$$

Now, regarding the ratio of the window sizes:

$$\begin{aligned} \frac{|B_k|}{|J_k|} &= \left(\frac{2^{2\times 3^k+1} - 1}{2^{2\times 3^{k+1}+1} - 1}\right)^2 \\ &\geq \left(\frac{2 \times 4^{3^k}}{2 \times 4^{3^{k+2}}}\right)^2 \times \left(1 - \frac{1}{2^{2\times 3^k+1}}\right)^2 \\ &= \frac{1}{256^{4\times 3^k}} \times \left(1 - \frac{1}{2 \times 4^{3^k}}\right)^2 . \end{aligned}$$

By putting together all of these bounds, we obtain:

$$\frac{\frac{\log_2(|G_k|)}{|J_k|}}{\frac{\log_2(|Q_k|)}{|B_k|}} \geq \left(1 - \frac{1}{t_k+1}\right)^2 \left(1 - \frac{\mathrm{per}_k \log_2\left(|\mathcal{A}_{6(M)}|\right) + o(1)}{h_k}\right)\left(1 - \frac{1}{2\times 4^{3^k}}\right)^2 .$$

One can check that the slowest factor here is by orders of magnitude the leftmost one, hence we conclude the announced $\kappa_k = \frac{2+o(1)}{\log_2(k)}$. □

In other words, using the family of marker sets (Q_k) will allow us to satisfy the entropy criterion of Theorem 5.20 with $\kappa_k \to 0$.

5.5.2 Injecting the Marker Bounds in the Equidistribution Result

Here, we provide a restating of Theorem 5.20 into Theorem 5.41, for this specific class of tilesets. First we give a conditional uniformity result which will serve us later on.

Proposition 5.40. *Let $\mu \in \mathcal{G}(\beta)$ be a shift-invariant Gibbs measure, and μ_k the induced conditional measure on $Q_k^{\mathtt{H}}$. Denote $\widetilde{\mu_k}$ be the push-forward measure on $\widetilde{Q_k^{\mathtt{H}}}$. Then $\widetilde{\mu_k}$ is the uniform distribution.*

Proof. We use here the factorisation $Q_k^{\mathtt{H}} = \widetilde{Q_k^{\mathtt{H}}} \times Q_k^{\mathtt{H},\text{sparse}}$ as in Lemma 5.35. Once the coordinate $Q_k^{\mathtt{H},\text{sparse}}$ is fixed, all possible choices of a marker in $Q_k^{\mathtt{H}}$ (*i.e.* of seed-signal pairs in $\widetilde{Q_k^{\mathtt{H}}}$) have the same energy content, so $\widetilde{\mu_k}$ is uniform conditionally to this choice. This is true regardless of the choice, hence $\widetilde{\mu_k}$ is uniform. □

Quite notably, this structural result follows from the way we defined the tileset, and holds at *any* temperature, regardless of the measure μ we consider. The role of Theorem 5.20 is then to provide us with a temperature range where μ_k represents some typical behaviour.

With the associated potential φ_M we use here, we can somewhat simplify some of the notations used in Theorem 5.20, for the condition on β in particular. As these interactions only take the values 0 and 1, with range $r = 1$, we have $\alpha = 1$ and $\|\varphi_M\| = 2$. Assuming we look at the marker set Q_k, then what was called A is here the window B_k. What was called I_n is here the window J_k. As $r = 1$, $|I_n| - |I_{n-2r}|$ is here the perimeter of $J_k = B_{k+2}$, which we denoted $\text{per}_{k+2} \leq 4 l_{n_{k+2}}$. It follows that we can replace $\frac{|I_n|}{|I_n| - |I_{n-2r}|}$ by the slightly tighter upper bound $\frac{l_{n_{k+2}}}{4}$. At last, if we denote $C := 2\log_2\left(|\mathcal{A}_{6(M)}|\right)$, we obtain the following statement:

Theorem 5.41. *Let M be a well-behaved Turing machine, and φ_M the associated potential on $\Omega_{\mathcal{A}_{6(M)}}$. Assume that ε_k and β are such that $\beta \in T_k := \left[\frac{C \times l_{n_k}^2}{\varepsilon_k}, \frac{\varepsilon_k \times l_{n_{k+2}}}{8}\right]$. Denote $0 \triangleleft Q_k^{\mathtt{H}}$ the event where the origin is covered by a $\mathtt{H}$ k-marker. Then, for any $\mu \in \mathcal{G}_\sigma(\beta)$:*

$$\mu\left(\left\{0 \triangleleft Q_k^{\mathtt{H}}\right\}\right) \geq 1 - O\left(4^k \varepsilon_k' \left(1 - \log_2\left(\varepsilon_k'\right)\right)\right) - \underset{k\to\infty}{o}(1)\,.$$

Proof. As proven in the previous section, Q_k satisfies Definition 5.13, it is a marker set with margin factor $\tau_k = \frac{6}{l_{n_k}}$. Hence, it satisfies the entropy criterion with the κ_k computed in Lemma 5.39, and the temperature criterion rewritten as $\beta \in T_k$, so that Theorem 5.20 applies here. Using the first part of the result of the theorem, we have:

$$\begin{aligned}\mu\left(\left\{0 \triangleleft Q_k^{\mathtt{H}}\right\}\right) &= \mu\left(\left\{0 \triangleleft Q_k\right\}\right) \times \mu_{Q_k}\left(Q_k^{\mathtt{H}}\right)\\ &\geq \left(1 - \varepsilon_k'\right) \times \mu_{Q_k}\left(Q_k^{\mathtt{H}}\right)\\ &\geq 1 - \varepsilon_k' - \left(1 - \mu_{Q_k}\left(Q_k^{\mathtt{H}}\right)\right)\,.\end{aligned}$$

To conclude, we now need a bound on this last term. We will use the fact that the convergence of the sequence $\lambda_{Q_k}\left(Q_k^{\mathtt{H}}\right) = p_k^{\mathtt{H}} \to 1$ happens *really* fast. We want to estimate $q_k^{\mathtt{H}} := \mu_{Q_k}\left(Q_k^{\mathtt{H}}\right)$. We can rewrite:

$$\begin{aligned}H\left(\mu_{Q_k}\right) &= H\left(\mathcal{B}\left(q_k^{\mathtt{H}}\right)\right) + q_k^{\mathtt{H}} H\left(\mu_{Q_k^{\mathtt{H}}}\right) &&+ \left(1 - q_k^{\mathtt{H}}\right) H\left(\mu_{Q_k^{\mathtt{B}} \sqcup Q_k^{\mathtt{F}}}\right)\\ &\leq O(1) &&+ q_k^{\mathtt{H}} \log_2\left(|Q_k^{\mathtt{H}}|\right) + \left(1 - q_k^{\mathtt{H}}\right)\left(\log_2\left(|Q_k^{\mathtt{B}}|\right) + o(1)\right)\\ &\leq O(1) &&+ q_k^{\mathtt{H}} \log_2\left(|Q_k^{\mathtt{H}}|\right) + \left(1 - q_k^{\mathtt{H}}\right)\log_2\left(|Q_k^{\mathtt{H}}|\right)\left(\frac{3}{4} + o(1)\right)\,.\end{aligned}$$

As $H\left(\lambda_{Q_k}\right) = \log_2\left(|Q_k|\right) \geq \log_2\left(|Q_k^{\mathtt{H}}|\right)$, we obtain $H\left(\mu_{Q_k}\right) \leq \frac{3 + q_k^{\mathtt{H}}}{4} H\left(\lambda_{Q_k}\right)(1 + o(1))$. On the other hand, using the second part of the result of Theorem 5.20:

$$\begin{aligned}\frac{H\left(\mu_{Q_k}\right)}{H\left(\lambda_{Q_k}\right)} &\geq 1 - 2\kappa_k - \frac{H\left(\kappa_k\right)}{H\left(\lambda_{Q_k}\right)} - \left(\varepsilon_k + O\left(\varepsilon_k'\right) + H\left(\varepsilon_k'\right)\right)\frac{l_{n_k}^2}{H\left(\lambda_{Q_k}\right)}\\ &\geq 1 - o(1) - O\left(\varepsilon_k' \times \left(1 - \log_2\left(\varepsilon_k'\right)\right)\right) \times O\left(4^k\right)\,,\end{aligned}$$

hence the announced bound. □

In particular, to obtain a full control on the low-temperature behaviour of the system, we need the intervals T_k from the temperature criterion for k-markers to overlap from one scale to the next (for this, the bigger ε_k is, the better) while still having $\varepsilon_k \to 0$ fast-enough to guarantee the bound goes to 0 too. A careful study allows us to propose the following tuning of ε:

Corollary 5.42. *Let $\varepsilon_k := \sqrt{\frac{1}{l_{n_k}}}$. Then $4^k \varepsilon'_k \left(1 - \log_2\left(\varepsilon'_k\right)\right) \to 0$. Moreover, the intervals T_k overlap in the sense that $\min T_{k+1} \leq \max T_k$ after a rank, so that $\bigcup T_k$ contains an asymptotic interval $[\beta_0, \infty[$, as illustrated in Figure 5.2.*

Proof. Let us remind that $\varepsilon'_k = 1 - \frac{1-\varepsilon_k}{(1+\tau_k)^2} \leq \varepsilon_k + 2\tau_k$. In particular, $\tau_k = \frac{6}{l_{n_k}} = o\left(\varepsilon_k\right)$ so $\varepsilon'_k \sim \varepsilon_k$, thus we can instead study the asymptotic behaviour of $-4^k \varepsilon_k \log_2\left(\varepsilon_k\right)$. We have:

$$-\log_2\left(\varepsilon_k\right) \sim \frac{1}{2} \log\left(l_{n_k}\right) \sim \frac{n_k}{2} \sim 3^k .$$

On the other hand, $\varepsilon_k = \frac{1}{\sqrt{2 \times 4^{3^k} + 1}} \sim \frac{1}{\sqrt{2} \times 2^{3^k}}$. Thus, the product is equivalent to $\frac{12^k}{\sqrt{2} \times 2^{3^k}} = o(1)$. For the other part of the result, we have:

$$\frac{\max T_k}{\min T_{k+1}} = \frac{\varepsilon_k \times l_{n_{k+2}}}{8} \times \frac{\varepsilon_{k+1}}{C \times l_{n_k}^2} = \frac{1}{8C} \times \varepsilon_k \varepsilon_{k+1} \frac{l_{n_{k+2}}}{l_{n_{k+1}}^2} \sim \frac{1}{8C} 2^{n_{k+2} - \frac{n_k + 5n_{k+1}}{2}} \sim \frac{4^{3^k}}{32C} .$$

This equivalent goes to ∞, so in particular $\max T_k \geq \min T_{k+1}$ after a rank. □

Remark 5.43. The asymptotic behaviour $\frac{\max T_k}{\min T_{k+1}} \to \infty$ (and thus the overlapping property) still holds if we change the common multiplicative constants used to define the intervals (*i.e.* α and $\|\varphi\|$). It follows that we can actually generalise the result to a larger class of potentials associated to the forbidden patterns $\mathcal{F}_{6(M)}$, as long as we still give a positive penalty to each forbidden pattern in the local interactions, but not necessarily zero-one valued. ■

We now want to use this theorem to prove that, in each of the overlapping temperature intervals T_k, we have an increasingly tight bound on the distance of Gibbs measures to some well-behaved measure $\lambda_k \in \mathcal{M}_\sigma\left(X_{\mathcal{F}_{6(M)}}\right)$, induced by the computations of the Turing machine M. This will give us a contracting "tube" around Gibbs measures as $\beta \to \infty$, so that $\mathcal{G}_\sigma(\infty) := \mathrm{Acc}_{\beta\to\infty}\left(\mathcal{G}_\sigma(\beta)\right) = \mathrm{Acc}\left(\lambda_k\right)$.

Before doing so, let's make a brief detour through the question of the induced measures on words, which will hopefully give along the way a mental picture of the multiple behaviours that co-exist at multiple scales in a tiling.

5.5.3 Frequencies of Signals in Markers in Markers in Markers

Denote $m_k = M^*\left(U_k\right)$ the image measure on $\{\pm 1\}^{b_{\mathrm{read}}(k)}$ induced by the uniform distribution U_k on $\{\pm 1\}^k$, which is notably a dyadic measure with precision $\frac{1}{2^k}$.

Theorem 5.33 tells us that, at temperature zero, ground states can be entirely described by a corresponding measure on infinite binary words $\{\pm 1\}^{\mathbb{N}}$. What we need now is a more general way to relate measures in $\mathcal{M}_\sigma\left(\Omega_{\mathcal{A}_{6(M)}}\right)$, and Gibbs states in particular, to measures in $\mathcal{M}_\sigma\left(X_{\mathcal{F}_0}\right)$ (*i.e.* measures on words).

In practice, we will end up with measures on *finite* binary words of various lengths. In order to compare such measures, we may either shorten the length of the encoded words (through projection) or lengthen it (by adding a deterministic string of `+1` symbols at the end).

More precisely, to follow a physical analogy, we are interested at what happens at the *microscopic* scale of the system. At this scale, what we typically observe is a highly ordered grid of small `F` markers, encoding the same signal, of length $b_{\mathrm{read}}(l)$, the distribution of which we want to quantify. At the *macroscopic* end of the spectrum, what we have is a kind of loosely structured (but still dense) arrangement of `H` k-markers. While such markers don't align or synchronise with their neighbours, they have a highly intricate structure, with a complex arrangement of `B` markers of various sizes at an intermediate *mesoscopic* scale. Within each such `B` area, the Turing machine forces the distribution m_k on the signal. There is a non-zero amount of non-`F` l-markers, for which the encoded signal isn't well-defined, but this won't be an issue asymptotically, as the scale k goes to ∞. This idea formalises as follows:

Proposition 5.44. *Let $w_{l,k}$ be the empirical signal distribution (on binary words of length $b_{\text{read}}(l)$) for l-markers in a* `H` *k-marker with uniformly distributed seeds. Note that $w_{l,k}$ is not a* probability *distribution, as there is a positive frequency of H l-markers that do not encode a well-defined signal. Then we initialise $w_{l,l} = 0$ (the* null *measure) and:*

$$w_{l,k+1} = \frac{1}{4(t_k+1)} m_k + \frac{4t_k+3}{4(t_k+1)} w_{l,k} .$$

Proof. The idea of the proof is the same as in Proposition 5.27: we see the `H` $(k+1)$-marker as a grid of k-markers, with a frequency $\frac{1}{t_k+1}$ of `B` ones, and only a quarter of their area actually being inside the `B` square where the computations force the distribution m_k induced by its uniform seed. □

We may translate $w_{l,k}$ into a well-defined probability distribution, of total mass 1, by adding a "trash" state of weight $1 - w_{l,k}\left(\{\pm 1\}^{b_{\text{read}}(l)}\right)$. However, as m_k is a probability distribution with total mass 1, we have $1 - w_{l,k+1}\left(\{\pm 1\}^{b_{\text{read}}(l)}\right) = \prod_{i=l}^{k} \frac{4t_i+3}{4(t_i+1)} \underset{k\to\infty}{\longrightarrow} 0$, *i.e.* the trash state plays no role asymptotically in any case.

Remark 5.45 (Diagonal Convergence)**.** Assume there is a diagonal extraction θ such that, for any $l \in \mathbb{N}$, $\left(w_{l,\theta(k)}\right)_k$ converges to a limit w_l. Then the family (w_l) is compatible with projections, and can uniquely be extended as a measure $w_\infty \in \mathcal{M}\left(\{\pm 1\}^{\mathbb{N}}\right)$. If the measures m_k are constant after some rank (*i.e.* are all projections of some measure on infinite words m_∞), then naturally $w_\infty = m_\infty$. ■

With this multi-scale picture in mind, we now want to actually relate the Gibbs measures to similar measures on the admissible tilings.

5.5.4 Distance Bounds for Gibbs Measures and Random Tilings

To prove the desired bounds, we use here the weak-$*$ distance $d^* = \sum_{n=0}^{\infty} \frac{d_n}{2^{n+1}}$ with:

$$d_n(\mu, \mu') = \sum_{\omega \in \mathcal{A}^{[\![0,n-1]\!]^d}} |\mu([\omega]) - \mu'([\omega])| .$$

In particular, as $d_n \leq d_{n+1} \leq 2$, we obtain $d \leq d_n + \frac{1}{2^n}$.

To obtain the desired tube of measures λ_k, we will step by step jump from Gibbs measures to measures on a grid of `H` macroscopic k-markers (the scale $k(\beta)$ of which will be given by Theorem 5.41, so that $\beta \in T_k$ for the Gibbs measure), then on a grid of `F` microscopic l-markers (at a much lower scale $l(k)$, such that $l \to \infty$), and finally on admissible tilings. In doing so, we will need several times the following lemma:

Lemma 5.46. *Let X be a family of disjoint events, and $d_X(\mu, \mu') := \sum_{A\in X} |\mu(A) - \mu'(A)|$. Consider two probability measures μ and μ', with events U and V such that, for any $A \in X$, $\mu(A|U) = \mu'(A|V)$. Then $d_X(\mu, \mu') \leq 2\left(\mu(U^c) + \mu'(V^c)\right)$.*

Proof. For any $A \in X$, we have:

$$\begin{aligned} |\mu(A) - \mu'(A)| &= |\mu(U)\mu(A|U) - \mu'(V)\mu'(A|V) + \mu(U^c)\mu(A|U^c) - \mu'(V^c)\mu(A|V^c)| \\ &\leq |\mu(U) - \mu'(V)|\mu(A|U) + \mu(U^c)\mu(A|U^c) + \mu(U^c)\mu(A|V^c) . \end{aligned}$$

As the events $A \in X$ are disjoint, by summing over X for each of the conditional distributions $\mu(\cdot|U) = \mu'(\cdot|V)$, $\mu(\cdot|U^c)$ and $\mu'(\cdot|V^c)$, we conclude that:

$$d(\mu, \mu') \leq |\mu(U) - \mu'(V)| + \mu(U^c) + \mu(U^c) = |\mu(U^c) - \mu'(V^c)| + \mu(U^c) + \mu'(V^c) ,$$

thus the announced result. □

The informal interpretation of this result is that, if two measures are identical conditionally to a pair of high-probability events, then they are weak-$*$ close.

Consider a Gibbs measure $\mu \in \mathcal{G}_\sigma(\beta)$. Define then $\overline{\mu_k}$ the probability distribution where we have a grid of well-aligned iid k-markers of law $\mu_k = \mu_{Q_k^{\mathtt{H}}}$, with tiles chosen at random in the one-tile thick grid around the markers, all of this being averaged over the periodic translational orbit.

Proposition 5.47. *Let $\mu \in \mathcal{G}_\sigma(\beta)$, and $\overline{\mu_k}$ the induced distribution on a grid of* `H` *k-markers. Then for any $i \leq l_{n_k}$ we have:*

$$d^*(\mu, \overline{\mu_k}) \leq \frac{1}{2^i} + 4\left(1 - \left(\frac{l_{n_k} - i + 1}{l_{n_k}}\right)^2 \mu\left(0 \lhd Q_k^{\mathtt{H}}\right)\right) .$$

Proof. Let us denote U_i^k the event where the whole window $I_i = [\![0, i-1]\!]^2$ is covered by a single marker from $Q_k^{\mathtt{H}}$. By shift-invariance of μ, conditionally to U_i^k, the content of the window I_i is a uniform compatible translation inside a k-marker (among the $\left(\frac{l_{n_k}-i+1}{l_{n_k}}\right)^2$ such that the square fits in the marker), which itself is chosen with law μ_k independently.

By shift-invariance, for any $0 \leq x, y < l_{n_k}$ (with l_{n_k} the length of a k-marker), we have the same probability of having the upper-right corner of the marker at position (x, y) conditionally to $0 \lhd Q_k^{\mathtt{H}}$, so by counting the number of positions such that $[\![0, i-1]\!]^2 \lhd Q_k^{\mathtt{H}}$ we conclude that, as long as $i \leq l_{n_k}$, we have $\mu\left(U_i^k\right) = \left(\frac{l_{n_k}-i+1}{l_{n_k}}\right)^2 \mu\left(U_1^k\right)$.

Likewise, for $\overline{\mu_k}$, the event U_i^k is realised whenever the window I_i doesn't overlap with the period grid, with probability $\left(\frac{l_{n_k}-i+1}{l_{n_k}}\right)^2$, and induces the same conditional distribution.

By using $X = \left\{[w], w \in \mathcal{A}_{6(M)}^{I_i}\right\}$ in Lemma 5.46, so that $d_X = d_i$, it follows that:

$$
\begin{aligned}
d^*\left(\mu, \overline{\mu_k}\right) &\leq \frac{1}{2^i} + d_i\left(\mu, \overline{\mu_k}\right) \\
&\leq \frac{1}{2^i} + 2\left(2 - \left(\left(\frac{l_{n_k}-i+1}{l_{n_k}}\right)^2 \mu\left(U_1^k\right) - \left(\frac{l_{n_k}-i+1}{l_{n_k}}\right)^2\right)\right) \\
&\leq \frac{1}{2^i} + 4\left(1 - \left(\frac{l_{n_k}-i+1}{l_{n_k}}\right)^2 \mu\left(U_1^k\right)\right),
\end{aligned}
$$

thus announced bound. ☐

Informally, if $\beta \in T_k$, then Theorem 5.41 applies, so that k-markers represent the appropriate macroscopic scale for $\mu \in \mathcal{G}_\sigma(\beta)$. Likewise, we then need a similar argument to compare μ to its microscopic scale, and forth.

Theorem 5.48. *Let $k(\beta)$ be a non-decreasing scale parameter such that $\beta \in T_{k(\beta)}$ in Theorem 5.41. Then we have a sequence of distributions $\lambda_k \in \mathcal{M}_\sigma\left(X_{\mathcal{F}_0}\right)$ such that $d^*\left(\mathcal{G}_\sigma(\beta), \lambda_{k(\beta)}\right) \underset{\beta\to\infty}{\longrightarrow} 0$. It follows that we have $\mathcal{G}_\sigma(\infty) = \mathrm{Acc}\left(\lambda_k\right) = \mathrm{Acc}_{\beta\to\infty}\left(\mu_\beta\right)$ for any trajectory $\left(\mu_\beta\right)_{\beta>0}$.*

Proof. First, consider $\beta \in T_k$, so that Theorem 5.41 applies to $\mu \in \mathcal{G}_\sigma(\beta)$ at the scale of k-markers. Following the notations of the previous proposition, the event U_1^k (*i.e.* $0 \lhd Q_k^{\mathtt{H}}$) has a high probability. In particular, if $i = \underset{k\to\infty}{o}\left(l_{n_k}\right)$ but $i \to \infty$ nevertheless, then the upper bound on $d\left(\mu, \overline{\mu_k}\right)$ goes to 0 as $k \to \infty$.

Likewise, denote μ_k^m the empirical distribution of $\mathtt{F}$ m-markers inside a $\mathtt{H}$ k-marker itself chosen at random with law μ_k. This time, we define the shift-invariant measure $\overline{\mu_k^m}$, with independent m-markers aligned on a smaller grid. Then, we consider the event V_i^m such that $I_i \lhd Q_m^{\mathtt{F}}$. According to Proposition 5.27, which computes the proportion of $\mathtt{F}$ m-markers in a $\mathtt{H}$ k-marker, we conclude that V_i^m has high probability for both $\overline{\mu_k}$ and $\overline{\mu_k^m}$ in the asymptotic regime $m = o\left(\frac{k}{\log_2(k)}\right)$ (we will use $m \approx \log_2\left(\log_2(k)\right)$, for reasons made clear in Proposition 5.50 in the last section). Then, with $i = m = o\left(l_{n_m}\right)$, we have an asymptotic regime such that $d^*\left(\overline{\mu_k}, \overline{\mu_k^m}\right) \to 0$.

Now, the important thing to notice is that, following Proposition 5.40, the law $\widetilde{\mu_k}$ induced by μ_k on $\widetilde{Q_k^{\mathtt{H}}}$ stays the same regardless of the initial measure μ. In turn, this implies that the measure on $\{\mathtt{\pm 1}\}^{b_{\mathrm{read}}(m)}$ induced by μ_k^m is always the same (renormalised) measure $w_{m,k}$ from Proposition 5.44. After extending $w_{m,k}$ as a probability measure on $\{\mathtt{\pm 1}\}^{\mathbb{N}}$ (by adding infinitely many $\mathtt{+1}$ bits), using Proposition 5.31, we can define the corresponding random tiling $\lambda_k^m = \gamma\left(w_{m,k}\right) \in \mathcal{M}_\sigma\left(X_{\mathcal{F}_0}\right)$, uniquely defined as a function of (m, k) regardless of the initial choice of μ.

Notice how the globally admissible tiling used for λ_k^m may affect the content of the sparse communication area of a m-marker in a different way than it does for $\overline{\mu_k^m}$: assuming both m-markers encode the same *word*, the only part where we can guarantee they are identical is inside their Red squares. Thankfully, Lemma 4.61 tells us that such Red squares form a Sierpiński-carpet-like shape, with a proportion of tiles outside of the Red squares of order $O\left(\left(\frac{\sqrt{3}}{2}\right)^n\right)$ in a Robinson n-macro-tile. It follows that, in the right asymptotic regime, a small square window I_i is inside a Red square of a m-marker with high probability for both $\overline{\mu_k^m}$ and λ_k^m, and conditionally to this event we observe the very same empirical behaviour once again. Thus, in the right asymptotic regime we have $d^*\left(\overline{\mu_k^m}, \lambda_k^m\right) \to 0$.

To conclude, in the asymptotic regime, with $\lambda_k := \lambda_k^m$ that doesn't depend on the initial choice of the Gibbs measure μ, we have $d^*\left(\mathcal{G}_\sigma(\beta), \lambda_{k(\beta)}\right) \to 0$. ☐

In this theorem, we used the equivalence between measures on words $w_{m,k}$ (which we will simply denote w_k from now on) and invariant measures on tilings $\lambda_k = \gamma(w_k)$. In particular, because γ is a continuous map, we can now focus on the accumulation set of random words $\text{Acc}(w_k) \subseteq \mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right)$ instead of $\text{Acc}(\lambda_k)$.

5.6 Forcing Π_2 Sets of Ground States to Appear

In Section 5.2, we established that for a computable potential inducing a uniform model, the ground states have a set complexity Π_2. Such sets can in particular be described as accumulation sets of computable sequences. In Section 5.5, we established that the class of simulating tilesets from Section 5.4 always induces a uniform model (stable if $\mathcal{G}_\sigma(\infty)$ is a singleton, uniformly chaotic otherwise).

In this final section, our goal is to use this class of simulating tilesets to realise any connected Π_2 set, through a corresponding computable sequence, as a set of ground states. Such matters are usually studied in the context of computability, not complexity, the question revolving around whether it *can* be done, not if it can be done *rapidly*. Hence, we will use several slowdown arguments, to guarantee that we can not only get increasingly close to the target measures, but also that they can be computed fast-enough, by a well-behaved Turing machine. After this, I will at last conclude with the main result of the chapter.

5.6.1 Slowdown Through Repetition

Remark that we define the weak-$*$ distance d^* on $\mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right)$ using the same formula as in Section 5.5.4 for $\mathcal{M}_\sigma(\Omega_\mathcal{A})$, so that the idea of Lemma 5.46 still applies here.

In a direct sense, the measures on words *simulated* by the Turing machine M are the measures m_k forced within $\texttt{B}$ squares. However, as proven in Theorem 5.48, the measure *induced* by M, the behaviour we typically observe in Gibbs measures at temperatures in the range T_k, is w_k (which is a convex combination of the computed measures) and we have no guarantee that $d^*(m_k, w_k) \to 0$. However, by adding repetitions of each instance of m_k, we obtain the following proposition.

Lemma 5.49. *Assume that* $d^*(m_k, m_{k+1}) \to 0$. *Then* $\text{Acc}([m_k, m_{k+1}]) = \text{Acc}(m_k)$.

Proposition 5.50. *Let M be a well-behaved Turing machine, simulating a sequence (m_k), and such that $d(m_k, m_{k+1}) \to 0$. Then there exists a well-behaved machine M', simulating (m'_k) and inducing (w'_k), such that* $\text{Acc}(m_k) = \text{Acc}(w'_k)$.

Proof. To do so, consider the sequence of scales $s_j = 2^{2^j}$, and define $j(k) := \lfloor \log_2(\lfloor \log_2(k) \rfloor) \rfloor$, such that $s_j \le k < s_{j+1}$. We want a machine M' for which the induced measures (m'_k) satisfy $m'_k = m_{j(k)}$ (on words of length $b_{\text{read}}(k)$, by adding $b_{\text{read}}(k) - b_{\text{read}}(j)$ bits $\texttt{+1}$ at the end). If we have an input seed s, then we have $|s| = k$, so s can act as a read-only unary input to compute j, which can be done in time $o\left(2^{3^k}\right)$ according to Section 3.3. After this, we use $s_{\le j}$ (the first j bits of the seed s) to compute $M(s_{\le j})$, and finally return $M'(s) := M(s_{\le j}) \cdot (\texttt{+1})^{b_{\text{read}}(k) - b_{\text{read}}(j)}$. The machine M' thus defined is indeed well-behaved and adds repetitions to the sequence (m_k).

With these repetitions, we now have $d^*\left(m_k, w'_{s_{k+1}}\right) \to 0$. Indeed, using the same computations as in Proposition 5.44, we conclude that on the first $b_{\text{read}}(j)$ bits we have the following measure inequality:

$$w'_{j,s_{j+1}} \ge \left(1 - \exp\left(-\frac{1}{4}\left(\frac{s_{j+1}}{2^{j+1}} - s_j\right)\right)\right) m_j \,.$$

As $\frac{s_{j+1}}{2^{j+1}} - s_j = s_j\left(2^{2^j - j} - 1\right) \to \infty$, this gives us a high-probability event where $w'_{j,s_{j+1}}$ behaves as m_j, thus using Lemma 5.46 we conclude that their distance goes to 0. Thus, in the asymptotic regime $m = j(k)$ in Theorem 5.48 (and $w'_k := w'_{j(k),k}$), $\text{Acc}_{j\to\infty}(m_j) = \text{Acc}_{j\to\infty}\left(w'_{s_j}\right) \subseteq \text{Acc}_{k\to\infty}(w_k)$.

What's more, the measures $(w'_k)_{s_j \le k \le s_{j+1}}$ are all in the interval $\left[w'_{s_j}, m_j\right]$ by induction, and we have $\left[w'_{s_j}, m_j\right] \approx [m_{j-1}, m_j]$, so $\text{Acc}_{k\to\infty}(w'_k) \subseteq \text{Acc}_{j\to\infty}([m_j, m_{j+1}])$.

By hypothesis, $d(m_k, m_{k+1}) \to 0$ so the previous lemma applies, the chain of inclusions collapses into equalities, thus $\text{Acc}_{k\to\infty}(w'_k) = \text{Acc}_{j\to\infty}(m_j)$. ∎

Remark 5.51. If we remove the $d(m_k, m_{k+1}) \to 0$ hypothesis from the previous proposition, then we conclude instead that $\mathrm{Acc}(w'_k) = \mathrm{Acc}([m_k, m_{k+1}])$, because as $j \to \infty$, the block $(w'_k)_{s_j \leq k \leq s_{j+1}}$ goes through increasingly more tiny steps across $\left[w'_{s_j}, m_j\right] \approx [m_{j-1}, m_j]$. ∎

In particular, as long as (m_j) doesn't converge, M' induces a uniformly chaotic model.

Theorem 5.52. *There exists a (computable, associated to a well-behaved Turing machine) potential inducing a uniformly chaotic model.*

Proof. Consider the sequence of words $u_{2n} = (\texttt{+1})^{b_{\text{read}}(2n)}$ and $u_{2n+1} = (\texttt{-1})^{b_{\text{read}}(2n+1)}$. By simply counting the parity of $|s|$ in linear time and then simply writing this parity bit $b_{\text{read}}(|s|)$ times, we obtain a well-behaved machine M that simulates $m_k = \delta_{u_k}$. Notably, we have $m_{2k} \to \delta_+ := \delta_{(\texttt{+1})^{\mathbb{N}}}$ and $m_{2k+1} \to \delta_- := \delta_{(\texttt{-1})^{\mathbb{N}}}$, and $d^*(\delta_-, \delta_+) > 1$, so we actually fall out of the scope of the previous proposition. However, by following the argument of the remark, we conclude that for the system induced by the machine M', we have a uniform accumulation (according to Theorem 5.48) to the non-singleton set $\mathcal{G}_\sigma(\infty) \cong [\delta_-, \delta_+]$. □

In some sense, this proof transposes the initial idea of Chazottes and Hochman [CH10], *i.e.* forcing the system to oscillate between two distant incompatible states, in another framework.

5.6.2 Complements of Computable Analysis on $\mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right)$

In Section 5.2, we established that any computable potential that induces a uniform model has a Π_2-computable compact and connected set of ground states $\mathcal{G}_\sigma(\infty)$. Then, we established the existence of a class of computable potentials $(\varphi_M)_{M \text{ well-behaved}}$ that induce uniform models, with their ground states as a connected subset of $\mathcal{M}_\sigma(X_{\mathcal{F}_0}) \cong \mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right)$.

To obtain a kind of completeness result, we would like to justify that *any* connected Π_2 subset of $\mathcal{M}_\sigma(X_{\mathcal{F}_0})$ can be realised as a set of ground states. However, by using well-behaved Turing machines, what we may be able to control is the limit set $\mathrm{Acc}(m_j) \subseteq \mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right)$.

In Corollary 3.31 it is established that, whenever two computable spaces are in a bicomputable bijection, they have the same classes of (connected) Π_2-computable compact subsets. Hence, if $\gamma : \mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right) \to \mathcal{M}_\sigma(X_{\mathcal{F}_0})$ (the map from Theorem 5.33) is bicomputable, then we will indeed be able to focus on connected Π_2 sets of measures on binary strings.

First, to avoid ambiguity, let's make the current notion of computation on $\mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right)$ explicit. We use the dyadic base $\mathfrak{D} = \bigcup \mathfrak{D}_k$, with $\mathfrak{D}_k$ the combinations of Dirac measures $\left(\delta_{w\cdot(\texttt{+1})^\infty}\right)_{w \in \{\texttt{±1}\}^{b_{\text{read}}(k)}}$, with weights $\frac{i}{2^k}$. We have a monotonous inclusion $\mathfrak{D}_k \subseteq \mathfrak{D}_{k+1}$ and, as both the length of the words $b_{\text{read}}(k)$ and the power k go to infinity, any dyadic measure is indeed in the sets $\mathfrak{D}_k$ after a rank. Thus defined, $\left(\mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right), d^*, \mathfrak{D}\right)$ is a computable space, and is (locally) recursively compact.

Proposition 5.53. *Let M be a well-behaved machine. The map $\gamma : \mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right) \to \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}_{6(M)}}\right)$ from Theorem 5.33 is computable. What's more, there exists $\gamma' : \mathcal{M}_\sigma\left(\Omega_{\mathcal{A}_{6(M)}}\right) \to \mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right)$ a computable pseudo-inverse, such that $\gamma' \circ \gamma$ is the identity map.*

Proof. As is often the case for computability, doing an exhaustive formal proof would be quite long without giving much insight, so we will focus on the key ideas here.

Notice that the distances d^* are vector norm, and γ is convex, so we just need to computably map Dirac measures $\delta_{w\cdot(\texttt{+1})^\infty} \in \mathfrak{D}_k$ (with $k = |w|$) to periodic points $f(w, r) \in \mathfrak{P}$ (such that $d^*\left(\gamma\left(\delta_{w\cdot(\texttt{+1})^\infty}\right), f(w, r)\right) = o(r)$ uniformly in w). Then, we extend $f(\cdot, r)$ by convexity on each set $\mathfrak{D}_k$. This will give us dyadic combinations of periodic points instead of elements of $\mathfrak{P}$ directly, but we can then computably send those onto elements of $\mathfrak{P}$ at distance r. Now, to compute $f(w, r)$, we just need to return a periodic grid of $\texttt{F}$ $(2|w|+1)$-macro-tiles encoding the signal w.

For the other direction, we want γ' to be a computable "projection" onto $\mathcal{M}_\sigma(X_{\mathcal{F}_0})$. To do so, starting from a periodic point $p \in \mathfrak{P}$, with a target precision ε in mind, we first need to find a value of k high-enough such that the ε-neighbourhood of $\mathfrak{D}_k$ covers $\mathcal{M}\left(\{\texttt{±1}\}^{\mathbb{N}}\right)$. Then, for each element $x \in \mathfrak{D}_k$, we can (uniformly) compute $d^*(\gamma(x), p)$, so we just need to find an approximate minimiser of this distance in $\mathfrak{D}_k$ and return it as an approximation of $\gamma'(p)$. □

The proposition also applies if we consider the common alphabet $\mathcal{A}_0$ and subshift $X_{\mathcal{F}_0}$ instead of those specific to a Turing machine M.

Now, using Corollary 3.31, we can equivalently consider connected Π_2-computable subsets of $\mathcal{M}\left(\{\pm1\}^{\mathbb{N}}\right)$, instead of connected Π_2-computable subsets of $\mathcal{M}_\sigma\left(X_{\mathcal{F}_{6(M)}}\right)$ (computable as subsets of $\mathcal{M}_\sigma\left(\Omega_{\mathcal{A}_{6(M)}}\right)$). Using Proposition 3.25, in both cases, we know that these connected Π_2-computable compact sets K can be realised as $K = \mathrm{Acc}\,(x_n)$ for a computable (x_n) such that $d^*\left(x_n, x_{n+1}\right) \to 0$.

In order to realise any Π_2-computable set, the only gap left is one of complexity. Indeed, we can realise a connected Π_2 subset $X \subseteq \mathcal{M}\left(\{\pm1\}^{\mathbb{N}}\right)$ as an accumulation set of a computable sequence $(m_k) \in \mathfrak{D}^{\mathbb{N}}$, but we haven't explained yet why this sequence may be simulated by a well-behaved machine M. That will be the topic of the next subsection.

5.6.3 Faster Computations Through Inductive Repetitions

Another subtlety we glossed over is the difference between a measure m being simulated into $\mathtt{B}$ squares using a well-behaved machine M, and m being explicitly computed, encoded as an element of $\mathfrak{D}$.

To encode $m \in \mathfrak{D}_k$, *i.e.* a dyadic measure on binary words of length $b_{\text{read}}(k)$ with precision $\frac{1}{2^k}$, we can use the lexicographical order on $\{\pm1\}^{b_{\text{read}}(k)} \equiv \left[\!\left[0, 2^{b_{\text{read}}(k)} - 1\right]\!\right]$ (with $2^{b_{\text{read}}(k)} \approx k$), and write down the corresponding sequence of integer weights in $\left[\!\left[0, 2^k\right]\!\right]$ (using a binary string of length $k+1$ bits for each). Hence, we can encode $m \in \mathfrak{D}_k$ using a string of length $|m| \approx k^2$.

Then, being given this binary string $m \in \mathfrak{D}_k$ and a seed s (with $|s| = k$), we can simply compute the partial sums of weights to compute a corresponding signal prefix $M(s)$, in polynomial time (thus a $o\left(2^{3^k}\right)$), thus simulating m when s is uniformly distributed in $\{\pm1\}^k$.

Proposition 5.54. *Let $X \subseteq \mathcal{M}\left(\{\pm1\}^{\mathbb{N}}\right)$ be a connected Π_2-computable compact. Then there is a well-behaved Turing machine M that simulates measures (m_k) (in $\mathtt{B}$ squares) such that $X = \mathrm{Acc}\,(m_k)$ and $d^*\left(m_k, m_{k+1}\right) \to 0$.*

Proof. As X is a connected Π_2 set, Proposition 3.25 tells us that there is a Turing machine $T : \mathbb{N} \to \mathfrak{D}$ such that $X = \mathrm{Acc}\,(T(k))$ and $d^*\left(T(k), T(k+1)\right) \to 0$.

Without loss of generality, we may assume that $T(k) \in \mathfrak{D}_k$. Indeed, if $T(k) \in \mathfrak{D}_j$ with $j < k$, as $\mathfrak{D}_j \subseteq \mathfrak{D}_k$, we just have to translate its encoding. If $j > k$ then we can computably "project" $T(k) \in \mathfrak{D}_j$ back into $\mathfrak{D}_k$, by first summing the weights of words in $\{\pm1\}^{b_{\text{read}}(j)}$ with the same prefix in $\{\pm1\}^{b_{\text{read}}(k)}$ and then finding the nearest distribution with precision $\frac{1}{2^k}$ instead of $\frac{1}{2^j}$). This projection to a measure in $\mathfrak{D}_k$ doesn't affect the accumulation set $\mathrm{Acc}\,(T(k))$ by density of $\mathfrak{D}$. Remark also that the both the re-encoding and the projection can be done in polynomial time.

Then, let $T'(k)$ be the machine defined inductively, initialised at $T'(0) = \left(\delta_{(+1)^\infty}, 0\right)$ (that returns the only measure of $\mathfrak{D}_0$). To compute $T'(k+1)$, we begin by computing $T'(k) = (m, j)$, and then simulate $T(j+1)$ for k steps. If the new computation terminates, we return $(T(j+1), j+1)$ and else $T'(k)$. This way, T' induces the same accumulation set T does (with distances going to 0), by adding repetitions to the sequence $(T(k))$ but still visiting at least one time each element, and now with a time complexity $O\left(k^2\right)$.

Thence, we can define M that, given a read-only binary seed s, first computes $k = |s|$, then $T'(k) \in \mathfrak{D}_j$ (with $j \leq k$), re-encodes it into $\mathfrak{D}_k$, and finally computes the corresponding signal-prefix $M(s)$ of length $b_{\text{read}}(k)$. It follows that M has a time complexity $o\left(2^{3^{|s|}}\right)$, *i.e.* it is a well-behaved machine that simulates the sequence $\left(m_k := T'(k)\right)_{k \in \mathbb{N}}$ such that $\mathrm{Acc}\,(m_k) = \mathrm{Acc}\,(T(k)) = X$ and $d^*\left(m_k, m_{k+1}\right) \to 0$. ☐

Theorem 5.55. *Let $K \subseteq \mathcal{M}_\sigma\left(X_{\mathcal{F}_0}\right) \cong \mathcal{M}\left(\{\pm1\}^{\mathbb{N}}\right)$ be a connected Π_2-computable compact subset of $\mathcal{M}_\sigma\left(\Omega_{\mathcal{A}_0}\right)$. Then there exists a well-behaved Turing machine M such that, on the extended alphabet $\mathcal{A}_{6(M)} \supset \mathcal{A}_0$, for the associated potential $\varphi_M : \Omega_{\mathcal{A}_{6(M)}} \to \mathbb{R}$, we have $K = \mathcal{G}_\sigma(\infty)$.*

Proof. First, by using the bicomputable map from Proposition 5.53, Corollary 3.31 tells us that the set $X := \gamma'(K) \subseteq \mathcal{M}\left(\{\pm1\}^{\mathbb{N}}\right)$ is also a connected Π_2-computable compact. Then, Proposition 5.54 says that $X = \mathrm{Acc}\,(m_j)$ is the accumulation of a sequence (m_j) simulated by a well-behaved machine M, such that $d^*\left(m_j, m_{j+1}\right) \to 0$. Likewise, Proposition 5.50 allows us to obtain the same statement but with the sequence of empirically induced measures (w_k) instead. Thus, with $\lambda_k = \gamma\left(w_k\right)$, we have $K = \gamma(X) = \mathrm{Acc}\,(\lambda_k)$. Finally, Theorem 5.48 from the previous tells us that, using the associated potential, we have $\mathcal{G}_\sigma(\infty) = \mathrm{Acc}\,(\lambda_k)$, which concludes the proof. ☐

5.6.4 Undecidability of Chaoticity

Let me conclude this section, and the whole chapter, by a small application of the previous framework, that of *undecidability* of the question of chaoticity. By the question of chaoticity, I here mean the decision problem of whether a given computable potential induces a chaotic model, as defined in Section 5.1.3. Among uniform models, such as our class of potentials induced by well-behaved machines, deciding if the system is chaotic is equivalent to deciding if $\mathcal{G}_\sigma(\infty)$ is not a singleton.

Theorem 5.56. *The question of chaoticity is Σ_3-complete.*

Proof. Let's begin with the upper bound. Consider $\varphi : \Omega_{\mathcal{A}} \to \mathbb{R}$ a computable potential. A ground state $\mu \in \mathcal{G}_\sigma(\infty)$ is stable *iff* we have $d^*(\mu, \mathcal{G}_\sigma(\beta)) \to 0$. Likewise, when $\mu \notin \mathcal{G}_\sigma(\infty)$ we clearly have $d^*(\mu, \mathcal{G}_\sigma(\beta)) \not\to 0$. Thus, a model is *chaotic iff* $\forall \mu, D(\mu) := \overline{\lim}_{\beta\to\infty} d^*(\mu, \mathcal{G}_\sigma(\beta)) > 0$. The maps $\mu \mapsto d^*(\mu, \mathcal{G}_\sigma(\beta))$ are all 1-Lipschitz, thus so is D. By compactness of the space of measures, if follows that the model is chaotic *iff* $\inf_{\mu\in\mathcal{M}_\sigma(\Omega_{\mathcal{A}})} D(\mu) > 0$, and we can then restrict this infimum to the dense family $\mathfrak{P}$:

$$\begin{aligned} & \exists\varepsilon > 0, \forall x \in \mathfrak{P}, \forall \beta_0 > 0, \exists \beta \geq \beta_0, \mathcal{G}_\sigma(\beta) \cap \overline{B(x,\varepsilon)} = \emptyset \\ \Leftrightarrow\quad & \exists \varepsilon \in \mathbb{Q}^{+*}, \forall x \in \mathfrak{P}, \forall \beta_0 \in \mathbb{N}, \exists \beta', \beta'' \in \mathbb{Q} \cap [\beta, +\infty[, \mathcal{G}_\sigma([\beta', \beta'']) \cap \overline{B(x,\varepsilon)} = \emptyset\,. \end{aligned}$$

The second formula comes from the fact that we can replace ε and β_0 by countable parameters by monotonicity, and $\mathcal{G}_\sigma(\beta) \cap \overline{B(x,\varepsilon)} = \emptyset$ *iff* it holds on a neighbourhood of β by Lemma 5.2. Since by Proposition 5.8 it is semi-decidable to know if $\mathcal{G}_\sigma([\beta', \beta'']) \cap \overline{B(x,\varepsilon)} = \emptyset$, we deduce that decide the chaoticity is Σ_3-computable.

Now, to prove that the problem is Σ_3-complete, I will make a computable reduction from the *cofinality* problem (*i.e.* whether M doesn't halt on *finitely* many inputs, which is Σ_3-complete) to chaoticity.

Let M a Turing machine (which takes encoded integers as inputs), and denote $\tau(k) \in \overline{\mathbb{N}}$ the time it takes M to compute $M(k)$. First, fix a computable bijective map $f : \mathbb{N} \to \mathbb{N}^3$ such that each coordinate stays bounded by i. Define the computable sequence of measure $(m_i) \in \mathfrak{D}^{\mathbb{N}}$ such that, with $f(i) = (k, \ell, t)$:

$$m_i = \begin{cases} \dfrac{1}{2^k}\delta_{(-1)^i(+1)^\infty} + \left(1 - \dfrac{1}{2^k}\right)\delta_{(+1)^\infty} & \text{if } \max\limits_{i\in[\![k,k+\ell]\!]} \tau(i) = t, \\ \delta_{(+1)^\infty} & \text{else.} \end{cases}$$

Then, we can replace each measure m_i by a path from $\delta_{(+1)^\infty}$ to m_i and forth, using incremental steps of weight $\frac{1}{2^i}$ (which divides $\frac{1}{2^k}$ as $k(i) \leq i$) for the case $m_i = \frac{1}{2^k}\delta_{(-1)^i(+1)^\infty} + \left(1 - \frac{1}{2^k}\right)\delta_{(+1)^\infty}$. This defines a computable sequence of measures $\left(m'_j\right)$ such that $d\left(m'_j, m'_{j+1}\right) \to 0$, and so that the accumulation set $K := \operatorname{Acc}\left(m'_j\right) = \operatorname{Acc}\left(\left[\delta_{(+1)^\infty}, m_i\right]\right) \subseteq \mathcal{M}_\sigma\left(\{\pm 1\}^{\mathbb{N}}\right)$ is a connected Π_2-computable compact. Notably, we always have $\delta_{(+1)^\infty} \in K$.

If M has a cofinite support, then after some rank k it always halt. Thus, for any interval length ℓ, with $i_\ell := f^{-1}\left(k, \ell, \max_{i\in[\![k,k+\ell]\!]} \tau(i)\right) \xrightarrow[\ell\to\infty]{} \infty$, we have $m_{i_\ell} \to \frac{1}{2^k}\delta_{(-1)^\infty} + \left(1 - \frac{1}{2^k}\right)\delta_{(+1)^\infty}$, so K is not a singleton.

Conversely, assume M doesn't have a cofinite support. Then, for any choice of j we have only finitely many intervals $[\![k, k+l]\!]$ such that $k \leq j$ and M halts on all the entries. After some scale i, we thus have $m_i \geq \left(1 - \frac{1}{2^j}\right)\delta_{(+1)^\infty}$. It follows that $m_i \to \delta_{(+1)^\infty}$ and K is a singleton.

The process that maps M to the sequence $\left(m'_j\right)$, and then to the corresponding simulating potential φ_M using Proposition 5.50, is thus a computable reduction from the cofinality problem to the question of whether K is *not* a singleton, *i.e.* of whether the corresponding (uniform) model is chaotic. □

References

[Alf71] Erik Alfsen. *Compact Convex Sets and Boundary Integrals.* Springer, 1971. 10.1007/978-3-642-65009-3.

[Bar+22] Sebastián Barbieri, Rodrigo Bissacot, Gregório Dalle Vedove, and Philippe Thieullen. *Chaos in Bidimensional Models With Short-Range.* 2022. 10.48550/arXiv.2208.10346.

[BM23] Sebastián Barbieri and Tom Meyerovitch. "The Lanford-Ruelle Theorem for Actions of Sofic Groups". In: *Transactions of the American Mathematical Society* 376.2 (2023), pp. 1299–1342. 10.1090/tran/8810.

[Bré03] Julien Brémont. "Gibbs Measures at Temperature Zero". In: *Nonlinearity* 16.2 (2003), pp. 419–426. 10.1088/0951-7715/16/2/303.

[Bur+22] Michael Burr, Suddhasattwa Das, Christian Wolf, and Yun Yang. "Computability of Topological Pressure on Compact Shift Spaces Beyond Finite Type". In: *Nonlinearity* 35.8 (2022), pp. 4250–4282. 10.1088/1361-6544/ac7702.

[CGU11] Jean-René Chazottes, Jean-Marc Gambaudo, and Edgardo Ugalde. "Zero-Temperature Limit of One-Dimensional Gibbs States via Renormalization: the Case of Locally Constant Potentials". In: *Ergodic Theory and Dynamical Systems* 31.4 (2011), pp. 1109–1161. 10.1017/S014338571000026X.

[CH10] Jean-René Chazottes and Michael Hochman. "On the Zero-Temperature Limit of Gibbs States". In: *Communications in Mathematical Physics* 297.1 (2010), pp. 265–281. 10.1007/s00220-010-0997-8.

[CS20] Jean-René Chazottes and Mao Shinoda. *On the Absence of Zero-Temperature Limit of Equilibrium States for Finite-Range Interactions on the Lattice* $\mathbb{Z}^2$. 2020. 10.48550/arXiv.2010.08998.

[CT06] Thomas Cover and Joy Thomas. *Elements of Information Theory*. 2nd edition. Wiley, 2006. 10.1002/047174882X.

[DGS76] Manfred Denker, Christian Grillenberger, and Karl Sigmund. *Ergodic Theory on Compact Spaces*. Vol. 527. Lecture Notes in Mathematics. Springer, 1976. 10.1007/BFb0082364.

[EFS93] Aernout van Enter, Roberto Fernández, and Alan Sokal. "Regularity Properties and Pathologies of Position-Space Renormalization-Group Transformations: Scope and Limitations of Gibbsian Theory". In: *Journal of Statistical Physics* 72.5–6 (1993), pp. 879–1167. 10.1007/BF01048183.

[Gal14] David Galvin. *Three Tutorial Lectures on Entropy and Counting*. 2014. 10.48550/arXiv.1406.7872.

[GST23] Léo Gayral, Mathieu Sablik, and Siamak Taati. *Characterisation of the Set of Ground States of Uniformly Chaotic Finite-Range Lattice Models*. 2023. 10.48550/arXiv.2302.07326.

[HM10] Michael Hochman and Tom Meyerovitch. "A Characterization of the Entropies of Multidimensional Shifts of Finite Type". In: *Annals of Mathematics* 171.3 (2010), pp. 2011–2038. 10.4007/annals.2010.171.2011.

[Lep05] Renaud Leplaideur. "A Dynamical Proof for the Convergence of Gibbs Measures at Temperature Zero". In: *Nonlinearity* 18.6 (2005), pp. 2847–2880. 10.1088/0951-7715/18/6/023.

[Mię90] Jacek Miękisz. "A Microscopic Model With Quasicrystalline Properties". In: *Journal of Statistical Physics* 58.5–6 (1990), pp. 1137–1149. 10.1007/BF01026568.

[Mię97] Jacek Miękisz. "Stable Quasicrystalline Ground States". In: *Journal of Statistical Physics* 88.3–4 (1997), pp. 691–711. 10.1023/B:JOSS.0000015168.25151.22.

[Mię98] Jacek Miękisz. "An Ultimate Frustration in Classical Lattice-Gas Models". In: *Journal of Statistical Physics* 90.1–2 (1998), pp. 285–300. 10.1023/A:1023264004272.

[Rad86] Charles Radin. "Crystals and Quasicrystals: a Lattice Gas Model". In: *Physics Letters. A* 114.7 (1986), pp. 381–383. 10.1016/0375-9601(86)90720-6.

[Rue04] David Ruelle. *Thermodynamic Formalism*. 2nd edition. Cambridge University Press, 2004. 10.1017/CBO9780511617546.

[Wal82] Peter Walters. *An Introduction to Ergodic Theory*. Vol. 79. Graduate Texts in Mathematics. Springer, 1982.

[Wol02] Stephen Wolfram. *A New Kind of Science*. Wolfram Media, 2002.

[Zha07] Zhengmin Zhang. "Estimating Mutual Information via Kolmogorov Distance". In: *IEEE Transactions on Information Theory* 53.9 (2007), pp. 3280–3282. 10.1109/TIT.2007.903122.

Chapter 6

Miscellaneous Musings and Leftovers

The goal of this chapter is to tie some loose ends that did not really fit elsewhere.

Of course, some "obvious" big questions were left open in the previous chapters, notably regarding complexity. In Chapter 4, I bound the complexity of Besicovitch-stability between Π_2 and Π_4, without managing to close this gap and obtain a completeness result. Still relating to this chapter, many questions revolve around the Robinson tiling, notably whether the stability bound on the enhanced tiling can be improved, or whether the canonical tiling is stable at all. In Chapter 5, I obtain an optimal Π_2 realisation result in the case of uniform models, but without a clue on how to proceed in the general case with a Π_3 upper bound. However, in each case, this next step seems quite out of grasp for now.

Now, I will discuss matters still related to these chapters, but with probably much more accessible results in the near future. Each of the following sections will discuss its own matter, mostly independently from the others.

6.1 Stability for the Dense Domino Tileset

Setting the case of the canonical Robinson tiling aside, I studied other examples that don't fit the almost-periodic scheme at all, to try and obtain different Besicovitch stability results.

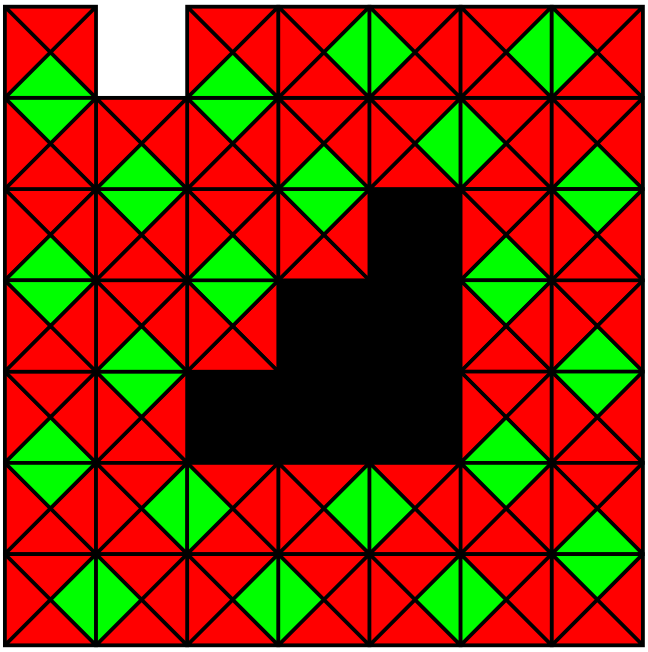

Figure 6.1: Non-tileable obscured area in a dense domino tiling.

One such example was the Kari-Culik tiling [Kar96]. This two-dimensional tileset enforces an aperiodic structure, that encodes a real number on each line, and simulates the transitions $x \mapsto f(x)$ of a continuous system from one line to the next, with positive entropy thus without a self-similar structure [DGG17]. Similar computable systems f on real numbers were studied [BGR12], in which the limit measure was not computable, but adding perturbations broke the asymptotic complexity of the set of invariant measures (such that $f_*(\mu) = \mu$, which corresponds to an invariance for vertical translations on tilings). From this point of view, the Kari-Culik would represent an interesting system to study, as the independent Bernoulli noise on tilings translates as an unusual kind of perturbation on real numbers. However, precisely because of the weird structure of the noise once translated to real numbers, no conclusive results were obtained.

Another much simpler example to define, with a much looser structure overall, but no less interesting to study, is that of the *dense* domino tileset. I have no conclusive results on stability for this tiling, but I will discuss some of my ideas nonetheless.

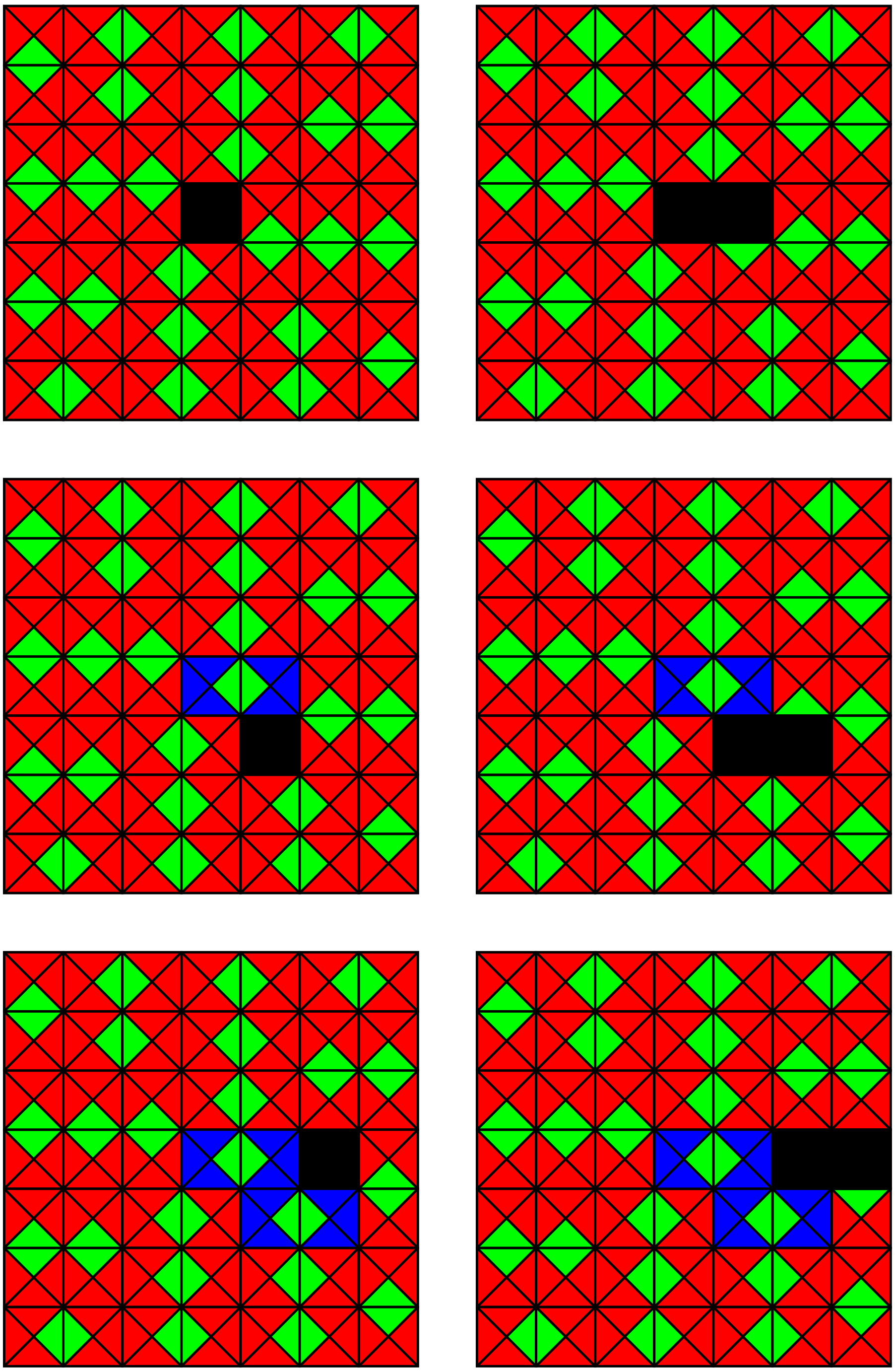

Figure 6.2: An obscured cell can be sent to infinity.

In Section 4.1, I introduced the diluted domino tileset, which uses four half dominos, and an “empty” tile that allows us to easily repair the neighbourhood of obscured cells, as illustrated in Figure 4.2. The dense domino case results from the removal of this empty tile, by requiring each and every cell to contain a half-domino that must pair with an adjacent one. Such admissible domino tilings are sometimes called *dimer* tilings, and can be for example seen as perfect matchings of the graph associated to the grid.

Because of these matching constraints, some windows simply cannot be tiled using only complete dominos. This is obviously the case when the window contains an odd number of cells, but the obstruction is much broader, as illustrated in Figure 6.1, where the non-tileable obscured area contains an even number of cells. Consequently, there is no natural way to "locally" repair mistakes on some fixed neighbourhood of obscured cells.

One may instead try to iteratively repair the configurations by moving the mistakes around until they cancel each other. However, the intricate structure of the configurations makes the task hard. For example, if we want to repair only one obscured hole, then the best we can do is decide in which direction we want to move it, which will ultimately form a zigzagy half-line that differs from the initial (noisy) tiling and send the obscured cell to infinity. For example, in Figure 6.2, I want to send the obscured cell to the right. On the top row of the figure, I pair the obscured cell with its neighbour. This pair can be tiled by a domino, highlighted with a blue background in the middle row, but I need to replace the other half of the broken domino by an obscured cell. Likewise, I can move this obscured cell on the middle row to the right, forming another highlighted domino on the bottom row, and so on. At the limit, the blue dominos form the announced half-line of null density where the repaired tiling differs from the initial perturbed one. Now, the issue is that this process cannot be realised in a shift-invariant way, as the correction processes of several obscured cells may interfere with each other, and choosing an order basically means fixing an origin point. Even worse, while one single zigzagy half-line may have null density in $\mathbb{Z}^2$, having an infinity of them may fill the plane with a high density of mismatches between the initial perturbed tiling and the "repaired" one.

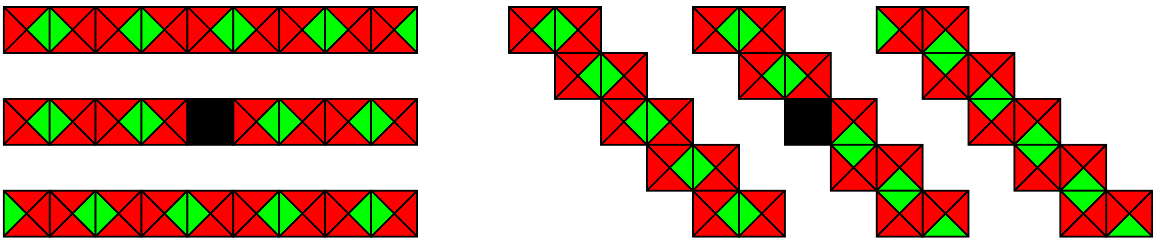

Figure 6.3: Each "one-dimensional" layer admits two globally admissible tilings using complete dominos, and behaves like a 2-periodic one-dimensional SFT.

Regarding stability of the dense domino tiling, my current intuition leans on the side of instability. My best proof-of-concept candidates so far are illustrated in Figure 6.3. In both cases, the key idea is to first partition $\mathbb{Z}^2$ into "one-dimensional" identical layers stacked onto each other. Then, within each layer, we create a noisy tiling just like we did in the one-dimensional periodic case in Section 4.3, by choosing one of the two admissible tilings independently at random inside each clear window. The end result of this process is a noisy tiling at distance around $\frac{1}{2}$ of any non-noisy tiling *with the same layered structure*. The next step of the argument should then be to conclude that, generically, a tiling with a different structure (layered or not) must have some density of mismatching tiles.

Assuming conclusive results are obtained for this dense domino tileset, a next logical step may be to study the case of dimer tilings *on the triangle lattice* instead, and see if the (in)stability adapts.

6.2 A Robust Variant of the Robinson Tiling

As already discussed, one of the main limitations of the notion of Besicovitch topology is that it doesn't account for the "aperiodic" low-density structure of a given tiling. Notably, the enhanced Robinson tiling is stable, but for any fixed $\varepsilon > 0$, at a high-enough scale of N-macro-tiles, the surrounding grid is just made of finite local patches independent from each other.

Consequently, phenomenons and computations that depend on what happens at arbitrarily high scales will be broken by any small amount of perturbations. In particular, Mozes-like substitutive structures [Moz89] cannot be implemented in the enhanced Robinson tiling in a stable way. The Red-Black instability result in Section 4.6 perfectly illustrates this phenomenon: it implements a basic exchange between Red and Black from one scale to the next, but synchronising the colour of two neighbouring N-macro-tiles may require to climb arbitrarily high in the hierarchical structure until we find a macro-tile they both belong to. Similar colour-flip instability arguments may probably hold when implementing a two-dimensional Thue-Morse substitution, with the supplementary property that the base non-perturbed subshift is now uniquely ergodic.

Durand, Romashchenko and Shen proved that, using a *robustness* assumption on the tileset, we can use an iterative repair procedure to correct increasingly bigger obscured areas (the so-called islands of errors), and ultimately obtain a globally admissible configuration nearby in the Besicovitch topology [DRS12]. In particular, in such tilesets, small perturbations aren't enough to break the "aperiodic structure" that can be uniquely deduced from this iterative repair procedure. Their article already provides a robust simulating tileset, but as the (non-robust) Robinson tiling has a simple known structure with many arguments already implemented, I tried to add a supplementary structure to obtain a *robustified* Robinson-like tileset, using a zoom factor argument similar to that of Durand, Romashchenko and Shen.

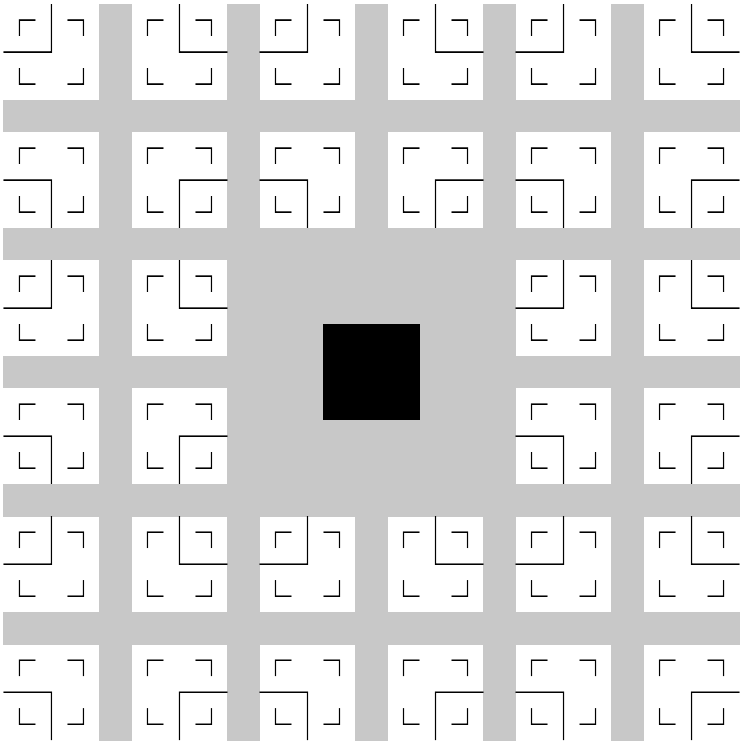

Figure 6.4: The locally admissible area around the obscured hole contains a grid of macro-tiles.

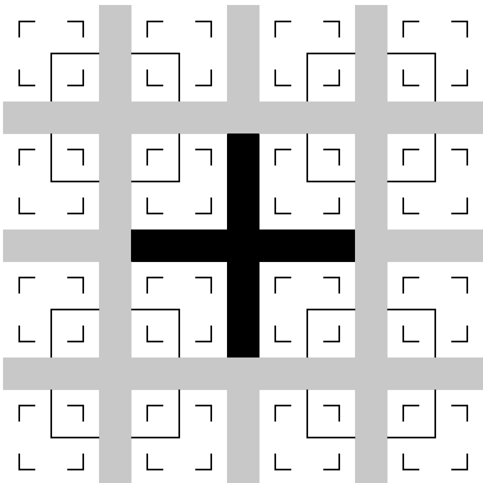

Figure 6.5: Once the macro-tiles are fixed too, we are left a possibly non-tileable cross.

Let's start with the enhanced Robinson tileset. In simple words, a robust tileset is such that we can fill the hole in any locally admissible tiling of a large-enough ring around the hole, up to the replacement of the neighbourhood directly surrounding the hole. As shown on Figure 6.4, if we have a hole that fits in an N-macro-tile, then using the reconstruction argument for the enhanced tileset from Proposition 4.48 we obtain a grid of well-aligned and well-oriented macro-tiles, without control on the neighbourhood of the obscured area.

Of course, assuming the whole area is tiled in a locally admissible way, then necessarily it must also contain well-aligned and well-oriented macro-tiles in the central region. Thus, we can obtain the structure shown in Figure 6.5, where the only unknown is this obscured central cross. For the enhanced Robinson tiling, this cross is not necessary fillable in an admissible way, as shown in Figure 4.4. However, we can notice that the grid is included in the 5×5 neighbourhood of any macro-tile directly around the hole in Figure 6.4 (here, neighbourhood must be understood with respect to the macro-tiles, at their own scale, not at the microscopic scale of tiles).

Figure 6.6: Admissible arrangements of two consecutive scales of macro-tiles.

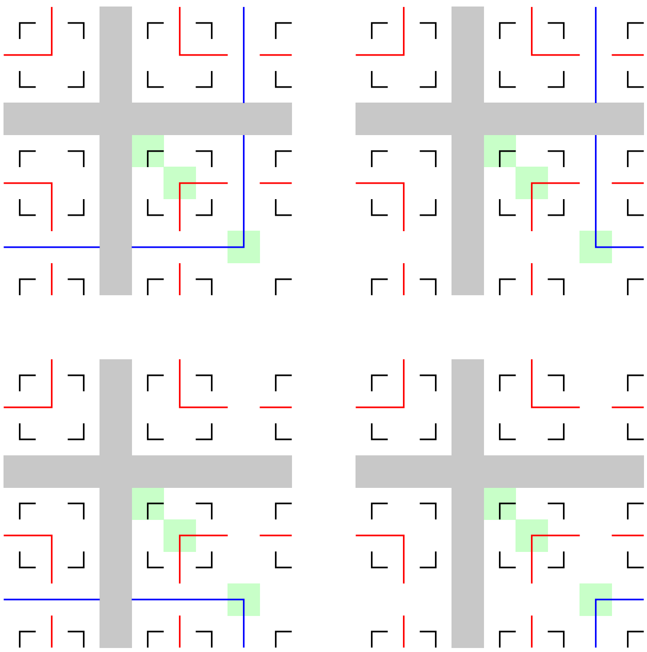

Figure 6.7: Admissible arrangements of three consecutive scales of macro-tiles.

Hence, to *robustify* the Robinson structure, we simply need to encode this macroscopic neighbourhood information inside the central cross of a given macro-tile (using a finite alphabet as there are only finitely many globally admissible tilings of this finite grid in a 5×5 block of macro-tiles). Then, this information is compared between neighbouring macro-tiles in the four directions.

Let me describe a bit more precisely this finite alphabet of admissible grids. Without loss of generality, I'll assume that we are considering a ◲ macro-tile, and the other orientations follow by rotating the corresponding figures. Naturally, this N-macro-tile must form a square with other N-macro-tiles, and contain the central cross of an $(N+1)$-macro-tile. Once the orientation of this new central cross is fixed, the position and orientation of the other crosses at the same scale directly follows, as illustrated in Figure 6.6. Likewise, in Figure 6.7, assuming we use the bottom-right choice of orientations in Figure 6.6 for both N-macro-tiles (the black lines) and $(N+1)$-macro-tiles (the red lines), we have four possible choices orientation for $(N+2)$-macro-tiles (the blue lines). For each of these $4^3 = 64$ triplets of orientations, we are left with one cross to fill. This cross may then be filled with the corner of an $(N+3)$-macro-tile, or lines for macro-tiles of non-specified scale $K \geq N+3$ that cross the 5×5 window from end to end (note that two lines corresponding to two different scales may intersect in this cross), hence *at most* four possible fillings of any such cross, and a total of around two hundreds of 5×5 neighbourhoods. The local rules naturally follow, in that the four arms of the central cross of a given N-macro-tile all agree on the same neighbourhood, and when matching two N-macro-tiles, they must encode the same information in their common 4×5 neighbourhood, as illustrated in Figure 6.8.

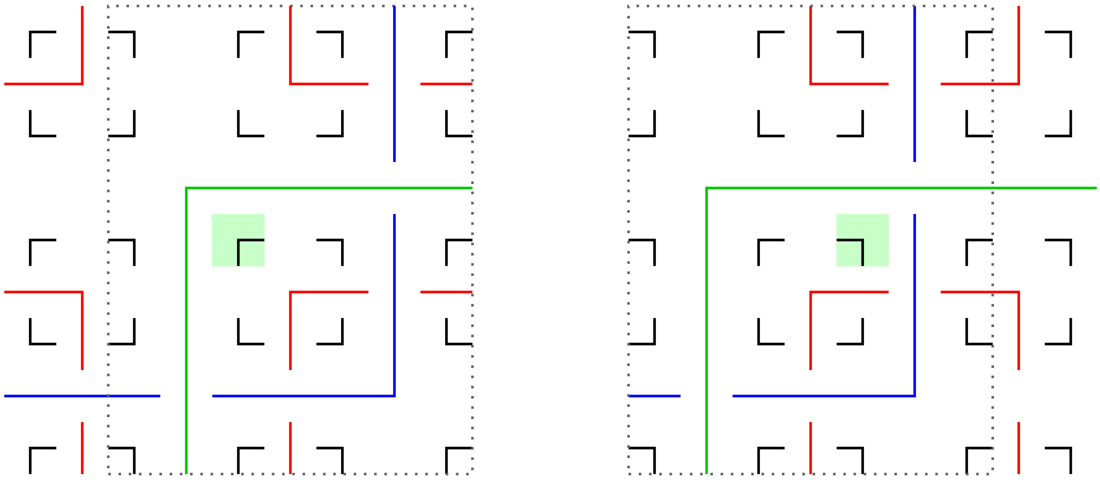

Figure 6.8: A pair of labellings is admissible if they contain the same information in their overlapping 4×5 neighbourhood.

Consequently, in a locally admissible setting, all the neighbouring macro-tiles must agree on what happens in the 2×2 area in the middle of Figure 6.4, on how to fill the obscured cross in Figure 6.5, so the corresponding tileset is robust.

As a robust tileset, it must induce a stable SFT, but without any good bound on the speed of convergence. In a globally admissible configuration, if we consider specifically the N-macro-tiles that contain some information on macro-tiles of scale *at most* $N+R$, then we still obtain an almost-periodic behaviour, with a density of ignored tiles that goes to 0 as $N, R \to \infty$. To obtain a polynomial bound on the speed of convergence, we would need to obtain a local almost-reconstruction argument, at the very least made much harder by this robust structure.

6.3 Quantifying the Stability of the Aperiodic Structure

Directly related to this conversation around robust tilesets lies the question of *how to define a topology that accounts for the stability of the (aperiodic) low-density structure?*

As underlined in Remark 4.51, even though the enhanced Robinson tiling is stable, the "aperiodic" structure cannot be recovered. The main reason, informally, is that the "aperiodic" structure is a zero-density phenomenon caused by (almost) periodic stability at increasingly large scales.

Whenever a tiling has a kind of substitutive structure, for the aperiodic part of the structure to be preserved, we want to be able to iterate the substitution in a noisy environment, for example by saying that if part of the window to substitute is perturbed at a given scale, in such a way that the substitution cannot be uniquely deduced anymore, then the whole window must be at the next one.

In some sense, robustness seems to address this issue, in that it forbids a local modification of the low-density aperiodic structure in the middle of the large area that behaves periodically. Thence, non-admissible patterns can be iteratively locally repaired in increasignly large obscured areas (the islands of errors). Of course, at the limit, the induced invariant map $f : X_{\widetilde{\mathcal{F}}} \to X_{\mathcal{F}}$ (from perturbed tilings to non-perturbed ones) is not a cellular automaton, but it still describes a measurable correction process. Hence, we might define the divergence from *global agreement* as:

$$D_A(\nu\|\mu) = \inf_{f_*(\nu)=\mu} \int d_H(\omega, f(\omega))\, \mathrm{d}\nu(\omega)\,,$$

by analogy with the Kullback–Leibler divergence. In particular, $D_A = \infty$ when no such function f exists. Naturally, $D_A(\nu\|\mu) \geq d_B(\nu, \mu)$, and it even satisfies the triangle inequality, but D_A is not symmetrical so it is not a distance. What's more, $f_*(\nu) = \mu$ imposes $h(\mu) \leq h(\nu)$, so f must really be thought as a kind of denoiser, and $D_A(\nu\|\mu) = \infty$ whenever $h(\mu) > h(\nu)$. Consequently, an SFT is stable with global agreement when:

$$D_A\left(\mathcal{M}^{\mathcal{B}}_{\mathcal{F}}(\varepsilon)\middle\|\mathcal{M}_\sigma\left(X_{\mathcal{F}}\right)\right) := \sup_{\nu\in\mathcal{M}^{\mathcal{B}}_{\mathcal{F}}(\varepsilon)} \inf_{\mu\in\mathcal{M}_\sigma(X_{\mathcal{F}})} D_A(\nu\|\mu) \underset{\varepsilon\to 0}{\longrightarrow} 0\,.$$

Note that stability with global agreement implies Besicovitch-stability.

The voting argument used in Section 4.4, for higher-dimensional periodic stability, doesn't really use the noise coordinate, so it directly gives global agreement stability. In a similar fashion, the correction process used by Durand, Romashchenko and Shen [DRS12] for robust tilesets doesn't really rely on having access to the obscured cells, as long as we identify the forbidden patterns, so it shall probably adapt to give global agreement stability. Other similar local reconstruction arguments using iterated cellular automata have been studied in recent years [FMT22].

The same cannot be said for the enhanced Robinson tiling. Indeed, in this case, the denoising process used in Proposition 4.49 gives us a grid of well-aligned and well-oriented N-macro-tile, with an explicit upper bound on the scale N that can be guaranteed for a fixed ε. Then, the argument for Besicovitch-stability relies on the introduction of an outside source of noise, of randomness, to fill the grid around these N-macro-tiles, which is forbidden here.

Thus, the global agreement divergence D_A seems like a good candidate to measure and quantify the stability of the aperiodic structure in a perturbed setting, with the idea of denoising baked into its definition. Even tighter restrictions may also be placed upon f, such as being a cellular automaton, or a computable function.

6.4 Moving Onto Other Groups

Another direction of investigation, instead of changing the kind of SFTs we study or the notion of stability itself, may be to change the underlying group structure G of the full-shift space $\Omega_{\mathcal{A}}$. In order to obtain interesting stability results, we would need a group G for which there are known percolation results *and* almost-periodic hierarchical SFTs.

Regarding invariant percolations on groups, some results were obtained two decades ago [Ben+99]. For any amenable group G and any $\varepsilon > 0$, there exists an invariant percolation with a density of obscured cells ε that separates the space into finite components (which corresponds for example to the periodic grid noise on $\mathbb{Z}^2$ in Section 4.4.1). Thus, for most SFTs, if we allow for any noise to be considered, then we will have instability.

If we push non-amenability to its limit, when $G = \mathbb{F}_n$ is a free group (with $n \geq 2$ generators), then *any* positive amount of obscured cells, regardless of the noise considered, is enough to separate the space into two or more connected components, and thus we have instability in an intuitive sense, with incompatible patches that cannot coexist in an admissible tiling (but the Besicovitch distance, which relies on amenability by default, cannot be formalised).

On the other hand, (strongly) aperiodic SFTs on other groups have been an active field in recent years [Jea15]. Notably, it may be interesting to look at the case of Baumslag-Solitar groups $BS(m, n) = \langle a, b | ab^m = b^n a\rangle$. Whenever $|m|, |n| \geq 2$ and $|n| \neq |m|$, the group is not amenable [Fim10], and likewise for $BS(n, n)$ that contains a free group $\mathbb{F}_n$ when $n \geq 2$ thus is not amenable either, so we can skip any conversation around the Besicovitch distance on those instances. This notably leaves out one case, that of $BS(1, n)$ (which is solvable thus amenable) where we have the existence of strongly aperiodic SFTs [EM22].

I must admit that the matter of Bernoulli percolation on general amenable groups seems quite complicated to me, so I will abstain to comment it for now, but it certainly warrants further discussion. In particular, I haven't found discussions around the Bernoulli percolation on Baumslag-Solitar groups in the literature, but I don't really know whether there are new things to be said here or if this is already covered by more general results. Assuming we have percolation results on $BS(1, n)$ similar to those on $\mathbb{Z}^2 = BS(1, 1)$, it may be interesting to see what this implies for the Besicovitch stability of perturbed tilings on these structures.

6.5 Robustness of Chaoticity to Potential Perturbations

So far, I have discussed many things related to the Besicovitch topology. Let me conclude the chapter with a few ideas about what more can be said in the thermodynamic framework of Chapter 5. The main perspective, in the near future, will be to see how this class of potentials, inducing uniform models, may be used to study properties related to chaoticity.

One such general question is the existence of phase transitions. Informally, a phase transition in a model can be understood as a discontinuity of some property with respect to β at positive temperatures. For example, the number of infinite connected components in a two-dimensional Ising model jumps from 0 to 1 at a critical value β_c, where a phase transition occurs.

In particular, the question of uniqueness of the Gibbs measure at a temperature β can be studied as a phase transition phenomenon. We already know that, using one-dimensional finite-range interactions, we always have uniqueness of the Gibbs measure. When using long-range interactions, however, we can realise any positive increasing sequence (β_n) as the set of values of β in $]\beta_0, \infty[$ for which the set $\mathcal{G}_\sigma(\beta)$ is not a singleton [KQW21]. It may be interesting to see if some kind of simulating tileset could adapt this argument using a finite-range higher-dimensional potential. However, as a concept, uniformity is deeply uninterested in the question of *uniqueness* of a Gibbs measure, so the previous class of simulating tilesets may not be adapted.

Another question, for which these finite-range potentials may give interesting results, is that of *robustness* of chaoticity and stability to perturbations *of the potential*. Chaoticity and stability are (mutually exclusive) yes/no discrete properties, and the space of potentials is a continuum, so these properties cannot behave continuously in the parameter φ. Consequently, we will say that a potential φ induces a *robust* chaotic (resp. stable) model if, for any potential ψ in a sufficiently small neighbourhood of φ, ψ induces a chaotic (resp. stable) model. Recently, the two-dimensional Robinson tiling was proven to be non-robust in a similar sense [GQS21]. This question of robustness is of interest for physicists in that, in real life, either due to theoretical constraints like Heisenberg's uncertainty principle or simply because of observational errors, we may never know the exact mathematical value of some parameters. Hence, ideally, our models for phenomena such as the formation of quasi-crystals would have to be robust to have any practical implication.

Using the simulating tilesets from Chapter 5, I can exhibit models that are *not* robust. As explained in that chapter, as long as all the forbidden patterns of a tileset $\mathcal{F}$ (with the right combinatorial properties) are given a positive weight in φ, then φ induces a uniform model with the same accumulation set. Thus, the idea will be to create a simulating tileset inducing a uniformly chaotic model, but in such a way that adding specific forbidden patterns induces a uniformly stable model.

To do so, consider a sequence of computed measures on words (m_k), inducing a chaotic model, as in Section 5.6. In the corresponding simulating tileset, at a given scale, if the input seed is equal to $(\texttt{+1})^k$, we change the computations so that we force $(\texttt{+1})^{b_{\text{read}}(k)}$ instead. Thus, we obtain computed measures (m'_k), such that $d(m_k, m'_k) = O\left(\frac{1}{2^k}\right)$, so that this new tileset $\mathcal{F}$ induces a uniformly chaotic model with the very same accumulation set $\mathcal{G}_\sigma(\infty)$.

Now, however, consider the purely local rule that imposes each bit of the seed to be equal to $\texttt{+1}$, and denote

$$\mathcal{F}' = \mathcal{F} \sqcup \{a \in \mathcal{A}, a \text{ encodes a } \texttt{-1} \text{ bit for the seed}\}$$

this new tileset. As a given Blocking square encodes strictly more than $b_{\text{read}}(k)$ bits, this means that there is still a source of entropy left in the tileset. Hence, the bounds such as Proposition 5.36 still adapt to this case, and we can conclude that the tileset $\mathcal{F}'$ still induces a uniform model for the corresponding potentials. The main difference, however, is that $\mathcal{F}'$ now computes the sequence of measures $\left(\delta_{(\texttt{+1})^{b_{\text{read}}(k)}}\right)$, so the limit set $\mathcal{G}_\sigma(\infty)$ corresponds to the measure $\delta_{(\texttt{+1})^\infty}$, the model is uniformly *stable*.

Denote φ a potential that corresponds to the forbidden patterns $\mathcal{F}$, and ψ that corresponds to the forbidden patterns $\{a \in \mathcal{A}, a \text{ encodes a } \texttt{-1} \text{ bit for the seed}\}$. Then, for any choice of $\varepsilon > 0$, the potential $\varphi + \varepsilon\psi$ corresponds to $\mathcal{F}'$, thus induces a uniformly stable model, even though the limit potential φ induces a uniformly chaotic model. In other words, φ induces a non-robust uniformly chaotic model.

Following a similar logic, for the non-robust stable case, consider a system that always computes $\left(\delta_{(\texttt{+1})^{b_{\text{read}}(k)}}\right)$ regardless of the seed. As discussed in Theorem 5.52, even without using the random seed, by computing a word w_k (that is a prefix of either $(\texttt{+1})^\infty$ or $(\texttt{-1})^\infty$ depending on k) we can obtain a uniformly chaotic model. Thus, consider the system $\mathcal{F}$ that computes $\frac{1}{2^k}\delta_{w_k} + \left(1 - \frac{1}{2^k}\right)\delta_{(\texttt{+1})^{b_{\text{read}}(k)}}$ (whith w_k corresponding to the case of the seed $(\texttt{+1})^k$), and $\mathcal{F}' \supset \mathcal{F}$ where the seed is fixed equal to $(\texttt{+1})^k$. If φ corresponds to $\mathcal{F}$ and ψ to the supplementary rules, then φ induces a uniformly stable model, but any $\varphi + \varepsilon\psi$ corresponds to $\mathcal{F}'$ thus induces a uniformly chaotic model, so that ultimately φ induces a non-robust stable model.

More broadly, we may perhaps even simulate U a non-deterministic universal Turing machine using a tileset $\mathcal{F}$, and add supplementary local rules $\mathcal{F}_M$ to force it to simulate the machine M specifically. Thus, for the corresponding potentials φ and ψ_M, and any $\varepsilon > 0$, we obtain a potential $\varphi + \varepsilon\psi_M$ that forces the uniform asymptotic behaviour corresponding to M. In doing so, we may thus obtain arbitrarily small neighbourhoods of φ inside which, for any Π_2-computable accumulation set, there is a potential ψ that realises it uniformly.

References

[Ben+99] Itai Benjamini, Russell Lyons, Yuval Peres, and Oded Schramm. "Group-Invariant Percolation on Graphs". In: *Geometric and Functional Analysis* 9.1 (1999), pp. 29–66. 10.1007/s000390050080.

[BGR12] Mark Braverman, Alexander Grigo, and Cristóbal Rojas. "Noise vs Computational Intractability in Dynamics". In: *ITCS*. ACM, 2012, pp. 128–141. 10.1145/2090236.2090247.

[DGG17] Bruno Durand, Guilhem Gamard, and Anaël Grandjean. "Aperiodic Tilings and Entropy". In: *Theoretical Computer Science* 666 (2017), pp. 36–47. 10.1016/j.tcs.2016.12.013.

[DRS12] Bruno Durand, Andrei Romashchenko, and Alexander Shen. "Fixed-Point Tile Sets and Their Applications". In: *Journal of Computer and System Sciences* 78.3 (2012), pp. 731–764. 10.1016/j.jcss.2011.11.001.

[EM22] Solène Esnay and Etienne Moutot. "Aperiodic SFTs on Baumslag-Solitar Groups". In: *Theoretical Computer Science* 917 (2022), pp. 31–50. 10.1016/j.tcs.2022.03.010.

[Fim10] Pierre Fima. *A Note on the von Neumann Algebra of a Baumslag-Solitar Group*. 2010. 10.48550/arXiv.1011.3160.

[FMT22] Nazim Fatès, Irène Marcovici, and Siamak Taati. "Self-Stabilisation of Cellular Automata on Tilings". In: *Fundamenta Informaticae* 185.1 (2022), pp. 27–82. 10.3233/fi-222103.

[GQS21] Juliano Gonschorowski, Anthony Quas, and Jason Siefken. "Support Stability of Maximizing Measures for Shifts of Finite Type". In: *Ergodic Theory and Dynamical Systems* 41.3 (2021), pp. 869–880. 10.1017/etds.2019.69.

[Jea15] Emmanuel Jeandel. *Aperiodic Subshifts of Finite Type on Groups*. 2015. 10.48550/arXiv.1501.06831.

[Kar96] Jarkko Kari. "A Small Aperiodic Set of Wang Tiles". In: *Discrete Mathematics* 160 (1996), pp. 259–264. 10.1016/0012-365X(95)00120-L.

[KQW21] Tamara Kucherenko, Anthony Quas, and Christian Wolf. "Multiple Phase Transitions on Compact Symbolic Systems". In: *Advances in Mathematics* 385.107768 (2021). 10.1016/j.aim.2021.107768.

[Moz89] Shahar Mozes. "Tilings, Substitution Systems and Dynamical Systems Generated by Them". In: *Journal d'Analyse Mathématique* 53 (1989), pp. 139–186. 10.1007/BF02793412.

Index

Complexity and Robustness of Tilings with Random Perturbations

Abstract:

The central question of this thesis is that of the robustness to perturbations of tilings defined by local rules. Here, a tiling refers to a labelling of the grid $\mathbb{Z}^d$ by a finite alphabet, with neighbourhood rules determining which symbols can be juxtaposed, like puzzle pieces. When the number of rules is finite, we obtain a Subshift of Finite Type (or SFT). Starting from dimension $d = 2$, complex structures (aperiodic, hierarchical, self-similar) can appear for certain choices of local rules. The challenge is then to know to what extent these structures are preserved when a small proportion of perturbations is allowed.

The challenge here is twofold. Firstly, for computer scientists, such structures can be used to encode computational models (*i.e.* computers) such as Turing machines. In information theory, it is natural to ask whether a signal can be transmitted in a noisy environment, and by extension whether it is physically possible to create a computational model that will produce correct results despite rare but inevitable localised faults. Secondly, for physicists, such structures are comparable to those of quasicrystals, materials that are highly ordered but have no crystalline (*i.e.* periodic) structure, and for which it is still difficult to propose a theoretical model explaining their formation. My work in this manuscript reflects and intertwines these two viewpoints.

After some introductory chapters, the first thematic block concerns localised noise, independent from one cell to another. Here, the question is to know whether, globally, an admissible structure of high density occurs in the perturbed tilings (we will then speak of a stable structure). As is often the case, the problem in dimension $d = 1$ is treated separately, in which case it is possible to decide (algorithmically) whether a choice of rules leads to a stable system. In dimension $d \geq 2$, examples of stable structures are proposed, first periodic, by exploiting their ultimately quite localised redundancy, then aperiodic, based on a variant of Robinson's aperiodic tiling (with a complex quasi-periodic structure emerging from simple rules). By implementing computations in this hierarchical structure, by reduction from other undecidable problems, it is shown that this question of stability is undecidable, and more precisely Π_2-hard in the arithmetic hierarchy. Finally, using computable analysis arguments on the space of probability measures on these perturbed tilings, a Π_4 upper bound on the complexity of stability is obtained.

The second block adopts a more physical outlook, by associating to the local rules a potential inducing Gibbs measures. The quantity of perturbations is then related to a temperature parameter, which will also go to 0. Here, the local behaviour is studied (in the weak-$*$ topology), and the question is no longer *whether* the system converges but *to what limit(s)*. In the case $d = 1$, such potentials induce a single Gibbs measurement for any temperature, and this family of measurements converges when the temperature goes to 0. In dimension $d \geq 2$, this is no longer the case, and there are such potentials for which the system is *chaotic* (it admits no convergent trajectory). Under an additional assumption of *uniformity* of the model, I prove here that the set of adherence values of the trajectories is at most Π_2-computable. Then, the optimality of this bound is established via a realisation result (up to a computable affine homeomorphism) of all Π_2-computable sets as accumulation points for some potential, encoding the necessary computations in structures that will vanish at the limit.

Keywords:
Subshift of Finity Type ;
Random Perturbations ;
Aperiodicity ;
Complexity ;
Robinson Tiling ;
Undecidability ;
Arithmetical Hierarchy ;
Gibbs Measures ;
Chaos.

Résumé :

La question centrale de ce manuscrit est celle de la robustesse aux perturbations de pavages définis par règles locales. Ici, un pavage désigne un étiquetage de la grille $\mathbb{Z}^d$ par un alphabet fini, avec des règles de voisinnage déterminant quels symboles peuvent être juxtaposés, à l'instar de pièces de puzzle. Lorsque le nombre de règles est fini, on parle de sous-décalage de type fini (ou *Subshift of Finite Type* en anglais, abrégé en SFT). Dès la dimension $d = 2$, des structures complexes (apériodiques, hiérarchiques, autosimilaires) peuvent apparaître pour certains choix de règles locales. L'enjeu est alors de savoir à quel point ces structures sont préservées lorsqu'une faible proportion de perturbations est autorisée.

L'enjeu est ici double. D'une part, pour les informaticiens, de telles structures peuvent servir à encoder des modèles de calcul (*i.e.* des ordinateurs) comme les machines de Turing. En théorie de l'information, il est naturel de se demander si un signal peut être transmis dans un milieu bruité, et par extension s'il est possible de réaliser physiquement un modèle de calcul qui produira des résultats corrects malgré des rares mais inévitables défauts localisés. D'autre part, pour les physiciens, de telles structures sont comparables à celles des quasicristaux, des matériaux hautement ordonnés mais sans structure cristalline (c'est-à-dire périodique), dont on peine encore à proposer un modèle théorique expliquant leur formation. Mon travail dans ce manuscrit vient refléter et entremêler ces deux points de vue.

Après quelques chapitres introductifs, le premier bloc thématique concerne un bruit localisé, indépendant d'une cellule à l'autre. Ici, la question est de savoir si, globalement, une structure admissible de forte densité est observable dans les pavages perturbés (on parlera alors de structure stable). Le problème en dimension $d = 1$ est comme souvent traité à part, auquel cas il est possible de décider (algorithmiquement) si un choix de règles induit un système stable. En dimension $d \geq 2$, des exemples de structures stables sont proposés, d'abord périodiques, en exploitant leur redondance en fin de compte assez localisée, puis apériodiques, en se basant sur une variante du pavage apériodique de Robinson (avec une structure quasi-périodique complexe émergeant de règles simples). En implémentant des calculs dans cette structure hiérarchique, par réduction depuis d'autres problèmes indécidables, il est montré que cette question de stabilité est indécidable, et plus précisément Π_2-dure dans la hiérarchie arithmétique. Enfin, par des arguments d'analyse calculable sur l'espace des mesures de probabilité sur ces pavages perturbés, une majoration Π_4 sur la complexité de la stabilité est obtenue.

Le second bloc adopte un point de vue plus physique, en associant aux règles locales un potentiel induisant des mesures de Gibbs. La quantité de perturbations découle alors d'un paramètre de température, qu'on fait tendre vers 0 également. Ici, le comportement local est étudié (dans la topologie faible-$*$), et la question n'est plus de savoir *si* le système converge mais *vers quelle(s) limite(s)*. Dans le cas $d = 1$, de tels potentiels induisent une unique mesure perturbée pour toute température, et cette famille de mesures converge lorsque la température tend vers 0. En dimension $d \geq 2$, ce n'est plus le cas, et il existe de tels potentiels pour lesquels le système est *chaotique* (il n'admet aucune trajectoire convergente). Sous une hypothèse supplémentaire d'*uniformité* du modèle, je prouve ici que l'ensemble des valeurs d'adhérence des trajectoires est au plus Π_2-calculable. Ensuite, l'optimalité de cette borne est établie via un résultat de réalisation (à homéomorphisme affine calculable près) de tous les ensembles Π_2-calculables comme ensembles d'accumulation pour un certain potentiel, en encodant les calculs nécessaires dans des structures qui s'effaceront à la limite.

Mots-clés :
Sous-shift de type fini ;
Perturbations aléatoires ;
Apériodicité ;
Complexité ;
Pavage de Robinson ;
Indécidabilité ;
Hiérarchie arithmétique ;
Mesures de Gibbs ;
Chaos.