Tullio Ceccherini-Silberstein
Michel Coornaert

SPRINGER
MONOGRAPHS IN MATHEMATICS

# Cellular Automata and Groups

Springer

**S***pringer* **M***onographs* in **M***athematics*

For further volumes:
www.springer.com/series/3733

Tullio Ceccherini-Silberstein • Michel Coornaert

# Cellular Automata and Groups

Springer

Tullio Ceccherini-Silberstein
Dipartimento di Ingegneria
Università del Sannio
C.so Garibaldi 107
82100 Benevento
Italy
tceccher@mat.uniroma1.it

Michel Coornaert
Institut de Recherche Mathématique Avancée
Université de Strasbourg
7 rue René-Descartes
67084 Strasbourg Cedex
France
coornaert@math.unistra.fr

ISSN 1439-7382
ISBN 978-3-642-14033-4 e-ISBN 978-3-642-14034-1
DOI 10.1007/978-3-642-14034-1
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2010934641

Mathematics Subject Classification (2010): 37B15, 68Q80, 20F65, 43A07, 16S34, 20C07



*Cover design*: deblik

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To Katiuscia, Giacomo, and Tommaso

To Martine and Nathalie

# Preface

Two seemingly unrelated mathematical notions, namely that of an amenable group and that of a cellular automaton, were both introduced by John von Neumann in the first half of the last century. Amenability, which originated from the study of the Banach-Tarski paradox, is a property of groups generalizing both commutativity and finiteness. Nowadays, it plays an important role in many areas of mathematics such as representation theory, harmonic analysis, ergodic theory, geometric group theory, probability theory, and dynamical systems. Von Neumann used cellular automata to serve as theoretical models for self-reproducing machines. About twenty years later, the famous cellular automaton associated with the Game of Life was invented by John Horton Conway and popularized by Martin Gardner. The theory of cellular automata flourished as one of the main branches of computer science. Deep connections with complexity theory and logic emerged from the discovery that some cellular automata are universal Turing machines.

A group $G$ is said to be amenable (as a discrete group) if the set of all subsets of $G$ admits a right-invariant finitely additive probability measure. All finite groups, all solvable groups (and therefore all abelian groups), and all finitely generated groups of subexponential growth are amenable. Von Neumann observed that the class of amenable groups is closed under the operation of taking subgroups and that the free group of rank two $F_2$ is non-amenable. It follows that a group which contains a subgroup isomorphic to $F_2$ is non-amenable. However, there are examples of groups which are non-amenable and contain no subgroups isomorphic to $F_2$ (the first examples of such groups were discovered by Alexander Y. Ol'shanskii and by Sergei I. Adyan).

Loosely speaking, a general cellular automaton can be described as follows. A configuration is a map from a set called the universe into another set called the alphabet. The elements of the universe are called cells and the elements of the alphabet are called states. A cellular automaton is then a map from the set of all configurations into itself satisfying the following local property: the state of the image configuration at a given cell only depends on

the states of the initial configuration on a finite neighborhood of the given cell. In the classical setting, for instance in the cellular automata constructed by von Neumann and the one associated with Conway's Game of Life, the alphabet is finite, the universe is the two dimensional infinite square lattice, and the neighborhood of a cell consists of the cell itself and its eight adjacent cells. By iterating a cellular automaton one gets a discrete dynamical system. Such dynamical systems have proved very useful to model complex systems arising from natural sciences, in particular physics, biology, chemistry, and population dynamics.

*
* *

In this book, the universe will always be a group $G$ (in the classical setting the corresponding group was $G = \mathbb{Z}^2$) and the alphabet may be finite or infinite. The left multiplication in $G$ induces a natural action of $G$ on the set of configurations which is called the $G$-shift and all cellular automata will be required to commute with the shift.

It was soon realized that the question whether a given cellular automaton is surjective or not needs a special attention. From the dynamical viewpoint, surjectivity means that each configuration may be reached at any time. The first important result in this direction is the celebrated theorem of Moore and Myhill which gives a necessary and sufficient condition for the surjectivity of a cellular automaton with finite alphabet over the group $G = \mathbb{Z}^2$. Edward F. Moore and John R. Myhill proved that such a cellular automaton is surjective if and only if it is pre-injective. As the term suggests it, pre-injectivity is a weaker notion than injectivity. More precisely, a cellular automaton is said to be pre-injective if two configurations are equal whenever they have the same image and coincide outside a finite subset of the group. Moore proved the "surjective $\Rightarrow$ pre-injective" part and Myhill proved the converse implication shortly after. One often refers to this result as to the Garden of Eden theorem. This biblical terminology is motivated by the fact that, regarding a cellular automaton as a dynamical system with discrete time, a configuration which is not in the image of the cellular automaton may only appear as an initial configuration, that is, at time $t = 0$.

The surprising connection between amenability and cellular automata was established in 1997 when Antonio Machì, Fabio Scarabotti and the first author proved the Garden of Eden theorem for cellular automata with finite alphabets over amenable groups. At the same time, and completely independently, Misha Gromov, using a notion of spacial entropy, presented a more general form of the Garden of Eden theorem where the universe is an amenable graph with a dense holonomy and cellular automata are called maps of bounded propagation. Machì, Scarabotti and the first author also showed that both implications in the Garden of Eden theorem become false if the underlying group contains a subgroup isomorphic to $F_2$. The question whether the Garden of Eden theorem could be extended beyond the class of

amenable groups remained open until 2008 when Laurent Bartholdi proved that the Moore implication fails to hold for non-amenable groups. As a consequence, the whole Garden of Eden theorem only holds for amenable groups. This gives a new characterization of amenable groups in terms of cellular automata. Let us mention that, up to now, the validity of the Myhill implication for non-amenable groups is still an open problem.

Following Walter H. Gottschalk, a group $G$ is said to be surjunctive if every injective cellular automaton with finite alphabet over $G$ is surjective. Wayne Lawton proved that all residually finite groups are surjunctive and that every subgroup of a surjunctive group is surjunctive. Since injectivity implies pre-injectivity, an immediate consequence of the Garden of Eden theorem for amenable groups is that every amenable group is surjunctive. Gromov and Benjamin Weiss introduced a class of groups, called sofic groups, which includes all residually finite groups and all amenable groups, and proved that every sofic group is surjunctive. Sofic groups can be defined in three equivalent ways: in terms of local approximation by finite symmetric groups equipped with their Hamming distance, in terms of local approximation of their Cayley graphs by finite labelled graphs, and, finally, as being the groups that can be embedded into ultraproducts of finite symmetric groups (this last characterization is due to Gábor Elek and Endre Szabó). The class of sofic groups is the largest known class of surjunctive groups. It is not known, up to now, whether all groups are surjunctive (resp. sofic) or not.

Stimulated by Gromov ideas, we considered cellular automata whose alphabets are vector spaces. In this framework, the space of configurations has a natural structure of a vector space and cellular automata are required to be linear. An analogue of the Garden of Eden theorem was proved for linear cellular automata with finite dimensional alphabets over amenable groups. In the proof, the role of entropy, used in the finite alphabet case, is now played by the mean dimension, a notion introduced by Gromov. Also, examples of linear cellular automata with finite dimensional alphabets over groups containing $F_2$ showing that the linear version of the Garden of Eden theorem may fail to hold in this case, were provided. It is not known, up to now, if the Garden of Eden theorem for linear cellular automata with finite dimensional alphabet only holds for amenable groups or not. We also introduced the notion of linear surjunctivity: a group $G$ is said to be L-surjunctive if every injective linear cellular automaton with finite dimensional alphabet over $G$ is surjective. We proved that every sofic group is L-surjunctive. Linear cellular automata over a group $G$ with alphabet of finite dimension $d$ over a field $\mathbb{K}$ may be represented by $d \times d$ matrices with entries in the group ring $\mathbb{K}[G]$. This leads to the following characterization of L-surjunctivity: a group is L-surjunctive if and only if it satisfies Kaplansky's conjecture on the stable finiteness of group rings (a ring is said to be stably finite if one-sided invertible finite dimensional square matrices with coefficients in that ring are in fact two-sided invertible). As a corollary, one has that group rings of sofic groups are stably finite, a result previously established by Elek and Szabó

using different methods. Moreover, given a group $G$ and a field $\mathbb{K}$, the pre-injectivity of all nonzero linear cellular automata with alphabet $\mathbb{K}$ over $G$ is equivalent to the absence of zero-divisors in $\mathbb{K}[G]$. As a consequence, another important problem on the structure of group rings also formulated by Irving Kaplansky may be expressed in terms of cellular automata. Is every nonzero linear cellular automaton with one-dimensional alphabet over a torsion-free group always pre-injective?

*
* *

The material presented in this book is entirely self-contained. In fact, its reading only requires some acquaintance with undergraduate general topology and abstract algebra. Each chapter begins with a brief overview of its contents and ends with some historical notes and a list of exercises at various difficulty levels. Some additional topics, such as subshifts and cellular automata over subshifts, are treated in these exercises. Hints are provided each time help may be needed. In order to improve accessibility, a few appendices are included to quickly introduce the reader to facts he might be not too familiar with.

In the first chapter, we give the definition of a cellular automaton. We present some basic examples and discuss general methods for constructing cellular automata. We equip the set of configurations with its prodiscrete uniform structure and prove the generalized Curtis-Hedlund theorem: a necessary and sufficient condition for a self-mapping of the configuration space to be a cellular automaton is that it is uniformly continuous and commutes with the shift.

Chapter 2 is devoted to residually finite groups. We give several equivalent characterizations of residual finiteness and prove that the class of residually finite groups is closed under taking subgroups and projective limits. We establish in particular the theorems, respectively due to Anatoly I. Mal'cev and Gilbert Baumslag, which assert that finitely generated residually finite groups are Hopfian and that their automorphism group is residually finite.

Surjunctive groups are introduced in Chap. 3. We show that every subgroup of a surjunctive group is surjunctive and that locally residually finite groups are surjunctive. We also prove a theorem of Gromov which says that limits of surjunctive marked groups are surjunctive.

The theory of amenable groups is developed in Chap. 4. The class of amenable groups is closed under taking subgroups, quotients, extensions, and inductive limits. We prove the theorems due to Erling Følner and Alfred Tarski which state the equivalence between amenability, the existence of a Følner net, and the non-existence of a paradoxical decomposition.

The Garden of Eden theorem is established in Chap. 5. It is proved by showing that both surjectivity and pre-injectivity of the cellular automaton are equivalent to the fact that the image of the configuration space has maxi-

mal entropy. We give an example of a cellular automaton with finite alphabet over $F_2$ which is pre-injective but not surjective. Following Bartholdi's construction, we also prove the existence of a surjective but not pre-injective cellular automaton with finite alphabet over any non-amenable group.

In Chap. 6 we present the basic elementary notions and results on growth of finitely generated groups. We prove that finitely generated nilpotent groups have polynomial growth. We then introduce the Grigorchuk group and show that it is an infinite finitely generated periodic group of intermediate growth. We show that every finitely generated group of subexponential growth is amenable. We also establish the Kesten-Day characterization of amenability which asserts that a group with a finite (not necessarily symmetric) generating subset is amenable if and only if 0 is in the $\ell^2$-spectrum of the associated Laplacian. Finally, we consider the notion of quasi-isometry for not necessarily countable groups and we show that amenability is a quasi-isometry invariant.

In Chap. 7 we consider the notion of local embeddability of groups into a class of groups. For the class of finite groups, this gives the class of LEF groups introduced by Anatoly M. Vershik and Edward I. Gordon. We discuss several stability properties of local embeddability and show that locally embeddable groups are closed in marked groups spaces. The remaining of the chapter is devoted to the class of sofic groups. We show that the three definitions, namely analytic, geometric, and algebraic, we alluded to before, are equivalent. We then prove the Gromov-Weiss theorem which states that every sofic group is surjunctive.

The last chapter is devoted to linear cellular automata. We prove the linear version of the Garden of Eden theorem and show that every sofic group is L-surjunctive. We end the chapter with a discussion on the stable finiteness and the zero-divisors conjectures of Kaplansky and their reformulation in terms of linear cellular automata.

Appendix A gives a quick overview of a few fundamental notions and results of topology (nets, compactness, product topology, and the Tychonoff product theorem). Appendix B is devoted to André Weil's theory of uniform spaces. It includes also a detailed exposition of the Hausdorff-Bourbaki uniform structure on subsets of a uniform space. In Appendix C, we establish some basic properties of symmetric groups and prove the simplicity of the alternating groups. The definition and the construction of free groups are given in Appendix D. The proof of Klein's ping-pong lemma is also included there. In Appendix E we shortly describe the constructions of inductive and projective limits of groups. Appendix F treats topological vector spaces, the weak-$*$ topology, and the Banach-Alaoglu theorem. The proof of the Markov-Kakutani fixed point theorem is presented in Appendix G. In the subsequent appendix, of a pure graph-theoretical and combinatorial flavour, we consider bipartite graphs and their matchings. We prove Hall's marriage theorem and its harem version which plays a key role in the proof of Tarski's theorem on amenability. The Baire theorem, the open mapping theorem, as well as other

complements of functional analysis including uniform convexity are treated in Appendix I. The last appendix deals with the notions of filters and ultrafilters.

We would like to express our deep gratitude to Dr. Catriona Byrne, Dr. Marina Reizakis and Annika Eling from Springer Verlag and to Donatas Akmanavičius for their constant and kindest help at all stages of the editorial process.

Rome and Strasbourg

Tullio Ceccherini-Silberstein
Michel Coornaert

# Contents

# Notation

Throughout this book, the following conventions are used:

- $\mathbb{N}$ is the set of nonnegative integers so that $0 \in \mathbb{N}$;
- the notation $A \subset B$ means that each element in the set $A$ is also in the set $B$ so that $A$ and $B$ may coincide;
- a countable set is a set which admits a bijection onto a subset of $\mathbb{N}$ so that finite sets are countable;
- all group actions are left actions;
- all rings are assumed to be associative (but not necessarily commutative) with a unity element;
- a field is a nonzero commutative ring in which each nonzero element is invertible.

# Chapter 1
# Cellular Automata

In this chapter we introduce the notion of a cellular automaton. We fix a group and an arbitrary set which will be called the alphabet. A configuration is defined as being a map from the group into the alphabet. Thus, a configuration is a way of attaching an element of the alphabet to each element of the group. There is a natural action of the group on the set of configurations which is called the shift action (see Sect. 1.1). A cellular automaton is a self-mapping of the set of configurations defined from a system of local rules commuting with the shift (see Definition 1.4.1). We equip the configuration set with the prodiscrete topology, that is, the topology of pointwise convergence associated with the discrete topology on the alphabet (see Sect. 1.2). It turns out that every cellular automaton is continuous with respect to the prodiscrete topology (Proposition 1.4.8) and commutes with the shift (Proposition 1.4.4). Conversely, when the alphabet is finite, every continuous self-mapping of the configuration space which commutes with the shift is a cellular automaton (Theorem 1.8.1). Another important fact in the finite alphabet case is that every bijective cellular automaton is invertible, in the sense that its inverse map is also a cellular automaton (Theorem 1.10.2). We give examples showing that, when the alphabet is infinite, a continuous self-mapping of the configuration space which commutes with the shift may fail to be a cellular automaton and a bijective cellular automaton may fail to be invertible. In Sect. 1.9, we introduce the prodiscrete uniform structure on the configuration space. We show that a self-mapping of the configuration space is a cellular automaton if and only if it is uniformly continuous and commutes with the shift (Theorem 1.9.1).

## 1.1 The Configuration Set and the Shift Action

Let $G$ be a group. For $g \in G$, denote by $L_g$ the left multiplication by $g$ in $G$, that is, the map $L_g \colon G \to G$ given by

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_1, 

$$L_g(g') = gg' \quad \text{for all } g' \in G.$$

Observe that for all $g_1, g_2, g' \in G$ one has

$$(L_{g_1} \circ L_{g_2})(g') = L_{g_1}(L_{g_2}(g')) = L_{g_1}(g_2 g') = g_1 g_2 g' = L_{g_1 g_2}(g')$$

which shows that

$$L_{g_1} \circ L_{g_2} = L_{g_1 g_2}. \tag{1.1}$$

Let $A$ be a set. Consider the set $A^G$ consisting of all maps from $G$ to $A$:

$$A^G = \prod_{g \in G} A = \{x \colon G \to A\}.$$

The set $A$ is called the *alphabet*. The elements of $A$ are called the *letters*, or the *states*, or the *symbols*, or the *colors*. The group $G$ is called the *universe*. The set $A^G$ is called the set of *configurations*.

Given an element $g \in G$ and a configuration $x \in A^G$, we define the configuration $gx \in A^G$ by

$$gx = x \circ L_{g^{-1}}. \tag{1.2}$$

Thus one has

$$gx(g') = x(g^{-1}g') \quad \text{for all } g' \in G.$$

The map

$$\begin{aligned} G \times A^G &\to A^G \\ (g, x) &\mapsto gx \end{aligned}$$

is a left action of $G$ on $A^G$. Indeed, for all $g_1, g_2 \in G$ and $x \in A^G$, one has

$$\begin{aligned} g_1(g_2 x) &= g_1(x \circ L_{g_2^{-1}}) = x \circ L_{g_2^{-1}} \circ L_{g_1^{-1}} = x \circ L_{g_2^{-1} g_1^{-1}} = x \circ L_{(g_1 g_2)^{-1}} \\ &= (g_1 g_2)x, \end{aligned}$$

where the third equality follows from (1.1). Also, denoting by $1_G$ the identity element of $G$ and by $\mathrm{Id}_G \colon G \to G$ the identity map, one has

$$1_G x = x \circ L_{1_G} = x \circ \mathrm{Id}_G = x.$$

This left action of $G$ on $A^G$ is called the *$G$-shift* on $A^G$.

A *pattern* over the group $G$ and the alphabet $A$ is a map $p \colon \Omega \to G$ defined on some finite subset $\Omega$ of $G$. The set $\Omega$ is then called the *support* of $p$.

## 1.2 The Prodiscrete Topology

Let $G$ be a group and let $A$ be a set.

We equip each factor $A$ of $A^G$ with the discrete topology (all subsets of $A$ are open) and $A^G$ with the associated product topology (see Sect. A.4). This topology is called the *prodiscrete* topology on $A^G$. This is the smallest topology on $A^G$ for which the projection map $\pi_g \colon A^G \to A$, given by $\pi_g(x) = x(g)$, is continuous for every $g \in G$ (cf. Sect. A.4). The elementary cylinders

$$C(g, a) = \pi_g^{-1}(\{a\}) = \{x \in A^G : x(g) = a\} \quad (g \in G, a \in A)$$

are both open and closed in $A^G$. A subset $U \subset A^G$ is open if and only if $U$ can be expressed as a (finite or infinite) union of finite intersections of elementary cylinders.

For a subset $\Omega \subset G$ and a configuration $x \in A^G$ let $x|_\Omega \in A^\Omega$ denote the restriction of $x$ to $\Omega$, that is, the map $x|_\Omega \colon \Omega \to A$ defined by $x|_\Omega(g) = x(g)$ for all $g \in \Omega$.

If $x \in A^G$, a neighborhood base of $x$ is given by the sets

$$V(x, \Omega) = \{y \in A^G : x|_\Omega = y|_\Omega\} = \bigcap_{g \in \Omega} C(g, x(g)), \tag{1.3}$$

where $\Omega$ runs over all finite subsets of $G$.

**Proposition 1.2.1.** *The space $A^G$ is Hausdorff and totally disconnected.*

*Proof.* The discrete topology on $A$ is Hausdorff and totally disconnected, and, by Proposition A.4.1 and Proposition A.4.2, a product of Hausdorff (resp. totally disconnected) topological spaces is Hausdorff (resp. totally disconnected). □

Recall that an action of a group $G$ on a topological space $X$ is said to be *continuous* if the map $\varphi_g \colon X \to X$ given by $\varphi_g(x) = gx$ is continuous on $X$ for each $g \in G$.

**Proposition 1.2.2.** *The action of $G$ on $A^G$ is continuous.*

*Proof.* Let $g \in G$ and consider the map $\varphi_g \colon A^G \to A^G$ defined by $\varphi_g(x) = gx$. The map $\pi_h \circ \varphi_g$ is equal to $\pi_{g^{-1}h}$ and is therefore continuous on $A^G$ for every $h \in G$. Consequently, $\varphi_g$ is continuous (cf. Sect. A.4). □

## 1.3 Periodic Configurations

Let $G$ be a group and let $A$ be a set. Let $H$ be a subgroup of $G$. A configuration $x \in A^G$ is called *$H$-periodic* if $x$ is fixed by $H$, that is, if one has

$$hx = x \quad \text{for all } h \in H.$$

Let $\mathrm{Fix}(H)$ denote the subset of $A^G$ consisting of all $H$-periodic configurations.

*Examples 1.3.1.* (a) One has $\mathrm{Fix}(\{1_G\}) = A^G$.

(b) The set $\mathrm{Fix}(G)$ consists of all constant configurations and may be therefore identified with $A$.

(c) For $G = \mathbb{Z}$ and $H = n\mathbb{Z}$, $n \geq 1$, the set $\mathrm{Fix}(H)$ is the set of sequences $x\colon \mathbb{Z} \to A$ which admit $n$ as a (not necessarily minimal) period, that is, such that $x(i+n) = x(i)$ for all $i \in \mathbb{Z}$.

**Proposition 1.3.2.** *Let $H$ be a subgroup of $G$. Then the set $\mathrm{Fix}(H)$ is closed in $A^G$ for the prodiscrete topology.*

*Proof.* We have

$$\mathrm{Fix}(H) = \bigcap_{h \in H} \{x \in A^G : hx = x\}. \tag{1.4}$$

The space $A^G$ is Hausdorff by Proposition 1.2.1 and the action of $G$ on $A^G$ is continuous by Proposition 1.2.2. Thus the set of fixed points of the map $x \mapsto gx$ is closed in $A^G$ for each $g \in G$. Therefore $\mathrm{Fix}(H)$ is closed in $A^G$ by (1.4). □

Consider the set $H\backslash G = \{Hg : g \in G\}$ consisting of all right cosets of $H$ in $G$ and the canonical surjective map

$$\begin{aligned} \rho\colon G &\to H\backslash G \\ g &\mapsto Hg. \end{aligned}$$

Given an element $y \in A^{H\backslash G}$, i.e., a map $y\colon H\backslash G \to A$, we can form the composite map $y \circ \rho\colon G \to A$ which is an element of $A^G$. In fact, we have $y \circ \rho \in \mathrm{Fix}(H)$ since

$$(h(y \circ \rho))(g) = y \circ \rho(h^{-1}g) = y(\rho(h^{-1}g)) = y(\rho(g)) = y \circ \rho(g)$$

for all $g \in G$ and $h \in H$.

**Proposition 1.3.3.** *Let $H$ be a subgroup of $G$ and let denote by $\rho\colon G \to H\backslash G$ the canonical surjection. Then the map $\rho^*\colon A^{H\backslash G} \to \mathrm{Fix}(H)$ defined by $\rho^*(y) = y \circ \rho$ for all $y \in A^{H\backslash G}$ is bijective.*

*Proof.* If $y_1, y_2 \in A^{H\backslash G}$ satisfy $y_1 \circ \rho = y_2 \circ \rho$, then $y_1 = y_2$ since $\rho$ is surjective. Thus $\rho^*$ is injective.

If $x \in \mathrm{Fix}(H)$, then $hx = x$ for all $h \in H$, that is,

$$x(h^{-1}g) = x(g) \quad \text{for all } h \in H, g \in G.$$

Thus, the configuration $x$ is constant on each right coset of $G$ modulo $H$, that is, $x$ is in the image of $\rho^*$. This shows that $\rho^*$ is surjective. □

**Corollary 1.3.4.** *If the set $A$ is finite and $H$ is a subgroup of finite index of $G$, then the set* $\mathrm{Fix}(H)$ *is finite and one has* $|\mathrm{Fix}(H)| = |A|^{[G:H]}$, *where* $[G:H]$ *denotes the index of $H$ in $G$.* □

*Example 1.3.5.* Let $G = \mathbb{Z}$ and $H = n\mathbb{Z}$, where $n \geq 1$. If $A$ is finite of cardinality $k$, then $|\mathrm{Fix}(H)| = k^n$.

Suppose now that $N$ is a normal subgroup of $G$, that is, $gN = Ng$ for all $g \in G$. Then, there is a natural group structure on $G/N = N\backslash G$ for which the canonical surjection $\rho\colon G \to G/N$ is a homomorphism.

**Proposition 1.3.6.** *Let $N$ be a normal subgroup of $G$. Then* $\mathrm{Fix}(N)$ *is a $G$-invariant subset of $A^G$.*

*Proof.* Let $x \in \mathrm{Fix}(N)$ and $g \in G$. Given $h \in N$, then there exists $h' \in N$ such that $hg = gh'$, since $N$ is normal in $G$. Thus, we have

$$h(gx) = g(h'x) = gx$$

which shows that $gx \in \mathrm{Fix}(N)$. □

Since every element of $\mathrm{Fix}(N)$ is fixed by $N$, the action of $G$ on $\mathrm{Fix}(N)$ induces an action of $G/N$ on $\mathrm{Fix}(N)$ which satisfies $\rho(g)x = gx$ for all $g \in G$ and $x \in \mathrm{Fix}(N)$.

Suppose that a group $\Gamma$ acts on two sets $X$ and $Y$. A map $\varphi\colon X \to Y$ is called *$\Gamma$-equivariant* if one has $\varphi(\gamma x) = \gamma\varphi(x)$ for all $\gamma \in \Gamma$ and $x \in X$.

**Proposition 1.3.7.** *Let $N$ be a normal subgroup of $G$ and let $\rho\colon G \to G/N$ denote the canonical epimorphism. Then the map $\rho^*\colon A^{G/N} \to \mathrm{Fix}(N)$ defined by $\rho^*(y) = y \circ \rho$ for all $y \in A^{G/N}$ is a $G/N$-equivariant bijection.*

*Proof.* We already know that $\rho^*$ is bijective (Proposition 1.3.3).

Let $g \in G$ and $y \in A^{G/N}$. For all $g' \in G$, we have

$$\begin{aligned}
\rho(g)\rho^*(y)(g') &= g\rho^*(y)(g') \\
&= \rho^*(y)(g^{-1}g') \\
&= (y \circ \rho)(g^{-1}g') \\
&= y(\rho(g^{-1}g')) \\
&= y((\rho(g))^{-1}\rho(g')) \\
&= \rho(g)y(\rho(g')) \\
&= \rho^*(\rho(g)y)(g').
\end{aligned}$$

Thus $\rho(g)\rho^*(y) = \rho^*(\rho(g)y)$. This shows that $\rho^*$ is $G/N$-equivariant. □

## 1.4 Cellular Automata

Let $G$ be a group and let $A$ be a set.

**Definition 1.4.1.** A *cellular automaton* over the group $G$ and the alphabet $A$ is a map $\tau\colon A^G \to A^G$ satisfying the following property: there exist a finite subset $S \subset G$ and a map $\mu\colon A^S \to A$ such that

$$\tau(x)(g) = \mu((g^{-1}x)|_S) \tag{1.5}$$

for all $x \in A^G$ and $g \in G$, where $(g^{-1}x)|_S$ denotes the restriction of the configuration $g^{-1}x$ to $S$.

Such a set $S$ is called a *memory set* and $\mu$ is called a *local defining map* for $\tau$.

Observe that formula (1.5) says that the value of the configuration $\tau(x)$ at an element $g \in G$ is the value taken by the local defining map $\mu$ at the pattern obtained by restricting to the memory set $S$ the shifted configuration $g^{-1}x$.

*Remark 1.4.2.* (a) Equality (1.5) may also be written

$$\tau(x)(g) = \mu((x \circ L_g)|_S) \tag{1.6}$$

by (1.2).

(b) For $g = 1_G$, formula (1.5) gives us

$$\tau(x)(1_G) = \mu(x|_S). \tag{1.7}$$

As the restriction map $A^G \to A^S$, $x \mapsto x|_S$, is surjective, this shows that if $S$ is a memory set for the cellular automaton $\tau$, then there is a unique map $\mu\colon A^S \to A$ which satisfies (1.5). Thus one says that this unique $\mu$ is *the* local defining map for $\tau$ *associated with* the memory set $S$.

*Examples 1.4.3.* (a) **The cellular automaton associated with the Game of Life.** Consider an infinite two-dimensional orthogonal grid of square cells, each of which is in one of two possible states, live or dead. Every cell $c$ interacts with its eight neighboring cells, namely the North, North-East, East, South-East, South, South-West, West and North-West cells (see Fig. 1.1).

At each step in time, the following rules for the evolution of the states of the cells are applied (in Figs. 1.2–1.5 we label with a "•" a live cell and with a "∘" a dead cell):

- **(birth):** a cell that is dead at time $t$ becomes alive at time $t + 1$ if and only if three of its neighbors are alive at time $t$ (cf. Fig. 1.2);
- **(survival):** a cell that is alive at time $t$ will remain alive at time $t + 1$ if and only if it has exactly two or three live neighbors at time $t$ (cf. Fig. 1.3);

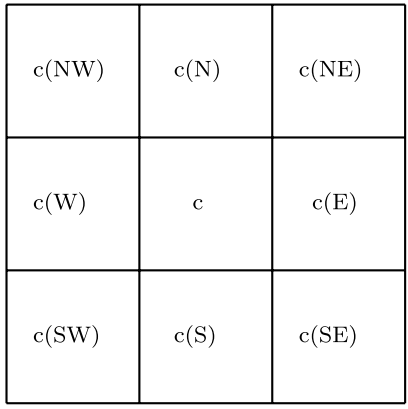



**Fig. 1.1** The cell $c$ and its eight neighboring cells

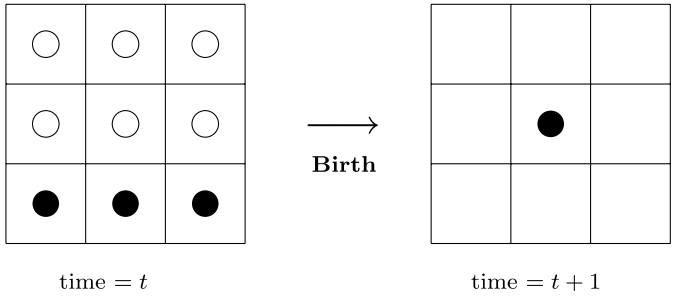



**Fig. 1.2** A cell that is dead at time $t$ becomes alive at time $t + 1$ if and only if three of its neighbors are alive at time $t$

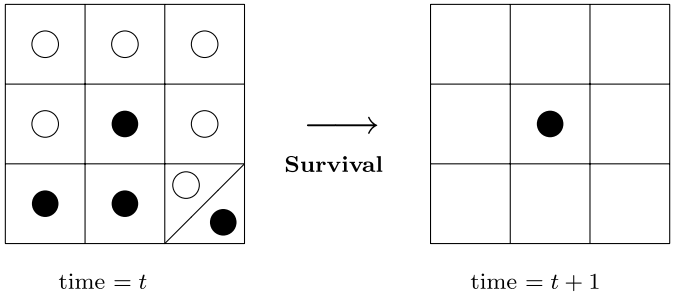



**Fig. 1.3** A cell that is alive at time $t$ will remain alive at time $t + 1$ if and only if it has exactly two or three live neighbors at time $t$

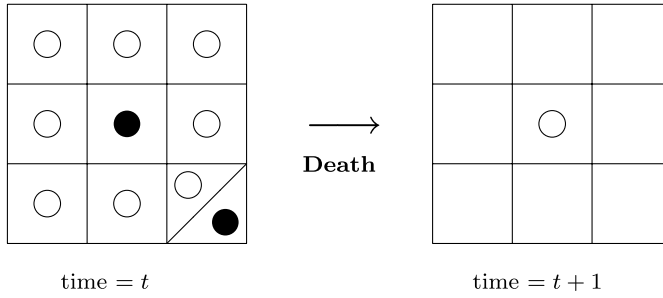



**Fig. 1.4** A live cell that has at most one live neighbor at time $t$ will be dead at time $t+1$

- **(death by loneliness):** a live cell that has at most one live neighbor at time $t$ will be dead at time $t+1$ (cf. Fig. 1.4);
- **(death by overcrowding):** a cell that is alive at time $t$ and has four or more live neighbors at time $t$, will be dead at time $t+1$ (cf. Fig. 1.5).

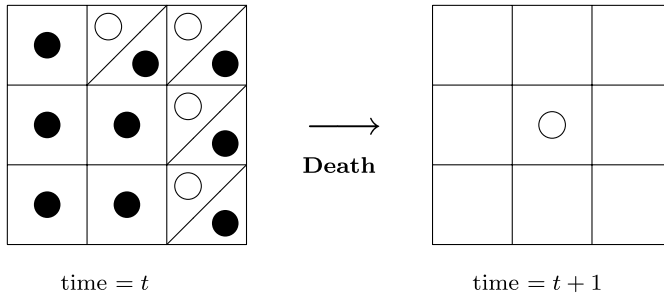



**Fig. 1.5** A cell that is alive at time $t$ and has four or more live neighbors at time $t$, will be dead at time $t+1$

Let us show that the map which transforms a configuration of cells at time $t$ into the configuration at time $t+1$ according to the above rules is indeed a cellular automaton.

Consider the group $G = \mathbb{Z}^2$ and the finite set $S = \{-1, 0, 1\}^2 \subset G$. Then there is a one-to-one correspondence between the cells in the grid and the elements in $G$ in such a way that the following holds. If $c$ is a given cell, then $c+(0,1)$ is the neighboring North cell, $c+(1,1)$ is the neighboring North-East cell, and so on; in other words, $c$ and its eight neighboring cells correspond to the group elements $c+s$ with $s \in S$ (see Fig. 1.6).

Consider the alphabet $A = \{0, 1\}$. The state 0 (resp. 1) corresponds to *absence* (resp. *presence*) *of life*. With each configuration of the states of the cells in the grid we associate a map $x \in A^G$ defined as follows. Given a cell $c$ we set $x(c) = 1$ (resp. 0) if the cell $c$ is alive (resp. dead).

| | | |
|---|---|---|
| c+(-1,1) | c+(0,1) | c+(1,1) |
| c+(-1,0) | c | c+(1,0) |
| c+(-1,-1) | c+(0,-1) | c+(1,-1) |

**Fig. 1.6** The cell $c$ and its eight neighboring cells $c + s$, $s \in S = \{-1, 0, 1\}^2$

Consider the map $\mu\colon A^S \to A$ given by

$$\mu(y) = \begin{cases} 1 & \text{if } \begin{cases} \sum_{s\in S} y(s) = 3 \\ \text{or} \\ \sum_{s\in S} y(s) = 4 \text{ and } y((0,0)) = 1, \end{cases} \\ 0 & \text{otherwise} \end{cases} \tag{1.8}$$

for all $y \in A^S$.

A moment of thought tells us that $\mu$ just expresses the rules for the Game of Life.

The cellular automaton $\tau\colon A^G \to A^G$ with memory set $S$ and local defining map $\mu$ is called the cellular automaton associated with the *Game of Life*.

(b) **The Discrete Laplacian.** Let $G = \mathbb{Z}$ and $A = \mathbb{R}$. Consider the map $\Delta\colon \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}^{\mathbb{Z}}$ defined by

$$\Delta(x)(n) = 2x(n) - x(n-1) - x(n+1).$$

Then $\Delta$ is the cellular automaton over $\mathbb{Z}$ with memory set $S = \{-1, 0, 1\}$ and local defining map $\mu\colon \mathbb{R}^S \to \mathbb{R}$ given by

$$\mu(y) = 2y(0) - y(-1) - y(1) \quad \text{for all } y \in \mathbb{R}^S.$$

This may be generalized in the following way. Let $G$ be an arbitrary group and let $S$ be a nonempty finite subset of $G$. Let $\mathbb{K}$ be a field. Consider the map $\Delta_S = \Delta_S^G\colon \mathbb{K}^G \to \mathbb{K}^G$ defined by

$$\Delta_S(x)(g) = |S|x(g) - \sum_{s\in S} x(gs).$$

Then $\Delta_S$ is a cellular automaton over $G$ with memory set $S \cup \{1_G\}$ and local defining map $\mu\colon \mathbb{K}^{S\cup\{1_G\}} \to \mathbb{K}$ given by

$$\mu(y) = |S|y(1_G) - \sum_{s\in S} y(s) \quad \text{for all } y \in \mathbb{K}^{S\cup\{1_G\}}.$$

This cellular automaton is called the *discrete Laplacian* over $\mathbb{K}$ associated with $G$ and $S$.

(c) **The Majority action cellular automaton.** Let $G$ be a group and let $S$ be a finite subset of $G$. Take $A = \{0, 1\}$ and consider the map $\tau\colon A^G \to A^G$ defined by

$$\tau(x)(g) = \begin{cases} 1 & \text{if } \sum_{s\in S} x(gs) > \frac{|S|}{2} \\ 0 & \text{if } \sum_{s\in S} x(gs) < \frac{|S|}{2} \\ x(g) & \text{if } \sum_{s\in S} x(gs) = \frac{|S|}{2} \end{cases}$$

for all $x \in A^G$. Then $\tau$ is a cellular automaton over $G$ with memory set $S \cup \{1_G\}$ and local defining map $\mu\colon A^{S\cup\{1_G\}} \to A$ given by

$$\mu(y) = \begin{cases} 1 & \text{if } \sum_{s\in S} y(s) > \frac{|S|}{2} \\ 0 & \text{if } \sum_{s\in S} y(s) < \frac{|S|}{2} \\ y(1_G) & \text{if } \sum_{s\in S} y(s) = \frac{|S|}{2} \end{cases}$$

for all $y \in A^{S\cup\{1_G\}}$.

The cellular automaton $\tau$ is called the *majority action* cellular automaton associated with $G$ and $S$ (see Figs. 1.7–1.8).

The terminology comes from the fact that given $x \in A^G$ and $g \in G$, the value $\tau(x)(g)$ is equal to $a \in \{0, 1\}$ if there is a strict majority of elements of $gS$ at which the configuration $x$ takes the value $a$, or to $x(g)$ if no such majority exists.

(d) Let $G$ be a group, $A$ a set, and $f\colon A \to A$ a map from $A$ into itself. Then the map $\tau\colon A^G \to A^G$ defined by $\tau(x) = f \circ x$ is a cellular automaton with memory set $S = \{1_G\}$ and local defining map $\mu\colon A^S \to A$ given by $\mu(y) = f(y(1_G))$. Note that, if $f$ is the identity map $\mathrm{Id}_A$ on $A$, then $\tau$ equals the identity map $\mathrm{Id}_{A^G}$ on $A^G$.

(e) Let $G$ be a group, $A$ a set, and $s_0$ an element of $G$. Let $R_{s_0}\colon G \to G$ denote the right multiplication by $s_0$ in $G$, that is, the map $R_{s_0}\colon G \to G$ defined by $R_{s_0}(g) = gs_0$. Then the map $\tau\colon A^G \to A^G$ defined by $\tau(x) = x \circ R_{s_0}$ is a cellular automaton with memory set $S = \{s_0\}$ and local defining map $\mu\colon A^S \to A$ given by $\mu(y) = y(s_0)$.

**Proposition 1.4.4.** *Let $G$ be a group and let $A$ be a set. Then every cellular automaton $\tau\colon A^G \to A^G$ is $G$-equivariant.*

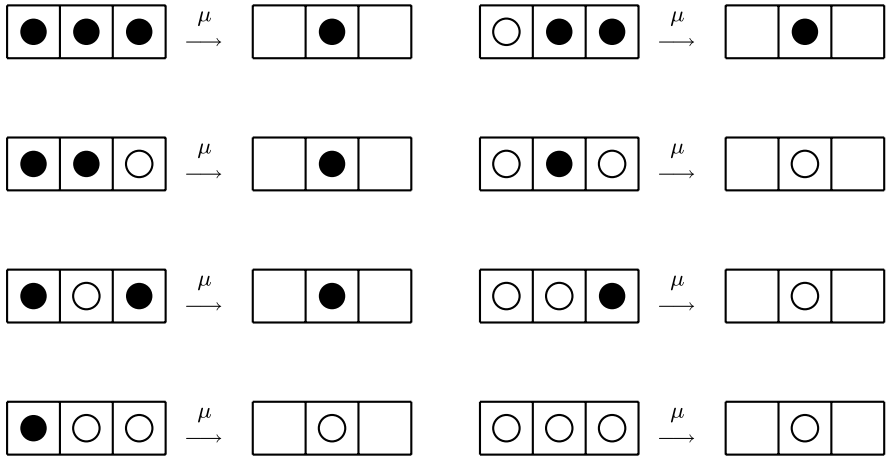


**Fig. 1.7** The local defining map $\mu$ for the majority action on $\mathbb{Z}$ associated with $S = \{+1, -1\}$

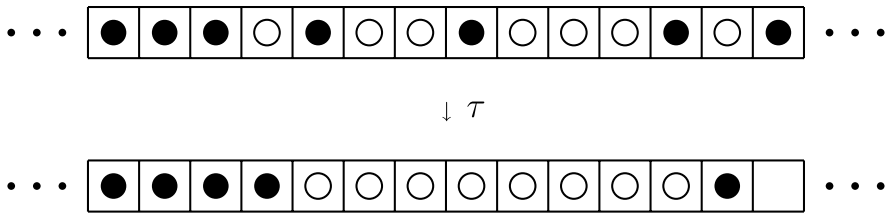


**Fig. 1.8** The majority action $\tau$ on $\mathbb{Z}$ associated with $S = \{+1, -1\}$

*Proof.* Let $S$ be a memory set for $\tau$ and let $\mu\colon A^S \to A$ be the associated local defining map. For all $g, h \in G$ and $x \in A^G$, we have

$$\tau(gx)(h) = \mu((h^{-1}gx)|_S) = \mu(((g^{-1}h)^{-1}x)|_S) = \tau(x)(g^{-1}h) = g\tau(x)(h).$$

Thus $\tau(gx) = g\tau(x)$. □

**Corollary 1.4.5.** *Let $\tau\colon A^G \to A^G$ be a cellular automaton and let $H$ be a subgroup of $G$. Then one has $\tau(\mathrm{Fix}(H)) \subset \mathrm{Fix}(H)$.*

*Proof.* Let $x \in \mathrm{Fix}(H)$. By the previous Proposition, we have, for every $h \in H$,

$$h\tau(x) = \tau(hx) = \tau(x).$$

Thus $\tau(x) \in \mathrm{Fix}(H)$. □

The following characterization of cellular automata will be useful in the sequel.

**Proposition 1.4.6.** *Let $G$ be a group and let $A$ be a set. Consider a map $\tau\colon A^G \to A^G$. Let $S$ be a finite subset of $G$ and let $\mu\colon A^S \to A$. Then the following conditions are equivalent:*

(a) *$\tau$ is a cellular automaton admitting $S$ as a memory set and $\mu$ as the associated local defining map;*
(b) *$\tau$ is $G$-equivariant and one has $\tau(x)(1_G) = \mu(x|_S)$ for every $x \in A^G$.*

*Proof.* The fact that (a) implies (b) follows from Proposition 1.4.4 and formula (1.7)

Conversely, suppose (b). Then, by using the $G$-equivariance of $\tau$, we get

$$\tau(x)(g) = \tau(g^{-1}x)(1_G) = \mu((g^{-1}x)|_S)$$

for all $x \in A^G$ and $g \in G$. Consequently, $\tau$ satisfies (a). □

An important feature of cellular automata is their continuity (with respect to the prodiscrete topology). In the proof of this property, we shall use the following.

**Lemma 1.4.7.** *Let $G$ be a group and let $A$ be a set. Let $\tau\colon A^G \to A^G$ be a cellular automaton with memory set $S$ and let $g \in G$. Then $\tau(x)(g)$ depends only on the restriction of $x$ to $gS$.*

*Proof.* This is an immediate consequence of (1.5) since $(g^{-1}x)(s) = x(gs)$ for all $s \in S$. □

**Proposition 1.4.8.** *Let $G$ be a group and let $A$ be a set. Then every cellular automaton $\tau\colon A^G \to A^G$ is continuous.*

*Proof.* Let $S$ be a memory set for $\tau$. Let $x \in A^G$ and let $W$ be a neighborhood of $\tau(x)$ in $A^G$. Then we can find a finite subset $\Omega \subset G$ such that (cf. equation (1.3))

$$V(\tau(x), \Omega) \subset W.$$

Consider the finite set $\Omega S = \{gs : g \in \Omega, s \in S\}$. If $y \in A^G$ coincide with $x$ on $\Omega S$, then $\tau(y)$ and $\tau(x)$ coincide on $\Omega$ by Lemma 1.4.7. Thus, we have

$$\tau(V(x, \Omega S)) \subset V(\tau(x), \Omega) \subset W.$$

This shows that $\tau$ is continuous. □

**Proposition 1.4.9.** *Let $G$ be a group and let $A$ be a set. Let $\sigma\colon A^G \to A^G$ and $\tau\colon A^G \to A^G$ be cellular automata. Then the composite map $\sigma \circ \tau\colon A^G \to A^G$ is a cellular automaton. Moreover, if $S$ (resp. $T$) is a memory set for $\sigma$ (resp. $\tau$), then $ST = \{st : s \in S, t \in T\}$ is a memory set for $\sigma \circ \tau$.*

*Proof.* It is clear that the map $\sigma \circ \tau$ is $G$-equivariant since $\sigma$ and $\tau$ are $G$-equivariant (by Proposition 1.4.4). Let $S$ (resp. $T$) be a memory set for

$\sigma$ (resp. $\tau$). For every $x \in A^G$, we have $\sigma \circ \tau(x)(1_G) = \sigma(\tau(x))(1_G)$. By Lemma 1.4.7, $\sigma(\tau(x))(1_G)$ depends only on the restriction of $\tau(x)$ to $S$. By using Lemma 1.4.7 again, we deduce that, for every $s \in S$, the element $\tau(x)(s)$ depends only on the restriction of $x$ to $sT$. Therefore, $\sigma \circ \tau(x)(1_G)$ depends only on the restriction of $x$ to $ST$. By applying Proposition 1.4.6, we conclude that $\sigma \circ \tau$ is a cellular automaton admitting $ST$ as a memory set. □

*Remark 1.4.10.* With the hypotheses and notation of the previous proposition, denote by $\mu\colon A^S \to A$ and $\nu\colon A^T \to A$ the local defining maps for $\sigma$ and $\tau$, respectively. Then, the local defining map $\kappa\colon A^{ST} \to A$ for $\sigma \circ \tau$ may be described in the following way.

For $y \in A^{ST}$ and $s \in S$ define $y_s \in A^T$ by setting $y_s(t) = y(st)$ for all $t \in T$. Also, denote by $\overline{y} \in A^S$ the map defined by $\overline{y}(s) = \nu(y_s)$ for all $s \in S$. We finally define the map $\kappa\colon A^{ST} \to A$ by setting

$$\kappa(y) = \mu(\overline{y}) \tag{1.9}$$

for all $y \in A^{ST}$.

Let $x \in A^G$, $g \in G$, $s \in S$, and $t \in T$. We then have

$$\begin{aligned}(s^{-1}g^{-1}x)|_T(t) &= s^{-1}g^{-1}x(t)\\ &= g^{-1}x(st)\\ &= (g^{-1}x)|_{ST}(st)\\ &= \left((g^{-1}x)|_{ST}\right)_s(t).\end{aligned}$$

This shows that

$$(s^{-1}g^{-1}x)|_T = \left((g^{-1}x)|_{ST}\right)_s$$

and therefore

$$\tau(g^{-1}x)(s) = \nu\left((s^{-1}g^{-1}x)|_T\right) = \nu\left(\left((g^{-1}x)|_{ST}\right)_s\right) = \overline{(g^{-1}x)|_{ST}}(s).$$

As a consequence,

$$\tau(g^{-1}x)|_S = \overline{(g^{-1}x)|_{ST}}. \tag{1.10}$$

Finally, one has

$$\begin{aligned}(\sigma \circ \tau)(x)(g) &= \sigma(\tau(x))(g)\\ &= \mu\left((g^{-1}\tau(x))|_S\right)\\ &= \mu(\tau(g^{-1}x)|_S)\\ \text{(by (1.10))} \quad &= \mu(\overline{(g^{-1}x)|_{ST}})\\ \text{(by (1.9))} \quad &= \kappa\left((g^{-1}x)|_{ST}\right).\end{aligned} \tag{1.11}$$

Recall that a *monoid* is a set equipped with an associative binary operation admitting an identity element. Denote by $\mathrm{CA}(G; A)$ the set consisting of all

cellular automata $\tau\colon A^G \to A^G$. In Example 1.4.3(d) we have seen that the identity map $\mathrm{Id}_{A^G}\colon A^G \to A^G$ is a cellular automaton. Thus we have:

**Corollary 1.4.11.** *The set* $\mathrm{CA}(G;A)$ *is a monoid for the composition of maps.* □

## 1.5 Minimal Memory

Let $G$ be a group and let $A$ be a set. Let $\tau\colon A^G \to A^G$ be a cellular automaton. Let $S$ be a memory set for $\tau$ and let $\mu\colon A^S \to A$ be the associated defining map. If $S'$ is a finite subset of $G$ such that $S \subset S'$, then $S'$ is also a memory set for $\tau$ and the local defining map associated with $S'$ is the map $\mu'\colon A^{S'} \to A$ given by $\mu' = \mu \circ p$, where $p\colon A^{S'} \to A^S$ is the canonical projection (restriction map). This shows that the memory set of a cellular automaton is not unique in general. However, we shall see that every cellular automaton admits a unique memory set of minimal cardinality. Let us first establish the following result.

**Lemma 1.5.1.** *Let* $\tau\colon A^G \to A^G$ *be a cellular automaton. Let* $S_1$ *and* $S_2$ *be memory sets for* $\tau$*. Then* $S_1 \cap S_2$ *is also a memory set for* $\tau$*.*

*Proof.* Let $x \in A^G$. Let us show that $\tau(x)(1_G)$ depends only on the restriction of $x$ to $S_1 \cap S_2$. To see this, consider an element $y \in A^G$ such that $x|_{S_1\cap S_2} = y|_{S_1\cap S_2}$. Let us choose an element $z \in A^G$ such that $z|_{S_1} = x|_{S_1}$ and $z|_{S_2} = y|_{S_2}$ (we may take for instance the configuration $z \in A^G$ which coincides with $x$ on $S_1$ and with $y$ on $G \setminus S_1$). We have $\tau(x)(1_G) = \tau(z)(1_G)$ since $x$ and $z$ coincide on $S_1$, which is a memory set for $\tau$. On the other hand, we have $\tau(y)(1_G) = \tau(z)(1_G)$ since $y$ and $z$ coincide on $S_2$, which is also a memory set for $\tau$. It follows that $\tau(x)(1_G) = \tau(y)(1_G)$.

Thus there exists a map $\mu\colon A^{S_1\cap S_2} \to A$ such that

$$\tau(x)(1_G) = \mu(x|_{S_1\cap S_2}) \quad \text{for all } x \in A^G.$$

As $\tau$ is $G$-equivariant (Proposition 1.4.4), we deduce that $S_1 \cap S_2$ is a memory set for $\tau$ by using Proposition 1.4.6. □

**Proposition 1.5.2.** *Let* $\tau\colon A^G \to A^G$ *be a cellular automaton. Then there exists a unique memory set* $S_0 \subset G$ *for* $\tau$ *of minimal cardinality. Moreover, if* $S$ *is a finite subset of* $G$*, then* $S$ *is a memory set for* $\tau$ *if and only if* $S_0 \subset S$*.*

*Proof.* Let $S_0$ be a memory set for $\tau$ of minimal cardinality. As we have seen at the beginning of this section, every finite subset of $G$ containing $S_0$ is also a memory set for $\tau$. Conversely, let $S$ be a memory set for $\tau$. As $S \cap S_0$ is a memory set for $\tau$ by Lemma 1.5.1, we have $|S \cap S_0| \geq |S_0|$. This implies

$S \cap S_0 = S_0$, that is, $S_0 \subset S$. In particular, $S_0$ is the unique memory set of minimal cardinality. □

The memory set of minimal cardinality of a cellular automaton is called its *minimal* memory set.

*Remark 1.5.3.* A map $F\colon A^G \to A^G$ is constant if there exists a configuration $x_0 \in A^G$ such that $F(x) = x_0$ for all $x \in A^G$. By $G$-equivariance, a cellular automaton $\tau\colon A^G \to A^G$ is constant if and only if there exists $a \in A$ such that $\tau(x)(g) = a$ for all $x \in A^G$ and $g \in G$. Observe that a cellular automaton $\tau\colon A^G \to A^G$ is constant if and only if its minimal memory set is the empty set.

## 1.6 Cellular Automata over Quotient Groups

Let $G$ be a group and let $A$ be a set. Let $\tau\colon A^G \to A^G$ be a cellular automaton. Suppose that $N$ is a normal subgroup of $G$ and let $\rho\colon G \to G/N$ denote the canonical epimorphism. It follows from Proposition 1.3.7 that the map $\rho^*\colon A^{G/N} \to \mathrm{Fix}(N)$, defined by $\rho^*(y) = y \circ \rho$ for all $y \in A^{G/N}$, is a bijection from the set $A^{G/N}$ of configurations over the group $G/N$ onto the set $\mathrm{Fix}(N) \subset A^G$ of $N$-periodic configurations over $G$. On the other hand, the set $\mathrm{Fix}(N)$ satisfies $\tau(\mathrm{Fix}(N)) \subset \mathrm{Fix}(N)$ by Corollary 1.4.5. Thus, we can define a map $\overline{\tau}\colon A^{G/N} \to A^{G/N}$ by setting

$$\overline{\tau} = (\rho^*)^{-1} \circ \tau|_{\mathrm{Fix}(N)} \circ \rho^*. \tag{1.12}$$

In other words, the map $\overline{\tau}$ is obtained by conjugating by $\rho^*$ the restriction of $\tau$ to $\mathrm{Fix}(N)$, so that the diagram

$$\begin{array}{ccc} A^{G/N} & \xrightarrow{\rho^*} & \mathrm{Fix}(N) \subset A^G \\ {\scriptstyle\overline{\tau}}\downarrow & & \downarrow{\scriptstyle\tau|_{\mathrm{Fix}(N)}} \\ A^{G/N} & \xrightarrow[\rho^*]{} & \mathrm{Fix}(N) \end{array}$$

is commutative.

Suppose that $S \subset G$ is a memory set for $\tau$ and that $\mu\colon A^S \to A$ is the associated local defining map. Consider the finite subset $\overline{S} = \rho(S) \subset G/N$ and the map $\overline{\mu}\colon A^{\overline{S}} \to A$ defined by $\overline{\mu} = \mu \circ \pi$, where $\pi\colon A^{\overline{S}} \to A^S$ is the injective map induced by $\rho$.

**Proposition 1.6.1.** *The map $\overline{\tau}\colon A^{G/N} \to A^{G/N}$ is a cellular automaton over the group $G/N$ admitting $\overline{S}$ as a memory set and $\overline{\mu}\colon A^{\overline{S}} \to A$ as the associated local defining map.*

*Proof.* Let $y \in A^{G/N}$, $g \in G$, and $\overline{g} = \rho(g)$. We have

$$\begin{aligned}\overline{\tau}(y)(\overline{g}) &= \tau(y \circ \rho)(g)\\ &= \mu((g^{-1}(y \circ \rho))|_S)\\ &= \overline{\mu}((\overline{g}^{-1}y)|_{\overline{S}}).\end{aligned}$$

Thus $\overline{\tau}$ is a cellular automaton with memory set $\overline{S}$ and local defining map $\overline{\mu}\colon A^{\overline{S}} \to A$. □

Consider now the map $\Phi\colon \mathrm{CA}(G; A) \to \mathrm{CA}(G/N; A)$ given by $\Phi(\tau) = \overline{\tau}$, where $\overline{\tau}$ is defined by (1.12). We have the following:

**Proposition 1.6.2.** *The map $\Phi\colon \mathrm{CA}(G; A) \to \mathrm{CA}(G/N; A)$ is a monoid epimorphism.*

*Proof.* Let $\sigma\colon A^{G/N} \to A^{G/N}$ be a cellular automaton over $G/N$ with memory set $T \subset G/N$ and local defining map $\nu\colon A^T \to A$. Let $S \subset G$ be a finite set such that $\rho$ induces a bijection $\phi\colon S \to T$. Consider the map $\mu\colon A^S \to A$ defined by $\mu(y) = \nu(y \circ \phi^{-1})$ for all $y \in A^S$. Let $\tau\colon A^G \to A^G$ be the cellular automaton over $G$ with memory set $S$ and local defining map $\mu$. We have

$$\overline{\mu}(z) = (\mu \circ \pi)(z) = \nu(\pi(z) \circ \phi^{-1}) = \nu(z)$$

for all $z \in A^{\overline{S}}$. It follows that $\overline{\mu} = \nu$ and $\overline{\tau} = \sigma$. This shows that $\Phi$ is surjective.

The fact that $\Phi$ is a monoid morphism immediately follows from (1.12). □

*Examples 1.6.3.* Let $G$ be a group, $S \subset G$ a finite subset, and $N$ a normal subgroup of $G$. Denote by $\rho\colon G \to G/N$ the canonical epimorphism and suppose that $\rho$ induces a bijection between $S$ and $\overline{S} = \rho(S) \subset G/N$.

(a) Consider the discrete laplacian $\Delta_S\colon \mathbb{R}^G \to \mathbb{R}^G$ associated with $G$ and $S$ (cf. Example 1.4.3(b)). Then $\Phi(\Delta_S)\colon \mathbb{R}^{G/N} \to \mathbb{R}^{G/N}$ is the discrete laplacian associated with $G/N$ and $\overline{S}$.

(b) Consider the majority action cellular automaton $\tau\colon \{0,1\}^G \to \{0,1\}^G$ associated with $G$ and $S$ (cf. Example 1.4.3(c)). Then $\Phi(\tau)\colon \{0,1\}^{G/N} \to \{0,1\}^{G/N}$ is the majority action cellular automaton associated with $G/N$ and $\overline{S}$.

## 1.7 Induction and Restriction of Cellular Automata

Let $G$ be a group and let $A$ be a set. Let $H$ be a subgroup of $G$.

Let $\mathrm{CA}(G, H; A)$ denote the set consisting of all cellular automata $\tau\colon A^G \to A^G$ admitting a memory set $S$ such that $S \subset H$. Thus, $\mathrm{CA}(G, H; A)$ is the

subset of $\mathrm{CA}(G;A)$ consisting of the cellular automata whose minimal memory set is contained in $H$.

Recall that a subset $N$ of a monoid $M$ is called a *submonoid* if the identity element $1_M$ is in $N$ and $N$ is stable under the monoid operation (that is, $xy \in N$ for all $x, y \in N$). If $N$ is a submonoid of a monoid $M$, then the monoid operation induces by restriction a monoid structure on $N$.

**Proposition 1.7.1.** *The set* $\mathrm{CA}(G,H;A)$ *is a submonoid of* $\mathrm{CA}(G;A)$.

*Proof.* The identity element of $\mathrm{CA}(G;A)$ is the identity map $\mathrm{Id}_{A^G}$. We have $\mathrm{Id}_{A^G} \in \mathrm{CA}(G,H;A)$ since $\{1_G\}$ is a memory set for $\mathrm{Id}_{A^G}$ and $\{1_G\} \subset H$. Let $\sigma, \tau \in \mathrm{CA}(G,H;A)$. Let $S$ (resp. $T$) be a memory set for $\sigma$ (resp. $\tau$) such that $S \subset H$ (resp. $T \subset H$). It follows from Proposition 1.4.9 that $ST$ is a memory set for $\sigma \circ \tau$. Since $ST \subset H$, this implies that $\sigma \circ \tau \in \mathrm{CA}(G,H;A)$. This shows that $\mathrm{CA}(G,H;A)$ is a submonoid of $\mathrm{CA}(G;A)$. □

Let $\tau \in \mathrm{CA}(G,H;A)$. Let $S$ be a memory set for $\tau$ such that $S \subset H$ and let $\mu\colon A^S \to A$ denote the associated local defining map. Then, the map $\tau_H\colon A^H \to A^H$ defined by

$$\tau_H(x)(h) = \mu((h^{-1}x)|_S) \quad \text{for all } x \in A^H, h \in H,$$

is a cellular automaton over the group $H$ with memory set $S$ and local defining map $\mu$. Observe that if $\widetilde{x} \in A^G$ is such that $\widetilde{x}|_H = x$, then

$$\tau_H(x)(h) = \tau(\widetilde{x})(h) \quad \text{for all } h \in H. \tag{1.13}$$

This shows in particular that $\tau_H$ does not depend on the choice of the memory set $S \subset H$. One says that $\tau_H$ is the *restriction* of the cellular automaton $\tau$ to $H$.

Conversely, let $\sigma\colon A^H \to A^H$ be a cellular automaton with memory set $S$ and local defining map $\mu\colon A^S \to A$. Then the map $\sigma^G\colon A^G \to A^G$ defined by

$$\sigma^G(\widetilde{x})(g) = \mu((g^{-1}\widetilde{x})|_S) \quad \text{for all } \widetilde{x} \in A^G, g \in G,$$

is a cellular automaton over $G$ with memory set $S$ and local defining map $\mu$. If $S_0$ is the minimal memory set of $\sigma$ and $\mu_0\colon A^{S_0} \to A$ is the associated local defining map then $\mu = \mu_0 \circ \pi$, where $\pi\colon A^S \to A^{S_0}$ is the restriction map (see Sect. 1.5). Thus, one has

$$\sigma^G(\widetilde{x})(g) = \mu((g^{-1}\widetilde{x})|_S) = \mu_0 \circ \pi((g^{-1}\widetilde{x})|_S) = \mu_0((g^{-1}\widetilde{x})|_{S_0})$$

for all $\widetilde{x} \in A^G$ and $g \in G$. This shows in particular that $\sigma^G$ does not depend on the choice of the memory set $S \subset H$. One says that $\sigma^G \in \mathrm{CA}(G,H;A)$ is the cellular automaton *induced* by $\sigma \in \mathrm{CA}(H;A)$.

**Proposition 1.7.2.** *The map* $\tau \mapsto \tau_H$ *is a monoid isomorphism from* $\mathrm{CA}(G,H;A)$ *onto* $\mathrm{CA}(H;A)$ *whose inverse is the map* $\sigma \mapsto \sigma^G$.

*Proof.* To simplify notation, denote by $\alpha\colon \mathrm{CA}(G,H;A) \to \mathrm{CA}(H;A)$ and $\beta\colon \mathrm{CA}(H;A) \to \mathrm{CA}(G,H;A)$ the maps defined by $\alpha(\tau) = \tau_H$ and $\beta(\sigma) = \sigma^G$ respectively. It is clear from the definitions given above that $\beta \circ \alpha$ and $\alpha \circ \beta$ are the identity maps. Therefore, $\alpha$ is bijective with inverse $\beta$.

It remains to show that $\alpha$ is a monoid homomorphism.

Let $x \in A^H$ and let $\widetilde{x} \in A^G$ extending $x$. By applying (1.13), we get

$$\alpha(\mathrm{Id}_{A^G})(x)(h) = \mathrm{Id}_{A^G}(\widetilde{x})(h) = \widetilde{x}(h) = x(h)$$

for all $h \in H$. This shows that $\alpha(\mathrm{Id}_{A^G})(x) = x$ for all $x \in A^H$, that is, $\alpha(\mathrm{Id}_{A^G}) = \mathrm{Id}_{A^H}$.

Let $\sigma, \tau \in \mathrm{CA}(G,H;A)$. Let $x \in A^H$ and let $\widetilde{x} \in A^G$ extending $x$. By applying (1.13) again, we have

$$\alpha(\sigma \circ \tau)(x)(h) = (\sigma \circ \tau)(\widetilde{x})(h) = \sigma(\tau(\widetilde{x}))(h) \tag{1.14}$$

for all $h \in H$. On the other hand, since $\tau(\widetilde{x})$ extends $\alpha(\tau)(x)$, we have

$$\alpha(\sigma)(\alpha(\tau)(x))(h) = \sigma(\tau(\widetilde{x}))(h)$$

that is,

$$(\alpha(\sigma) \circ \alpha(\tau))(x)(h) = \sigma(\tau(\widetilde{x}))(h) \tag{1.15}$$

for all $h \in H$. From (1.14) and (1.15), we deduce that $\alpha(\sigma \circ \tau)(x) = (\alpha(\sigma) \circ \alpha(\tau))(x)$ for all $x \in A^H$, that is, $\alpha(\sigma \circ \tau) = \alpha(\sigma) \circ \alpha(\tau)$. □

Let $\tau \in \mathrm{CA}(G,H;A)$. In order to analyze the way $\tau$ transforms a configuration $\widetilde{x} \in A^G$, we now introduce the set $G/H = \{gH : g \in G\}$ consisting of all left cosets of $H$ in $G$. Since the cosets $c \in G/H$ form a partition of $G$, we have a natural identification $A^G = \prod_{c \in G/H} A^c$. With this identification, we have

$$\widetilde{x} = (\widetilde{x}|_c)_{c \in G/H}$$

for each $\widetilde{x} \in A^G$, where $\widetilde{x}|_c \in A^c$ denotes the restriction of $\widetilde{x}$ to $c$. Observe now that if $c \in G/H$ and $g \in c$, then $\tau(\widetilde{x})(g)$ depends only on $\widetilde{x}|_c$ (this directly follows from Lemma 1.4.7 since if $S$ is a memory set for $\tau$ with $S \subset H$, then $gS \subset c$). This implies that $\tau$ may be written as a product

$$\tau = \prod_{c \in G/H} \tau_c, \tag{1.16}$$

where $\tau_c\colon A^c \to A^c$ is the unique map which satisfies $\tau_c(\widetilde{x}|_c) = (\tau(\widetilde{x}))|_c$ for all $\widetilde{x} \in A^G$. Note that the notation is coherent when $c = H$, since, in this case, $\tau_c = \tau_H\colon A^H \to A^H$ is the cellular automaton obtained by restriction of $\tau$ to $H$.

Given a coset $c \in G/H$ and an element $g \in c$, denote by $\phi_g\colon H \to c$ the bijective map defined by $\phi_g(h) = gh$ for all $h \in H$. Then $\phi_g$ induces a bijective map $\phi_g^*\colon A^c \to A^H$ given by

$$\phi_g^*(x) = x \circ \phi_g \tag{1.17}$$

for all $x \in A^c$. It turns out that the maps $\tau_c$ and $\tau_H$ are conjugate by $\phi_g^*$:

**Proposition 1.7.3.** *With the above notation, we have,*

$$\tau_c = (\phi_g^*)^{-1} \circ \tau_H \circ \phi_g^*. \tag{1.18}$$

*In other words, the following diagram*

$$\begin{array}{ccc} A^c & \xrightarrow{\tau_c} & A^c \\ {\scriptstyle \phi_g^*}\downarrow & & \downarrow{\scriptstyle \phi_g^*} \\ A^H & \xrightarrow[\tau_H]{} & A^H \end{array}$$

*is commutative.*

*Proof.* Let $x \in A^c$ and let $\widetilde{x} \in A^G$ extending $x$. For all $h \in H$, we have

$$\begin{aligned} (\phi_g^* \circ \tau_c)(x)(h) &= \phi_g^*(\tau_c(x))(h) \\ &= (\tau_c(x) \circ \phi_g)(h) \\ &= \tau_c(x)(gh) \\ &= \tau(\widetilde{x})(gh) \\ &= g^{-1}\tau(\widetilde{x})(h) \\ &= \tau(g^{-1}\widetilde{x})(h), \end{aligned}$$

where the last equality follows from the $G$-equivariance of $\tau$ (Proposition 1.4.4). Now observe that the configuration $g^{-1}\widetilde{x} \in A^G$ extends $x \circ \phi_g \in A^H$. Thus, we have

$$(\phi_g^* \circ \tau_c)(x)(h) = \tau_H(x \circ \phi_g)(h) = \tau_H(\phi_g^*(x))(h) = (\tau_H \circ \phi_g^*)(x)(h).$$

This shows that $\phi_g^* \circ \tau_c = \tau_H \circ \phi_g^*$, which gives (1.18) since $\phi_g^*$ is bijective. □

The following statement will be used in the proof of Proposition 3.2.1:

**Proposition 1.7.4.** *Let $G$ be a group and let $A$ be a set. Let $H$ be a subgroup of $G$ and let $\tau \in \mathrm{CA}(G, H; A)$. Let $\tau_H \colon A^H \to A^H$ denote the cellular automaton obtained by restriction of $\tau$ to $H$. Then the following hold:*

(i) *$\tau$ is injective if and only if $\tau_H$ is injective;*
(ii) *$\tau$ is surjective if and only if $\tau_H$ is surjective;*
(iii) *$\tau$ is bijective if and only if $\tau_H$ is bijective.*

*Proof.* It immediately follows from (1.16) that $\tau$ is injective (resp. surjective, resp. bijective) if and only if $\tau_c$ is injective (resp. surjective, resp. bijective) for all $c \in G/H$.

Now, (1.18) says that, given $c \in G/H$ and $g \in G$, the map $\tau_c$ and $\tau_H$ are conjugate by the bijection $\phi_g$. We deduce that $\tau_c$ is injective (resp. surjective, resp. bijective) if and only if $\tau_H$ is injective (resp. surjective, resp. bijective).

Thus, $\tau$ is injective (resp. surjective, resp. bijective) if and only if $\tau_H$ is injective (resp. surjective, resp. bijective). □

## 1.8 Cellular Automata with Finite Alphabets

Let $G$ be a group and let $A$ be a finite alphabet. As a product of finite spaces is compact by Tychonoff theorem (see Corollary A.5.3), it follows that $A^G$ is compact. This topological property is very useful in the study of cellular automata over finite alphabets. In particular, it may be used to prove the following:

**Theorem 1.8.1 (Curtis-Hedlund theorem).** *Let $G$ be a group and let $A$ be a finite set. Let $\tau\colon A^G \to A^G$ be a map and equip $A^G$ with its prodiscrete topology. Then the following conditions are equivalent:*

(a) *the map $\tau$ is a cellular automaton;*
(b) *the map $\tau$ is continuous and $G$-equivariant.*

*Proof.* The fact that (a) implies (b) directly follows from Proposition 1.4.4 and Proposition 1.4.8 (this implication does not require the finiteness assumption on the alphabet $A$).

Conversely, suppose (b). Let us show that $\tau$ is a cellular automaton. As the map $\varphi\colon A^G \to A$ defined by $\varphi(x) = \tau(x)(1_G)$ is continuous, we can find, for each $x \in A^G$, a finite subset $\Omega_x \subset G$ such that if $y \in A^G$ coincide with $x$ on $\Omega_x$, that is, if $y \in V(x, \Omega_x)$, then $\tau(y)(1_G) = \tau(x)(1_G)$. The sets $V(x, \Omega_x)$ form an open cover of $A^G$. As $A^G$ is compact, there is a finite subset $F \subset A^G$ such that the sets $V(x, \Omega_x)$, $x \in F$, cover $A^G$. Let us set $S = \cup_{x\in F}\Omega_x$ and suppose that two configurations $y, z \in A^G$ coincide on $S$. Let $x_0 \in F$ be such that $y \in V(x_0, \Omega_{x_0})$, that is, $y|_{\Omega_{x_0}} = x_0|_{\Omega_{x_0}}$. As $S \supset \Omega_{x_0}$ we have $y|_{\Omega_{x_0}} = z|_{\Omega_{x_0}}$ and therefore $\tau(y)(1_G) = \tau(x_0)(1_G) = \tau(z)(1_G)$. Thus there is a map $\mu\colon A^S \to A$ such that $\tau(x)(1_G) = \mu(x|_S)$ for all $x \in A^G$. As $\tau$ is $G$-equivariant, it follows from Proposition 1.4.6 that $\tau$ is a cellular automaton with memory set $S$ and local defining map $\mu$. □

When the alphabet $A$ is infinite, a continuous and $G$-equivariant map $\tau\colon A^G \to A^G$ may fail to be a cellular automaton. In other words, the implication (b) $\Rightarrow$ (a) in Theorem 1.8.1 becomes false if we suppress the finiteness hypothesis on $A$. This is shown by the following example.

*Example 1.8.2.* Let $G$ be an arbitrary infinite group and take $A = G$ as the alphabet set. To avoid confusion, we denote by $g \cdot h$ the product of two elements $g$ and $h$ in $G$. Consider the map $\tau\colon A^G \to A^G$ defined by

$$\tau(x)(g) = x(g \cdot x(g))$$

for all $x \in A^G$ and $g \in G$.

Given $x \in A^G$ and $g, h \in G$ we have

$$\begin{aligned} g(\tau(x))(h) &= \tau(x)(g^{-1} \cdot h) \\ &= x(g^{-1} \cdot h \cdot x(g^{-1} \cdot h)) \\ &= x(g^{-1} \cdot h \cdot [gx](h)) \\ &= gx(h \cdot [gx](h)) \\ &= \tau(gx)(h). \end{aligned}$$

This shows that $g(\tau(x)) = \tau(gx)$ for all $x \in A^G$ and $g \in G$. Therefore, $\tau$ is $G$-equivariant.

Moreover, $\tau$ is continuous. Indeed, given $x \in A^G$ and a finite set $K \subset G$, let us show that there exists a finite set $F \subset G$ such that, if $y \in A^G$ and $y \in V(x, F)$, then $\tau(y) \in V(\tau(x), K)$. Set $F = K \cup \{k \cdot x(k) : k \in K\}$. Then, if $y \in V(x, F)$, then, for all $k \in K$ one has

$$\tau(x)(k) = x(k \cdot x(k)) = y(k \cdot x(k)) = y(k \cdot y(k)) = \tau(y)(k).$$

This shows that $\tau(y) \in V(\tau(x), K)$. Thus, $\tau$ is continuous.

However, $\tau$ is not a cellular automaton. Indeed, fix $g_0 \in G \setminus \{1_G\}$ and, for all $g \in G$, consider the configurations $x_g$ and $y_g$ in $A^G$ defined by

$$x_g(h) = \begin{cases} g & \text{if } \ h = 1_G \\ g_0 & \text{if } \ h = g \\ 1_G & \text{otherwise} \end{cases}$$

and

$$y_g(h) = \begin{cases} g & \text{if } \ h = 1_G \\ 1_G & \text{otherwise} \end{cases}$$

for all $h \in G$.

Note that $x_g|_{G\setminus\{g\}} = y_g|_{G\setminus\{g\}}$. Let $F \subset G$ be a finite set and choose $g \in G \setminus F$ (this is possible because $G$ is infinite). Then one has $x_g|_F = y_g|_F$ while

$$\tau(x_g)(1_G) = x_g(x_g(1_G)) = x_g(g) = g_0$$

and

$$\tau(y_g)(1_G) = y_g(y_g(1_G)) = y_g(g) = 1_G,$$

so that $\tau(x_g)(1_G) \neq \tau(y_g)(1_G)$. It follows that there is no finite set $F \subset G$ such that, for all $x \in A^G$, the value of $\tau(x)$ at $1_G$ only depends on the values of $x|_F$. This shows that $\tau$ is not a cellular automaton (cf. Remark 1.4.2(b)).

## 1.9 The Prodiscrete Uniform Structure

Let $G$ be a group and let $A$ be a set. The *prodiscrete* uniform structure on $A^G$ is the product uniform structure obtained by taking the discrete uniform structure on each factor $A$ of $A^G = \prod_{g\in G} A$ (see Appendix B for definition and basic facts about uniform structures).

A base of entourages for the prodiscrete uniform structure on $A^G$ is given by the sets $W_\Omega \subset A^G \times A^G$, where

$$W_\Omega = \{(x, y) \in A^G \times A^G : x|_\Omega = y|_\Omega\} \tag{1.19}$$

and $\Omega$ runs over all finite subsets of $G$.

Observe that, using the notation introduced in (1.3), we have

$$V(x, \Omega) = \{y \in A^G : (x, y) \in W_\Omega\}$$

for all $x \in A^G$.

The following statement gives a global characterization of cellular automata in terms of the prodiscrete uniform structure and the $G$-shift on $A^G$.

**Theorem 1.9.1.** *Let $A$ be a set and let $G$ be a group. Let $\tau\colon A^G \to A^G$ be a map and equip $A^G$ with its prodiscrete uniform structure. Then the following conditions are equivalent:*

(a) *$\tau$ is a cellular automaton;*
(b) *$\tau$ is uniformly continuous and $G$-equivariant.*

*Proof.* Suppose that $\tau\colon A^G \to A^G$ is a cellular automaton. We already know that $\tau$ is $G$-equivariant by Proposition 1.4.4. Let us show that $\tau$ is uniformly continuous. Let $S$ be a memory set for $\tau$. It follows from Lemma 1.4.7 that if two configurations $x, y \in A^G$ coincide on $gS$ for some $g \in G$, then $\tau(x)(g) = \tau(y)(g)$. Consequently, if the configurations $x$ and $y$ coincide on $\Omega S = \{gs : g \in \Omega, s \in S\}$ for some subset $\Omega \subset G$, then $\tau(x)$ and $\tau(y)$ coincide on $\Omega$. Observe that $\Omega S$ is finite whenever $\Omega$ is finite. Using the notation introduced in (1.19), we deduce that

$$(\tau \times \tau)(W_{\Omega S}) \subset W_\Omega$$

for every finite subset $\Omega$ of $G$. As the sets $W_\Omega$, where $\Omega$ runs over all finite subsets of $G$, form a base of entourages for the prodiscrete uniform structure

on $A^G$, it follows that $\tau$ is uniformly continuous. This shows that (a) implies (b).

Conversely, suppose that $\tau$ is uniformly continuous and $G$-equivariant. Let us show that $\tau$ is a cellular automaton. Consider the subset $\Omega = \{1_G\} \subset G$. Since $\tau$ is uniformly continuous, there exists a finite subset $S \subset G$ such that $(\tau \times \tau)(W_S) \subset W_\Omega$. This means that $\tau(x)(1_G)$ only depends on the restriction of $x$ to $S$. Thus, there is a map $\mu\colon A^S \to A$ such that

$$\tau(x)(1_G) = \mu(x|_S)$$

for all $x \in A^G$. Using the $G$-equivariance of $\tau$, we get

$$\tau(x)(g) = [g^{-1}\tau(x)](1_G) = \tau(g^{-1}x)(1_G) = \mu((g^{-1}x)|_S)$$

for all $x \in A^G$ and $g \in G$. This shows that $\tau$ is a cellular automaton with memory set $S$ and local defining map $\mu$. Consequently, (b) implies (a). □

Every uniformly continuous map between uniform spaces is continuous with respect to the associated topologies, and the converse is true when the source space is compact (Theorem B.2.3). The topology defined by the prodiscrete uniform structure on $A^G$ is the prodiscrete topology (see Example (1) in Sect. B.3). In the case when $A$ is finite, the prodiscrete topology on $A^G$ is compact by Tychonoff theorem (Theorem A.5.2). Thus Theorem 1.9.1 reduces to the Curtis-Hedlund theorem (Theorem 1.8.1) in this case.

*Remark 1.9.2.* Suppose that $G$ is countable and $A$ is an arbitrary set. Then the prodiscrete uniform structure (and hence the prodiscrete topology) on $A^G$ is metrizable. To see this, choose an increasing sequence

$$\varnothing = E_0 \subset E_1 \subset \cdots \subset E_n \subset \cdots$$

of finite subsets of $G$ such that $\bigcup_{n\geq 0} E_n = G$. Then the sets $W_{E_n}$, $n \geq 0$, form a base of entourages for the prodiscrete uniform structure on $A^G$. Consider now the metric $d$ on $A^G$ defined by

$$d(x,y) = \begin{cases} 0 & \text{if } x = y, \\ 2^{-\max\{n\geq 0\colon\, x|_{E_n} = y|_{E_n}\}} & \text{if } x \neq y. \end{cases}$$

for all $x, y \in A^G$. Then we have

$$W_{E_n} = \{(x,y) \in A^G \times A^G : d(x,y) < 2^{-n+1}\}$$

for every $n \geq 0$. Consequently, $d$ defines the prodiscrete uniform structure on $A^G$.

Let $G$ be a group and let $A$ be a set. Let $H$ be a subgroup of $G$. Let us equip $A^{H\backslash G}$ with its prodiscrete uniform structure and $\mathrm{Fix}(H) \subset A^G$ with the

uniform structure induced by the prodiscrete uniform structure on $A^G$. Recall from Proposition 1.3.7 that there is a natural bijection $\rho^*\colon A^{H\backslash G} \to A^G$ defined by $\rho^*(y) = y \circ \rho$, where $\rho\colon G \to H\backslash G$ is the canonical surjection.

**Proposition 1.9.3.** *The map $\rho^*\colon A^{H\backslash G} \to \mathrm{Fix}(H)$ is a uniform isomorphism.*

*Proof.* For $g \in G$, let $\pi_g\colon A^G \to A$ and $\pi'_g\colon A^{H\backslash G} \to A$ denote the projection maps given by $x \mapsto x(g)$ and $y \mapsto y(\rho(g))$ respectively. Observe that $\pi_g \circ \rho^* = \pi'_g$ is uniformly continuous for all $g \in G$. This shows that $\rho^*$ is uniformly continuous. Similarly, the uniform continuity of $(\rho^*)^{-1}$ follows from the fact that $\pi'_g \circ (\rho^*)^{-1} = \pi_g|_{\mathrm{Fix}(H)}$ is uniformly continuous for each $g \in G$. Consequently, $\rho^*$ is a uniform isomorphism. □

## 1.10 Invertible Cellular Automata

Let $G$ be a group and let $A$ be a set. One says that a cellular automaton $\tau\colon A^G \to A^G$ is *invertible* (or *reversible*) if $\tau$ is bijective and the inverse map $\tau^{-1}\colon A^G \to A^G$ is also a cellular automaton. This is equivalent to the existence of a cellular automaton $\sigma\colon A^G \to A^G$ such that $\tau \circ \sigma = \sigma \circ \tau = \mathrm{Id}_{A^G}$. Thus, the set of invertible cellular automata over the group $G$ and the alphabet $A$ is exactly the group $\mathrm{ICA}(G; A)$ consisting of all invertible elements of the monoid $\mathrm{CA}(G; A)$.

**Theorem 1.10.1.** *Let $A$ be a set and let $G$ be a group. Let $\tau\colon A^G \to A^G$ be a map and equip $A^G$ with its prodiscrete uniform structure. Then the following conditions are equivalent:*

(a) *$\tau$ is an invertible cellular automaton;*
(b) *$\tau$ is a $G$-equivariant uniform automorphism of $A^G$.*

*Proof.* It is clear that the inverse map of a bijective $G$-equivariant map from $A^G$ onto itself is also $G$-equivariant. Therefore, the equivalence of conditions (a) and (b) follows from the characterization of cellular automata given in Theorem 1.9.1. □

Bijective cellular automata over finite alphabets are always invertible:

**Theorem 1.10.2.** *Let $G$ be a group and let $A$ be a finite set. Then every bijective cellular automaton $\tau\colon A^G \to A^G$ is invertible.*

*Proof.* Let $\tau\colon A^G \to A^G$ be a bijective cellular automaton. The map $\tau^{-1}$ is $G$-equivariant since $\tau$ is $G$-equivariant. On the other hand, $\tau^{-1}$ is continuous with respect to the prodiscrete topology by compactness of $A^G$. Consequently, $\tau^{-1}$ is a cellular automaton by Theorem 1.8.1. □

The following example shows that Theorem 1.10.2 becomes false if we omit the finiteness hypothesis on the alphabet set $A$.

*Example 1.10.3.* Let $\mathbb{K}$ be a field. Let us take as the alphabet set the ring $A = \mathbb{K}[[t]]$ of all formal power series in one indeterminate $t$ with coefficients in $\mathbb{K}$. Thus, an element of $A$ is just a sequence $a = (k_i)_{i\in\mathbb{N}}$ of elements of $\mathbb{K}$ written in the form

$$a = k_0 + k_1 t + k_2 t^2 + k_3 t^3 + \cdots = \sum_{i\in\mathbb{N}} k_i t^i,$$

and the addition and multiplication of two elements $a = \sum_{i\in\mathbb{N}} k_i t^i$ and $b = \sum_{i\in\mathbb{N}} k'_i t^i$ are respectively given by $a+b = \sum_{i\in\mathbb{N}} (k_i + k'_i) t^i$ and $ab = \sum_{i\in\mathbb{N}} k''_i t^i$ with $k''_i = \sum_{i_1+i_2=i} k_{i_1} k'_{i_2}$ for all $i \in \mathbb{N}$. We take $G = \mathbb{Z}$. Thus, a configuration $x \in A^G$ is a map $x\colon \mathbb{Z} \to \mathbb{K}[[t]]$. Consider the map $\tau\colon A^G \to A^G$ defined by

$$\tau(x)(n) = x(n) - tx(n+1)$$

for all $x \in A^G$ and $n \in \mathbb{Z}$. Clearly $\tau$ is a cellular automaton admitting $S = \{0, 1\}$ as a memory set (the local defining map associated with $S$ is the map $\mu\colon A^S \to A$ defined by $\mu(x_0, x_1) = x_0 - tx_1$ for all $x_0, x_1 \in A$).

Let us show that $\tau$ is bijective. Consider the map $\sigma\colon A^G \to A^G$ given by

$$\sigma(x)(n) = x(n) + tx(n+1) + t^2 x(n+2) + t^3 x(n+3) + \cdots$$

for all $x \in A^G$ and $n \in \mathbb{Z}$. Observe that $\sigma(x)(n) \in \mathbb{K}[[t]]$ is well defined by the preceding formula. In fact, if we develop $x(n) \in \mathbb{K}[[t]]$ in the form

$$x(n) = \sum_{i\in\mathbb{N}} x_{n,i} t^i \qquad (n \in \mathbb{Z}, x_{n,i} \in \mathbb{K}),$$

then

$$\sigma(x)(n) = \sum_{i\in\mathbb{N}} \left( \sum_{j=0}^{i} x_{n+j,i-j} \right) t^i.$$

One immediately checks that $\sigma \circ \tau = \tau \circ \sigma = \mathrm{Id}_{A^G}$. Therefore, $\tau$ is bijective with inverse map $\tau^{-1} = \sigma$.

Let us show that the map $\sigma\colon A^G \to A^G$ is not a cellular automaton.

Let $F$ be a finite subset of $\mathbb{Z}$ and choose an integer $M \geq 0$ such that $F \subset (-\infty, M]$. Consider the configuration $y$ defined by $y(n) = 0$ if $n \leq M$ and $y(n) = 1$ if $n \geq M+1$, and the configuration $z$ defined by $z(n) = 0$ for all $n \in \mathbb{Z}$. Then $y$ and $z$ coincide on $F$. However, the value at 0 of $\sigma(y)$ is

$$\sigma(y)(0) = t^{M+1} + t^{M+2} + t^{M+3} + \cdots$$

while the value of $\sigma(z)$ at 0 is $\sigma(z)(0) = 0$. It follows that there is no finite subset $F \subset \mathbb{Z}$ such that $\sigma(x)(0)$ only depends on the restriction of $x \in A^G$

to $F$. This shows that $\sigma$ is not a cellular automaton. Consequently, $\tau$ is a bijective cellular automaton which is not invertible.

In the next proposition we show that invertibility is preserved under the operations of induction and restriction.

**Proposition 1.10.4.** *Let $G$ be a group and let $A$ be a set. Let $H$ be a subgroup of $G$ and let $\tau \in \mathrm{CA}(G, H; A)$. Let $\tau_H \in \mathrm{CA}(H; A)$ denote the cellular automaton obtained by restriction of $\tau$ to $H$. Then the following conditions are equivalent:*

(a) *$\tau$ is invertible;*
(b) *$\tau_H$ is invertible.*

*Moreover, if $\tau$ is invertible, then $\tau^{-1} \in \mathrm{CA}(G, H; A)$ and one has*

$$(\tau^{-1})_H = (\tau_H)^{-1}. \tag{1.20}$$

*Proof.* First recall from (1.16) the factorizations

$$A^G = \prod_{c \in G/H} A^c \quad \text{and} \quad \tau = \prod_{c \in G/H} \tau_c, \tag{1.21}$$

where $\tau_c \colon A^c \to A^c$ satisfies $\tau_c(\widetilde{x}|_c) = (\tau(\widetilde{x}))|_c$ for all $\widetilde{x} \in A^G$.

Suppose that $\tau$ is invertible. Denote by $\sigma \in \mathrm{CA}(G; A)$ the inverse cellular automaton $\tau^{-1}$. It follows from (1.21) that the map $\tau_c \colon A^c \to A^c$ is bijective for each $c \in G/H$ and that

$$\sigma = \prod_{c \in G/H} (\tau_c)^{-1}, \tag{1.22}$$

where $(\tau_c)^{-1} \colon A^c \to A^c$ is the inverse map of $\tau_c$. Let us show that $\sigma \in \mathrm{CA}(G, H; A)$. Let $S \subset G$ be a memory set for $\sigma$. Let $\widetilde{x} \in A^G$. It follows from (1.22) that $(\sigma(\widetilde{x}))|_H = (\tau_H)^{-1}(\widetilde{x}|_H)$. Thus, we have

$$\sigma(\widetilde{x})(1_G) = (\sigma(\widetilde{x}))|_H(1_G) = (\tau_H)^{-1}(\widetilde{x}|_H)(1_G).$$

This shows that $\sigma(\widetilde{x})(1_G)$ only depends on $\widetilde{x}|_H$. Arguing as in the proof of Lemma 1.5.1, we deduce that $S \cap H$ is a memory set for $\sigma$. Indeed, suppose that two configurations $\widetilde{x}, \widetilde{y} \in A^G$ coincide on $S \cap H$. Consider the configuration $\widetilde{z} \in A^G$ which coincide with $\widetilde{x}$ on $S$ and with $\widetilde{y}$ on $G \setminus S$. We have $\sigma(\widetilde{x})(1_G) = \sigma(\widetilde{z})(1_G)$ since $\widetilde{x}$ and $\widetilde{z}$ coincide on $S$. On the other hand, we have $\sigma(\widetilde{y})(1_G) = \sigma(\widetilde{z})(1_G)$ since $\widetilde{y}$ and $\widetilde{z}$ coincide on $H$. This implies $\sigma(\widetilde{x})(1_G) = \sigma(\widetilde{y})(1_G)$. Thus, there is a map $\mu \colon A^{S \cap H} \to A$ such that

$$\sigma(\widetilde{x})(1_G) = \mu(\widetilde{x}|_{S \cap H})$$

for all $\widetilde{x} \in A^G$. By applying Proposition 1.4.6, it follows that $S \cap H$ is a memory set for $\sigma$. Since $S \cap H \subset H$, this shows that $\tau^{-1} = \sigma \in \mathrm{CA}(G, H; A)$.

Moreover, it follows from (1.22) that

$$(\tau^{-1})_H = \sigma_H = (\tau_H)^{-1}$$

which gives us (1.20).

The equivalence (a) $\Leftrightarrow$ (b) is then an immediate consequence of Proposition 1.7.2 which tells us that the restriction map $\mathrm{CA}(G, H; A) \to \mathrm{CA}(H; A)$ is a monoid isomorphism. □

## Notes

Cellular automata were introduced by J. von Neumann (see [vNeu2]) who used them to describe theoretical models of self-reproducing machines. He first attempted to get such models by means of partial differential equations in $\mathbb{R}^3$. Later he changed the perspective and tried to use ideas and methods coming from robotics and electrical engineering. Eventually, in 1952, following a suggestion of S. Ulam, his former colleague at the Los Alamos Laboratories, he constructed a cellular automaton over the group $\mathbb{Z}^2$ with an alphabet consisting of 29 states. He then outlined the construction of a pattern, containing approximatively 200,000 cells, which would reproduce itself. The details were later filled in by A.W. Burks in the 1960s [Bur].

The branch of mathematics which is concerned with the study of the dynamical properties of the shift action is known as *symbolic dynamics*. Many authors trace the birth of symbolic dynamics back to a paper published in 1898 by J. Hadamard [Had] in which words on two letters were used to code geodesics on certain surfaces with negative curvature. However, as it was pointed out by E.M. Coven and Z.W. Nitecki [CovN], Hadamard's symbolic description of geodesics is purely static and involves only finite words. According to the authors of [CovN], the beginning of symbolic dynamics should be placed in a paper by G.A. Hedlund [Hed-1] published in 1944. Symbolic dynamics has important applications in dynamical systems, especially in the study of hyperbolic dynamical systems for which symbolic codings may be obtained from Markov partitions. One of the first examples of such an application was the use of the properties of the Thue-Morse sequence (see Exercise 3.41) by M. Morse [Mors] in 1921 to prove the existence of non-periodic recurrent geodesics on surfaces with negative curvature. A detailed exposition of symbolic dynamics over $\mathbb{Z}$ may be found for example in the books by B. Kitchens [Kit], by P. Kůrka [Kur], and by D. Lind and B. Marcus [LiM].

In the mid-1950s, Hedlund studied the so-called shift-commuting block maps which turn out to be exactly cellular automata over the group $\mathbb{Z}$. The Curtis-Hedlund theorem (Theorem 1.8.1), also called Curtis-Hedlund-Lyndon's theorem or Hedlund's theorem, is named after Hedlund [Hed-3]

who proved it in 1969. Its generalization to infinite alphabets, as stated as in Theorem 1.9.1, was proved by the authors in [CeC7].

Cellular automata were intensely studied from the 1960s, both by pure and applied mathematicians, under different names such as *tessellation automata*, *parallel maps*, *cellular spaces*, *iterative automata*, *homogeneous structures*, *universal spaces*, and *sliding block codes* (cf. [LiM, Section 1.1]). In most cases, these researches focused on cellular automata with finite alphabet over the groups $\mathbb{Z}$ or $\mathbb{Z}^2$ (see the surveys [BanMS], [BKM], [Kar3], [Wolfr3]).

The Game of Life was invented by the British mathematician J.H. Conway. This cellular automaton was described for the first time by M. Gardner [Gar-1] in the October 1970 issue of the *Scientific American*. From a theoretical computer science point of view, it is important because it has the power of a universal Turing machine, that is, anything that can be computed algorithmically can be computed by using the Game of Life.

In the 1980s, S. Wolfram [Wolfr1], [Wolfr2] started a systematic study and empirical classification of *elementary cellular automata*, that is, of cellular automata over $\mathbb{Z}$ with alphabet $A = \{0, 1\}$ and memory set $S = \{-1, 0, 1\}$. There are $2^{(2^3)} = 256$ such elementary cellular automata. Wolfram introduced a naming scheme for them which is nowadays widely used. Each elementary cellular automaton $\tau\colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ is uniquely determined by the eight bit sequence

$$\mu(111)\mu(110)\mu(101)\mu(100)\mu(011)\mu(010)\mu(001)\mu(000) \in A^8,$$

where $\mu\colon A^S \to A$ is the associated local defining map. This bit sequence is the binary expansion of an integer in the interval $[0, 255]$, called the *Wolfram number* of $\tau$. For example, the majority action $\tau\colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ associated with $S = \{-1, 0, 1\}$ (cf. Example 1.4.3(c)) is an elementary cellular automaton. Its local defining map $\mu$ gives $\mu(111)\mu(110)\mu(101)\mu(100)\mu(011)\mu(010)\mu(001)\mu(000) = 11101000$ (cf. Fig. 1.7). It follows that the Wolfram number of $\tau$ is 232. Let us also mention that the elementary cellular automaton with Wolfram number 110 was recently proved computationally universal by M. Cook.

Wolfram introduced an empirical classification of elementary cellular automata into four classes according to the behavior of random initial configurations under iterations. These are known as *Wolfram classes* and are defined as follows:

(W1) Almost all initial configurations lead to the same uniform fixed-point configuration,

(W2) Almost all initial configurations lead to a periodic configuration,

(W3) Almost all initial configurations lead to chaos,

(W4) Localized structures with complex interactions emerge.

The survey paper [BKM] contains other dynamical classifications of cellular automata over $\mathbb{Z}$ with finite alphabet due to R. Gilman and to M. Hurley and P. Kůrka.

Invertible cellular automata are used to model time-reversible processes occurring in physics and biology. A group $G$ is called *periodic* if every element $g \in G$ has finite order. In [CeC11] it is shown that if $G$ is a non-periodic group, then for every infinite set $A$ there exists a bijective cellular automaton $\tau\colon A^G \to A^G$ which is not invertible (cf. Theorem 1.10.2 and Example 1.10.3). It was shown by Amoroso and Patt in 1972 [Amo] that it is decidable whether a given cellular automaton with finite alphabet over $\mathbb{Z}$ is invertible. This means that there exists an algorithm which establishes, after a finite number of steps, whether the cellular automaton corresponding to a given local defining map is invertible or not. On the other hand, J. Kari [Kar1], [Kar2], [Kar3] proved that the similar problem for cellular automata with finite alphabet over $\mathbb{Z}^d$, $d \geq 2$, is undecidable. Its proof is based on R. Berger's undecidability result for the *tiling problem of Wang tiles*.

## Exercises

**1.1.** An action of a group $\Gamma$ on a topological space $X$ is said to be *topologically mixing* if for each pair of nonempty subsets $U$ and $V$ of $X$ there exists a finite subset $F \subset \Gamma$ such that $U \cap \gamma V \neq \varnothing$ for all $\gamma \in \Gamma \setminus F$. Show that if $G$ is a group and $A$ is a set then the $G$-shift on $A^G$ is topologically mixing for the prodiscrete topology on $A^G$.

**1.2.** Let $G$ be a group and let $A$ be a set. Let $x \in A^G$ and let $\Omega_1$ and $\Omega_2$ be two subsets of $G$. Show that $V(x, \Omega_1 \cup \Omega_2) = V(x, \Omega_1) \cap V(x, \Omega_2)$ and $W_{\Omega_1 \cup \Omega_2} = W_{\Omega_1} \cap W_{\Omega_2}$ (see (1.3) and (1.19) for the definition of $V(x, \Omega)$ and $W_\Omega$).

**1.3.** Let $G$ be a countable group and let $A$ be a set. Show that the metric $d$ on $A^G$ introduced in Remark 1.9.2 is complete.

**1.4.** Let $G$ be an uncountable group and let $A$ be a set having at least two elements. Prove that the prodiscrete topology on $A^G$ is not metrizable. Hint: Prove that this topology does not satisfy the first axiom of countability.

**1.5.** Let $G$ be a group. Let $A$ and $B$ be two sets. Let $\tau_A\colon A^G \to A^G$ and $\tau_B\colon B^G \to B^G$ be cellular automata. For $x \in (A \times B)^G$, let $x_A \in A^G$ and $x_B \in B^G$ be the configurations defined by $x(g) = (x_A(g), x_B(g))$ for all $g \in G$. Show that the map $\tau\colon (A \times B)^G \to (A \times B)^G$ given by $\tau(x)(g) = (\tau_A(x_A)(g), \tau_B(x_B)(g))$ for all $g \in G$ is a cellular automaton.

**1.6.** Let $G$ be a group and let $S$ be a finite subset of $G$ of cardinality $k$. Let $A$ be a finite set of cardinality $n$. Show that there are exactly $n^{n^k}$ cellular automata $\tau\colon A^G \to A^G$ admitting $S$ as a memory set.

**1.7.** Let $G = \mathbb{Z}^2$ and $A = \{0, 1\}$. Let $\tau\colon A^G \to A^G$ denote the cellular automaton associated with the Game of Life. Let $y \in A^G$ be the constant configuration defined by $y(g) = 1$ for all $g \in G$ (all cells are alive). Find a configuration $x \in A^G$ such that $y = \tau(x)$.

**1.8.** Let $A$ be a set and suppose that $G$ is a trivial group. Show that the monoid $\mathrm{CA}(G; A)$ is canonically isomorphic to the monoid consisting of all maps from $A$ to $A$ (with composition of maps as the monoid operation). Also show that the group $\mathrm{ICA}(G; A)$ is canonically isomorphic to the symmetric group of $A$.

**1.9.** Let $G$ be a group and let $A$ be a set. Let $\tau\colon A^G \to A^G$ be a cellular automaton. Show that $\tau$ admits a memory set which is reduced to a single element if and only if there exist an element $s \in G$ and a map $f\colon A \to A$ such that one has $\tau(x)(g) = f(x(gs))$ for all $x \in A^G$ and $g \in G$.

**1.10.** Prove that there are exactly 218 cellular automata $\tau\colon \{0, 1\}^{\mathbb{Z}} \to \{0, 1\}^{\mathbb{Z}}$ whose minimal memory set is $\{-1, 0, 1\}$.

**1.11.** Let $\tau \in \mathrm{CA}(\mathbb{Z}^2; \{0, 1\})$ denote the cellular automaton associated with the Game of Life. Show that the minimal memory set of $\tau$ is the set $\{-1, 0, 1\}^2$.

**1.12.** Let $G$ be a group and let $A$ be a set. Let $\sigma, \tau \in \mathrm{CA}(G; A)$. Let $S_0$ (resp. $T_0$, resp. $C_0$) denote the minimal memory set of $\sigma$ (resp. $\tau$, resp. $\sigma \circ \tau$). Prove that $C_0 \subset S_0 T_0$. Give an example showing that this inclusion may be strict.

**1.13.** Let $G$ be a group and let $A$ be a set. Let $H$ be a subgroup of $G$ and let $\tau \in \mathrm{CA}(G, H; A)$. Show that $\tau$ and $\tau_H$ have the same minimal memory set.

**1.14.** Prove Proposition 1.4.9 by applying Theorem 1.9.1. Hint: Observe that the composite of two uniformly continuous maps is a uniformly continuous map.

**1.15.** Let $G$ be a group and let $A$ be a set. Let $F$ be a nonempty finite subset of $G$ and set $B = A^F$. The sets $A^G$ and $B^G$ are equipped with their prodiscrete uniform structures and with the $G$-shift action. Show that the map $\Phi_F\colon A^G \to B^G$ defined by $\Phi_F(x)(g) = (g^{-1}x)|_F$ for all $x \in A^G$ and $g \in G$ is a $G$-equivariant uniform embedding.

**1.16.** Let $G$ be a group, $H \subset G$ a subgroup of $G$, and let $A$ be a set. Let $H\backslash G = \{Hg : g \in G\}$ be the set of all right cosets of $H$ in $G$ and set $B = A^{H\backslash G}$. The set $A^G$ (resp. $B^H$) is equipped with its prodiscrete uniform structure and with the $G$-shift (resp. $H$-shift) action. Let $T \subset G$ be a complete set of representatives for the right cosets of $H$ in $G$ so that $G = \coprod_{t \in T} Ht$. Show that the map $\Psi = \Psi(H, T)\colon A^G \to B^H$ defined by $\Psi(x)(h)(Ht) = x(ht)$ for all $x \in A^G$, $h \in H$ and $t \in T$ is an $H$-equivariant uniform isomorphism.

**1.17.** Let $G$ be a group and let $A$ be a set. For each $s \in G$, let $\tau_s \colon A^G \to A^G$ be the cellular automaton defined by $\tau_s(x)(g) = x(gs)$ for all $x \in A^G$, $g \in G$ (cf. Example 1.4.3(e)).

(a) Show that $\tau_s \in \mathrm{ICA}(G; A)$ for every $s \in G$.

(b) Prove that the map $\Phi \colon G \to \mathrm{ICA}(G; A)$ defined by $\phi(s) = \tau_s$ for all $s \in G$ is a group homomorphism.

(c) Prove that if $A$ has at least two elements, then $\Phi$ is injective but not surjective.

**1.18.** Let $G$ be a group and let $A$ be a set. Prove that the set consisting of all invertible cellular automata $\tau \colon A^G \to A^G$ admitting a memory set which is reduced to a single element is a subgroup of $\mathrm{ICA}(G; A)$ isomorphic to the direct product $G \times \mathrm{Sym}(A)$.

**1.19.** (cf. [Amo]) Let $G = \mathbb{Z}$ and $A = \{0, 1\}$. Fix an integer $n \geq 3$ and let $S = \{-1, 0, 1, \ldots, n\}$. Consider the element $\alpha \in A^S$ (resp. $\beta \in A^S$) defined by $\alpha(-1) = \alpha(n) = 0$ and $\alpha(k) = 1$ for $0 \leq k \leq n-1$ (resp. $\beta(-1) = \beta(0) = \beta(n) = 0$ and $\beta(k) = 1$ for $1 \leq k \leq n-1$) and the map $\mu \colon A^S \to A$ defined by $\mu(\alpha) = 0$, $\mu(\beta) = 1$ and $\mu(y) = y(0)$ for $y \in A^S \backslash \{\alpha, \beta\}$. Let $\tau \colon A^G \to A^G$ be the cellular automaton with memory set $S$ and local defining map $\mu$.

(a) Show that $S$ is the minimal memory set of $\tau$.

(b) Show that $\tau$ is an invertible cellular automaton and that $\tau^{-1} = \tau$.

**1.20.** Show that the inverse map of the bijective cellular automaton $\tau \colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ studied in Example 1.10.3 is discontinuous, with respect to the prodiscrete topology on $A^{\mathbb{Z}}$, at every configuration $x \in A^{\mathbb{Z}}$.

**1.21.** Let $G$ be a group and let $A$ be a set. A *subshift* of the configuration space $A^G$ is a subset $X \subset A^G$ which is $G$-invariant (i.e., such that $gx \in X$ for all $x \in X$ and $g \in G$) and closed in $A^G$ with respect to the prodiscrete topology.

(a) Show that $\varnothing$ and $A^G$ are subshifts of $A^G$.

(b) Show that if $x \in A^G$ then its orbit closure $\overline{Gx} \subset A^G$ is a subshift.

(c) Show that if $(X_i)_{i \in I}$ is a family of subshifts of $A^G$ then $\bigcap_{i \in I} X_i$ is a subshift of $A^G$.

(d) Show that if $(X_i)_{i \in I}$ is a finite family of subshifts of $A^G$ then $\bigcup_{i \in I} X_i$ is a subshift of $A^G$.

(e) Suppose that $A$ is finite. Show that if $X \subset A^G$ is a subshift and $\tau \colon A^G \to A^G$ is a cellular automaton then $\tau(X)$ is a subshift of $A^G$. Note: This last statement becomes false when $A$ is infinite (see Example 3.3.3).

**1.22.** Let $G$ be a group and let $A$ and $B$ be two sets. Let $f \colon A \to B$ be a map and consider the map $f_* \colon A^G \to B^G$ defined by $f_*(x) = f \circ x$ for all $x \in A^G$.

(a) Show that if $A$ is finite and $X$ is a subshift of $A^G$ then $f_*(X)$ is a subshift of $B^G$. Hint: Use the compactness of the configuration space $A^G$.

(b) Let $G = \mathbb{Z}$, $A = \mathbb{Z}$, $B = \{0, 1\}$ and let $f\colon A \to B$ be defined by $f(n) = 0$ if $n < 0$ and $f(n) = 1$ otherwise. Let $X = \{x_n : n \in \mathbb{Z}\} \subset A^{\mathbb{Z}}$, where $x_n(m) = n + m$ for all $n, m \in \mathbb{Z}$. In other words, $X$ is the $\mathbb{Z}$-orbit $\mathbb{Z}x_0$ of the configurations $x_0$. Show that $X$ is a subshift of $A^{\mathbb{Z}}$ but $f^*(X)$ is not a subshift of $B^{\mathbb{Z}}$.

**1.23.** Let $G$ be a group and let $A$ be a set. Given a set of patterns $\mathcal{P} \subset \bigcup A^{\Omega}$, where the union runs over all finite subsets $\Omega$ of $G$, we set

$$X_{\mathcal{P}} = \{x \in A^G : (gx)|_{\Omega} \notin \mathcal{P} \text{ for all } g \in G \text{ and all finite subsets } \Omega \subset G\},$$

where $(gx)|_{\Omega}$ denotes the restriction of the configuration $gx$ to $\Omega$.

(a) Let $\mathcal{P} \subset \bigcup A^{\Omega}$ be a set of patterns. Show that $X_{\mathcal{P}}$ is a subshift of $A^G$.

(b) Conversely, show that if $X \subset A^G$ is a subshift, then there exists a subset $\mathcal{P} \subset \bigcup A^{\Omega}$ such that $X = X_{\mathcal{P}}$. Such a set $\mathcal{P}$ is called a *defining set* of *forbidden patterns* for $X$.

**1.24.** Let $G$ be a group and let $A$ be a set. Given a finite subset $\Omega \subset G$ and a finite subset $\mathcal{A} \subset A^{\Omega}$ we set

$$X(\Omega, \mathcal{A}) = \{x \in A^G : (gx)|_{\Omega} \in \mathcal{A} \text{ for all } g \in G\}.$$

(a) Let $\Omega \subset G$ and $\mathcal{A} \subset A^{\Omega}$ be finite subsets. Show that $X(\Omega, \mathcal{A})$ is a subshift of $A^G$. A subshift $X \subset A^G$ is said to be of *finite type* if there exists a finite subset $\Omega \subset G$ and a finite subset $\mathcal{A} \subset A^{\Omega}$ such that $X = X(\Omega, \mathcal{A})$. Such a set $\mathcal{A}$ is then called a *defining set* of *admissible patterns* for $X$ and the subset $\Omega$ is called a *memory set* for $X$.

(b) Suppose that $A$ is finite. Show that a subshift $X \subset A^G$ is of finite type if and only if it admits a finite defining set of forbidden patterns.

**1.25.** Let $G$ be a countable group and let $A$ be a finite set. Show that there are at most countably many distinct subshifts $X \subset A^G$ of finite type.

**1.26.** Let $G$ be a group and let $A$ be a finite set.

(a) Let $\tau_1, \tau_2\colon A^G \to A^G$ be two cellular automata. Show that the set $\{x \in A^G : \tau_1(x) = \tau_2(x)\} \subset A^G$ is a subshift of finite type.

(b) Deduce from (a) that if $\tau\colon A^G \to A^G$ is a cellular automaton then the set $\mathrm{Fix}(\tau) = \{x \in A^G : \tau(x) = x\} \subset A^G$ is a subshift of finite type.

(c) Conversely, show that if $X \subset A^G$ is a subshift of finite type then there exists a cellular automaton $\tau\colon A^G \to A^G$ such that $X = \mathrm{Fix}(\tau)$.

**1.27.** Suppose that a group $\Gamma$ acts continuously on a topological space $Z$. One says that the action of $\Gamma$ on $Z$ is *topologically transitive* if for any pair of nonempty open subsets $U$ and $V$ of $Z$ there exists an element $\gamma \in \Gamma$ such that $U \cap \gamma V \neq \varnothing$.

Let $G$ be a group and let $A$ be a set. A subshift $X \subset A^G$ is said to be *irreducible* if for any finite subset $\Omega$ of $G$ and any two elements $x_1, x_2 \in X$,

there exist a configuration $x \in X$ and an element $g \in G$ such that $x|_\Omega = x_1|_\Omega$ and $(gx)|_\Omega = x_2|_\Omega$.

Suppose that $X \subset A^G$ is a subshift. Show that the action of $G$ on $X$ induced by the $G$-shift is topologically transitive if and only if $X$ is irreducible.

**1.28.** Let $G$ be a group and let $A$ be a set. Let $B \subset A$ and consider the subsets $X, Y \subset A^G$ defined by $X = \{\overline{b} : b \in B\} \subset A^G$, where $\overline{b}$ denotes the constant configuration given by $\overline{b}(g) = b$ for all $g \in G$, and $Y = \{y \in A^G : y(g) \in B \text{ for all } g \in G\}$.

(a) Show that $X$ and $Y$ are subshifts of $A^G$.

(b) Show that if $G$ is infinite then $Y$ is irreducible.

(c) Suppose that $B$ has at least two distinct elements. Show that $X$ is not irreducible.

**1.29.** Let $G$ be a group acting continuously on a nonempty complete metric space $X$ whose topology satisfies the second axiom of countability (i.e., admitting a countable base of open subsets). Show that the following conditions are equivalent:

(i) the action of $G$ on $X$ is topologically transitive;

(ii) there is a point $x \in X$ whose $G$-orbit is dense in $X$;

(iii) there is a dense subset $D \subset X$ such that the $G$-orbit of each point $x \in D$ is dense in $X$.

Hint: The implications (iii) $\Rightarrow$ (ii) and (ii) $\Rightarrow$ (i) are straightforward. To prove (i) $\Rightarrow$ (iii), consider a sequence $(U_n)_{n\in\mathbb{N}}$ of nonempty open subsets of $X$ which form a base of the topology and denote by $\Omega_n$ the set of points $x \in X$ whose $G$-orbit meets $U_n$. Then observe that each $\Omega_n$ is and open dense subset of $X$ if (i) is satisfied and apply Baire's theorem (Theorem I.1.1, also cf. Remark I.1.2(ii)).

**1.30.** Let $G$ be a countable group and let $A$ be a countable (e.g. finite) set. Let $X \subset A^G$ be a nonempty subshift. Show that the following conditions are equivalent:

(i) the subshift $X$ is irreducible;

(ii) there is a configuration $x \in X$ whose $G$-orbit is dense in $X$;

(iii) there is a dense subset $D \subset X$ such that the $G$-orbit of each configuration $x \in D$ is dense in $X$.

Hint: Use the results of Exercises 1.3 and 1.29.

**1.31.** Let $G$ be a group and let $A$ be a set. One says that a subshift $X \subset A^G$ is *topologically mixing* if the action of $G$ on $X$ induced by the $G$-shift is topologically mixing (cf. Exercise 1.1).

(a) Let $X \subset A^G$ be a subshift. Show that $X$ is topologically mixing if and only if for any finite subset $\Omega$ of $G$ and any two configurations $x_1, x_2 \in X$, there exists a finite subset $F \subset G$ such that, for all $g \in G \setminus F$, there exists a configuration $x \in X$ satisfying $x|_\Omega = x_1|_\Omega$ and $(gx)|_\Omega = x_2|_\Omega$.

(b) Show that if $G$ is infinite then every topologically mixing subshift $X \subset A^G$ is irreducible.

**1.32.** Let $G$ be a group and let $A$ be a set. Let $\Delta \subset G$ be a finite subset. A subshift $X \subset A^G$ is said to be *$\Delta$-irreducible* if it satisfies the following condition: if $\Omega_1$ and $\Omega_2$ are two finite subsets of $G$ such that $\Omega_1$ and $\Omega_2\delta$ are disjoint for all $\delta \in \Delta$, then, given any two configurations $x_1, x_2 \in X$, there exists a configuration $x \in X$ which satisfies $x|_{\Omega_1} = x_1|_{\Omega_1}$ and $x|_{\Omega_2} = x_2|_{\Omega_2}$. A subshift $X \subset A^G$ is said to be *strongly irreducible* if there exists a finite subset $\Delta \subset G$ such that $X$ is $\Delta$-irreducible.

(a) Show that the subshift $Y \subset A^G$ described in Exercise 1.28 is $\{1_G\}$-irreducible and therefore strongly irreducible.

(b) Show that every strongly irreducible subshift is topologically mixing.

**1.33.** Let $G$ be a group and let $A$ be a set. Let $H$ be a subgroup of $G$. For $x \in A^G$ and $g \in G$, denote by $x_g^H \in A^H$ the configuration defined by $x_g^H = (gx)|_H$.

(a) Check that $hx_g^H = x_{hg}^H$ for all $x \in A^G$, $h \in H$ and $g \in G$.

(b) Let $X \subset A^H$ be a subshift. Show that the set $X^{(G)}$ defined by $X^{(G)} = \{x \in A^G : x_g^H \in X \text{ for all } g \in G\}$ is a subshift of $A^G$.

(c) Show that if $H \neq G$ then the subshift $X^{(G)} \subset A^G$ is irreducible for any subshift $X \subset A^H$.

(d) Let $X \subset A^H$ be a subshift. Show that $X^{(G)} \subset A^G$ is of finite type (resp. topologically mixing, resp. strongly irreducible) if and only if $X$ is of finite type (resp. topologically mixing, resp. strongly irreducible).

(e) Suppose that $\sigma\colon A^H \to A^H$ is a cellular automaton and let $\sigma^G\colon A^G \to A^G$ be the induced cellular automaton (cf. Sect. 1.7). Check that one has $\sigma^G(x)_g^H = \sigma(x_g^H)$ for all $x \in A^G$ and $g \in G$.

(f) Show that if $X \subset A^H$ is a subshift such that $\sigma(X) \subset X$ then one has $\sigma^G(X^{(G)}) \subset X^{(G)}$.

**1.34.** Let $G$ be a group and let $A$ be a set. Let $F$ be a nonempty finite subset of $G$ and consider the map $\Phi_F\colon A^G \to B^G$ defined in Exercise 1.15, where $B = A^F$. Let $X \subset A^G$ be a subshift and set $X^{[F]} = \Phi_F(X)$.

(a) Show that $X^{[F]}$ is a subshift of $B^G$.

(b) Show that $X$ is irreducible (resp. topologically mixing, resp. strongly irreducible) if and only if $X^{[F]}$ is irreducible (resp. topologically mixing, resp. strongly irreducible).

(c) Show that if $X$ is of finite type then $X^{[F]}$ is of finite type.

**1.35.** Let $G$ be a group, $H \subset G$ a subgroup of $G$, and let $A$ be a set. Let also $T \subset G$ be a complete set of representatives for the right cosets of $H$ in $G$ and consider the map $\Psi\colon A^G \to B^H$ defined in Exercise 1.16, where $B = A^{H\backslash G}$. Let $X \subset A^G$ be a subshift and set $X^{(H,T)} = \Psi(X)$.

(a) Show that $X^{(H,T)}$ is a subshift of $B^H$.

(b) Show that if $X$ is of finite type then $X^{(H,T)}$ is of finite type.

**1.36.** Let $A$ be a set. Let $A^*$ denote the monoid consisting of all words in the alphabet $A$ (cf. Sect. D.1). Recall that any word $w \in A^*$ can be uniquely

written in the form $w = a_1a_2\cdots a_n$, where $n \geq 0$ and $a_i \in A$ for $1 \leq i \leq n$. The integer $n$ is called the *length* of the word $w$ and it is denoted by $\ell(w)$. In the sequel, we shall identify the word $w = a_1a_2\cdots a_n$ with the pattern $p\colon \{1, 2, \ldots, n\} \to A$ defined by $p(i) = a_i$ for $1 \leq i \leq n$. Given a subshift $X \subset A^{\mathbb{Z}}$ and an integer $n \geq 0$, we denote by $L_n(X) \subset A^*$ the set consisting of all words $w \in A^*$ for which there exists an element $x \in X$ such that $w = x(1)x(2)\cdots x(n)$. The set $L(X) = \cup_{n\in\mathbb{N}} L_n(X)$ is called the *language* of $X$. The elements $w \in L(X)$ are called the *admissible words* of $X$ (or, simply, the $X$-admissible words). The elements $w \in A^* \setminus L(X)$ are called the *forbidden words* of $X$.

(a) Let $X$ and $Y$ be two subshifts of $A^{\mathbb{Z}}$. Show that one has $X \subset Y$ (resp. $X = Y$) if and only if $L(X) \subset L(Y)$ (resp. $L(X) = L(Y)$).

(b) One says that a word $u \in A^*$ is a *subword* of a word $w \in A^*$ if there exist $v_1, v_2 \in A^*$ such that $w = v_1uv_2$. Let $X \subset A^{\mathbb{Z}}$ be a subshift and let $L = L(X)$. Show that $L$ satisfies the following conditions:

(i) if $w \in L$, then $u \in L$ for every subword $u$ of $w$;

(ii) if $w \in L$, then there exist $a, a' \in A$ such that $awa' \in L$.

(c) Conversely, show that if a subset $L \subset A^*$ satisfies conditions (i) and (ii) in (b), then there exists a unique subshift $X \subset A^{\mathbb{Z}}$ such that $L = L(X)$.

**1.37.** Let $A$ be a set and let $X \subset A^{\mathbb{Z}}$ be a subshift.

(a) Show that $X$ is of finite type if and only if the following holds: there exists an integer $n_0 \geq 0$ such that if the words $u, v, w \in A^*$ satisfy $\ell(v) \geq n_0$ and $uv, vw \in L(X)$, then one has $uvw \in L(X)$.

(b) Show that $X$ is irreducible if and only if for every pair of words $u$ and $v$ in $L(X)$, there exists a word $w \in A^*$ such that $uwv \in L(X)$.

(c) Show that $X$ is topologically mixing if and only if the following holds: for every pair of words $u$ and $v$ in $L(X)$, there exists an integer $n_0 \geq 0$ such that for every integer $n \geq n_0$ there exists a word $w \in A^*$ of length $\ell(w) = n$ satisfying $uwv \in L(X)$.

(d) Show that $X$ is strongly irreducible if and only if the following holds: there exists an integer $n_0 \geq 0$ such that, for every pair of words $u$ and $v$ in $L(X)$, and for every integer $n \geq n_0$, there exists a word $w \in A^*$ of length $\ell(w) = n$ such that $uwv \in L(X)$.

**1.38.** Let $A = \{0, 1\}$ and let $X \subset A^{\mathbb{Z}}$ be the set of all $x \in A^{\mathbb{Z}}$ such that the following holds: if $x(n) = 1, x(n+1) = x(n+2) = \cdots = x(n+k) = 0, x(n+k+1) = 1$, for some $n \in \mathbb{Z}$ and $k \in \mathbb{N}$, then $k$ is even.

(a) Show that $X$ is a subshift (it is called the *even subshift*).

(b) Show that $X$ is not of finite type.

(c) Show that $X$ is strongly irreducible (and therefore topologically mixing and irreducible).

**1.39.** Let $A = \{0, 1\}$ and consider the subshift of finite type $X \subset A^{\mathbb{Z}}$ defined by $X = X_{\{11\}} = \{x \in A^{\mathbb{Z}} : (x(n), x(n+1)) \neq (1, 1) \text{ for all } n \in \mathbb{Z}\}$. Show that $X$ is strongly irreducible (and therefore topologically mixing and irreducible). The subshift $X$ is called the *golden mean* subshift.

**1.40.** Let $A = \{0, 1\}$ and let $X \subset A^{\mathbb{Z}}$ be the set consisting of the two configurations $x, y \in A^{\mathbb{Z}}$ defined by

$$x(n) = \begin{cases} 0 & \text{if } n \text{ is even} \\ 1 & \text{otherwise} \end{cases} \quad \text{and} \quad y(n) = \begin{cases} 1 & \text{if } n \text{ is even} \\ 0 & \text{otherwise,} \end{cases}$$

for all $n \in \mathbb{Z}$. Show that $X$ is an irreducible subshift of finite type which is not topologically mixing (and therefore not strongly irreducible either).

**1.41.** Let $A$ be a set. Let $X \subset A^{\mathbb{Z}}$ be a subshift of finite type. Show that $X$ is topologically mixing if and only if $X$ is strongly irreducible.

**1.42.** Let $A = \{0, 1\}$ and let $X \subset A^{\mathbb{Z}}$ be the subshift with defining set of forbidden words $\{01^k0^h1 : 1 \leq h \leq k, k = 1, 2, \ldots\}$. Show that $X$ is topologically mixing (and therefore irreducible) but not strongly irreducible.

**1.43.** *Cellular automata between subshifts.* Let $G$ be a group and let $A$ be a set. The set $A^G$ is equipped with its prodiscrete uniform structure and with the $G$-shift action. Let $X, Y \subset A^G$ be two subshifts and $\tau\colon X \to Y$ a map. Then the following are equivalent:

(i) there exists a cellular automaton $\overline{\tau}\colon A^G \to A^G$ such that $\overline{\tau}(x) = \tau(x)$ for all $x \in X$;

(ii) $\tau$ is $G$-equivariant and uniformly continuous.

One says that $\tau\colon X \to Y$ is a cellular automaton if the two equivalent conditions above are satisfied.

**1.44.** Let $G$ be a group and let $A$ be a finite set. Let $\tau\colon X \to Y$ be a map between subshifts $X, Y \subset A^G$. Show that the following conditions are equivalent:

(i) $\tau$ is a cellular automaton,

(ii) $\tau$ is $G$-equivariant and continuous (with respect to the topologies induced on $X$ and $Y$ by the prodiscrete topology on $A^G$).

**1.45.** Let $G$ be a group and let $A$ be a finite set. Let $\tau\colon A^G \to A^G$ be a cellular automaton and let $X \subset A^G$ be an irreducible (resp. topologically mixing, resp. strongly irreducible) subshift. Show that $\tau(X)$ is an irreducible (resp. topologically mixing, resp. strongly irreducible) subshift of $A^G$.

# Chapter 2
# Residually Finite Groups

This chapter is devoted to the study of residually finite groups, which form a class of groups of special importance in several branches of mathematics. As their name suggests it, residually finite groups generalize finite groups. They are defined as being the groups whose elements can be distinguished after taking finite quotients (see Sect. 2.1). There are many other equivalent definitions. For example, a group is residually finite if and only if it can be embedded into the direct product of a family of finite groups (Corollary 2.2.6). The class of residually finite groups is closed under taking subgroups and taking projective limits. It contains in particular all finite groups, all finitely generated abelian groups, and all free groups (Theorem 2.3.1). Every finitely generated residually finite group is Hopfian (Theorem 2.4.3) and the automorphism group of a finitely generated residually finite group is itself residually finite (Theorem 2.5.1). Examples of finitely generated groups which are not residually finite are presented in Sect. 2.6. The following dynamical characterization of residually finite groups is given in Sect. 2.7: a group is residually finite if and only if there is a Hausdorff topological space on which the group acts continuously and faithfully with a dense subset of points with finite orbit.

## 2.1 Definition and First Examples

**Definition 2.1.1.** A group $G$ is called *residually finite* if for each element $g \in G$ with $g \neq 1_G$, there exist a finite group $F$ and a homomorphism $\phi\colon G \to F$ such that $\phi(g) \neq 1_F$.

**Proposition 2.1.2.** *Let $G$ be a group. Then the following conditions are equivalent:*

(a) *$G$ is residually finite;*

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_2, 

(b) *for all $g, h \in G$ with $g \neq h$, there exist a finite group $F$ and a homomorphism $\phi\colon G \to F$ such that $\phi(g) \neq \phi(h)$.*

*Proof.* The fact that (b) implies (a) is obvious, since (b) gives (a) by taking $h = 1_G$. Conversely, suppose that $G$ is residually finite. Let $g, h \in G$ with $g \neq h$. As $gh^{-1} \neq 1_G$, there exist a finite group $F$ and a homomorphism $\phi\colon G \to F$ such that $\phi(gh^{-1}) \neq 1_F$. Since $\phi(gh^{-1}) = \phi(g)(\phi(h))^{-1}$, it follows that $\phi(g) \neq \phi(h)$. Therefore, (a) implies (b). □

**Proposition 2.1.3.** *Every finite group is residually finite.*

*Proof.* Let $G$ be a finite group and consider $g \in G$ such that $g \neq 1_G$. If $\phi = \mathrm{Id}_G\colon G \to G$ is the identity map, we have $\phi(g) = g \neq 1_G$. □

**Proposition 2.1.4.** *The additive group $\mathbb{Z}$ is residually finite.*

*Proof.* Consider $k \in \mathbb{Z}$ such that $k \neq 0$. Choose an integer $m$ such that $|k| < m$. Then the canonical homomorphism $\phi\colon \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ (reduction modulo $m$) satisfies $\phi(k) \neq 0$. □

A similar argument gives us the following:

**Proposition 2.1.5.** *The group $\mathrm{GL}_n(\mathbb{Z})$ is residually finite for every $n \geq 1$.*

*Proof.* Let $A = (a_{ij}) \in \mathrm{GL}_n(\mathbb{Z})$ with $A \neq I_n = 1_{\mathrm{GL}_n(\mathbb{Z})}$. Choose an integer $m$ such that $|a_{ij}| < m$ for all $i, j$. Then the homomorphism $\phi\colon \mathrm{GL}_n(\mathbb{Z}) \to \mathrm{GL}_n(\mathbb{Z}/m\mathbb{Z})$ given by reduction modulo $m$ satisfies $\phi(A) \neq 1_{\mathrm{GL}_n(\mathbb{Z}/m\mathbb{Z})}$. □

We shall now give examples of groups which are not residually finite.

A group $G$ is called *divisible* if for each $g \in G$ and each integer $n \geq 1$, there is an element $h \in G$ such that $h^n = g$.

*Example 2.1.6.* The additive groups $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are divisible. More generally, every $\mathbb{Q}$-vector space, with its additive underlying group structure, is divisible. In particular, every field of characteristic 0, with its underlying additive group structure, is divisible.

**Lemma 2.1.7.** *Let $G$ be a divisible group and let $F$ be a finite group. Then every homomorphism $\phi\colon G \to F$ is trivial.*

*Proof.* Set $n = |F|$. Let $g \in G$. As $G$ is divisible, we can find $h \in G$ such that $g = h^n$. If $\phi\colon G \to F$ is a homomorphism, then we have $\phi(g) = \phi(h^n) = \phi(h)^n = 1_F$. □

The preceding lemma immediately yields the following result.

**Proposition 2.1.8.** *A nontrivial divisible group cannot be residually finite.* □

*Example 2.1.9.* The additive group underlying a field of characteristic 0 is not residually finite. In particular, the additive group $\mathbb{Q}$ is not residually finite.

In order to give another characterization of residually finiteness, we shall use the following:

**Lemma 2.1.10.** *Let $G$ be a group and let $H$ be a subgroup of $G$. Let $K = \bigcap_{g\in G} gHg^{-1}$. Then $K$ is a normal subgroup of $G$ contained in $H$. Moreover, if $H$ is of finite index in $G$, then $K$ is of finite index in $G$.*

*Proof.* Since $gHg^{-1} = H$ for $g = 1_G$, we have $K \subset H$. Let $\mathrm{Sym}(G/H)$ denote the group of permutations of $G/H$. Consider the action of $G$ on $G/H$ given by left multiplication and let $\rho\colon G \to \mathrm{Sym}(G/H)$ be the associated homomorphism. For each $g \in G$, the stabilizer of $gH$ is $gHg^{-1}$. Consequently, we have $K = \mathrm{Ker}(\rho)$, which shows that $K$ is a normal subgroup of $G$. The group $G/K$ is isomorphic to $\mathrm{Im}(\rho) \subset \mathrm{Sym}(G/H)$. Suppose that $H$ is of finite index in $G$. This means that the set $G/H$ is finite. This implies that the group $\mathrm{Sym}(G/H)$ is finite. We deduce that the group $G/K$ is finite, that is, $K$ is of finite index in $G$. □

Given a group $G$, the intersection of all subgroups of finite index of $G$ is called the *residual subgroup* (or *profinite kernel*) of $G$.

**Proposition 2.1.11.** *Let $G$ be a group and let $N$ denote the residual subgroup of $G$. Then:*

(i) *$N$ is equal to the intersection of all normal subgroups of finite index in $G$;*
(ii) *$N$ is a normal subgroup of $G$;*
(iii) *$G$ is residually finite if and only if $N = \{1_G\}$.*

*Proof.* Denote by $N'$ the intersection of all normal subgroups of finite index of $G$. The inclusion $N \subset N'$ is trivial. If $H$ is a subgroup of finite index of $G$, then $K = \cap_{g\in G} gHg^{-1}$ is a normal subgroup of finite index of $G$ contained in $H$, by Lemma 2.1.10. This implies $N' \subset K \subset H$. It follows that $N' \subset N$. This shows (i).

Assertion (ii) follows from (i) since the intersection of a family of normal subgroups of $G$ is a normal subgroup of $G$.

Suppose that $G$ is residually finite. Let $g \in G$ such that $g \neq 1_G$. Then there exist a finite group $F$ and a homomorphism $\phi\colon G \to F$ such that $\phi(g) \neq 1_F$. Thus $g \notin \mathrm{Ker}(\phi)$. As the group $G/Ker(\phi)$ is isomorphic to $F$, the subgroup $\mathrm{Ker}(\phi)$ is of finite index in $G$. This shows that $N = \{1_G\}$.

Conversely, suppose that $N = \{1_G\}$. Let $g \in G$ such that $g \neq 1_G$. By (i), we can find a normal subgroup of finite index $K \subset G$ such that $g \notin K$. If $\phi\colon G \to G/K$ is the canonical homomorphism, we have $\phi(g) \neq 1_{G/K}$. This shows that $G$ is residually finite. □

## 2.2 Stability Properties of Residually Finite Groups

**Proposition 2.2.1.** *Every subgroup of a residually finite group is residually finite.*

*Proof.* Let $G$ be a residually finite group and let $H$ be a subgroup of $G$. Let $h \in H$ such that $h \neq 1_G$. Since $G$ is residually finite, there exist a finite group $F$ and a homomorphism $\phi\colon G \to F$ such that $\phi(h) \neq 1_F$. If $\phi'\colon H \to F$ is the restriction of $\phi$ to $H$, we have $\phi'(h) = \phi(h) \neq 1_F$. Consequently, $H$ is residually finite. □

**Proposition 2.2.2.** *Let $(G_i)_{i\in I}$ be a family of residually finite groups. Then their direct product $G = \prod_{i\in I} G_i$ is residually finite.*

*Proof.* Let $g = (g_i)_{i\in I} \in G$ such that $g \neq 1_G$. Then there exists $i_0 \in I$ such that $g_{i_0} \neq 1_{G_{i_0}}$. Since $G_{i_0}$ is residually finite, we can find a finite group $F$ and a homomorphism $\phi\colon G_{i_0} \to F$ such that $\phi(g_{i_0}) \neq 1_F$. Consider the homomorphism $\phi'\colon G \to F$ defined by $\phi' = \pi \circ \phi$, where $\pi\colon G \to G_{i_0}$ is the projection onto $G_{i_0}$. We have $\phi'(g) = \phi(g_{i_0}) \neq 1_F$. Consequently, $G$ is residually finite. □

**Corollary 2.2.3.** *Let $(G_i)_{i\in I}$ be a family of residually finite groups. Then their direct sum $G = \bigoplus_{i\in I} G_i$ is residually finite.*

*Proof.* This follows immediately from Proposition 2.2.1 and Proposition 2.2.2, since $G$ is the subgroup of the direct product $P = \prod_{i\in I} G_i$ consisting of all $g = (g_i) \in P$ for which $g_i = 1_{G_i}$ for all but finitely many $i \in I$. □

**Corollary 2.2.4.** *Every finitely generated abelian group is residually finite.*

*Proof.* If $G$ is a finitely generated abelian group, then there exist an integer $r \geq 0$ and a finite abelian group $T$ such that $G$ is isomorphic to $\mathbb{Z}^r \times T$. By using Proposition 2.1.3 and Proposition 2.1.4, we deduce that $G$ is residually finite. □

*Remark 2.2.5.* An arbitrary abelian group need not be residually finite. For example, the additive group $\mathbb{Q}$ is not residually finite (Example 2.1.9).

**Corollary 2.2.6.** *Let $G$ be a group. Then the following conditions are equivalent:*

(a) *the group $G$ is residually finite;*
(b) *there exist a family $(F_i)_{i\in I}$ of finite groups such that the group $G$ is isomorphic to a subgroup of the direct product group $\prod_{i\in I} F_i$.*

*Proof.* The fact that (b) implies (a) follows from Proposition 2.1.3, Proposition 2.2.2 and Proposition 2.2.1. Conversely, suppose that $G$ is residually finite. Then, for each $g \in G \setminus \{1_G\}$, we can find a finite group $F_g$ and a homomorphism $\phi_g\colon G \to F_g$ such that $\phi_g(g) \neq 1_{F_g}$. Consider the group

$$H = \prod_{g \in G \setminus \{1_G\}} F_g.$$

The homomorphism $\psi \colon G \to H$ defined by

$$\psi = \prod_{g \in G \setminus \{1_G\}} \phi_g$$

is injective. Therefore, $G$ is isomorphic to a subgroup of $H$. This shows that (a) implies (b). □

The class of residually finite groups is closed under taking projective limits (see Sect. E.2 for the definition of the limit of a projective system of groups):

**Proposition 2.2.7.** *If a group $G$ is the limit of a projective system of residually finite groups, then $G$ is residually finite.*

*Proof.* Let $(G_i)_{i \in I}$ be a projective system of residually finite groups such that $G = \varprojlim G_i$. By construction of a projective limit (see Appendix E), $G$ is a subgroup of the group $\prod_{i \in I} G_i$. We deduce that $G$ is residually finite by using Proposition 2.2.2 and Proposition 2.2.1. □

A group $G$ is called *profinite* if $G$ is the limit of some projective system of finite groups. An immediate consequence of Proposition 2.2.7 is the following:

**Corollary 2.2.8.** *Every profinite group is residually finite.* □

*Example 2.2.9.* Let $p$ be a prime number. Given integers $n \geq m \geq 0$, let $\phi_{n,m} \colon \mathbb{Z}/p^m\mathbb{Z} \to \mathbb{Z}/p^n\mathbb{Z}$ denote reduction modulo $p^n$. Then $(\mathbb{Z}/p^n\mathbb{Z}, \phi_{n,m})$ is a projective system of groups over $\mathbb{N}$. The limit of this projective system is called the group of *$p$-adic integers* and is denoted by $\mathbb{Z}_p$. Since $\mathbb{Z}_p$ is the projective limit of finite groups, it is profinite and hence residually finite by Corollary 2.2.8.

*Remark 2.2.10.* A group which is the limit of an inductive system of residually finite groups need not be residually finite. For instance, we have seen in Example 2.1.9 that the additive group underlying a field of characteristic 0 is not residually finite. However, such a group is the limit of the inductive system formed by its finitely generated subgroups, which are all residually finite by Corollary 2.2.4.

If $\mathcal{P}$ is a property of groups, one says that a group $G$ is *virtually* $\mathcal{P}$ if $G$ contains a subgroup of finite index which satisfies $\mathcal{P}$.

**Lemma 2.2.11.** *Let $G$ be a group. Let $H$ be a subgroup of finite index of $G$ and let $K$ be a subgroup of finite index of $H$. Then $K$ is a subgroup of finite index of $G$.*

*Proof.* Let $h_1, \dots, h_n$ be a complete set of representatives of the left cosets of $G$ modulo $H$ and let $k_1, \dots, k_p$ be a complete set of representatives of the left cosets of $H$ modulo $K$. Observe that the elements $h_i k_j$, $1 \leq i \leq n$, $1 \leq j \leq p$, form a complete set of representatives of the left cosets of $G$ modulo $K$. Therefore $[G : K] = np = [G : H][H : K] < \infty$. □

**Proposition 2.2.12.** *Every virtually residually finite group is residually finite.*

*Proof.* Let $G$ be a group and let $H$ be a subgroup of finite index of $G$. By Lemma 2.2.11, the intersection of the subgroups of finite index of $G$ is contained in the intersection of the subgroups of finite index of $H$. Since a group is residually finite if and only if the intersection of its subgroups of finite index is reduced to the identity element (Proposition 2.1.11), we deduce that $G$ is residually finite if $H$ is residually finite. □

## 2.3 Residual Finiteness of Free Groups

The goal of this section is to establish the following result:

**Theorem 2.3.1.** *Every free group is residually finite.*

To prove this theorem, we shall use the following:

**Lemma 2.3.2.** *The subgroup of* $\mathrm{SL}_2(\mathbb{Z})$ *generated by the matrices*

$$a = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$$

*is a free group of rank* 2.

*Proof.* The group $\mathrm{GL}_2(\mathbb{R})$ naturally acts on the set of lines of $\mathbb{R}^2$ passing through the origin, that is, on the projective line $P_1(\mathbb{R})$. In nonhomogeneous coordinates this action is given by

$$gt = \frac{g_{11}t + g_{12}}{g_{21}t + g_{22}}$$

for $g = (g_{ij})_{1 \leq i,j \leq 2} \in \mathrm{GL}_2(\mathbb{R})$ and $t \in P_1(\mathbb{R}) = \mathbb{R} \cup \{\infty\}$ representing the line of $\mathbb{R}^2$ with slope $1/t$ passing through $(0, 0)$ (see Fig. 2.1). Note that we have

$$a^k t = t + 2k \quad \text{and} \quad b^k t = \frac{t}{2kt + 1}$$

for all $k \in \mathbb{Z}$. Consider the subsets $Y$ and $Z$ of $P_1(\mathbb{R})$ defined by

$$Y = ]-1, 1[ \quad \text{and} \quad Z = P_1(\mathbb{R}) \setminus [-1, 1].$$

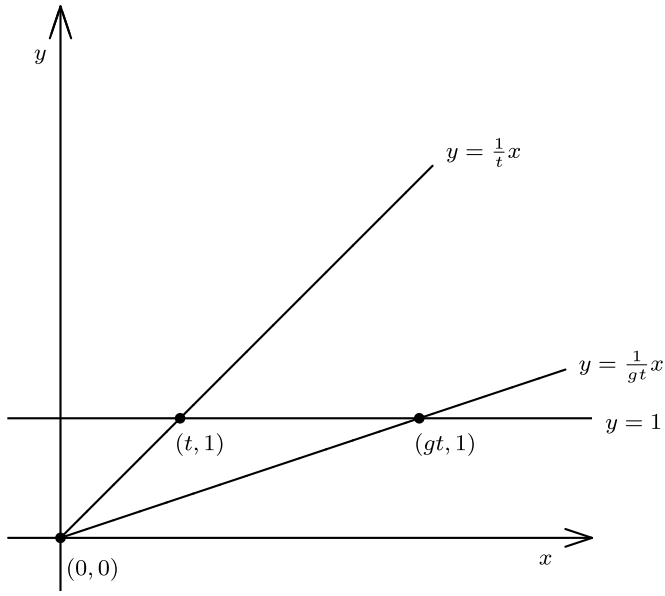



**Fig. 2.1** The action of $\mathrm{GL}_2(\mathbb{R})$ on $P_1(\mathbb{R})$. Here, $t = 1$ and $g = a$

One immediately checks that, for all $k \in \mathbb{Z} \setminus \{0\}$, one has

$$a^k Y = ]2k-1, 2k+1[ \subset Z \quad \text{and} \quad b^k Z = ]1/(2k+1), 1/(2k-1)[ \subset Y.$$

By applying the Klein Ping-Pong theorem (Theorem D.5.1), we deduce that $a$ and $b$ generate a free group of rank 2. □

*Proof of Theorem 2.3.1.* Since $\mathrm{GL}_2(\mathbb{Z})$ is residually finite by Proposition 2.1.5, it follows from Lemma 2.3.2, Corollary D.5.3, and Proposition 2.2.1 that every free group of finite rank is residually finite.

Consider now an arbitrary set $X$ and let $F(X)$ denote the free group based on $X$. Let $g \in F(X)$ such that $g \neq 1_{F(X)}$. Let $Y \subset X$ denote the set of elements $x \in X$ such that $x^k$ appear in the reduced form of $g$ for some $k \in \mathbb{Z} \setminus \{0\}$. The subgroup $F(Y) \subset F(X)$ generated by $Y$ is a free group with base $Y$ (see Proposition D.2.4). We have $g \in F(Y)$. As the group $F(Y)$ is free of finite rank, it is residually finite by the first part of the proof. Thus we can find a finite group $H$ and a homomorphism $\phi\colon F(Y) \to H$ such that $\phi(g) \neq 1_H$. Consider the unique homomorphism $\pi\colon F(X) \to F(Y)$ such that $\pi(y) = y$ for every $y \in Y$ and $\pi(x) = 1_{F(Y)}$ for every $x \in X \setminus Y$. Then the homomorphism $\phi \circ \pi\colon F(X) \to H$ satisfies $\phi \circ \pi(g) = \phi(\pi(g)) = \phi(g) \neq 1_H$. This shows that $F(X)$ is residually finite. □

## 2.4 Hopfian Groups

**Definition 2.4.1.** A group $G$ is called *Hopfian* if every surjective endomorphism of $G$ is injective.

*Examples 2.4.2.* (a) Every finite group is Hopfian.

(b) The additive group $\mathbb{Q}$ is Hopfian. Indeed, every endomorphism of the group $\mathbb{Q}$ is of the form $x \mapsto ax$ for some $a \in \mathbb{Q}$, and it is clear that such an endomorphism is surjective (resp. injective) if and only if $a \neq 0$.

(c) Every simple group is Hopfian. (we recall that a *simple* group is a nontrivial group $G$ such that the only normal subgroups of $G$ are $\{1_G\}$ and $G$).

(d) The additive group $\mathbb{Q}/\mathbb{Z}$ is not Hopfian. Indeed, the endomorphism $\psi\colon \mathbb{Q}/\mathbb{Z} \to \mathbb{Q}/\mathbb{Z}$ defined by $\psi(x) = 2x$ is surjective but not injective.

**Theorem 2.4.3.** *Every finitely generated residually finite group is Hopfian.*

Given groups $G_1$ and $G_2$, we shall denote by $\mathrm{Hom}(G_1, G_2)$ the set of all homomorphisms from $G_1$ to $G_2$. We begin by establishing the following result.

**Lemma 2.4.4.** *Let $G$ be a finitely generated group and let $F$ be a finite group. Then the set $\mathrm{Hom}(G, F)$ is finite.*

*Proof.* Let $A$ be a finite generating subset of $G$. Let us set $n = |A|$ and $p = |F|$. As $A$ generates $G$, any homomorphism $u\colon G \to F$ is completely determined by the elements $u(a)$, $a \in A$. Thus the set $\mathrm{Hom}(G, F)$ contains at most $p^n$ elements. □

*Proof of Theorem 2.4.3.* Let $G$ be a finitely generated residually finite group. Suppose that $\psi\colon G \to G$ is a surjective endomorphism of $G$. Let $K$ be a normal subgroup of finite index of $G$ and let $\rho\colon G \to G/K$ denote the canonical homomorphism. Consider the map

$$\Phi\colon \mathrm{Hom}(G, G/K) \to \mathrm{Hom}(G, G/K)$$

defined by $\Phi(u) = u \circ \psi$ for all $u \in \mathrm{Hom}(G, G/K)$. The map $\Phi$ is injective since $\psi$ is surjective by our hypothesis. As the set $\mathrm{Hom}(G, G/K)$ is finite by Lemma 2.4.4, we deduce that $\Phi$ is also surjective. In particular, there exists a homomorphism $u_0 \in \mathrm{Hom}(G, G/K)$ such that $\rho = u_0 \circ \psi$. This implies $\mathrm{Ker}(\psi) \subset \mathrm{Ker}(\rho) = K$. It follows that $\mathrm{Ker}(\psi)$ is contained in the intersection of all normal subgroups of finite index of $G$. As $G$ is residually finite, we deduce that $\mathrm{Ker}(\psi) = \{1_G\}$ by Proposition 2.1.11(iii). Thus $\psi$ is injective. This shows that $G$ is Hopfian. □

Since every free group is residually finite by Theorem 2.3.1, an immediate consequence of Theorem 2.4.3 is the following:

**Corollary 2.4.5.** *Every free group of finite rank is Hopfian.*

*Remarks 2.4.6.* (a) Let $X$ be an infinite set and let $F(X)$ denote the free group based on $X$. Every surjective but non injective map $f\colon X \to X$ extends to a surjective endomorphism $\psi\colon F(X) \to F(X)$ which is not injective. Consequently, the group $F(X)$ is not Hopfian. Since $F(X)$ is residually finite by Theorem 2.3.1, this shows that we cannot suppress the hypothesis that $G$ is finitely generated in Theorem 2.4.3.

(b) An example of a finitely generated Hopfian group which is not residually finite will be given in Sect. 2.6 (see Proposition 2.6.1).

## 2.5 Automorphism Groups of Residually Finite Groups

Let $G$ be a group. Recall that an *automorphism* of $G$ is a bijective homomorphism $\alpha\colon G \to G$. Clearly, the set $\mathrm{Aut}(G)$ consisting of all automorphisms of $G$ is a subgroup of the symmetric group of $G$. The group $\mathrm{Aut}(G)$ is called the *automorphism group* of $G$.

**Theorem 2.5.1.** *Let $G$ be a finitely generated residually finite group. Then the group $\mathrm{Aut}(G)$ is residually finite.*

Let us first establish the following result.

**Lemma 2.5.2.** *Let $G$ be a group. Let $H_1$ and $H_2$ be subgroups of finite index of $G$. Then the subgroup $H = H_1 \cap H_2$ is of finite index in $G$.*

*Proof.* Two elements in $G$ are left congruent modulo $H$ if and only if they are both left congruent modulo $H_1$ and left congruent modulo $H_2$. Therefore, there is an injective map from $G/H$ into $G/H_1 \times G/H_2$ given by $gH \mapsto (gH_1, gH_2)$. As the sets $G/H_1$ and $G/H_2$ are finite by hypothesis, we deduce that $G/H$ is finite, that is, $H$ is of finite index in $G$. □

*Proof of Theorem 2.5.1.* Let $\alpha_0 \in \mathrm{Aut}(G)$ such that $\alpha_0 \neq \mathrm{Id}_G$. Then we can find an element $g_0 \in G$ such that $\alpha_0(g_0) \neq g_0$. As $G$ is residually finite, there exist a finite group $F$ and a homomorphism $\phi\colon G \to F$ satisfying $\phi(\alpha_0(g_0)) \neq \phi(g_0)$. Consider the set $H$ defined by

$$H = \bigcap_{\psi \in \mathrm{Hom}(G,F)} \mathrm{Ker}(\psi),$$

where $\mathrm{Hom}(G, F)$ denotes, as above, the set of all homomorphisms from $G$ to $F$. Observe that $H$ is a normal subgroup of $G$ since it is the intersection of a family of normal subgroups of $G$. On the other hand, for every $\alpha \in \mathrm{Aut}(G)$, one has

$$\begin{aligned}\alpha(H) &= \alpha\left(\bigcap_{\psi\in\mathrm{Hom}(G,F)} \mathrm{Ker}(\psi)\right)\\ &= \bigcap_{\psi\in\mathrm{Hom}(G,F)} \alpha(\mathrm{Ker}(\psi))\\ &= \bigcap_{\psi\in\mathrm{Hom}(G,F)} \mathrm{Ker}(\psi\circ\alpha^{-1}).\end{aligned}$$

As the map from $\mathrm{Hom}(G,F)$ to itself defined by $\psi \mapsto \psi\circ\alpha^{-1}$ is bijective (with $\psi \mapsto \psi\circ\alpha$ as inverse map), we get

$$\alpha(H) = \bigcap_{\psi\in\mathrm{Hom}(G,F)} \mathrm{Ker}(\psi) = H.$$

Therefore $\alpha$ induces an automorphism $\overline{\alpha}$ of $G/H$, given by $\overline{\alpha}(gH) = \alpha(g)H$ for all $g \in G$. The map $\alpha \mapsto \overline{\alpha}$ is clearly a homomorphism from $\mathrm{Aut}(G)$ to $\mathrm{Aut}(G/H)$. Let us show that the group $\mathrm{Aut}(G/H)$ is finite and that $\overline{\alpha_0} \neq 1_{\mathrm{Aut}(G/H)} = \mathrm{Id}_{G/H}$. Observe first that the set $\mathrm{Hom}(G,F)$ is finite by Lemma 2.4.4. As $\mathrm{Ker}(\psi)$ is of finite index in $G$ for every $\psi \in \mathrm{Hom}(G,F)$, we deduce that $H$ is of finite index in $G$ by applying Lemma 2.5.2. This implies that the group $\mathrm{Aut}(G/H)$ is finite. On the other hand, we have $\overline{\alpha_0}(g_0H) = \alpha_0(g_0)H \neq g_0H$ since $H$ is a subgroup of $\mathrm{Ker}(\phi)$ and $\phi(g_0) \neq g_0$. Therefore $\overline{\alpha_0} \neq \mathrm{Id}_{G/H}$. This shows that the group $\mathrm{Aut}(G)$ is residually finite. □

Every free group is residually finite by Theorem 2.3.1. Therefore we deduce from Theorem 2.5.1 the following result.

**Corollary 2.5.3.** *The group* $\mathrm{Aut}(F_n)$ *is residually finite for every* $n \geq 1$. □

Every finitely generated abelian group is residually finite by Corollary 2.2.4. Thus we have:

**Corollary 2.5.4.** *The automorphism group of a finitely generated abelian group is residually finite.* □

*Remark 2.5.5.* The automorphism group of a free abelian group of finite rank $n$ is isomorphic to $\mathrm{GL}_n(\mathbb{Z})$. Thus, the residual finiteness of $\mathrm{GL}_n(\mathbb{Z})$ (Proposition 2.1.5) may also be deduced from Corollary 2.5.4.

**Corollary 2.5.6.** *Let* $R$ *be a ring and let* $M$ *be a left (or right) module over* $R$*. Suppose that* $M$ *is finitely generated as a* $\mathbb{Z}$*-module. Then the automorphism group* $\mathrm{Aut}_R(M)$ *of the* $R$*-module* $M$ *is residually finite.*

*Proof.* The group $\mathrm{Aut}_R(M)$ is a subgroup of $\mathrm{Aut}_{\mathbb{Z}}(M)$. Since $\mathrm{Aut}_{\mathbb{Z}}(M)$ is residually finite by Corollary 2.5.4, we deduce that $\mathrm{Aut}_R(M)$ is residually finite by applying Proposition 2.2.1. □

**Corollary 2.5.7.** *Let $R$ be a ring. Suppose that $R$ is finitely generated as a $\mathbb{Z}$-module. Then the group $\mathrm{GL}_n(R)$ is residually finite for every $n \geq 1$.*

*Proof.* This is an immediate consequence of the preceding corollary since the group $\mathrm{GL}_n(R)$ is isomorphic to $\mathrm{Aut}_R(R^n)$, where $R^n$ is viewed as a left module over $R$. □

*Example 2.5.8.* The ring of Gaussian integers $\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$ is a free abelian $\mathbb{Z}$-module of rank 2. Thus, the group $\mathrm{GL}_n(\mathbb{Z}[i])$ is residually finite for every $n \geq 1$ by Corollary 2.5.7. More generally, consider a number field $K$, that is, a field extension of $\mathbb{Q}$ such that $d = \dim_{\mathbb{Q}} K < \infty$. Let $A$ denote the ring of algebraic integers of $K$ (we recall that an element $x \in K$ is called an *algebraic integer* of $K$ if $x$ is a root of a monic polynomial with integral coefficients). It is a standard fact in algebraic number theory that $A$ is a free $\mathbb{Z}$-module of rank $d$. Thus, the group $\mathrm{GL}_n(A)$ is residually finite for every $n \geq 1$ by Corollary 2.5.7.

## 2.6 Examples of Finitely Generated Groups Which Are Not Residually Finite

We have seen in Example 2.1.9 that the additive group $\mathbb{Q}$ is not residually finite. Observe that the group $\mathbb{Q}$ is not finitely generated. In fact, any finitely generated abelian group is residually finite by Corollary 2.2.4. The purpose of this section is to give two examples of finitely generated groups $G_1$ and $G_2$ which are not residually finite.

Our first example is a subgroup of the symmetric group $\mathrm{Sym}(\mathbb{Z})$ generated by two elements:

**Proposition 2.6.1.** *Let $G_1$ denote the subgroup of $\mathrm{Sym}(\mathbb{Z})$ generated by the translation $T\colon n \mapsto n + 1$ and the transposition $S = (0\ 1)$. Then $G_1$ is a finitely generated Hopfian group which is not residually finite.*

Let us first establish the following lemmas:

**Lemma 2.6.2.** *Let $G$ be an infinite simple group. Then $G$ is not residually finite.*

*Proof.* The only normal subgroup of finite index of $G$ is $G$ itself. Therefore $G$ is not residually finite by Proposition 2.1.11(iii). □

Given a set $X$, we recall that $\mathrm{Sym}_0(X)$ denotes the subgroup of $\mathrm{Sym}(X)$ consisting of all permutations of $X$ whose support is finite (see Appendix C).

**Lemma 2.6.3.** *Let $X$ be an infinite set. Then the group $\mathrm{Sym}_0(X)$ is not residually finite.*

*Proof.* The subgroup $\mathrm{Sym}_0^+(X) \subset \mathrm{Sym}(X)$ consisting of all permutations of $X$ with finite support and signature 1 is an infinite simple group by Theorem C.4.3. Therefore $\mathrm{Sym}_0^+(X)$ is not residually finite by Lemma 2.6.2. We deduce that $\mathrm{Sym}_0(X)$ is not residually finite by using Proposition 2.2.1. □

**Lemma 2.6.4.** *The group $G_1$ contains $\mathrm{Sym}_0(\mathbb{Z})$ as a normal subgroup. Moreover, $G_1$ is the semidirect product of $\mathrm{Sym}_0(\mathbb{Z})$ with the infinite cyclic subgroup of $G_1$ generated by $T$.*

*Proof.* For all $i \in \mathbb{Z}$, we have

$$T^i S T^{-i} = (i \;\; i+1), \tag{2.1}$$

which shows that $(i \;\; i+1) \in G_1$. If $i < j$, we also have

$$(i \; j+1) = (j \;\; j+1)(i \;\; j)(j \;\; j+1)$$

which, by induction on $j$, shows that $(i \; j) \in G_1$ for all $i, j \in \mathbb{Z}$ with $i < j$. Since the transpositions $(i \; j)$, with $i, j \in \mathbb{Z}$ and $i < j$, generate $\mathrm{Sym}_0(\mathbb{Z})$ by Corollary C.2.4, it follows that $G_1$ contains $\mathrm{Sym}_0(\mathbb{Z})$. The fact that $\mathrm{Sym}_0(\mathbb{Z})$ is normal in $\mathrm{Sym}(\mathbb{Z})$ (Proposition C.2.2) implies that $\mathrm{Sym}_0(\mathbb{Z})$ is normal in $G_1$. Let $H$ denote the subgroup of $G_1$ generated by $T$. It is clear that $H \cap \mathrm{Sym}_0(\mathbb{Z})$ is reduced to the identity map $\mathrm{Id}_{\mathbb{Z}}$. On the other hand, by using (2.1), we see that every element $g \in G_1$ may be written in the form $g = h\sigma$, where $h \in H$ and $\sigma \in \mathrm{Sym}_0(\mathbb{Z})$. Therefore, $G_1$ is the semidirect product of $\mathrm{Sym}_0(\mathbb{Z})$ and $H$. □

*Proof of Proposition 2.6.1.* The group $G_1$ contains $\mathrm{Sym}_0(\mathbb{Z})$ as a subgroup by Lemma 2.6.4. As the group $\mathrm{Sym}_0(\mathbb{Z})$ is not residually finite by Lemma 2.6.3, we conclude that $G_1$ is not residually finite by applying Proposition 2.2.1.

It remains to show that $G_1$ is Hopfian. We start by observing that there are exactly two surjective homomorphisms from $G_1$ onto $\mathbb{Z}$. Indeed, as $G_1$ is the semidirect product of $\mathrm{Sym}_0(\mathbb{Z})$ with the subgroup generated by $T$ (Lemma 2.6.4), every element $g \in G_1$ can be uniquely written in the form $g = T^k\sigma$, where $k \in \mathbb{Z}$ and $\sigma \in \mathrm{Sym}_0(\mathbb{Z})$, and the map $u\colon G_1 \to \mathbb{Z}$ defined by $u(g) = k$ is a surjective homomorphism. As all elements of $\mathrm{Sym}_0(\mathbb{Z})$ have finite order, it immediately follows that the only homomorphisms from $G_1$ onto $\mathbb{Z}$ are $u$ and $-u$. Now, let $\phi\colon G_1 \to G_1$ be a surjective homomorphism. Then $u\circ\phi\colon G_1 \to \mathbb{Z}$ is a surjective homomorphism so that, by our preceding observation, we have $u\circ\phi = u$ or $-u$. As $\mathrm{Ker}(u) = \mathrm{Sym}_0(\mathbb{Z})$ and $\phi\colon G_1 \to G_1$ is onto, it follows that $\mathrm{Ker}(\phi) \subset \mathrm{Sym}_0(\mathbb{Z})$ and $\phi(\mathrm{Sym}_0(\mathbb{Z})) = \mathrm{Sym}_0(\mathbb{Z})$. The simplicity of $\mathrm{Sym}_0^+(\mathbb{Z})$ (Theorem C.4.3) implies that $\mathrm{Ker}(\phi) \cap \mathrm{Sym}_0^+(\mathbb{Z})$ is either equal to $\mathrm{Sym}_0^+(\mathbb{Z})$ or reduced to the identity. We cannot have $\mathrm{Ker}(\phi) \cap \mathrm{Sym}_0^+(\mathbb{Z}) = \mathrm{Sym}_0^+(\mathbb{Z})$, that is, $\mathrm{Sym}_0^+(\mathbb{Z}) \subset \mathrm{Ker}(\phi)$, since this would imply that $\phi(\mathrm{Sym}_0(\mathbb{Z}))$ is either reduced to the identity or cyclic of order 2, which contradicts $\phi(\mathrm{Sym}_0(\mathbb{Z})) = \mathrm{Sym}_0(\mathbb{Z})$. Consequently, we have $\mathrm{Ker}(\phi) \cap \mathrm{Sym}_0^+(\mathbb{Z}) = \{\mathrm{Id}_{\mathbb{Z}}\}$. If $\mathrm{Ker}(\phi)$ is not

reduced to the identity, it follows that $\mathrm{Ker}(\phi)$ is a normal subgroup of order 2 of $\mathrm{Sym}_0(\mathbb{Z})$. But this is impossible since Proposition C.2.3 combined with Proposition C.3.2 implies that, if $X$ is an infinite set, then every non-trivial element in $\mathrm{Sym}_0(X)$ has an infinite number of conjugates in $\mathrm{Sym}_0(X)$. Thus $\mathrm{Ker}(\phi)$ must be reduced to the identity, that is, $\phi$ is injective. This shows that $G_1$ is Hopfian. □

Our second example of a finitely generated but not residually finite group will also show that the class of residually finite groups is not closed under extensions. In other words, a group $G$ admitting a normal subgroup $N$ such that both $N$ and $G/N$ are residually finite may fail to be residually finite.

In order to construct it we consider first the group

$$H = \bigoplus_{i\in\mathbb{Z}} H_i,$$

where each $H_i$ is a copy of the alternating group $\mathrm{Sym}_5^+$. Let $\psi\colon \mathbb{Z} \to \mathrm{Aut}(H)$ be the homomorphism defined by $\psi(n) = \alpha^n$, where $\alpha \in \mathrm{Aut}(H)$ is the one-step shift given by

$$\alpha(h) = (h_{i-1})_{i\in\mathbb{Z}}$$

for all $h = (h_i)_{i\in\mathbb{Z}} \in H$.

**Proposition 2.6.5.** *The semidirect product $G_2 = H \ltimes_\psi \mathbb{Z}$ is a finitely generated but not residually finite group.*

*Proof.* By definition of a semidirect product, $G_2$ is the group with underlying set $H \times \mathbb{Z}$ and group operation given by

$$(h, n)(h', n') = (h\alpha^n(h'), n + n')$$

for all $(h, n), (h', n') \in H \times \mathbb{Z}$.

Let $t$ be the element of $G_2$ defined by $t = (1_H, 1)$ and identify each element $h \in H$ with the element $(h, 0) \in G$. Then $H$ is a normal subgroup of $G_2$ and the quotient group $G_2/H$ is an infinite cyclic group generated by the class of $t$. One has

$$(h, n) = ht^n \tag{2.2}$$

and

$$t^n h t^{-n} = \alpha^n(h) \tag{2.3}$$

for all $h \in H$ and $n \in \mathbb{Z}$. It follows from (2.2) that $G_2$ is generated by $t$ and the elements of $H$. In fact, since $H$ is the direct sum of the groups $H_i$ and

$$H_i = t^i H_0 t^{-i} \tag{2.4}$$

by (2.3), we deduce that $G_2$ is generated by $t$ and the elements of $H_0$. As the group $H_0$ is finite, this shows that $G$ is finitely generated.

Let us prove now that $G_2$ is not residually finite. Suppose on the contrary that $G_2$ is residually finite. Let $g$ be an element in $H_0$ such that $g \neq 1_{H_0}$. The residual finiteness of $G_2$ implies the existence of a finite group $F$ and a homomorphism $\phi\colon G_2 \to F$ such that $\phi(g) \neq 1_F$. As $H_0$ is a simple group, the restriction of $\phi$ to $H_0$ is injective. Therefore, the group $\phi(H_0)$ is isomorphic to $H_0$ and hence to $\mathrm{Sym}_5^+$. Let $m = |F|$. We have $\phi(t^m) = \phi(t)^m = 1_F$. Thus, it follows from (2.4) that $\phi(H_m) = \phi(H_0)$. Since $xy = yx$ for all $x \in H_0$ and $y \in H_m$, this implies that $\phi(H_0)$ is abelian. This contradicts the fact that $\phi(H_0)$ is isomorphic to $\mathrm{Sym}_5^+$, which is not abelian. □

*Remark 2.6.6.* Observe that the group $H$ is residually finite by Corollary 2.2.3. The quotient group $G_2/H$ is also residually finite since it is isomorphic to $\mathbb{Z}$. This shows that an extension of a residually finite group by a residually finite group need not to be residually finite.

## 2.7 Dynamical Characterization of Residual Finiteness

Let $G$ be a group and let $A$ be a set. Recall that the set $A^G = \{x\colon G \to A\}$ is equipped with the prodiscrete topology and that $G$ acts on $A^G$ by the left shift defined by (1.2).

**Theorem 2.7.1.** *Let $G$ be a group. Then the following conditions are equivalent:*

(a) *the group $G$ is residually finite;*
(b) *for every set $A$, the set of points of $A^G$ which have a finite $G$-orbit is dense in $A^G$;*
(c) *there exists a set $A$ having at least two elements such that the set of points of $A^G$ which have a finite $G$-orbit is dense in $A^G$;*
(d) *there exists a Hausdorff topological space $X$ equipped with a continuous and faithful action of $G$ such that the set of points of $X$ which have a finite $G$-orbit is dense in $X$.*

We recall that an action of a group $G$ on a set $X$ is called *faithful* if $1_G$ is the only element of $G$ fixing all points of $X$.

**Lemma 2.7.2.** *Let $G$ be a group and let $A$ be a set having at least two elements. Then the action of $G$ on $A^G$ is faithful.*

*Proof.* Let $a$ and $b$ be two distinct elements in $A$. Consider an element $g_0$ in $G$ such that $g_0 \neq 1_G$. Let $x \in A^G$ be the configuration defined by $x(g) = a$ if $g = 1_G$ and $x(g) = b$ otherwise. We have $g_0 x \neq x$ since $g_0 x(1_G) = x(g_0^{-1}) = b$ and $x(1_G) = a$. Consequently, the action of $G$ on $A^G$ is faithful. □

**Lemma 2.7.3.** *Let $G$ be a residually finite group and let $\Omega$ be a finite subset of $G$. Then there exists a normal subgroup of finite index $K$ of $G$ such that the restriction of the canonical homomorphism $\rho\colon G \to G/K$ to $\Omega$ is injective.*

*Proof.* Consider the finite subset

$$S = \{g^{-1}h : g, h \in \Omega \text{ and } g \neq h\} \subset G.$$

Since $G$ is residually finite, we can find, for every $s \in S$, a normal subgroup of finite index $N_s \subset G$ such that $s \notin N_s$. The set $K = \cap_{s \in S} N_s$ is a normal subgroup of finite index in $G$ by Lemma 2.5.2. Let $\rho\colon G \to G/K$ be the canonical homomorphism. If $g$ and $h$ are distinct elements in $\Omega$, then $g^{-1}h \notin K$ and hence $\rho(g) \neq \rho(h)$. □

*Proof of Theorem 2.7.1.* Suppose that $G$ is residually finite. Let $A$ be a set and let $W$ be a neighborhood of a point $x$ in $A^G$. Let us show that $W$ contains a configuration with finite $G$-orbit.

Consider a finite subset $\Omega \subset G$ such that

$$V(x, \Omega) = \{y \in A^G : y|_\Omega = x|_\Omega\} \subset W.$$

By Lemma 2.7.3, we can find a normal subgroup of finite index $K \subset G$ such that the restriction to $\Omega$ of the canonical homomorphism $\rho\colon G \to G/K (= K\backslash G)$ is injective. This implies that the map $\Phi\colon A^{G/K} \to A^\Omega$ defined by $\Phi(z) = (z \circ \rho)|_\Omega$ is surjective. Thus we can find an element $z_0 \in A^{G/K}$ such that the configurations $z_0 \circ \rho$ and $x$ coincide on $\Omega$, that is, such that $z_0 \circ \rho \in V(x, \Omega)$. On the other hand, the configuration $z_0 \circ \rho$ is $K$-periodic by Proposition 1.3.3 (observe that $K\backslash G = G/K$ as $K$ is normal in $G$). As $K$ is of finite index in $G$, we deduce that the $G$-orbit of $z_0 \circ \rho$ is finite. Thus $W$ contains a configuration whose $G$-orbit is finite. This shows that (a) implies (b).

Implication (b) ⇒ (c) is trivial.

The fact that (c) implies (d) follows from Proposition 1.2.1, Proposition 1.2.2, and Lemma 2.7.2.

Let us show that (d) implies (a). Suppose that the group $G$ acts continuously and faithfully on a Hausdorff topological space $X$ and let $E$ denote the set of points of $X$ whose $G$-orbit is finite. For $x \in E$, let $\mathrm{Stab}(x) = \{g \in G : gx = x\}$ denote the stabilizer of $x$ in $G$. Observe that $x \in E$ if and only if $\mathrm{Stab}(x)$ is of finite index in $G$. If $E$ is dense in $X$, then $\cap_{x \in E} \mathrm{Stab}(x) = \{1_G\}$ since the action of $G$ on $X$ is continuous and faithful. Thus, by Proposition 2.1.11 we have that $G$ is residually finite. □

## Notes

A survey article on residually finite groups has been written by W. Magnus [Mag].

A group $G$ is called *linear* if there exist an integer $n \geq 1$ and a field $K$ such that $G$ is isomorphic to a subgroup of $\mathrm{GL}_n(K)$. A theorem of A.I. Mal'cev [Mal1] asserts that every finitely generated linear group is residually finite.

An example of a finitely presented group which is residually finite but not linear was recently given by C. Druţu and M. Sapir [DrS].

The residual finiteness of free groups was established by F. Levi [Lev]. The proof presented in Sect. 2.3 is based on the fact that the matrices $\left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)$ generate a free subgroup of $\mathrm{GL}_2(\mathbb{Z})$. This last result is due to I.N. Sanov [San]. The reader may find other proofs of the residual finiteness of free groups in [RobD, p. 158] and [MaKS, p. 116].

Hopfian groups are named after H. Hopf who used topological methods to prove that fundamental groups of closed orientable surfaces are Hopfian (see [MaKS, p. 415]) and raised the question of the existence of finitely generated non-Hopfian groups. The fact that every finitely generated residually finite group is Hopfian (Theorem 2.4.3) was discovered by Mal'cev [Mal1]. The first example of a finitely generated non-Hopfian group was given by B.H. Neumann [Neu2]. Shortly after, a finitely presented non-Hopfian group was found by G. Higman [Hig]. The simplest example of a finitely presented non-Hopfian group is certainly provided by the Baumslag-Solitar group $BS(2,3) = \langle x, y : yx^2y^{-1} = x^3 \rangle$, that is, the quotient of the free group $F_2$ based on two generators $x$ and $y$ by the smallest normal subgroup of $F_2$ containing the element $x^{-3}yx^2y^{-1}$ (see [BaS], [MaKS], [LS]). On the other hand, the Baumslag-Solitar group $BS(2,4) = \langle x, y : yx^2y^{-1} = x^4 \rangle$ is an example of a finitely presented Hopfian group which is not residually finite (the incorrect statement in [BaS] about the residual finiteness of $BS(2,4)$ was corrected by S. Meskin in [Mes]).

The residual finiteness of the automorphism group of a finitely generated residually finite group (Theorem 2.5.1) was proved by G. Baumslag in [Bau].

The group $G_1$ of Sect. 2.6 was considered by Mal'cev in [Mal1]. Given groups $A$ and $G$, the *wreath product* $A \wr G$ is the semidirect product $H \ltimes_\psi G$, where $H = \oplus_{g \in G} A$ and $\psi \colon G \to \mathrm{Aut}(A)$ is the group homomorphism associated with the action of $G$ on $H \subset A^G$ induced by the $G$-shift (see [RobD], [Rot]). Thus, the group $G_2$ of Sect. 2.6 is the wreath product $G_2 = \mathrm{Sym}_5^+ \wr \mathbb{Z}$. The proof of Proposition 2.6.5 shows that, if $A$ is a finitely generated nonabelian simple group and $G$ is a finitely generated infinite group, then the wreath product $A \wr G$ is a finitely generated group which is not residually finite. Examples of finitely generated nonabelian simple groups are provided by the (finite) groups $\mathrm{Sym}_n^+$, where $n \geq 5$, or $\mathrm{PSL}_n(K)$, where $n \geq 2$ and $K$ a finite field having at least 4 elements, or one of the famous infinite finitely presented simple Thompson groups $T$ or $V$ (see [CFP]).

## Exercises

**2.1.** Show that every quotient of a divisible group is a divisible group.

**2.2.** Show that every torsionfree divisible abelian group $G$ is isomorphic to a direct sum of copies of $\mathbb{Q}$. Hint: Prove that there is a natural $\mathbb{Q}$-vector space structure on $G$.

**2.3.** Let $G$ be a group.

(a) Show that it is possible to define a topology on $G$ by taking as open sets the subsets $\Omega \subset G$ which satisfy the following property: for each $g \in \Omega$ there is a subgroup of finite index $H \subset G$ such that $gH \subset \Omega$. This topology is called the *profinite topology* on $G$.

(b) Show that $G$ is residually finite if and only if the profinite topology on $G$ is Hausdorff.

**2.4.** Show that the class of residually finite groups is not closed under taking quotients. Hint: Any group is isomorphic to a quotient of a free group (see Corollary D.4.2).

**2.5.** Let $X$ be an infinite set. Show that the symmetric group $\mathrm{Sym}(X)$ is not residually finite. Hint: Use Cayley's theorem (Theorem C.1.2) to prove that $\mathrm{Sym}(X)$ contains a subgroup isomorphic to $\mathbb{Q}$ or use Lemma 2.6.3.

**2.6.** Let $G$ be a residually finite group and let $A$ be a finite set.

(a) Show that there exists a canonical injective homomorphism of the group $\mathrm{ICA}(G; A)$ (cf. Sect. 1.10) into the group $\prod_H \mathrm{Sym}(\mathrm{Fix}(H))$, where $H$ runs over all finite index subgroups of $G$. Hint. Use the fact that the configurations with finite $G$-orbit are dense in $A^G$ (cf. Theorem 2.7.1).

(b) Deduce from (a) that $\mathrm{ICA}(G; A)$ is residually finite.

**2.7.** Let $m$ be an integer such that $|m| \geq 2$. Let $G$ denote the quotient of the free group $F_2$ on two generators $x$ and $y$ by the normal closure of the single element $xyx^{-1}y^{-m}$. Thus, $G$ is the group given by the presentation $G = \langle a, b : aba^{-1} = b^m \rangle$, where $a$ and $b$ denote the images in $G$ of $x$ and $y$ by the quotient homomorphism.

(a) Show that every element $g \in G$ may be (not uniquely) written in the form $g = a^i b^j a^k$ for some $i, j, k \in \mathbb{Z}$.

(b) Show that there is a unique group homomorphism $\phi\colon G \to \mathrm{GL}_2(\mathbb{Q})$ which satisfies

$$\phi(a) = \begin{pmatrix} m & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \phi(b) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

(c) Use (a) to show that $\phi$ is injective and that the image of $\phi$ is the subgroup $\phi(G) \subset \mathrm{GL}_2(\mathbb{Q})$ given by

$$\phi(G) = \left\{ \begin{pmatrix} m^n & r \\ 0 & 1 \end{pmatrix} : n \in \mathbb{Z}, r \in \mathbb{Z}[1/m] \right\},$$

where $\mathbb{Z}[1/m]$ denotes the set of all rationals $r \in \mathbb{Q}$ which can be written in the form $r = um^v$ for some $u, v \in \mathbb{Z}$.

(d) Use (c) to prove that $G$ is torsionfree.

(e) Prove that $G$ is residually finite. Hint: You have to show that, for any element $g \in G \setminus \{1_G\}$, there exist a finite group $F$ and a homomorphism $\psi\colon G \to F$ such that $\psi(g) \neq 1_F$. Write $g = a^i b^j a^k$, where $i, j, k \in \mathbb{Z}$, and

treat first the case $i + k \neq 0$ by using determinants. If $i + k = 0$, choose a prime number $p$ which is not a divisor of $m$ nor a divisor of $j$, and consider the homomorphism $\psi\colon G \to \mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$ defined by

$$\psi(a) = \begin{pmatrix} \overline{m} & \overline{0} \\ \overline{0} & \overline{1} \end{pmatrix} \quad \text{and} \quad \psi(b) = \begin{pmatrix} \overline{1} & \overline{1} \\ \overline{0} & \overline{1} \end{pmatrix},$$

where $\overline{n} \in \mathbb{Z}/p\mathbb{Z}$ denotes the class of $n \in \mathbb{Z}$ modulo $p$.

**2.8.** Let $m \geq 2$ be an integer. Denote by $\mathbb{Z}[1/m]$ the subgroup of the additive group $\mathbb{Q}$ consisting of all rationals $r$ which can be written in the form $r = km^n$ for some $k, n \in \mathbb{Z}$. Show that the group $\mathbb{Z}[1/m]$ is residually finite. Hint: Observe that $\mathbb{Z}[1/m]$ is isomorphic to a subgroup of the group $G$ studied in Exercise 2.7.

**2.9.** Let $K$ be a field of characteristic $p > 0$. Show that the additive group $K$ is residually finite. Hint: Use a base of $K$ seen as a $\mathbb{Z}/p\mathbb{Z}$ vector space and apply Corollary 2.2.3.

**2.10.** Show that all elements of the additive group $\mathbb{Q}/\mathbb{Z}$ have finite order but that $\mathbb{Q}/\mathbb{Z}$ is not residually finite.

**2.11.** Show that the multiplicative group $\mathbb{Q}^*$ of nonzero rational numbers is residually finite. Hint: Use the unique factorization of elements of $\mathbb{Q}^*$ into powers of prime numbers and apply Corollary 2.2.3.

**2.12.** Show that the automorphism group $\mathrm{Aut}(\mathbb{Q})$ of the additive group $\mathbb{Q}$ is isomorphic to the multiplicative group $\mathbb{Q}^*$.

**2.13.** Show that the additive group $\mathbb{R}$ is neither residually finite nor Hopfian.

**2.14.** Let $p$ be a prime. Show that the additive group $\mathbb{Z}_p$ of $p$-adic integers is not Hopfian.

**2.15.** Show that the multiplicative group $\mathbb{R}^*$ of nonzero real numbers is neither residually finite nor Hopfian.

**2.16.** Let $F$ be a free group of finite rank $n$. Suppose that $S$ is a finite generating subset of $F$ of cardinality $|S| \leq n$. Use the fact that $F$ is Hopfian (Corollary 2.4.5) to prove that $|S| = n$ and that $S$ is a base for $F$.

**2.17.** (M. Hall [Hal-M-2]) Let $G$ be a finitely generated group and let $n \geq 1$ be an integer. Show that $G$ contains only a finite number of subgroups of index $n$. Hint: Use Lemma 2.1.10 and Lemma 2.4.4.

**2.18.** Show that the group $\mathrm{GL}_n(\mathbb{Z})$ is finitely generated. Hint: For $1 \leq i, j \leq n$, let $E_{ij}$ denote the $n \times n$ matrix all of whose entries are 0 except the entry located on the $i$th row and the $j$th column which is equal to 1. Use Euclidean division in $\mathbb{Z}$ to show that $\mathrm{GL}_n(\mathbb{Z})$ is generated by the $n^2$ matrices $I_n - 2E_{ii}$, $I_n + E_{ij}$, where $1 \leq i, j \leq n$ and $i \neq j$.

**2.19.** Show that $\mathrm{GL}_n(\mathbb{Z})$ is Hopfian.

**2.20.** Give a direct proof of the fact that the automorphism group of a finitely generated abelian group $G$ is residually finite (Corollary 2.5.4). Hint: Show that there is a finite group $F$ and an integer $n \geq 0$ such that $\mathrm{Aut}(G)$ is isomorphic to $F \times \mathrm{GL}_n(\mathbb{Z})$.

**2.21.** Show that the group $G_1$ considered in Sect. 2.6 contains no normal subgroup $H$ such that both $H$ and $G/H$ are residually finite.

**2.22.** Show that the group $G_2$ considered in Sect. 2.6 is Hopfian.

**2.23.** A group $G$ is called *almost perfect* if all nontrivial finite quotients of $G$ are nonabelian. Show that, if $A$ is a finitely generated almost perfect nontrivial group and $G$ is a finitely generated infinite group, then the wreath product group $A \wr G$ is finitely generated but not residually finite. Hint: Follow the proof of Proposition 2.6.5.

**2.24.** Let $G$ be a group. Denote by $\mathcal{N}_{fq}$ the set of all normal subgroups of finite index of $G$, partially ordered by reverse inclusion.

(a) Show that $\mathcal{N}_{fq}$ is a directed set.

(b) For $H, K \in \mathcal{N}_{fq}$ with $H \subset K$, let $\varphi_{K,H} \colon G/H \to G/K$ denote the canonical homomorphism. Show that the directed set $\mathcal{N}_{fq}$ together with the homomorphisms $\varphi_{K,H}$ form a projective system of groups. The limit of this projective system is called the *profinite completion* of the group $G$ and is denoted by $\widehat{G}$.

(c) Show that there is a canonical homomorphism $\eta \colon G \to \widehat{G}$ and that the kernel of $\eta$ is the residual subgroup of $G$.

(d) Prove that $G$ is residually finite if and only if the canonical homomorphism $\eta \colon G \to \widehat{G}$ is injective.

# Chapter 3
# Surjunctive Groups

Surjunctive groups are defined in Sect. 3.1 as being the groups on which all injective cellular automata with finite alphabet are surjective. In Sect. 3.2 it is shown that every subgroup of a surjunctive group is a surjunctive group and that every locally surjunctive group is surjunctive. Every locally residually finite group is surjunctive (Corollary 3.3.6). The class of locally residually finite groups is quite large and includes in particular all finite groups, all abelian groups, and all free groups (a still wider class of surjunctive groups, namely the class of sofic groups, will be described in Chap. 7). In Sect. 3.4, given an arbitrary group $\Gamma$, we introduce a natural topology on the set of its quotient groups. In Sect. 3.7, it is shown that the set of surjunctive quotients is closed in the space of all quotients of $\Gamma$.

## 3.1 Definition

A set $X$ is finite if and only if every injective map $f\colon X \to X$ is surjective. The definition given below is related to this characterization of finite sets.

**Definition 3.1.1.** A group $G$ is said to be *surjunctive* if it satisfies the following condition: if $A$ is a finite set, then every injective cellular automaton $\tau\colon A^G \to A^G$ is surjective (and hence bijective).

*Remark 3.1.2.* Given a group $G$ and a finite set $A$, it follows from Theorem 1.8.1 that a map $f\colon A^G \to A^G$ is a cellular automaton if and only if $f$ is $G$-equivariant (with respect to the $G$-shift) and continuous (with respect to the prodiscrete topology on $A^G$). Thus, the definition of a surjunctive group may be reformulated as follows: a group $G$ is surjunctive if and only if, for any finite set $A$, every injective $G$-equivariant continuous map $f\colon A^G \to A^G$ is surjective.

**Proposition 3.1.3.** *Every finite group is surjunctive.*

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_3, 

*Proof.* If $G$ is a finite group and $A$ is a finite set, then the set $A^G$ is finite. Therefore, every injective cellular automaton $\tau\colon A^G \to A^G$ is surjective. □

## 3.2 Stability Properties of Surjunctive Groups

**Proposition 3.2.1.** *Every subgroup of a surjunctive group is surjunctive.*

*Proof.* Suppose that $H$ is a subgroup of a surjunctive group $G$. Let $A$ be a finite set and let $\tau\colon A^H \to A^H$ be an injective cellular automaton over $H$. Consider the cellular automaton $\tau^G\colon A^G \to A^G$ over $G$ obtained from $\tau$ by induction (see Sect. 1.7). The fact that $\tau$ is injective implies that $\tau^G$ is injective by Proposition 1.7.4(i). Since $G$ is surjunctive, it follows that $\tau^G$ is surjective. By applying Proposition 1.7.4(ii), we deduce that $\tau$ is surjective. □

**Proposition 3.2.2.** *Let $G$ be a group. Then the following conditions are equivalent:*

(a) *$G$ is surjunctive;*
(b) *every finitely generated subgroup of $G$ is surjunctive.*

*Proof.* The fact that (a) implies (b) follows from Proposition 3.2.1. Conversely, let $G$ be a group all of whose finitely generated subgroups are surjunctive. Let $A$ be a finite set and let $\tau\colon A^G \to A^G$ be an injective cellular automaton with memory set $S$. Let $H$ denote the subgroup of $G$ generated by $S$ and consider the cellular automaton $\tau_H\colon A^H \to A^H$ obtained by restriction of $\tau$ (see Sect. 1.7). The fact that $\tau$ is injective implies that $\tau_H$ is injective by Proposition 1.7.4(i). As $H$ is finitely generated, it is surjunctive by our hypothesis on $G$. It follows that $\tau_H$ is surjective. By applying Proposition 1.7.4(ii), we deduce that $\tau$ is also surjective. This shows that (b) implies (a). □

If $\mathcal{P}$ is a property of groups (e.g. being finite, being nilpotent, being solvable, being free, etc.), a group $G$ is called *locally $\mathcal{P}$* if all finitely generated subgroups of $G$ satisfy $\mathcal{P}$. With this terminology, Proposition 3.2.2 may be rephrased by saying that the class of surjunctive groups and the class of locally surjunctive groups coincide.

**Corollary 3.2.3.** *Every locally finite group is surjunctive.*

*Proof.* This immediately follows from Proposition 3.2.2 since every finite group is surjunctive by Proposition 3.1.3. □

*Example 3.2.4.* Let $X$ be a set and consider the group $\mathrm{Sym}_0(X)$ consisting of all permutations of $X$ which have finite support (see Sect. C.2). The group $\mathrm{Sym}_0(X)$ is locally finite. Indeed, if $\Sigma$ is a finite subset of $\mathrm{Sym}_0(X)$, then

the subgroup $H \subset \mathrm{Sym}_0(X)$ generated by $\Sigma$ is finite since it is isomorphic to a subgroup of $\mathrm{Sym}(A)$, where $A$ denotes the union of the supports of the elements of $\Sigma$. Consequently, the group $\mathrm{Sym}_0(X)$ is surjunctive. Observe that $\mathrm{Sym}_0(X)$ is infinite when $X$ is infinite. This yields our first examples of infinite surjunctive groups. Finally, note that it follows from Lemma 2.6.3 that $\mathrm{Sym}_0(X)$ is not residually finite whenever $X$ is infinite.

## 3.3 Surjunctivity of Locally Residually Finite Groups

**Theorem 3.3.1.** *Every residually finite group is surjunctive.*

Let us first establish an important property of cellular automata with finite alphabet, namely the fact that they always have a closed image with respect to the prodiscrete topology on the set of configurations:

**Lemma 3.3.2.** *Let $G$ be a group and let $A$ be a finite set. Let $\tau\colon A^G \to A^G$ be a cellular automaton. Then the set $\tau(A^G)$ is closed in $A^G$ for the prodiscrete topology.*

*Proof.* Since $A$ is finite, the space $A^G$ is compact by Tychonoff theorem (see Corollary A.5.3). As $\tau$ is continuous by Proposition 1.4.8, we deduce that $\tau(A^G)$ is a compact subset of $A^G$. This implies that $\tau(A^G)$ is closed in $A^G$ since $A^G$ is Hausdorff by Proposition 1.2.1. □

The following example shows that Lemma 3.3.2 becomes false if we suppress the hypothesis that the alphabet is finite.

*Example 3.3.3.* Consider the map $\tau\colon \mathbb{N}^{\mathbb{Z}} \to \mathbb{N}^{\mathbb{Z}}$ given by

$$\tau(x)(n) = \max(0, x(n) - x(n+1))$$

for all $x \in \mathbb{N}^{\mathbb{Z}}$ and $n \in \mathbb{Z}$. Clearly, $\tau$ is a cellular automaton over the group $\mathbb{Z}$ and the alphabet $\mathbb{N}$ with memory set $\{0, 1\}$. Consider the configuration $y\colon \mathbb{Z} \to \mathbb{N}$ defined by $y(n) = 1$ if $n \geq 0$ and $y(n) = 0$ otherwise. Let $F$ be a finite subset of $\mathbb{Z}$ and choose an integer $M \geq 1$ such that $F \subset [-M+1, M-1]$. Consider the configuration $x_M\colon \mathbb{Z} \to \mathbb{N}$ defined by $x_M(n) = \max(0, M - n)$ if $n \geq 0$ and $x_M(n) = M$ otherwise. Observe that $y_M = \tau(x_M)$ is such that $y_M(n) = 1$ if $0 \leq n \leq M - 1$ and $y_M(n) = 0$ otherwise. Therefore, the configurations $y_M$ and $y$ coincide on $[-M+1, M-1]$ and hence on $F$. Thus $y$ is in the closure of $\tau(\mathbb{N}^{\mathbb{Z}})$ in $\mathbb{N}^{\mathbb{Z}}$. On the other hand, it is clear that $y$ is not in the image of $\tau$. This shows that $\tau(\mathbb{N}^{\mathbb{Z}})$ is not closed in $\mathbb{N}^{\mathbb{Z}}$ for the prodiscrete topology.

In the proof of the surjunctivity of residually finite groups, we shall also use the following result:

**Lemma 3.3.4.** *Let $G$ be a group. Suppose that $G$ satisfies the following property: for each finite subset $\Omega \subset G$, there exist a surjunctive group $\Gamma$ and a homomorphism $\phi\colon G \to \Gamma$ such that the restriction of $\phi$ to $\Omega$ is injective. Then $G$ is surjunctive.*

*Proof.* Let $A$ be a finite set and equip $A^G$ with its prodiscrete topology. Suppose that $\tau\colon A^G \to A^G$ is an injective cellular automaton. Let us show that $\tau$ is surjective, that is, $\tau(A^G) = A^G$. Since $\tau(A^G)$ is closed in $A^G$ by Lemma 3.3.2, it suffices to prove that $\tau(A^G)$ is dense in $A^G$. Consider a configuration $x \in A^G$ and a neighborhood $W$ of $x$ in $A^G$. Then there is a finite subset $\Omega \subset G$ such that

$$V(x, \Omega) = \{y \in A^G : y|_\Omega = x|_\Omega\} \subset W.$$

By our hypothesis, we may find a surjunctive group $\Gamma$ and a homomorphism $\phi\colon G \to \Gamma$ such that the restriction of $\phi$ to $\Omega$ is injective. Let $K$ denote the image of $\phi$ and let $N$ denote its kernel. By Proposition 1.3.3, there is a bijective map $\psi\colon A^K \to \mathrm{Fix}(N)$ given by $\psi(z) = z \circ \phi$ for all $z \in A^K$. Moreover, if we identify $A^K$ with $\mathrm{Fix}(N)$ via $\psi$, the restriction of $\tau$ to $\mathrm{Fix}(N)$ yields a cellular automaton $\overline{\tau}\colon A^K \to A^K$ over the group $K$ (see Proposition 1.6.1). Since the group $\Gamma$ is surjunctive, the group $K$ is surjunctive by Proposition 3.2.1. We deduce that $\overline{\tau}$ is surjective and hence $\tau(\mathrm{Fix}(N)) = \mathrm{Fix}(N)$. On the other hand, as the restriction of $\phi$ to $\Omega$ is injective, we may find $z_0 \in A^K$ such that $\psi(z_0) = x|_\Omega$. We then have

$$\psi(z_0) \in \mathrm{Fix}(N) \cap V(x, \Omega) = \tau(\mathrm{Fix}(N)) \cap V(x, \Omega).$$

This shows that $W$ meets the image of $\tau$. Thus $\tau(A^G)$ is dense in $A^G$. □

*Proof of Theorem 3.3.1.* Let $G$ be a residually finite group. If $\Omega$ is a finite subset of $G$, then there exist, By Lemma 2.7.3, a finite group $\Gamma$ and a homomorphism $\phi\colon G \to \Gamma$ such that the restriction of $\phi$ to $\Omega$ is injective. As finite groups are surjunctive by Proposition 3.1.3, it follows that $G$ satisfies the hypothesis of Lemma 3.3.4. Consequently, $G$ is surjunctive. □

**Corollary 3.3.5.** *Every free group is surjunctive.*

*Proof.* Free groups are residually finite by Theorem 2.3.1. □

**Corollary 3.3.6.** *Every locally residually finite group is surjunctive.*

*Proof.* This immediately follows from Theorem 3.3.1 by using Proposition 3.2.2. □

**Corollary 3.3.7.** *Every abelian group is surjunctive.*

*Proof.* This is an immediate consequence of Corollary 3.3.6 since every abelian group is locally residually finite by Corollary 2.2.4. □

When $X$ is a finite set, every surjective map $f\colon X \to X$ is injective. Therefore, a surjective cellular automaton with finite alphabet over a finite group is necessarily injective. However, a surjective cellular automaton with finite alphabet $\tau\colon A^G \to A^G$ may fail to be injective when the group $G$ is infinite as shown by the following example.

*Example 3.3.8.* Take $A = \mathbb{Z}/2\mathbb{Z} = \{0,1\}$ and $G = \mathbb{Z}$. Let $\tau\colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ be the map defined by $\tau(x)(n) = x(n) + x(n+1)$ for all $x \in A^{\mathbb{Z}}$ and $n \in \mathbb{Z}$. Clearly, $\tau$ is a cellular automaton with memory set $S = \{0,1\} \subset \mathbb{Z}$ and local defining map $\mu\colon A^S \to A$ given by $\mu(y) = y(0) + y(1)$ for all $y \in A^S$. The cellular automaton $\tau$ is surjective. Indeed, given an element $y \in A^G$, the configuration $x\colon \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$ defined by

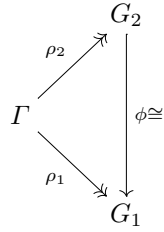
$$x(n) = \begin{cases} 0 & \text{if } n = 0, \\ y(0) + y(1) + \cdots + y(n-1) & \text{if } n > 0, \\ y(n) + y(n+1) + \cdots + y(-1) & \text{if } n < 0, \end{cases}$$

satisfies $\tau(x) = y$. However $\tau$ is not injective since the two constant configurations have the same image under $\tau$.

## 3.4 Marked Groups

Let $\Gamma$ be a group.

A *$\Gamma$-quotient* is a pair $(G, \rho)$, where $G$ is a group and $\rho\colon \Gamma \to G$ is a surjective homomorphism. We define an equivalence relation on the class of all $\Gamma$-quotients by declaring that two $\Gamma$-quotients $(G_1, \rho_1)$ and $(G_2, \rho_2)$ are *equivalent* when there exists a group isomorphism $\phi\colon G_2 \to G_1$ such that the following diagram is commutative:




that is, such that $\rho_1 = \phi \circ \rho_2$. An equivalence class of $\Gamma$-quotients is called a *$\Gamma$-marked group*. Observe that two $\Gamma$-quotients $(G_1, \rho_1)$ and $(G_2, \rho_2)$ are equivalent if and only if $\mathrm{Ker}(\rho_1) = \mathrm{Ker}(\rho_2)$. Thus, the set of $\Gamma$-marked groups may be identified with the set $\mathcal{N}(\Gamma)$ consisting of all normal subgroups of $\Gamma$.

Let us identify the set $\mathcal{P}(\Gamma)$ consisting of all subsets of $\Gamma$ with the set $\{0,1\}^\Gamma$ by means of the bijection from $\mathcal{P}(\Gamma)$ onto $\{0,1\}^\Gamma$ given by $A \mapsto \chi_A$,

where $\chi_A \colon \Gamma \to \{0,1\}$ is the characteristic map of $A \subset \Gamma$. We equip the set $\mathcal{P}(\Gamma) = \{0,1\}^\Gamma$ with its prodiscrete uniform structure and $\mathcal{N}(\Gamma) \subset \mathcal{P}(\Gamma)$ with the induced uniform structure. Thus, a base of entourages of $\mathcal{N}(\Gamma)$ is provided by the sets

$$V_F = \{(N_1, N_2) \in \mathcal{N}(\Gamma) \times \mathcal{N}(\Gamma) : N_1 \cap F = N_2 \cap F\},$$

where $F$ runs over all finite subsets of $\Gamma$. Intuitively, two normal subgroups of $\Gamma$ are "close" in $\mathcal{N}(\Gamma)$ when their intersection with a large finite subset of $\Gamma$ coincide.

**Proposition 3.4.1.** *Let $\Gamma$ be a group. Then the space $\mathcal{N}(\Gamma)$ of $\Gamma$-marked groups is a totally disconnected compact Hausdorff topological space.*

*Proof.* The space $\mathcal{P}(\Gamma) = \{0,1\}^\Gamma$ is totally disconnected and Hausdorff by Proposition 1.2.1. Moreover, $\mathcal{P}(\Gamma)$ is compact by Corollary A.5.3 since it is a product of finite spaces. Thus, it suffices to show that $\mathcal{N}(\Gamma)$ is closed in $\mathcal{P}(\Gamma)$. To see this, observe that a subset $E \in \mathcal{P}(\Gamma)$ is a normal subgroup of $\Gamma$ if and only if it satisfies

(1) $1_\Gamma \in E$;
(2) $\alpha\beta^{-1} \in E$ for all $\alpha, \beta \in E$;
(3) $\gamma\alpha\gamma^{-1} \in E$ for all $\alpha \in E$ and $\gamma \in \Gamma$.

Denoting, for each $\gamma \in \Gamma$, by $\pi_\gamma \colon \mathcal{P}(\Gamma) = \{0,1\}^\Gamma \to \{0,1\}$ the projection map corresponding to the $\gamma$-factor, these conditions are equivalent to

(1') $\pi_{1_\Gamma}(E) = 1$;
(2') $\pi_\alpha(E)\pi_\beta(E)(\pi_{\alpha\beta^{-1}}(E) - 1) = 0$ for all $\alpha, \beta \in \Gamma$;
(3') $\pi_\alpha(E)(\pi_{\gamma\alpha\gamma^{-1}}(E) - 1) = 0$ for all $\alpha, \gamma \in \Gamma$.

As all projection maps are continuous on $\mathcal{P}(\Gamma)$, this shows that $\mathcal{N}(\Gamma)$ is closed in $\mathcal{P}(\Gamma)$. □

*Remark 3.4.2.* If $\Gamma$ is countable then the uniform structure on $\mathcal{N}(\Gamma)$ is metrizable. Indeed, the uniform structure on $\mathcal{P}(\Gamma) = \{0,1\}^\Gamma$ is metrizable when $\Gamma$ is countable (cf. Remark 1.9.2).

Let $\mathcal{P}$ be a property of groups. A group $\Gamma$ is called *residually $\mathcal{P}$* if for each element $\gamma \in \Gamma$ with $\gamma \neq 1_\Gamma$, there exist a group $\Gamma'$ satisfying $\mathcal{P}$ and an epimorphism $\phi \colon \Gamma \to \Gamma'$ such that $\phi(\gamma) \neq 1_{\Gamma'}$. Observe that every group which satisfies $\mathcal{P}$ is residually $\mathcal{P}$.

**Proposition 3.4.3.** *Let $\Gamma$ be a group and let $\mathcal{P}$ be a property of groups. Suppose that the class of groups which satisfy $\mathcal{P}$ is closed under taking finite products and subgroups. Then the following conditions are equivalent:*

(a) *$\Gamma$ is residually $\mathcal{P}$;*
(b) *there exists a net $(N_i)_{i \in I}$ in $\mathcal{N}(\Gamma)$ which converges to $\{1_\Gamma\}$ such that $\Gamma/N_i$ satisfies $\mathcal{P}$ for all $i \in I$.*

*Proof.* Suppose that $\Gamma$ is residually $\mathcal{P}$. For all $\gamma \in \Gamma \backslash \{1_\Gamma\}$ we can find a group $\Gamma_\gamma$ satisfying $\mathcal{P}$ and an epimorphism $\phi_\gamma \colon \Gamma \to \Gamma_\gamma$ such that $\phi_\gamma(\gamma) \neq 1_{\Gamma_\gamma}$. Let $I \subset \mathcal{P}(\Gamma)$ be the directed set consisting of all finite subsets of $\Gamma$ not containing $1_\Gamma$ partially ordered by inclusion. For $i \in I$ we denote by $\phi_i = \prod_{\gamma \in i} \phi_\gamma \colon \Gamma \to \prod_{\gamma \in i} \Gamma_\gamma$ the homomorphism defined by $\phi_i(\gamma') = (\phi_\gamma(\gamma'))_{\gamma \in I}$ for all $\gamma' \in \Gamma$. Set $N_i = \ker(\phi_i)$ and let us show that the net $(N_i)_{i \in I}$ in $\mathcal{N}(\Gamma)$ converges to $\{1_\Gamma\}$. Given a finite set $F \subset \Gamma$ set $i_F = F \setminus \{1_\Gamma\}$. Let $f \in F \setminus \{1_\Gamma\}$ and $i \in I$ such that $i_F \subset i$. Then $\phi_i(f) \neq 1_{\Gamma_i}$ since $\phi_f(f) \neq 1_{\Gamma_f}$. Thus

$$f \notin N_i \quad \text{and} \quad f \notin \{1_\Gamma\} \cap F. \tag{3.1}$$

It follows that if $1_\Gamma \notin F$ then $N_i \cap F = \varnothing = \{1_\Gamma\} \cap F$. On the other hand, if $1_\Gamma \in F$, from (3.1) we deduce $N_i \cap F = \{1_\Gamma\} = \{1_\Gamma\} \cap F$. In either cases we have $N_i \cap F = \{1_\Gamma\} \cap F$. We deduce that $\lim_i N_i = \{1_\Gamma\}$. Finally, by our assumptions on $\mathcal{P}$, we have that $\prod_{\gamma \in i} \Gamma_\gamma$ and its subgroup $\phi_i(\Gamma) \cong \Gamma / N_i$ satisfy $\mathcal{P}$ for all $i \in I$. This shows (a) $\Rightarrow$ (b).

Suppose now that (b) holds. Let $\gamma \in \Gamma \setminus \{1_\Gamma\}$. Consider the finite set $F = \{\gamma\}$. Then there exists $i_F \in I$ such that $N_{i_F} \cap F = \{1_\Gamma\} \cap F$ and $\Gamma / N_{i_F}$ satisfy $\mathcal{P}$. As $\{1_\Gamma\} \cap F = \varnothing$ we have that $\gamma \notin N_{i_F}$. In other words, if $\phi_F \colon \Gamma \to \Gamma / N_{i_F}$ denotes the canonical epimorphism, we have $\phi_F(\gamma) \neq 1_{\Gamma / N_{i_F}}$. We deduce that $\Gamma$ is residually $\mathcal{P}$. This shows (b) $\Rightarrow$ (a). □

Note that in the above proposition, the assumptions on $\mathcal{P}$ are not needed for the implication (b) $\Rightarrow$ (a)

Let $A$ be a set. Consider the set $A^\Gamma$ equipped with its prodiscrete uniform structure and the $\Gamma$-shift action. For each $N \in \mathcal{N}(\Gamma)$, let

$$\mathrm{Fix}(N) = \{x \in A^\Gamma : \gamma x = x \text{ for all } \gamma \in N\} \subset A^\Gamma$$

denote the set of configurations which are fixed by $N$. Recall from Proposition 1.3.6 that $\mathrm{Fix}(N)$ is a closed $\Gamma$-invariant subset of $A^\Gamma$.

Let us equip $\mathcal{P}(A^\Gamma)$ with the Hausdorff-Bourbaki uniform structure associated with the prodiscrete uniform structure on $A^\Gamma$ (see Sect. B.4 for the definition of the Hausdorff-Bourbaki uniform structure on the set of subsets of a uniform space).

**Theorem 3.4.4.** *Let $\Gamma$ be a group and let $A$ be a set. Then the map $\Psi \colon \mathcal{N}(\Gamma) \to \mathcal{P}(A^\Gamma)$ defined by $\Psi(N) = \mathrm{Fix}(N)$ is uniformly continuous. Moreover, if $A$ contains at least two elements then $\Psi$ is a uniform embedding.*

*Proof.* Let $N_0 \in \mathcal{N}(\Gamma)$ and let $W$ be an entourage of $\mathcal{P}(A^\Gamma)$. Let us show that there exists an entourage $V$ of $\mathcal{N}(\Gamma)$ such that

$$\Psi(V[N_0]) \subset W[\Psi(N_0)]. \tag{3.2}$$

This will prove that $\Psi$ is continuous. By definition of the Hausdorff-Bourbaki uniform structure on $\mathcal{P}(A^\Gamma)$, there is an entourage $T$ of $A^\Gamma$ such that

$$\widehat{T} = \{(X, Y) \in \mathcal{P}(A^\Gamma) \times \mathcal{P}(A^\Gamma) : Y \subset T[X] \text{ and } X \subset T[Y]\} \subset W. \quad (3.3)$$

Since $A^\Gamma$ is endowed with its prodiscrete uniform structure, there is a finite subset $F \subset \Gamma$ such that

$$U = \{(x, y) \in A^\Gamma \times A^\Gamma : \pi_F(x) = \pi_F(y)\} \subset T, \quad (3.4)$$

where $\pi_F \colon A^\Gamma \to A^F$ is the projection map. Consider now the finite subset $E \subset \Gamma$ defined by

$$E = F \cdot F^{-1} = \{\gamma\eta^{-1} : \gamma, \eta \in F\},$$

and the entourage $V$ of $\mathcal{N}(\Gamma)$ given by

$$V = \{(N_1, N_2) \in \mathcal{N}(\Gamma) \times \mathcal{N}(\Gamma) : N_1 \cap E = N_2 \cap E\}.$$

We claim that $V$ satisfies (3.2). To prove our claim, suppose that $N \in V[N_0]$. Let $x \in \mathrm{Fix}(N)$. The fact that $N \cap E = N_0 \cap E$ implies that if $\gamma$ and $\eta$ are elements of $F$ with $\gamma = \nu\eta$ for some $\nu \in N$, then $\nu \in N_0$ and therefore $x(\gamma) = x(\eta)$. Denoting by $\rho_0 \colon \Gamma \to \Gamma/N_0$ the canonical epimorphism, we deduce that we may find an element $x_0 \in A^{\Gamma/N_0}$ such that $x(\gamma) = x_0 \circ \rho_0(\gamma)$ for all $\gamma \in F$. We have $x_0 \circ \rho_0 \in \mathrm{Fix}(N_0)$ and $(x, x_0 \circ \rho_0) \in U$. Since $U \subset T$ by (3.4), this shows that $\mathrm{Fix}(N) \subset T[\mathrm{Fix}(N_0)]$. Therefore $\Psi$ is continuous. As $\mathcal{N}(\Gamma)$ is compact by Proposition 3.4.1, we deduce that $\Psi$ is uniformly continuous by applying Theorem B.2.3.

Suppose now that $A$ has at least two elements. Let us show that $\Psi$ is injective. Let $N_1, N_2 \in \mathcal{N}(\Gamma)$. Fix two elements $a, b \in A$ with $a \neq b$ and consider the map $x \colon \Gamma \to A$ defined by $x(\gamma) = a$ if $\gamma \in N_1$ and $x(\gamma) = b$ otherwise. We clearly have $x \in \mathrm{Fix}(N_1)$. Suppose that $\Psi(N_1) = \Psi(N_2)$, that is, $\mathrm{Fix}(N_1) = \mathrm{Fix}(N_2)$. Then for all $\nu \in N_2$, we have $\nu^{-1}x = x$ since $x \in \mathrm{Fix}(N_1) = \mathrm{Fix}(N_2)$, and hence $x(\nu) = \nu^{-1}x(1_\Gamma) = x(1_\Gamma) = a$. This implies $N_2 \subset N_1$. By symmetry, we also have $N_1 \subset N_2$. Therefore $N_1 = N_2$. This shows that $\Psi$ is injective. As $\mathcal{N}(\Gamma)$ is compact and $\mathcal{P}(A^\Gamma)$ is Hausdorff, we conclude that $\Psi$ is a uniform embedding by applying Proposition B.2.5.

□

## 3.5 Expansive Actions on Uniform Spaces

Let $X$ be a uniform space and let $\Gamma$ be a group acting on $X$. We consider the diagonal action of $\Gamma$ on $X \times X$ defined by $\gamma(x, y) = (\gamma x, \gamma y)$ for all $\gamma \in \Gamma$ and $x, y \in X$.

One says that the action of $\Gamma$ on $X$ is *uniformly continuous* if the orbit map $X \to X$, $x \mapsto \gamma x$, is uniformly continuous for each $\gamma \in \Gamma$. This is

equivalent to saying that $\gamma^{-1}V$ is an entourage of $X$ for all entourage $V$ of $X$ and $\gamma \in \Gamma$.

The action of $\Gamma$ on $X$ is said to be *expansive* if there exists an entourage $W_0$ of $X$ such that

$$\bigcap_{\gamma \in \Gamma} \gamma^{-1} W_0 = \Delta_X, \tag{3.5}$$

where $\Delta_X = \{(x, x) : x \in X\}$ denotes the diagonal in $X \times X$. Equality (3.5) means that if $x, y \in X$ satisfy $(\gamma x, \gamma y) \in W_0$ for all $\gamma \in \Gamma$, then $x = y$. An entourage $W_0$ satisfying (3.5) is then called an *expansivity entourage* for the action of $\Gamma$ on $X$.

*Remarks 3.5.1.* (a) If a uniform space $X$ admits an expansive uniformly continuous action of a group $\Gamma$, then the topology on $X$ must be Hausdorff. Indeed, if $W_0$ is an expansivity entourage for such an action then $\Delta_X$ is equal to the intersection of the entourages $\gamma^{-1}W_0$, $\gamma \in \Gamma$, by (3.5).

(b) Suppose that $\mathcal{B}$ is a base of the uniform structure on $X$. Then an action of a group $\Gamma$ on $X$ is expansive if and only if it admits an expansivity entourage $B_0 \in \mathcal{B}$.

(c) Let $(X, d)$ be a metric space. Then an action of a group $\Gamma$ on $X$ is expansive if and only if there exists a real number $\varepsilon_0 > 0$ with the following property: if $x, y \in X$ satisfy $d(\gamma x, \gamma y) < \varepsilon_0$ for all $\gamma \in \Gamma$, then $x = y$. Such an $\varepsilon$ is called an *expansivity constant* for the action of $\Gamma$ on $X$.

Our basic example of a uniformly continuous and expansive action is provided by the following:

**Proposition 3.5.2.** *Let $G$ be a group and let $A$ be a set. Then the $G$-shift on $A^G$ is uniformly continuous and expansive with respect to the prodiscrete uniform structure on $A^G$.*

*Proof.* Uniform continuity follows from the fact that $G$ acts on $A^G$ by permuting coordinates. Expansiveness is due to the fact that the action of $G$ on itself by left multiplication is transitive. Indeed, consider the entourage $W_0$ of $A^G$ defined by

$$W_0 = \{(x, y) \in A^G \times A^G : x(1_G) = y(1_G)\}.$$

Given $g \in G$, we have $(x, y) \in g^{-1}W_0$ if and only if $x(g^{-1}) = y(g^{-1})$. Thus $\bigcap_{g \in G} g^{-1}W_0$ is equal to the diagonal in $A^G \times A^G$. □

## 3.6 Gromov's Injectivity Lemma

The following result will be used in the next section to prove that surjunctive groups define a closed subset of the space of $\Gamma$-marked groups for any group $\Gamma$.

**Theorem 3.6.1 (Gromov's injectivity lemma).** *Let $X$ be a uniform space endowed with a uniformly continuous and expansive action of a group $\Gamma$. Let $f\colon X \to X$ be a uniformly continuous and $\Gamma$-equivariant map. Suppose that $Y$ is a subset of $X$ such that the restriction of $f$ to $Y$ is a uniform embedding. Then there exists an entourage $V$ of $X$ satisfying the following property: if $Z$ is a $\Gamma$-invariant subset of $X$ such that $Z \subset V[Y]$, then the restriction of $f$ to $Z$ is injective.*

(We recall that the notation $Z \subset V[Y]$ means that for each $z \in Z$ there exists $y \in Y$ such that $(z, y) \in V$.)

*Proof.* By expansivity of the action of $\Gamma$, there is an entourage $W_0$ of $X$ such that

$$\bigcap_{\gamma \in \Gamma} \gamma^{-1} W_0 = \Delta_X. \tag{3.6}$$

It follows from the axioms of a uniform structure that we can find a symmetric entourage $S$ of $X$ such that

$$S \circ S \circ S \subset W_0. \tag{3.7}$$

Since the restriction of $f$ to $Y$ is a uniform embedding, we can find an entourage $T$ of $X$ such that

$$(f(y_1), f(y_2)) \in T \Rightarrow (y_1, y_2) \in S \tag{3.8}$$

for all $y_1, y_2 \in Y$. Let $U$ be a symmetric entourage of $X$ such that

$$U \circ U \subset T. \tag{3.9}$$

Since $f$ is uniformly continuous, we can find an entourage $E$ of $X$ such that

$$(x_1, x_2) \in E \Rightarrow (f(x_1), f(x_2)) \in U \tag{3.10}$$

for all $x_1, x_2 \in X$.

Let us show that the entourage $V = S \cap E$ has the required property. So let $Z$ be a $\Gamma$-invariant subset of $X$ such that $Z \subset V[Y]$ and let us show that the restriction of $f$ to $Z$ is injective.

Let $z'$ and $z''$ be points in $Z$ such that $f(z') = f(z'')$. Since $f$ is $\Gamma$-equivariant, we have

$$f(\gamma z') = f(\gamma z'') \tag{3.11}$$

for all $\gamma \in \Gamma$. As the points $\gamma z'$ and $\gamma z''$ stay in $Z$, the fact that $Z \subset V[Y]$ implies that there are points $y'_\gamma$ and $y''_\gamma$ in $Y$ such that $(\gamma z', y'_\gamma) \in V$ and $(\gamma z'', y''_\gamma) \in V$. Since $V \subset E$, it follows from (3.10) that $(f(\gamma z'), f(y'_\gamma))$ and $(f(\gamma z''), f(y''_\gamma))$ are in $U$. As $U$ is symmetric, we also have $(f(y'_\gamma), f(\gamma z')) \in U$. We deduce that $(f(y'_\gamma), f(y''_\gamma)) \in U \circ U \subset T$ by using (3.9) and (3.11). This implies $(y'_\gamma, y''_\gamma) \in S$ by (3.8). On the other hand, we also have $(\gamma z', y'_\gamma) \in S$ and $(y''_\gamma, \gamma z'') \in S$ since $V \subset S$ and $S$ is symmetric. It follows that

$$(\gamma z', \gamma z'') \in S \circ S \circ S \subset W_0$$

by (3.7). This gives us

$$(z', z'') \in \bigcap_{\gamma \in \Gamma} \gamma^{-1} W_0,$$

and hence $z' = z''$ by (3.6). Thus the restriction of $f$ to $Z$ is injective. □

## 3.7 Closedness of Marked Surjunctive Groups

**Theorem 3.7.1.** *Let $\Gamma$ be a group. Then the set of normal subgroups $N \subset \Gamma$ such that the quotient group $\Gamma/N$ is surjunctive is closed in $\mathcal{N}(\Gamma)$.*

*Proof.* Let $N \in \mathcal{N}(\Gamma)$ and let $(N_i)_{i \in I}$ be a net in $\mathcal{N}(\Gamma)$ converging to $N$. Suppose that the groups $\Gamma/N_i$ are surjunctive for all $i \in I$. Let us show that the group $\Gamma/N$ is also surjunctive.

Let $A$ be a finite set and let $\tau \colon A^{\Gamma/N} \to A^{\Gamma/N}$ be an injective cellular automaton over the group $\Gamma/N$. Let $S \subset \Gamma/N$ be a memory set for $\tau$ with associated local defining map $\mu \colon A^S \to A$. Choose a subset $\widetilde{S} \subset \Gamma$ such that the canonical epimorphism $\rho \colon \Gamma \to \Gamma/N$ gives a bijection $\psi \colon \widetilde{S} \to S$, and let $\pi \colon A^S \to A^{\widetilde{S}}$ denote the bijective map induced by $\psi$. Consider the cellular automaton $\widetilde{\tau} \colon A^\Gamma \to A^\Gamma$ over $\Gamma$ with memory set $\widetilde{S}$ and local defining map $\widetilde{\mu} = \mu \circ \pi^{-1} \colon A^{\widetilde{S}} \to A$. To simplify notation, let us set $X = A^\Gamma$, $f = \widetilde{\tau}$, $Y = \mathrm{Fix}(N)$ and $Z_i = \mathrm{Fix}(N_i)$.

We claim that the hypotheses of Theorem 3.6.1 are satisfied by $X$, $f$ and $Y$. Indeed, we first observe that the action of $\Gamma$ on $X$ is uniformly continuous and expansive by Proposition 3.5.2. On the other hand, the cellular automaton $f \colon X \to X$ is uniformly continuous and $\Gamma$-equivariant by Theorem 1.9.1. Finally, the restriction of $f$ to $Y$ is injective since this restriction is conjugate to $\tau$ by Proposition 1.6.1. As $Y$ is a closed subset of $X$ and hence compact, it follows from Proposition B.2.5 that the restriction of $f$ to $Y$ is a uniform embedding. By applying Theorem 3.6.1, it follows that there exists an entourage $V$ of $X$ such that if $Z$ is a $\Gamma$-invariant subset of $X$ with $Z \subset V[Y]$, then the restriction of $f$ to $Z$ is injective.

Since the net $(Z_i)_{i \in I}$ converges to $Y$ for the Hausdorff-Bourbaki topology on $\mathcal{P}(X)$ by Theorem 3.4.4, there is an element $i_0 \in I$ such that $Z_i \subset V[Y]$ for all $i \geq i_0$. As the sets $Z_i$ are $\Gamma$-invariant by Proposition 1.3.6, it follows that the restriction of $f$ to $Z_i$ is injective for all $i \geq i_0$. On the other hand, $f(Z_i) \subset Z_i$ and the restriction of $f$ to $Z_i$ is conjugate to a cellular automaton $\tau_i \colon A^{\Gamma/N_i} \to A^{\Gamma/N_i}$ over the group $\Gamma/N_i$ for all $i \in I$ by Proposition 1.6.1. As the groups $\Gamma/N_i$ are surjunctive by our hypotheses, we deduce that $f(Z_i) = Z_i$ for all $i \geq i_0$. Now, it follows from Proposition B.4.6 that the net $(f(Z_i))_{i \in I}$ converges to $f(Y)$ in $\mathcal{P}(X)$. Thus, the net $(Z_i)_{i \in I}$ converges to both $Y$ and $f(Y)$. As $Y$ and $f(Y)$ are closed in $X$ (by compactness of $Y$), we deduce that

$Y = f(Y)$ by applying Proposition B.4.3. This shows that $\tau$ is surjective since $\tau$ is conjugate to the restriction of $f$ to $Y$. Consequently, the group $\Gamma/N$ is surjunctive. □

## Notes

Surjunctive groups were introduced by W. Gottschalk. The first results on these groups are due to W. Lawton who proved in particular Proposition 3.2.1, Corollary 3.2.3, Theorem 3.3.1, and Corollary 3.3.7 (see [Got], [Law]).

The characterization of infinite sets by the existence of injective but not surjective self-maps is known as *Dedekind's definition* of infinite sets.

Proposition 3.3.6 together with Mal'cev theorem [Mal1] which says that every finitely generated linear group is residually finite implies that every linear group is surjunctive.

The problem of the existence of a non surjunctive group was raised by Gottschalk in [Got] and remains open up to now. Note that to prove that all groups are surjunctive it would suffice to prove that the symmetric group $\mathrm{Sym}(\mathbb{N})$ is surjunctive (this immediately follows from Proposition 3.2.1, Proposition 3.2.2, and the fact that every finitely generated group is countable and hence isomorphic to a subgroup of $\mathrm{Sym}(\mathbb{N})$ by Cayley's theorem).

In [CeC11] it is shown that if $G$ is a non-periodic group, then for every infinite set $A$ there exists a cellular automaton $\tau\colon A^G \to A^G$ whose image $\tau(A^G)$ is not closed in $A^G$ with respect to the prodiscrete topology (cf. Lemma 3.3.2 and Example 3.3.3).

Theorem 3.6.1 is a uniform version of Lemma 4.H" in [Gro5]. The proof presented in this chapter of the surjunctivity of limits of surjunctive groups (Theorem 3.7.1) closely follows Sect. 4 of [Gro5] (see also [CeC10]). There is another proof based on techniques from model theory (see [Gro5] and [GlG]).

## Exercises

**3.1.** Let $\mathbb{K}$ be an algebraically closed field. Show that every injective polynomial map $f\colon \mathbb{K} \to \mathbb{K}$ is surjective.

**3.2.** Show that every injective polynomial map $f\colon \mathbb{R} \to \mathbb{R}$ is surjective.

**3.3.** Show that the polynomial map $f\colon \mathbb{Q} \to \mathbb{Q}$ defined by $f(x) = x^3$ is injective but not surjective.

**3.4.** Show that every injective holomorphic map $f\colon \mathbb{C} \to \mathbb{C}$ is surjective.

**3.5.** Give an example of a real analytic map $f\colon \mathbb{R} \to \mathbb{R}$ which is injective but not surjective.

**3.6.** Let $\mathbb{K}$ be a field and let $V$ be a vector space over $\mathbb{K}$. Show that $V$ is finite-dimensional if and only if every injective endomorphism $f\colon V \to V$ is surjective.

**3.7.** Let $R$ be a ring and let $M$ be a left (or right) $R$-module. One says that $M$ is *Artinian* if every descending chain $N_0 \supset N_1 \supset N_2 \supset \dots$ eventually stabilizes (i.e., there is an integer $n_0 \geq 0$ such that $N_N = N_{n+1}$ for all $n \geq n_0$). Show that if $M$ is Artinian then every injective endomorphism $f\colon M \to M$ is surjective. Hint: Consider the sequence of submodules defined by $N_n = \mathrm{Im}(f^n)$ for $n \geq 0$.

**3.8.** Let $U = \{z \in \mathbb{C} : |z| = 1\}$. Show that every injective continuous map $f\colon U \to U$ is surjective.

**3.9.** Let $n \geq 0$ be an integer. An *$n$-dimensional topological manifold* is a nonempty Hausdorff topological space $X$ such that each point in $X$ admits a neighborhood homeomorphic to $\mathbb{R}^n$. Show that if $X$ is a compact $n$-dimensional topological manifold, then every injective continuous map $f\colon X \to X$ is surjective. Hint: Use Brouwer's invariance of domain to prove that $f(X)$ is open in $X$.

**3.10.** Let $\mathcal{P}$ be a property of groups. Show that every subgroup of a locally $\mathcal{P}$ group is itself locally $\mathcal{P}$.

**3.11.** Let $\mathcal{P}$ be a property of groups. Let $G$ be a group. Show that $G$ is locally $\mathcal{P}$ if and only if all its finitely generated subgroups are locally $\mathcal{P}$.

**3.12.** Show that the additive group $\mathbb{Q}$ is locally cyclic.

**3.13.** Show that in a locally finite group every element has finite order.

**3.14.** Let $G$ be an abelian group. Show that $G$ is locally finite if and only if every element of $G$ has finite order.

**3.15.** Show that every subgroup and every quotient of a locally finite group is a locally finite group.

**3.16.** Let $G$ be a group. Suppose that $G$ contains a normal subgroup $H$ such that both $H$ and $G/H$ are locally finite. Show that $G$ is locally finite.

**3.17.** Let $G$ be a group which is the limit of an inductive system of locally finite groups. Show that $G$ is locally finite.

**3.18.** Show that the direct sum of any family of locally finite groups is a locally finite group.

**3.19.** Let $G$ be a group. Let $\mathcal{S}$ denote the set consisting of all normal locally finite subgroups of $G$.

(a) Show that if $H \in \mathcal{S}$ and $K \in \mathcal{S}$ then $HK \in \mathcal{S}$.

(b) Show that $M = \bigcup_{H \in \mathcal{S}} H$ is a normal locally finite subgroup of $G$ and that every normal locally finite subgroup of $G$ is contained in $M$.

**3.20.** Let $G$ be a locally finite group and let $A$ be a set. Show that every bijective cellular automaton $\tau\colon A^G \to A^G$ is invertible.

**3.21.** Let $G$ be a locally finite group and let $A$ be a finite set. Show that every surjective cellular automaton $\tau\colon A^G \to A^G$ is injective.

**3.22.** Let $G$ be a locally finite group and let $A$ be a set. Let $\tau\colon A^G \to A^G$ be a cellular automaton. Show that $\tau(A^G)$ is closed in $A^G$ with respect to the prodiscrete topology. Hint: First treat the case when $G$ is finite, then use Sect. 1.7 and Proposition A.4.3.

**3.23.** Let $G$ be a group and let $A$ be a nonzero abelian group. Suppose that $s_0 \in G$ is an element of infinite order. Show that the map $\tau\colon A^G \to A^G$ defined by $\tau(x)(g) = x(gs_0) - x(g)$, for all $x \in A^G$ and $g \in G$, is a cellular automaton which is surjective but not injective.

**3.24.** Let $X$ be an infinite set. Show that the symmetric group $\mathrm{Sym}(X)$ is not locally residually finite.

**3.25.** Let $G$ be a group. Suppose that there exists a family $(N_i)_{i\in I}$ of normal subgroups of $G$ satisfying the following properties:
(1) for all $i$ and $j$ in $I$, there exists $k$ in $I$ such that $N_k \subset N_i \cap N_j$;
(2) $\bigcap_{i\in I} N_i = \{1_G\}$;
(3) the group $G/N_i$ is surjunctive for each $i \in I$.
Show that $G$ is surjunctive. Hint: Apply Lemma 3.3.4.

**3.26.** It follows from Theorem 3.3.1 that every residually finite group is surjunctive. The goal of this exercise is to present an alternative proof of this result. Let $G$ be a residually finite group and let $A$ be a finite set. Let $\tau\colon A^G \to A^G$ be an injective cellular automaton. Fix a family $(\Gamma_i)_{i\in I}$ of subgroups of finite index of $G$ such that $\bigcap_{i\in I} \Gamma_i = \{1_G\}$ (the existence of such a family follows from the residual finiteness of $G$).
(a) Show that, for each $i \in I$, the set $\mathrm{Fix}(\Gamma_i) = \{x \in A^G : gx = x \text{ for all } g \in \Gamma_i\}$ is finite.
(b) Show that $\tau(\mathrm{Fix}(\Gamma_i)) = \mathrm{Fix}(\Gamma_i)$ for all $i \in I$.
(c) Prove that $\bigcup_{i\in I} \mathrm{Fix}(\Gamma_i)$ is dense in $A^G$ and conclude.

**3.27.** Let $X$ be a set. Let $(A_i)_{i\in I}$ be a net of subsets of $X$.
(a) Show that

$$\bigcup_i \bigcap_{j\geq i} A_j \subset \bigcap_i \bigcup_{j\geq i} A_j.$$

(b) Let $B$ be a subset of $X$. Show that the net $(A_i)_{i\in I}$ converges to $B$ with respect to the prodiscrete topology on $\mathcal{P}(X) = \{0,1\}^X$ if and only if

$$B = \bigcup_i \bigcap_{j\geq i} A_j = \bigcap_i \bigcup_{j\geq i} A_j.$$

**3.28.** Let $G$ be a group and let $A$ be a set. Let $H$ be a subgroup of $G$. In Exercise 1.33, we associated with each subshift $X \subset A^H$ the subshift $X^{(G)} = \{x \in A^G : x_g^H \in X \text{ for all } g \in G\} \subset A^G$, where $x_g^H = (gx)|_H$ for all $x \in A^G$, $h \in H$ and $g \in G$. Let $\sigma\colon A^H \to A^H$ be a cellular automaton and denote by $\sigma^G\colon A^G \to A^G$ the induced cellular automaton. Show that if $X, Y \subset A^H$ are two subshifts such that $\sigma(X) \subset Y$, then the cellular automaton $\sigma^G|_{X^{(G)}}\colon X^{(G)} \to Y^{(G)}$ is injective (resp. surjective) if and only if the cellular automaton $\sigma|_X\colon X \to Y$ is injective (resp. surjective).

**3.29.** Let $G$ be a group and let $A$ be a finite set. Given a subshift $Z \subset A^G$ we denote by $Z_f$ the set of all configurations in $Z$ whose $G$-orbit is finite (cf. Example 1.3.1(c)). We say that $Z$ is *surjunctive* if every injective cellular automaton $\sigma\colon Z \to Z$ is surjective. Let $X \subset A^G$ be a subshift such that $X_f$ is dense in $X$.

(a) Let $\tau\colon A^G \to A^G$ be a cellular automaton. Consider the subshift $Y = \tau(X)$. Show that $Y_f$ is dense in $Y$.

(b) Deduce from (a) that $X$ is surjunctive.

**3.30.** Let $A$ be a finite set and let $X \subset A^{\mathbb{Z}}$ be an irreducible subshift of finite type.

(a) Show that $X_f$ is dense in $X$.

(b) Deduce from (a) and Exercise 3.29(b) that $X$ is surjunctive (compare with its stronger version in Exercise 6.36).

(c) Let $A = \{0, 1\}$ and consider the subshift of finite type $Y \subset A^{\mathbb{Z}}$ defined by $Y = \{x \in A^{\mathbb{Z}} : (x(n), x(n+1)) \neq (0, 1) \text{ for all } n \in \mathbb{Z}\}$. Show that $Y_f$ consists of the two constant configurations $x_0$ and $x_1$ and therefore it is closed but not dense in $Y$.

(d) Show that the subshift $Y$ is surjunctive.

(e) Let $A = \{0, 1\}$ and let $Y$ as in (c). Consider the associated subshift $Z = Y^{(\mathbb{Z}^2)} \subset A^{\mathbb{Z}^2}$ defined in Exercise 1.33. Show that $Z$ is an irreducible subshift of finite type but that $Z_f$ is not dense in $Z$.

(f) Show that the subshift $Z$ is surjunctive.

**3.31.** Show that the golden mean subshift is surjunctive. Hint: Use Exercise 1.39 and Exercise 3.30(b).

**3.32.** Show that the even subshift is surjunctive. Hint: Use Exercise 3.30.

**3.33.** *A subshift of finite type which is not surjunctive* (cf. [Weiss, Sect. 4]). Let $A = \{0, 1, 2\}$ and consider the subshift of finite type $X = X(\Omega, \mathcal{A}) \subset A^{\mathbb{Z}}$ with memory set $\Omega = \{0, 1\} \subset \mathbb{Z}$ and defining set of admissible words $\mathcal{A} = \{00, 01, 11, 12, 22\} \subset A^{\Omega}$.

(a) Show that $X$ is not irreducible.

(b) Consider the cellular automaton $\sigma\colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ with memory set $S = \{0, 1\}$ and local defining map $\mu\colon A^S \to A$ defined by

$$\mu(y) = \begin{cases} 1 & \text{if } y(0)y(1) = 01 \\ y(0) & \text{otherwise.} \end{cases}$$

Show that $\sigma(X) \subset X$.

(c) Show that the cellular automaton $\tau = \sigma|_X \colon X \to X$ is injective but not surjective. Deduce that $X$ is not surjunctive.

**3.34.** Let $G$ be a group and let $A$ be a finite set. Suppose that $G$ contains an element of infinite order and that $A$ contains at least two distinct elements.

(a) Show that there exists a subshift $X \subset A^G$ which is not surjunctive.

(b) Show that if, in addition, $G$ is not infinite cyclic then there exists an irreducible subshift $X \subset A^G$ which is not surjunctive. Hint: Use Exercise 3.33 and Exercise 1.33(c).

**3.35.** Given a group $\Gamma$ acting continuously on a topological space $Z$, a subset $M \subset Z$ is called a *minimal set* if $M$ is a nonempty $\Gamma$-invariant subset of $Z$ and the $\Gamma$-orbit of every point $m \in M$ is dense in $M$.

Let $G$ be a group and let $A$ be a set. A subshift $X \subset A^G$ is called *minimal* if $X$ is a minimal set with respect to the $G$-shift action on $A^G$.

(a) Show that a subshift $X \subset A^G$ is minimal if and only if $X \neq \varnothing$ and there is no subshift $Y$ in $A^G$ such that $\varnothing \neq Y \subsetneqq X$.

(b) Show that every minimal subshift $X \subset A^G$ is irreducible.

(c) Show that the finite minimal subshifts in $A^G$ are precisely the finite $G$-orbits.

(d) Show that if $X \subset A^G$ is an infinite minimal subshift then the $G$-orbit of every configuration $x \in X$ is infinite.

(e) Show that if $X$ and $Y$ are minimal subshifts in $A^G$ then either $X = Y$ or $X \cap Y = \varnothing$.

**3.36.** Let $G$ be a group and let $A$ be a finite set. Let $X, Y \subset A^G$ be two subshifts. Show that if $X$ is nonempty and $Y$ is minimal then every cellular automaton $\tau \colon X \to Y$ is surjective.

**3.37.** Let $G$ be a group and let $A$ be a finite set. Show that every minimal subshift $X \subset A^G$ is surjunctive.

**3.38.** Let $G$ be a group and let $A$ be a finite set. Show that every nonempty subshift $X \subset A^G$ contains a minimal subshift. Hint: Apply Zorn's lemma to the set of nonempty subshifts contained in $X$.

**3.39.** Let $G$ be a group and let $A$ be a set. A subset $R \subset G$ is called *syndetic* if there exists a finite subset $K \subset G$ such that the set $Kg$ meets $R$ for every $g \in G$. A configuration $x \in A^G$ is called *almost periodic* if for every finite subset $\Omega \subset G$ the set $R(x, \Omega) = \{g \in G : (gx)|_\Omega = x|_\Omega\}$ is syndetic in $G$.

(a) Show that if $x \in A^G$ is an almost periodic configuration then its orbit closure $X = \overline{Gx} \subset A^G$ is a minimal subshift. Hint: Suppose that the orbit closure $X = \overline{Gx}$ of a configuration $x \in A^G$ is not minimal. Then there exists

a nonempty subshift $Y \subset X$ with $x \notin Y$. This implies that there is a finite subset $\Omega \subset G$ such that $x|_\Omega \neq y|_\Omega$ for all $y \in Y$. Let $K$ be a finite subset of $G$ and consider the set $\Omega' = K^{-1}\Omega$. Choose an arbitrary configuration $y_0 \in Y$. Since $y_0 \in X$, there is an element $g_0 \in G$ such that $(g_0 x)|_{\Omega'} = y_0|_{\Omega'}$. This implies that $(kg_0 x)|_\Omega = (ky_0)|_\Omega \neq x|_\Omega$ for all $k \in K$. Thus $Kg_0$ does not meet $R(x, \Omega)$ so that $x$ is not almost-periodic.

(b) Suppose that the set $A$ is finite. Show that if $X \subset A^G$ is a minimal subshift, then every configuration $x \in X$ is almost-periodic. Hint: Observe that if $\Omega \subset G$ is a finite subset and $x \in X$, then the open sets $gV$, where $V = \{y \in A^G : y|_\Omega = x|_\Omega\}$ and $g$ runs over $G$, cover $X$ by minimality and use the compactness of $X$.

(c) Suppose that the set $A$ is finite. Show that a nonempty subshift $X \subset A^G$ is minimal if and only if $X$ is irreducible and every configuration $x \in X$ is almost-periodic.

**3.40.** Let $A$ be a set. The *language* of a configuration $x \in A^{\mathbb{Z}}$ is the subset $L(x) \subset A^*$ consisting of all words $w \in A^*$ which can be written in the form $w = x(n+1)x(n+2)\dots x(n+m)$ for some integers $n \in \mathbb{Z}$ and $m \geq 0$.

(a) Let $X \subset A^{\mathbb{Z}}$ be a subshift and let $x \in X$. Show that $L(x) \subset L(X)$ and that one has $L(x) = L(X)$ if and only if the $\mathbb{Z}$-orbit of $x$ is dense in $X$.

(b) Show that a nonempty subshift $X \subset A^{\mathbb{Z}}$ is minimal if and only if one has $L(x) = L(X)$ for all $x \in X$.

(c) Show that a configuration $x \in A^{\mathbb{Z}}$ is almost periodic if and only if it satisfies the following condition: for every word $u \in L(x)$, there exists an integer $n = n(u) \geq 0$ such that $u$ is a subword of any word $v \in L(x)$ of length $n$.

**3.41.** *The Thue-Morse sequence.* Let $A = \{0, 1\}$. The *Thue-Morse sequence* is the map $x\colon \mathbb{N} \to A$ defined by

$$x(n) = \begin{cases} 1 & \text{if the number of ones in the binary expansion of } n \text{ is odd} \\ 0 & \text{otherwise.} \end{cases}$$

(a) Check that

$$x(0)x(1)\dots x(16) = 0110100110010110.$$

(b) Show that one has $x(k + 3m2^n) = x(k)$ for all $k, m, n \in \mathbb{N}$ such that $0 \leq k \leq 2^n$.

(c) Show that $x$ satisfies the recurrence relations

$$\begin{cases} x(0) = 0 \\ x(2n) = x(n) \\ x(2n+1) = 1 - x(n) \end{cases}$$

and that these relations uniquely determine $x$.

(d) Consider the unique monoid homomorphism $\varphi\colon A^* \to A^*$ which satisfies

$$\varphi(0) = 01 \quad \text{and} \quad \varphi(1) = 10.$$

Show that

$$\varphi(x(0)x(1)\dots x(2^n-1)) = x(0)x(1)\dots x(2^{n+1}-1)$$

for all $n \in \mathbb{N}$.

(e) Consider the unique monoid isomorphism $\iota\colon A^* \to A^*$ which satisfies

$$\iota(0) = 1 \quad \text{and} \quad \iota(1) = 0.$$

Show that

$$x(0)x(1)\dots x(2^{n+1}-1) = x(0)x(1)\dots x(2^n-1)\iota(x(0)x(1)\dots x(2^n-1))$$

for all $n \in \mathbb{N}$.

Note: The Thue-Morse sequence was first studied by E. Prouhet in a 1851 paper dealing with problems in number theory. It is an example of an *automatic sequence* (see [AlS, Sect. 5.1]).

**3.42.** *The Morse subshift.* Let $A = \{0,1\}$ and let $L = \{(x(n), x(n+1), \dots, x(n+m)) : n, m \in \mathbb{N}\} \subset A^*$, where $x\colon \mathbb{N} \to A$ is the Thue-Morse sequence.

(a) Show that the set $L$ satisfies the conditions (i) and (ii) in Exercise 1.36(b).

(b) Let $X \subset A^{\mathbb{Z}}$ denote the unique subshift whose language is $L(X) = L$ (cf. Exercise 1.36(c)). Show that $X$ is an infinite minimal subshift (it is called the *Morse subshift*). Hint: Use Exercises 3.39(a), 3.40 and 3.41(b).

**3.43.** *Toeplitz subshifts.* Let $A$ be a finite set. A configuration $x \in A^{\mathbb{Z}}$ is said to be a *Toeplitz configuration* if for every $n \in \mathbb{Z}$ there exists $k \geq 1$ such that $x(n) = x(n+kr)$ for all $r \in \mathbb{Z}$. In other words, $x$ is a Toeplitz configuration if there exists a partition of $\mathbb{Z}$ into arithmetic progressions such that $x$ is constant on each element of the partition. Given a Toeplitz configuration $x \in A^{\mathbb{Z}}$ and $n \in \mathbb{Z}$ we set

$$k(x,n) = \min\{k \geq 1 : x(n) = x(n+kr) \text{ for all } r \in \mathbb{Z}\}.$$

One says that a subshift $X \subset A^{\mathbb{Z}}$ is a *Toeplitz subshift* if it is the orbit closure of some Toeplitz configuration $x \in A^{\mathbb{Z}}$.

(a) Let $x \in A^{\mathbb{Z}}$ be a Toeplitz configuration. Show that the $\mathbb{Z}$-orbit of $x$ is finite if and only if the set $\{k(x,n) : n \in \mathbb{Z}\}$ is finite.

(b) Show that every Toeplitz configuration $x \in A^{\mathbb{Z}}$ is almost-periodic. Hint: Use Exercise 3.40.

(c) Show that every Toeplitz subshift $X \subset A^{\mathbb{Z}}$ is minimal.

**3.44.** Let $A$ be a finite set and let $X \subset A^{\mathbb{Z}}$ be a Toeplitz subshift.
(a) Show that $X$ is irreducible. Hint: Use Exercises 3.43(d) and 3.35.
(b) Show that $X$ is surjunctive. Hint: Use Exercises 3.43(d) and 3.37.

# Chapter 4
# Amenable Groups

This chapter is devoted to the class of amenable groups. This is a class of groups which plays an important role in many areas of mathematics such as ergodic theory, harmonic analysis, representation theory, dynamical systems, geometric group theory, probability theory and statistics. As residually finite groups, amenable groups generalize finite groups but there are residually finite groups which are not amenable and there are amenable groups which are not residually finite. An amenable group is a group whose subsets admit an invariant finitely additive probability measure (see Sect. 4.4). The class of amenable groups contains in particular all finite groups, all abelian groups and, more generally, all solvable groups (Theorem 4.6.3). It is closed under the operations of taking subgroups, taking quotients, taking extensions, and taking inductive limits (Sect. 4.5). The notion of a Følner net, roughly speaking, a net of almost invariant finite subsets of a group, is introduced in Sect. 4.7. Paradoxical decompositions are defined in Sect. 4.8. The Følner-Tarski theorem (Theorem 4.9.1) asserts that amenability, existence of a Følner net, and non-existence of a paradoxical decomposition are three equivalent conditions for groups. Another characterization of amenable groups is given in Sect. 4.10: a group is amenable if and only if every continuous affine action of the group on a nonempty compact convex subset of a Hausdorff topological vector space admits a fixed point (Corollary 4.10.2).

## 4.1 Measures and Means

Let $E$ be a set. We shall denote by $\mathcal{P}(E)$ the set of all subsets of $E$.

**Definition 4.1.1.** A map $\mu\colon \mathcal{P}(E) \to [0,1]$ is called a *finitely additive probability measure* on $E$ if it satisfies the following properties:

(1) $\mu(E) = 1$,
(2) $\mu(A \cup B) = \mu(A) + \mu(B)$ for all $A, B \in \mathcal{P}(E)$ such that $A \cap B = \varnothing$.

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_4, 

*Example 4.1.2.* Let $F$ be a nonempty finite subset of $E$. Then the map $\mu_F\colon \mathcal{P}(E) \to [0,1]$ defined by

$$\mu_F(A) = \frac{|A \cap F|}{|F|}$$

for all $A \subset E$, is a finitely additive probability measure on $E$.

**Proposition 4.1.3.** *Let $\mu\colon \mathcal{P}(E) \to [0,1]$ be a finitely additive probability measure on $E$. Then one has:*

(i) $\mu(\varnothing) = 0$;
(ii) $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$;
(iii) $\mu(A \cup B) \leq \mu(A) + \mu(B)$;
(iv) $A \subset B \Rightarrow \mu(B \setminus A) = \mu(B) - \mu(A)$;
(v) $A \subset B \Rightarrow \mu(A) \leq \mu(B)$;

*for all $A, B \in \mathcal{P}(E)$.*

*Proof.* We have $\mu(E) = \mu(E \cup \varnothing) = \mu(E) + \mu(\varnothing)$, which implies (i). For all $A, B \in \mathcal{P}(E)$, we have $\mu(A \cup B) = \mu(A) + \mu(B \setminus A)$ and $\mu(B) = \mu(A \cap B) + \mu(B \setminus A)$, which gives (ii). Since $\mu(A \cap B) \geq 0$, equality (ii) implies (iii). If $A \subset B$, we have $\mu(B) = \mu(B \setminus A) + \mu(A)$, which implies (iv). Finally, inequality (v) follows from (iv) since $\mu(B \setminus A) \geq 0$. □

Consider now the real vector space $\ell^\infty(E)$ consisting of all bounded functions $x\colon E \to \mathbb{R}$. Recall that $\ell^\infty(E)$ is a Banach space for the norm $\|\cdot\|_\infty$ defined by

$$\|x\|_\infty = \sup_{a \in E} |x(a)|.$$

We equip $\ell^\infty(E)$ with the partial ordering $\leq$ given by

$$x \leq y \iff (x(a) \leq y(a) \text{ for all } a \in E).$$

For each $\lambda \in \mathbb{R}$, we shall also denote by $\lambda$ the element of $\ell^\infty(E)$ which is identically equal to $\lambda$ on $E$.

**Definition 4.1.4.** A *mean* on $E$ is a linear map $m\colon \ell^\infty(E) \to \mathbb{R}$ such that:

(1) $m(1) = 1$,
(2) $x \geq 0 \Rightarrow m(x) \geq 0$

for all $x \in \ell^\infty(E)$.

*Example 4.1.5.* Let $S$ be a countable (finite or infinite) subset of $E$ and let $f\colon S \to \mathbb{R}$ such that:

(C1) $f(s) > 0$ for all $s \in S$;
(C2) $\sum_{s \in S} f(s) = 1$.

Then the map $m_f\colon \ell^\infty(E) \to \mathbb{R}$ defined by

$$m_f(x) = \sum_{s\in S} f(s)x(s)$$

is a mean on $E$.

One says that a mean $m$ on $E$ has *finite* (resp. *countable*) *support* if there exists a finite (resp. countable) subset $S \subset E$ and a map $f\colon S \to \mathbb{R}$ satisfying conditions (C1) and (C2) above such that $m = m_f$.

**Proposition 4.1.6.** *Let $m\colon \ell^\infty(E) \to \mathbb{R}$ be a mean on $E$. Then one has:*

(i) $m(\lambda) = \lambda$,
(ii) $x \le y \Rightarrow m(x) \le m(y)$,
(iii) $\inf_E x \le m(x) \le \sup_E x$,
(iv) $|m(x)| \le \|x\|_\infty$,

*for all $\lambda \in \mathbb{R}$ and $x, y \in \ell^\infty(E)$.*

*Proof.* (i) By linearity, we have $m(\lambda) = \lambda m(1) = \lambda$.

(ii) If $x \le y$ then $m(y) - m(x) = m(y - x) \ge 0$ and thus $m(x) \le m(y)$.

(iii) We have $\inf_E x \le x \le \sup_E x$ and therefore $\inf_E x \le m(x) \le \sup_E x$ by applying (i) and (ii).

(iv) From (iii) we get $-\|x\|_\infty \le m(x) \le \|x\|_\infty$, that is, $|m(x)| \le \|x\|_\infty$. □

Consider the topological dual of $\ell^\infty(E)$, that is, the vector space $(\ell^\infty(E))^*$ consisting of all continuous linear maps $u\colon \ell^\infty(E) \to \mathbb{R}$.

Recall that $(\ell^\infty(E))^*$ is a Banach space for the operator norm $\|\cdot\|$ defined by

$$\|u\| = \sup_{\|x\|_\infty \le 1} |u(x)| \tag{4.1}$$

for all $u \in (\ell^\infty(E))^*$.

**Proposition 4.1.7.** *Let $m\colon \ell^\infty(E) \to \mathbb{R}$ be a mean on $E$. Then $m \in (\ell^\infty(E))^*$ and $\|m\| = 1$.*

*Proof.* By definition $m$ is linear. Inequality (iv) in Proposition 4.1.6 shows that $m$ is continuous and satisfies $\|m\| \le 1$. Since $m(1) = 1$, we have $\|m\| = 1$. □

Let $\mathcal{PM}(E)$ (resp. $\mathcal{M}(E)$) denote the set of all finitely additive probability measures (resp. of all means) on $E$. We are going to show that there is a natural bijection between the sets $\mathcal{PM}(E)$ and $\mathcal{M}(E)$. This natural bijection, which is analogous to the Riesz representation in measure theory, may be used to view finitely additive measures on $E$ as points in the dual space $(\ell^\infty(E))^*$ where classical techniques of functional analysis may be applied.

For each subset $A \subset E$, we denote by $\chi_A$ the *characteristic map* of $A$, that is, the map $\chi_A\colon E \to \mathbb{R}$ defined by $\chi_A(x) = 1$ if $x \in A$ and $\chi_A(x) = 0$ otherwise.

Let $m \in \mathcal{M}(E)$. Consider the map $\widehat{m}\colon \mathcal{P}(E) \to \mathbb{R}$ given by

$$\widehat{m}(A) = m(\chi_A).$$

Observe that $0 \leq \chi_A \leq 1$ and therefore $\widehat{m}(A) \in [0,1]$ by Proposition 4.1.6(ii). We have $\widehat{m}(E) = m(\chi_E) = m(1) = 1$. On the other hand, if $A$ and $B$ are disjoint subsets of $E$, we have $\chi_{A\cup B} = \chi_A + \chi_B$ and thus $\widehat{m}(A \cup B) = \widehat{m}(A) + \widehat{m}(B)$ by linearity of $m$. It follows that $\widehat{m} \in \mathcal{PM}(E)$.

**Theorem 4.1.8.** *The map $\Phi\colon \mathcal{M}(E) \to \mathcal{PM}(E)$ defined by $\Phi(m) = \widehat{m}$ is bijective.*

The proof will be divided in several steps.

Denote by $\mathcal{E}(E)$ the set of all maps $x\colon E \to \mathbb{R}$ which take only finitely many values. Observe that $\mathcal{E}(E)$ is the vector subspace of $\ell^\infty(E)$ spanned by the set of characteristic maps $\{\chi_A : A \subset E\}$.

**Lemma 4.1.9.** *The vector subspace $\mathcal{E}(E)$ is dense in $\ell^\infty(E)$.*

*Proof.* Let $x \in \ell^\infty(E)$ and $\varepsilon > 0$. Set $\alpha = \inf_E x$ and $\beta = \sup_E x$. Choose an integer $n \geq 1$ such that $(\beta - \alpha)/n < \varepsilon$ and set $\lambda_i = \alpha + i(\beta - \alpha)/n$ for $i = 1, 2, \ldots, n$. Consider the map $y\colon E \to \mathbb{R}$ defined by

$$y(a) = \min\{\lambda_i : x(a) \leq \lambda_i\}$$

for all $a \in E$. The map $y$ take its values in the set $\{\lambda_1, \ldots, \lambda_n\}$. Thus $y \in \mathcal{E}(E)$. We have $\|x - y\|_\infty \leq (\beta - \alpha)/n < \varepsilon$ by construction. Consequently, $\mathcal{E}(E)$ is dense in $\ell^\infty(E)$. □

Let $\mu \in \mathcal{PM}(E)$. Let us set, for all $x \in \mathcal{E}(E)$,

$$\overline{\mu}(x) = \sum_{\lambda\in\mathbb{R}} \mu(x^{-1}(\lambda))\lambda. \tag{4.2}$$

Observe that there is only a finite number of nonzero terms in the right-hand side of (4.2) since $x \in \mathcal{E}(E)$ and $\mu(\varnothing) = 0$.

**Lemma 4.1.10.** *Let $x \in \mathcal{E}(E)$. Suppose that there is a finite partition $(A_i)_{i\in I}$ of $E$ such that the restriction of $x$ to $A_i$ is constant equal to $\alpha_i$ for each $i \in I$. Then one has*

$$\overline{\mu}(x) = \sum_{i\in I} \mu(A_i)\alpha_i.$$

*Proof.* Since $\mu$ is finitely additive, we have

$$\sum_{i\in I} \mu(A_i)\alpha_i = \sum_{\lambda\in\mathbb{R}} \left( \sum_{\alpha_i=\lambda} \mu(A_i) \right) \lambda = \sum_{\lambda\in\mathbb{R}} \mu(x^{-1}(\lambda))\lambda = \overline{\mu}(x).$$

□

**Lemma 4.1.11.** *The map $\overline{\mu}\colon \mathcal{E}(E) \to \mathbb{R}$ is linear and continuous.*

*Proof.* Let $x, y \in \mathcal{E}(E)$ and $\xi, \eta \in \mathbb{R}$. Denote by $V$ (resp. $W$) the set of values taken by $x$ (resp. $y$). The subsets $x^{-1}(\alpha) \cap y^{-1}(\beta)$, $(\alpha, \beta) \in V \times W$, form a finite partition of $E$. By applying Lemma 4.1.10, we get

$$\begin{aligned}\overline{\mu}(\xi x + \eta y) &= \sum_{(\alpha,\beta)\in V\times W} \mu(x^{-1}(\alpha) \cap y^{-1}(\beta))(\xi\alpha + \eta\beta) \\ &= \xi \left( \sum_{(\alpha,\beta)\in V\times W} \mu(x^{-1}(\alpha) \cap y^{-1}(\beta))\alpha \right) \\ &\quad + \left( \eta \sum_{(\alpha,\beta)\in V\times W} \mu(x^{-1}(\alpha) \cap y^{-1}(\beta))\beta \right) \\ &= \xi\overline{\mu}(x) + \eta\overline{\mu}(y).\end{aligned}$$

Consequently, $\overline{\mu}$ is linear.

Formula (4.2) implies $\inf_E x \le \overline{\mu}(x) \le \sup_E x$ since

$$\sum_{\lambda\in\mathbb{R}} \mu(x^{-1}(\lambda)) = \mu(E) = 1.$$

It follows that $|\overline{\mu}(x)| \le \|x\|_\infty$ for all $x \in \mathcal{E}(E)$. This shows that the linear map $\overline{\mu}$ is continuous. □

**Lemma 4.1.12.** *Let $X$ be a normed vector space and let $Y$ be a dense vector subspace of $X$. Suppose that $\varphi\colon Y \to \mathbb{R}$ is a continuous linear map. Then there exists a continuous linear map $\widetilde{\varphi}\colon X \to \mathbb{R}$ extending $\varphi$ (that is, such that $\widetilde{\varphi}|_Y = \varphi$).*

*Proof.* Let $x \in X$. Since $Y$ is dense in $X$, we can find a sequence $(y_n)_{n\ge0}$ of points of $Y$ which converges to $x$. The sequence $(\varphi(y_n))$ is a Cauchy sequence since

$$|\varphi(y_p) - \varphi(y_q)| = |\varphi(y_p - y_q)| \le \|\varphi\|\|y_p - y_q\|$$

for all $p, q \ge 0$. As $\mathbb{R}$ is complete, this sequence converges. Set $\widetilde{\varphi}(x) = \lim\varphi(y_n)$. If $(y'_n)$ is another sequence of points in $Y$ converging to $x$, then $|\varphi(y_n) - \varphi(y'_n)| \le \|\varphi\|\|y_n - y'_n\|$. Thus we have $\lim\varphi(y_n) = \lim\varphi(y'_n)$. This shows that $\widetilde{\varphi}(x)$ does not depend of the choice of the sequence $(y_n)$. The linearity of $\varphi$ gives the linearity of $\widetilde{\varphi}$ by taking limits. We also get $\|\widetilde{\varphi}(x)\| \le \|\varphi\|\|x\|$ for all $x \in X$ by taking limits. This shows that the linear map $\widetilde{\varphi}$ is continuous. If $x \in Y$, we have $\widetilde{\varphi}(x) = \varphi(x)$ since we can take as $(y_n)$ the constant sequence $y_n = x$ in this case. Thus $\widetilde{\varphi}$ extends $\varphi$. □

*Proof of Theorem 4.1.8.* Let $\mu \in \mathcal{PM}(E)$. By Lemma 4.1.9, Lemma 4.1.11 and Lemma 4.1.12, the map $\overline{\mu}\colon \mathcal{E}(E) \to \mathbb{R}$ can be extended to a continuous linear map $\widetilde{\mu}\colon \ell^\infty(E) \to \mathbb{R}$. We have

$$\widetilde{\mu}(1) = \overline{\mu}(1) = \mu(E) = 1.$$

If $y \in \mathcal{E}(E)$ and $y \geq 0$, then $\widetilde{\mu}(y) = \overline{\mu}(y) \geq 0$ by (4.2). Consider now an element $x \in \ell^\infty(E)$ such that $x \geq 0$. Let $(y_n)_{n\geq 0}$ be a sequence of elements of $\mathcal{E}(E)$ converging to $x$. Let us set $z_n = |y_n|$. Since $z_n \in \mathcal{E}(E)$ and $z_n \geq 0$, we have $\widetilde{\mu}(z_n) \geq 0$. On the other hand, the sequence $(z_n)$ converges to $x$ since, by the triangle inequality, $\|x - z_n\|_\infty \leq \|x - y_n\|_\infty$. It follows that

$$\widetilde{\mu}(x) = \lim_{n\to\infty} \overline{\mu}(z_n) \geq 0.$$

Thus $\widetilde{\mu} \in \mathcal{M}(E)$.

For every subset $A \subset E$, we have

$$\widetilde{\mu}(\chi_A) = \overline{\mu}(\chi_A) = \mu(A).$$

Therefore $\Phi(\widetilde{\mu}) = \mu$. This proves that $\Phi$ is surjective.

It remains to show that $\Phi$ is injective. To see this, consider $m_1, m_2 \in \mathcal{M}(E)$ such that $\Phi(m_1) = \Phi(m_2)$. This means that $m_1(\chi_A) = m_2(\chi_A)$ for every subset $A \subset E$. By linearity, this implies that $m_1$ and $m_2$ coincide on $\mathcal{E}(E)$. Since $\mathcal{E}(E)$ is dense in $\ell^\infty(E)$ by Lemma 4.1.9, we deduce that $m_1 = m_2$ by continuity of $m_1$ and $m_2$. This shows that $\Phi$ is injective. □

## 4.2 Properties of the Set of Means

Let $E$ be a set.

The topology on $(\ell^\infty(E))^*$ associated with the operator norm $\|\cdot\|$ defined in (4.1) is called the *strong topology* on $(\ell^\infty(E))^*$. Another topology that is commonly used is the *weak-∗ topology* on $(\ell^\infty(E))^*$. Recall that the weak-∗ topology on $(\ell^\infty(E))^*$ is by definition the smallest topology for which the evaluation map

$$\begin{aligned} \psi_x \colon (\ell^\infty(E))^* &\to \mathbb{R} \\ u &\mapsto u(x) \end{aligned}$$

is continuous for each $x \in \ell^\infty(E)$ (see Sect. F.2).

By Proposition 4.1.7, the set $\mathcal{M}(E)$ is contained in the unit sphere

$$\{u : \|u\| = 1\} \subset (\ell^\infty(E))^*.$$

The following result will play an important role in the sequel.

**Theorem 4.2.1.** *The set $\mathcal{M}(E)$ is a convex compact subset of $(\ell^\infty(E))^*$ with respect to the weak-∗ topology.*

*Proof.* Let $m_1, m_2 \in \mathcal{M}(E)$ and $t \in [0,1]$. Then $(tm_1 + (1-t)m_2)(1) = tm_1(1) + (1-t)m_2(1) = t + (1-t) = 1$ and $(tm_1 + (1-t)m_2)(x) = tm_1(x) + (1-t)m_2(x) \geq 0$ for all $x \in \ell^\infty(E)$ such that $x \geq 0$. Thus $tm_1 + (1-t)m_2 \in \mathcal{M}(E)$. This shows that $\mathcal{M}(E)$ is convex. Equip now $(\ell^\infty(E))^*$ with the weak-$*$ topology and suppose that $(m_i)_{i\in I}$ is a net in $\mathcal{M}(E)$ converging to $u \in (\ell^\infty(E))^*$. Then, for every $i \in I$, we have $\psi_1(m_i) = m_i(1) = 1$ and $\psi_x(m_i) = m_i(x) \geq 0$ for all $x \in \ell^\infty(E)$ such that $x \geq 0$. By taking limits, we get $u(1) = \psi_1(u) = 1$ and $u(x) = \psi_x(u) \geq 0$ for all $x \in \ell^\infty(E)$ such that $x \geq 0$. Thus $u \in \mathcal{M}(E)$. This shows that $\mathcal{M}(E)$ is closed in $(\ell^\infty(E))^*$ (Proposition A.2.1). As $\mathcal{M}(E)$ is contained in the unit ball of $(\ell^\infty(E))^*$, which is compact for the weak-$*$ topology by the Banach-Alaoglu Theorem (Theorem F.3.1), it follows that $\mathcal{M}(E)$ is compact. □

## 4.3 Measures and Means on Groups

In this section, we shall see that the set of finitely additive measures and the set of means carry additional structure when the underlying set is a group. Indeed, in this case, the group naturally acts on both sets. Moreover, there are involutions coming from the operation of taking inverses in the group.

More precisely, let $G$ be a group.

The group $G$ naturally acts on the left and on the right on each of the sets $\mathcal{PM}(G)$ and $\mathcal{M}(G)$ in the following way.

Firstly, for $\mu \in \mathcal{PM}(G)$ and $g \in G$, we define the maps $g\mu\colon \mathcal{P}(G) \to [0,1]$ and $\mu g\colon \mathcal{P}(G) \to [0,1]$ by

$$g\mu(A) = \mu(g^{-1}A) \quad \text{and} \quad \mu g(A) = \mu(Ag^{-1})$$

for all $A \in \mathcal{P}(G)$. One clearly has $g\mu \in \mathcal{PM}(G)$ and $\mu g \in \mathcal{PM}(G)$. Moreover, it is straightforward to check that the map $(g,\mu) \mapsto g\mu$ (resp. $(\mu, g) \mapsto \mu g$) defines a left (resp. right) action of $G$ on $\mathcal{PM}(G)$. Note that these two actions commute in the sense that $g(\mu h) = (g\mu)h$ for all $g, h \in G$ and $\mu \in \mathcal{PM}(G)$.

On the other hand, recall that we introduced in Sect. 1.1 a left action of $G$ on $\mathbb{R}^G$ (the $G$-shift) by defining, for all $g \in G$ and $x \in \mathbb{R}^G$, the element $gx \in \mathbb{R}^G$ by

$$gx(g') = x(g^{-1}g') \quad \text{for all } g' \in G.$$

Similarly, we make $G$ act on the right on $\mathbb{R}^G$ by defining the element $xg \in \mathbb{R}^G$ by

$$xg(g') = x(g'g^{-1}) \quad \text{for all } g' \in G.$$

These two actions of $G$ on $\mathbb{R}^G$ are linear and commute. Moreover, the vector subspace $\ell^\infty(G) \subset \mathbb{R}^G$ is left invariant by both actions. Observe that $\|gx\|_\infty = \|xg\|_\infty = \|x\|_\infty$ for all $g \in G$ and $x \in \ell^\infty(G)$. Thus the left (resp. right) action of $G$ on $\ell^\infty(G)$ is isometric (and therefore continuous).

By duality, we also get a left and a right action of $G$ on $(\ell^\infty(G))^*$. More precisely, for $g \in G$ and $u \in (\ell^\infty(G))^*$, we define the elements $gu$ and $ug$ by

$$gu(x) = u(g^{-1}x) \quad \text{and} \quad ug(x) = u(xg^{-1}),$$

for all $x \in \ell^\infty(G)$. We have $gu \in (\ell^\infty(G))^*$ and $ug \in (\ell^\infty(G))^*$ since the left and right actions of $G$ on $\ell^\infty(G)$ are linear and continuous.

Note that the set $\mathcal{M}(G)$ is invariant under both actions of $G$ on $(\ell^\infty(G))^*$.

**Proposition 4.3.1.** *The left (resp. right) action of $G$ on $\mathcal{M}(G)$ is affine and continuous with respect to the weak-$*$ topology on $\mathcal{M}(G)$.*

*Proof.* It is clear that the left (resp. right) action of $G$ on $(\ell^\infty(G))^*$ is linear. On the other hand, these actions are continuous if we equip $(\ell^\infty(G))^*$ with the weak-$*$ topology. Indeed, if we fix $g \in G$, the map $u \mapsto gu$ is continuous on $(\ell^\infty(G))^*$ since, for each $x \in \ell^\infty(G)$, the map $u \mapsto gu(x)$ is the evaluation map at $g^{-1}x$, which is continuous by definition of the weak-$*$ topology. Similarly, the map $u \mapsto ug$ is continuous. Since $\mathcal{M}(G)$ is a convex subset of $(\ell^\infty(G))^*$, we deduce that the restrictions of both actions to $\mathcal{M}(G)$ are affine and continuous. □

Given $\mu \in \mathcal{PM}(G)$, we define the map $\mu^* \colon \mathcal{P}(G) \to [0,1]$ by

$$\mu^*(A) = \mu(A^{-1}) \text{ for all } A \in \mathcal{P}(G).$$

It is clear that $\mu^* \in \mathcal{PM}(G)$ and that the map $\mu \mapsto \mu^*$ is an involution of $\mathcal{PM}(G)$.

For $x \in \ell^\infty(G)$, define $x^* \in \ell^\infty(G)$ by

$$x^*(g) = x(g^{-1}) \text{ for all } g \in G.$$

The map $x \mapsto x^*$ is an isometric involution of $\ell^\infty(G)$. By duality, it gives an isometric involution $u \mapsto u^*$ of $(\ell^\infty(G))^*$ defined by

$$u^*(x) = u(x^*) \quad \text{for all } x \in \ell^\infty(G).$$

Note that $m^* \in \mathcal{M}(G)$ for all $m \in \mathcal{M}(G)$.

**Proposition 4.3.2.** *Let $g \in G$, $x \in \ell^\infty(G)$, $\mu \in \mathcal{PM}(G)$ and $u \in (\ell^\infty(G))^*$. Then one has*

(i) $(gx)^* = x^* g^{-1}$;
(ii) $(xg)^* = g^{-1}x^*$;
(iii) $(g\mu)^* = \mu^* g^{-1}$;
(iv) $(\mu g)^* = g^{-1}\mu^*$;
(v) $(gu)^* = u^* g^{-1}$;
(vi) $(ug)^* = g^{-1}u^*$.

*Proof.* For every $h \in G$, we have

$$(gx)^*(h) = gx(h^{-1}) = x(g^{-1}h^{-1}) = x^*(hg) = x^*g^{-1}(h),$$

which gives $(gx)^* = x^*g^{-1}$. The proofs of the other properties are similar. □

**Proposition 4.3.3.** *Let $g \in G$ and $m \in \mathcal{M}(G)$. Then one has:*

(i) $\widehat{gm} = g\widehat{m}$;
(ii) $\widehat{mg} = \widehat{m}g$;
(iii) $\widehat{m^*} = \widehat{m}^*$.

We need the following result.

**Lemma 4.3.4.** *Let $g \in G$ and $A \subset G$. Then one has:*

(i) $(\chi_A)^* = \chi_{A^{-1}}$;
(ii) $g\chi_A = \chi_{gA}$.

*Proof.* (i) For $h \in G$ one has $(\chi_A)^*(h) = \chi_A(h^{-1}) = 1 \Leftrightarrow h^{-1} \in A \Leftrightarrow h \in A^{-1} \Leftrightarrow \chi_{A^{-1}}(h) = 1$. This shows that $(\chi_A)^* = \chi_{A^{-1}}$. (ii) For $h \in G$ one has $(g\chi_A)(h) = \chi_A(g^{-1}h) = 1 \Leftrightarrow g^{-1}h \in A \Leftrightarrow h \in gA \Leftrightarrow \chi_{gA}(h) = 1$. This shows that $g\chi_A = \chi_{gA}$. □

*Proof of Proposition 4.3.3.* For every $A \in \mathcal{P}(G)$, we have, using Lemma 4.3.4(ii),

$$\widehat{gm}(A) = gm(\chi_A) = m(g^{-1}\chi_A) = m(\chi_{g^{-1}A}) = \widehat{m}(g^{-1}A) = g\widehat{m}(A).$$

Thus $\widehat{gm} = g\widehat{m}$. Similarly, we get $\widehat{mg} = \widehat{m}g$.

On the other hand, for every $A \in \mathcal{P}(G)$, using Lemma 4.3.4(i) one obtains

$$\widehat{m}^*(A) = \widehat{m}(A^{-1}) = m(\chi_{A^{-1}}) = m((\chi_A)^*) = m^*(\chi_A) = \widehat{m^*}(A)$$

and this shows that $\widehat{m}^* = \widehat{m^*}$. □

*Remark 4.3.5.* Properties (i) and (ii) in Proposition 4.3.3 say that the bijective map $\Phi\colon \mathcal{M}(G) \to \mathcal{PM}(G)$ given by $m \mapsto \widehat{m}$ (see Theorem 4.1.8) is bi-equivariant.

## 4.4 Definition of Amenability

Let $G$ be a group.

A finitely additive probability measure $\mu \in \mathcal{PM}(G)$ is called *left-invariant* (resp. *right-invariant*) if $\mu$ is fixed under the left (resp. right) action of $G$ on $\mathcal{PM}(G)$, that is, if it satisfies $g\mu = \mu$ (resp. $\mu g = \mu$) for all $g \in G$. One says that $\mu$ is *bi-invariant* if $\mu$ is both left and right-invariant.

Similarly, a mean $m \in \mathcal{M}(G)$ is called *left-invariant* (resp. *right-invariant*) if $m$ is fixed under the left (resp. right) action of $G$ on $\mathcal{M}(G)$. One says that $m$ is *bi-invariant* if $m$ is both left and right invariant.

**Proposition 4.4.1.** *Let $\mu \in \mathcal{PM}(G)$. Then $\mu$ is left-invariant (resp. right-invariant) if and only if $\mu^*$ is right-invariant (resp. left-invariant).*

*Proof.* This immediately follows from Proposition 4.3.2(ii). □

**Proposition 4.4.2.** *Let $m \in \mathcal{M}(G)$. Then $m$ is left-invariant (resp. right-invariant) if and only if $m^*$ is right-invariant (resp. left-invariant).*

*Proof.* This immediately follows from Proposition 4.3.2(iii). □

**Proposition 4.4.3.** *Let $m \in \mathcal{M}(G)$. Then $m$ is left-invariant (resp. right-invariant, resp. bi-invariant) if and only if the associated finitely additive probability measure $\widehat{m} \in \mathcal{PM}(G)$ is left-invariant (resp. right-invariant, resp. bi-invariant).*

*Proof.* This immediately follows from assertions (i) and (ii) of Proposition 4.3.3. □

**Proposition 4.4.4.** *Let $G$ be a group. Then the following conditions are equivalent:*

(a) *there exists a left-invariant finitely additive probability measure $\mu\colon \mathcal{P}(G) \to [0,1]$ on $G$;*
(b) *there exists a right-invariant finitely additive probability measure $\mu\colon \mathcal{P}(G) \to [0,1]$ on $G$;*
(c) *there exists a bi-invariant finitely additive probability measure $\mu\colon \mathcal{P}(G) \to [0,1]$ on $G$;*
(d) *there exists a left-invariant mean $m\colon \ell^\infty(G) \to \mathbb{R}$ on $G$;*
(e) *there exists a right-invariant mean $m\colon \ell^\infty(G) \to \mathbb{R}$ on $G$;*
(f) *there exists a bi-invariant mean $m\colon \ell^\infty(G) \to \mathbb{R}$ on $G$.*

*Proof.* Equivalences (a) $\Leftrightarrow$ (d), (b) $\Leftrightarrow$ (e) and (c) $\Leftrightarrow$ (f) immediately follow from Proposition 4.4.3. The equivalence between (a) and (b) follows from Proposition 4.4.1. Implication (f) $\Rightarrow$ (d) is trivial.

Thus it suffices to show (d) $\Rightarrow$ (f) to conclude. Suppose that there exists a left-invariant mean $m\colon \ell^\infty(G) \to \mathbb{R}$. For each $x \in \ell^\infty(G)$, define the map $\widetilde{x}\colon G \to \mathbb{R}$ by

$$\widetilde{x}(g) = m(xg) \quad \text{for all } g \in G.$$

By Proposition 4.1.6(iv), we have

$$|\widetilde{x}(g)| = |m(xg)| \leq \|xg\|_\infty = \|x\|_\infty.$$

Therefore $\widetilde{x} \in \ell^\infty(G)$ for all $x \in \ell^\infty(G)$. Consider now the map $M\colon \ell^\infty(G) \to \mathbb{R}$ defined by

$$M(x) = m(\widetilde{x}) \text{ for all } x \in \ell^\infty(G).$$

Clearly $M$ is a mean on $G$.

Let $h \in G$ and $x \in \ell^\infty(G)$. For all $g \in G$, we have

$$\widetilde{hx}(g) = m(hxg) = m(xg) = \widetilde{x}(g),$$

since $m$ is left-invariant. We deduce that $\widetilde{hx} = \widetilde{x}$. This implies

$$M(hx) = m(\widetilde{hx}) = m(\widetilde{x}) = M(x).$$

Consequently $M$ is left-invariant. On the other hand, for all $g \in G$, we have

$$\widetilde{xh}(g) = m((xh)g) = m(x(hg)) = \widetilde{x}(hg) = h^{-1}\widetilde{x}(g).$$

It follows that $\widetilde{xh} = h^{-1}\widetilde{x}$. Therefore, by using again the fact that $m$ is left-invariant, we get

$$M(xh) = m(\widetilde{xh}) = m(h^{-1}\widetilde{x}) = m(\widetilde{x}) = M(x).$$

Therefore the mean $M$ is also right-invariant. This shows that (d) implies (f). □

**Definition 4.4.5.** A group $G$ is called *amenable* if it satisfies one of the equivalent conditions of Proposition 4.4.4.

**Proposition 4.4.6.** *Every finite group is amenable.*

*Proof.* If $G$ is a finite group, then the map $\mu\colon \mathcal{P}(G) \to [0,1]$ defined by

$$\mu(A) = \frac{|A|}{|G|} \quad \text{for all } A \in \mathcal{P}(G)$$

is a bi-invariant finitely additive probability measure on $G$. □

**Theorem 4.4.7.** *The free group on two generators $F_2$ is not amenable.*

*Proof.* Suppose that there exists a left-invariant finitely additive probability measure $\mu\colon \mathcal{P}(F_2) \to [0,1]$. Denote by $a$ and $b$ the canonical generators of $F_2$. Let $A \subset F_2$ be the set of elements of $F_2$ whose reduced form begins by a nonzero (positive or negative) power of $a$. We have $F_2 = A \cup aA$ and hence $\mu(F_2) \leq \mu(A) + \mu(aA) = 2\mu(A)$. Since $\mu(F_2) = 1$, this implies

$$\mu(A) \geq \frac{1}{2}. \tag{4.3}$$

On the other hand, the subsets $A$, $bA$ and $b^2A$ are pairwise disjoint. Thus one has $\mu(F_2) \geq \mu(A) + \mu(bA) + \mu(b^2A) = 3\mu(A)$. This gives

$$\mu(A) \leq \frac{1}{3}$$

contradicting (4.3). □

## 4.5 Stability Properties of Amenable Groups

**Proposition 4.5.1.** *Every subgroup of an amenable group is amenable.*

*Proof.* Let $G$ be an amenable group and let $H$ be a subgroup of $G$. Let $\mu\colon \mathcal{P}(G) \to [0,1]$ be a left-invariant finitely additive probability measure on $G$. Choose a complete set of representatives of the right cosets of $G$ modulo $H$, that is, a subset $R \subset G$ such that, for each $g \in G$, there exists a unique element $(h,r) \in H \times R$ satisfying $g = hr$. Let us check that the map $\widetilde{\mu}\colon \mathcal{P}(H) \to [0,1]$ defined by

$$\widetilde{\mu}(A) = \mu\left(\bigcup_{r\in R} Ar\right) \quad \text{for all } A \in \mathcal{P}(H)$$

is a left-invariant finitely additive probability measure on $H$.

Firstly, we have

$$\widetilde{\mu}(H) = \mu\left(\bigcup_{r\in R} Hr\right) = \mu(G) = 1.$$

On the other hand, if $A$ and $B$ are disjoint subsets of $H$, then

$$\begin{aligned}
\widetilde{\mu}(A \cup B) &= \mu\left(\bigcup_{r\in R} (A\cup B)r\right) \\
&= \mu\left(\left(\bigcup_{r\in R} Ar\right) \cup \left(\bigcup_{r\in R} Br\right)\right) \\
&= \mu\left(\bigcup_{r\in R} Ar\right) + \mu\left(\bigcup_{r\in R} Br\right) \\
&= \widetilde{\mu}(A) + \widetilde{\mu}(B),
\end{aligned}$$

since the sets

$$\bigcup_{r\in R} Ar \quad \text{and} \quad \bigcup_{r\in R} Br$$

are disjoint.

Finally, for all $h \in H$ and $A \in \mathcal{P}(H)$, we have

$$\widetilde{\mu}(hA) = \mu\left(\bigcup_{r\in R} hAr\right) = \mu\left(h\bigcup_{r\in R} Ar\right) = \mu\left(\bigcup_{r\in R} Ar\right) = \widetilde{\mu}(A).$$

This shows that $H$ is amenable. □

Combining Proposition 4.5.1 with Theorem 4.4.7, we get:

**Corollary 4.5.2.** *If $G$ is a group containing a subgroup isomorphic to the free group on two generators $F_2$, then $G$ is not amenable.* □

*Examples 4.5.3.* (a) Let $X$ be a set having at least two elements. Then the free group $F(X)$ based on $X$ is not amenable. Indeed, the subgroup of $F(X)$ generated by two distinct elements $a, b \in X$ is isomorphic to $F_2$.

(b) The group $\mathrm{SL}(n, \mathbb{Z})$ is not amenable for $n \geq 2$. Indeed, as we have seen in Lemma 2.3.2, the subgroup of $\mathrm{SL}(2, \mathbb{Z})$ generated by the matrices $\left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)$ is isomorphic to $F_2$.

**Proposition 4.5.4.** *Every quotient of an amenable group is amenable.*

*Proof.* Let $G$ be an amenable group and let $H$ be a normal subgroup of $G$. Let $\rho\colon G \to G/H$ denote the canonical homomorphism. Consider a left-invariant finitely additive probability measure $\mu\colon \mathcal{P}(G) \to [0, 1]$ on $G$. Let us show that the map $\widehat{\mu}\colon \mathcal{P}(G/H) \to [0, 1]$ defined by

$$\widehat{\mu}(A) = \mu(\rho^{-1}(A)) \quad \text{for all } A \in \mathcal{P}(G/H)$$

is a left-invariant finitely additive probability measure on $G/H$.

Firstly, we have

$$\widehat{\mu}(G/H) = \mu((\rho^{-1}(G/H)) = \mu(G) = 1.$$

On the other hand, if $A$ and $B$ are disjoint subsets of $G/H$, then

$$\widehat{\mu}(A \cup B) = \mu(\rho^{-1}(A \cup B)) = \mu(\rho^{-1}(A) \cup \rho^{-1}(B)) = \mu(\rho^{-1}(A)) + \mu(\rho^{-1}(B))$$

since $\rho^{-1}(A) \cap \rho^{-1}(B) = \varnothing$. Thus we have $\widehat{\mu}(A \cup B) = \widehat{\mu}(A) + \widehat{\mu}(B)$.

Finally, if $g \in G$ and $A \subset G/H$, we have

$$\widehat{\mu}(\rho(g)A) = \mu(\rho^{-1}(\rho(g)A)) = \mu(g\rho^{-1}(A)) = \mu(\rho^{-1}(A)) = \widehat{\mu}(A).$$

As $\rho$ is surjective, it follows that $\widehat{\mu}$ is left-invariant with respect to $G/H$. This shows that $G/H$ is amenable. □

In the next proposition we show that the class of amenable groups is closed under the operation of taking extensions of amenable groups by amenable groups.

**Proposition 4.5.5.** *Let $G$ be a group and let $H$ be a normal subgroup of $G$. Suppose that the groups $H$ and $G/H$ are amenable. Then the group $G$ is amenable.*

*Proof.* Since $H$ is amenable, we can find a $H$-left-invariant mean $m_0 \colon \ell^\infty(H) \to \mathbb{R}$. Let $x \in \ell^\infty(G)$. Consider the map $\widetilde{x} \colon G/H \to \mathbb{R}$ defined by

$$\widetilde{x}(\overline{g}) = m_0((g^{-1}x)|_H) \quad \text{for all } g \in G,$$

where $\overline{g} = gH = Hg \in G/H$ denote the class of $g$ modulo $H$. The map $\widetilde{x}$ is well defined. Indeed, if $g_1$ and $g_2$ are elements of $G$ such that $\overline{g_1} = \overline{g_2}$, then $g_2 = g_1 h$ for some $h \in H$, and hence

$$m_0((g_2^{-1}x)|_H) = m_0((h^{-1}g_1^{-1}x)|_H) = m_0(h^{-1}(g_1^{-1}x)|_H) = m_0((g_1^{-1}x)|_H)$$

since $m_0$ is $H$-left-invariant. Observe also that $\widetilde{x} \in \ell^\infty(G/H)$ since, by Proposition 4.1.6(iv),

$$|\widetilde{x}(\overline{g})| = |m_0((g^{-1}x)|_H)| \leq \sup_{h \in H} |(g^{-1}x)(h)| \leq \|g^{-1}x\|_\infty = \|x\|_\infty$$

for all $g \in G$.

As the group $G/H$ is amenable, there exists a $G/H$-left-invariant mean $m_1 \colon \ell^\infty(G/H) \to \mathbb{R}$. Let us set, for each $x \in \ell^\infty(G)$,

$$m(x) = m_1(\widetilde{x}).$$

Clearly $m$ is a mean on $G$. Let us show that it is $G$-left-invariant.

Let $g \in G$ and $x \in \ell^\infty(G)$. For all $g' \in G$, we have

$$\begin{aligned}\widetilde{gx}(\overline{g'}) &= m_0((g'^{-1}gx)|_H) = m_0(((g^{-1}g')^{-1}x)|_H) = \widetilde{x}(\overline{g^{-1}g'}) = \widetilde{x}(\overline{g}^{-1}\overline{g'}) \\ &= \overline{g}\widetilde{x}(\overline{g'}).\end{aligned}$$

Thus $\widetilde{gx} = \overline{g}\widetilde{x}$. As $m_1$ is $G/H$-left-invariant, it follows that

$$m(gx) = m_1(\widetilde{gx}) = m_1(\overline{g}\widetilde{x}) = m_1(\widetilde{x}) = m(x),$$

which shows that $m$ is $G$-left-invariant.Therefore $G$ is amenable. □

**Corollary 4.5.6.** *Suppose that $G_1$ and $G_2$ are amenable groups. Then the group $G = G_1 \times G_2$ is amenable.*

*Proof.* The set $H = \{(g_1, 1_{G_2}) : g_1 \in G_1\}$ is a normal subgroup of $G$ isomorphic to $G_1$ with quotient $G/H$ isomorphic to $G_2$. □

*Remark 4.5.7.* It immediately follows from the preceding corollary that every direct product of a finite number of amenable groups is amenable. However, a direct product of infinitely many amenable groups is not necessarily amenable. For example, it follows from Theorem 2.3.1 and Corollary 2.2.6

that there exists a (countable) family of finite, and hence amenable, groups $(G_i)_{i\in I}$ such that the group $G = \prod_{i\in I} G_i$ contains a subgroup isomorphic to $F_2$. Such a group $G$ is not amenable by Corollary 4.5.2.

**Corollary 4.5.8.** *Every virtually amenable group is amenable.*

*Proof.* Let $G$ be a virtually amenable group. Let $H$ be an amenable subgroup of finite index of $G$. By Lemma 2.1.10, the set $K = \cap_{g\in G} gHg^{-1}$ is a normal subgroup of finite index of $G$ contained in $H$. The group $K$ is amenable by Proposition 4.5.1. On the other hand the group $G/K$ is finite and hence amenable. Consequently, $G$ is amenable by Proposition 4.5.5. □

Our next goal is to show that an inductive limit of amenable groups is amenable. In the proof we shall use the following:

**Lemma 4.5.9.** *Let $G$ be a group. Suppose that there is a net $(m_i)_{i\in I}$ in $\mathcal{M}(G)$ such that, for each $g \in G$, the net $(gm_i - m_i)_{i\in I}$ converges to 0 in $(\ell^\infty(G))^*$ for the weak-$*$ topology. Then $G$ is amenable.*

*Proof.* Since $\mathcal{M}(G)$ is compact for the weak-$*$ topology by Theorem 4.2.1, we may assume, after taking a subnet if necessary, that the net $(m_i)$ converges to a mean $m \in \mathcal{M}(G)$. Let $g \in G$. Since the left action of $G$ on $\mathcal{M}(G)$ is continuous by Proposition 4.3.1, we deduce that, for every $x \in \ell^\infty(G)$ the net $(gm_i - m_i)(x) = gm_i(x) - m_i(x)$ converges to 0. By taking limits, we get $gm(x) - m(x) = 0$. Therefore $gm = m$. This shows that the mean $m$ is left-invariant. Thus $G$ is amenable. □

**Proposition 4.5.10.** *Every group which is the limit of an inductive system of amenable groups is amenable.*

*Proof.* Let $(G_i)_{i\in I}$ be an inductive system of amenable groups and set $G = \varinjlim G_i$. Consider the family $(H_i)_{i\in I}$ of subgroups $G$ defined by $H_i = h_i(G_i)$, where $h_i \colon G_i \to G$ is the canonical homomorphism. As $H_i$ is amenable by Proposition 4.5.4, we can find, for each $i \in I$, a $H_i$-left-invariant mean $\widetilde{m_i} \colon \ell^\infty(H_i) \to \mathbb{R}$. Consider the family $m_i \colon \ell^\infty(G) \to \mathbb{R}$ of means on $G$ defined by

$$m_i(x) = \widetilde{m_i}(x|_{H_i}) \quad \text{for all } x \in \ell^\infty(G).$$

Let $g \in G$. Since $G = \varinjlim G_i$, there exists $i_0(g) \in I$ such that $g \in H_i$ for all $i \geq i_0(g)$. For $x \in \ell^\infty(G)$ and $i \geq i_0(g)$, we have

$$(gm_i - m_i)(x) = gm_i(x) - m_i(x) = g\widetilde{m_i}(x|_{H_i}) - \widetilde{m_i}(x|_{H_i}) = 0,$$

since $\widetilde{m_i}$ is $H_i$-left-invariant. Thus $gm_i - m_i = 0$ for all $i \geq i_0(g)$. By applying Lemma 4.5.9, we deduce that $G$ is amenable. □

Every group is the inductive limit of its finitely generated subgroups. Therefore we have:

**Corollary 4.5.11.** *Every locally amenable group is amenable.* □

Since every finite group is amenable, we obtain in particular the following:

**Corollary 4.5.12.** *Every locally finite group is amenable.* □

*Example 4.5.13.* Let $X$ be a set. Then the group $\mathrm{Sym}_0(X)$ consisting of all permutations of $X$ with finite support is locally finite (see Example 3.2.4). Consequently, the group $\mathrm{Sym}_0(X)$ is amenable. The groups $\mathrm{Sym}_0(X)$, where $X$ is a infinite set, are our first examples of infinite amenable groups. Recall from Lemma 2.6.3 that $\mathrm{Sym}_0(X)$ is not residually finite whenever $X$ is infinite.

**Corollary 4.5.14.** *Let $(G_i)_{i\in I}$ be a family of amenable groups. Then their direct sum $G = \bigoplus_{i\in I} G_i$ is amenable.*

*Proof.* Every finitely generated subgroup of $G$ is a subgroup of some finite product of the groups $G_i$ and hence amenable by Corollary 4.5.6 and Proposition 4.5.1. Thus $G$ is locally amenable and therefore amenable by Corollary 4.5.11. □

## 4.6 Solvable Groups

**Theorem 4.6.1.** *Every abelian group is amenable.*

*Proof.* Let $G$ be an abelian group. Equip $(\ell^\infty(G))^*$ with the weak-$*$ topology. By Theorem 4.2.1, the set $\mathcal{M}(G)$ is a nonempty convex compact subset of $(\ell^\infty(G))^*$. On the other hand, it follows from Proposition 4.3.1 that the action of $G$ on $\mathcal{M}(G)$ is affine and continuous (note that the left and the right actions coincide since $G$ is Abelian). By applying the Markov-Kakutani fixed-point Theorem (Theorem G.1.1), we deduce that $G$ has at least one fixed point in $\mathcal{M}(G)$. Such a fixed point is clearly a bi-invariant mean on $G$. This shows that $G$ is amenable. □

Let $G$ be a group. Recall the following definitions. The *commutator* of two elements $h$ and $k$ in $G$ is the element $[h,k] \in G$ defined by $[h,k] = hkh^{-1}k^{-1}$. If $H$ and $K$ are subgroups of $G$, we denote by $[H,K]$ the subgroup of $G$ generated by all commutators $[h,k]$, where $h \in H$ and $k \in K$. Note that $[H,K] \subset K$ if $K$ is normal in $G$. Note also that $[H,K]$ is normal in $G$ if $H$ and $K$ are both normal in $G$.

The subgroup $D(G) = [G,G]$ is called the *derived subgroup*, or *commutator subgroup*, of $G$. The subgroup $D(G)$ is normal in $G$ and the quotient group $G/D(G)$ is abelian. Observe that $G$ is abelian if and only if $[G,G] = \{1_G\}$. A group is called *metabelian* if its derived subgroup is abelian.

The *derived series* of a group $G$ is the sequence $(D^i(G))_{i\geq 0}$ of subgroups of $G$ inductively defined by $D^0(G) = G$ and $D^{i+1}(G) = D(D^i(G))$ for all $i \geq 0$. One has

$$G = D^0(G) \supset D^1(G) \supset D^2(G) \supset \dots$$

with $D^{i+1}(G)$ normal in $D^i(G)$ and $D^i(G)/D^{i+1}(G)$ abelian for all $i \geq 0$.

The group $G$ is said to be *solvable* if there is an integer $i \geq 0$ such that $D^i(G) = \{1_G\}$. The smallest integer $i \geq 0$ such that $D^i(G) = \{1_G\}$ is then called the *solvability degree* of $G$.

*Examples 4.6.2.* (a) A group is solvable of degree 0 if and only if it is reduced to the identity element.

(b) The solvable groups of degree 1 are the nontrivial abelian groups.

(c) The solvable groups of degree 2 are the nonabelian metabelian groups.

(d) Let $\mathbb{K}$ be a field. The *affine group* over $\mathbb{K}$ is the subgroup $\mathrm{Aff}(\mathbb{K})$ of $\mathrm{Sym}(\mathbb{K})$ consisting of all permutations of $\mathbb{K}$ which are of the form $x \mapsto ax+b$, where $a, b \in \mathbb{K}$ and $a \neq 0$. The group $D^1(\mathrm{Aff}(\mathbb{K}))$ is the group of translations $x \mapsto x + b$, $b \in \mathbb{K}$. Since $D^1(\mathrm{Aff}(\mathbb{K}))$ is abelian, we have $D^i(\mathrm{Aff}(\mathbb{K})) = \{1_{\mathrm{Aff}(\mathbb{K})}\}$ for all $i \geq 2$. Consequently, the group $\mathrm{Aff}(\mathbb{K})$ is solvable of degree 2.

(e) The alternating groups $\mathrm{Sym}_2^+$ and $\mathrm{Sym}_3^+$ are abelian and hence solvable of degree 1. The group $\mathrm{Sym}_4^+$ is solvable of degree 2. For $n \geq 5$, the alternating group $\mathrm{Sym}_n^+$ is simple (see Remark C.4.4) and therefore $D^i(\mathrm{Sym}_n^+) = \mathrm{Sym}_n^+$ for all $i \geq 0$. Thus $\mathrm{Sym}_n^+$ is not solvable for $n \geq 5$.

(f) The symmetric group $\mathrm{Sym}_n$ is solvable of degree 1 for $n = 2$, solvable of degree 2 for $n = 3$, solvable of degree 3 for $n = 4$, and not solvable for $n \geq 5$.

**Theorem 4.6.3.** *Every solvable group is amenable.*

*Proof.* We proceed by induction on the solvability degree $i$ of the group. For $i = 0$, the group is reduced to the identity element and there is nothing to prove. Suppose now that the statement is true for solvable groups of degree $i$ for some $i \geq 0$. Let $G$ be a solvable group of degree $i + 1$. Then its derived subgroup $D(G)$ is solvable of degree $i$ and hence amenable by our induction hypothesis. As $G/D(G)$ is abelian and therefore amenable by Theorem 4.6.1, we deduce that $G$ is amenable by applying Proposition 4.5.5. □

*Remark 4.6.4.* The alternating group $\mathrm{Sym}_5^+$ is amenable since it is finite. However, as mentioned above, $\mathrm{Sym}_5^+$ is not solvable. An example of an infinite amenable group which is not solvable is provided by the group $\mathbb{Z} \times \mathrm{Sym}_5^+$.

Let $G$ be a group. The *lower central series* of $G$ is the sequence $(C^i(G))_{i \geq 0}$ of subgroups of $G$ defined by $C^0(G) = G$ and $C^{i+1}(G) = [C^i(G), G]$ for all $i \geq 0$. An easy induction shows that $C^i(G)$ is normal in $G$ and that $C^{i+1}(G) \subset C^i(G)$ for all $i$. The group $G$ is said to be *nilpotent* if there is an integer $i \geq 0$ such that $C^i(G) = \{1_G\}$. The smallest integer $i \geq 0$ such that $C^i(G) = \{1_G\}$ is then called the *nilpotency degree* of $G$.

*Example 4.6.5.* Let $R$ be a nontrivial commutative ring. The *Heisenberg group* with coefficients in $R$ is the subgroup $H_R$ of $\mathrm{GL}_3(R)$ consisting of all matrices of the form

$$M(x,y,z) = \begin{pmatrix} 1 & y & z \\ 0 & 1 & x \\ 0 & 0 & 1 \end{pmatrix} \quad (x,y,z \in R).$$

One easily checks that the center $Z(H_R)$ of $H_R$ consists of all matrices of the form $M(0,0,z)$, $z \in R$, and that $Z(H_R)$ is isomorphic to the additive group $(R,+)$. Moreover, one has $D(H_R) = Z(H_R)$. It follows that $H_R$ is nilpotent of degree 2.

**Proposition 4.6.6.** *Every nilpotent group is solvable.*

*Proof.* An easy induction yields $D^i(G) \subset C^i(G)$ for all $i \geq 0$. □

*Remark 4.6.7.* We have seen in Example 4.6.2(d) that the affine group $\mathrm{Aff}(\mathbb{K})$ over a field $\mathbb{K}$ is solvable. However, the group $G = \mathrm{Aff}(\mathbb{K})$ is not nilpotent. Indeed, the lower central series satisfies $C^i(G) = D(G) \neq \{1_G\}$ for all $i \geq 1$.

From Theorem 4.6.3 and Proposition 4.6.6, we get:

**Corollary 4.6.8.** *Every nilpotent group is amenable.* □

## 4.7 The Følner Conditions

**Proposition 4.7.1.** *Let $G$ be a group. Then the following conditions are equivalent:*

(a) *for every finite subset $K \subset G$ and every real number $\varepsilon > 0$, there exists a nonempty finite subset $F \subset G$ such that*

$$\frac{|F \setminus kF|}{|F|} < \varepsilon \quad \text{for all } k \in K; \tag{4.4}$$

(b) *there exists a net $(F_j)_{j \in J}$ of nonempty finite subsets of $G$ such that*

$$\lim_j \frac{|F_j \setminus gF_j|}{|F_j|} = 0 \quad \text{for all } g \in G; \tag{4.5}$$

(c) *for every finite subset $K \subset G$ and every real number $\varepsilon > 0$, there exists a nonempty finite subset $F \subset G$ such that*

$$\frac{|F \setminus Fk|}{|F|} < \varepsilon \quad \text{for all } k \in K; \tag{4.6}$$

(d) *there exists a net* $(F_j)_{j\in J}$ *of nonempty finite subsets of* $G$ *such that,*

$$\lim_j \frac{|F_j \setminus F_j g|}{|F_j|} = 0 \quad \text{for all } g \in G. \tag{4.7}$$

*Proof.* First observe that the equivalences (a) $\Leftrightarrow$ (c) and (b) $\Leftrightarrow$ (d) are obvious. Indeed, for any finite subset $F \subset G$ and any element $k \in G$, we have

$$\frac{|F \setminus kF|}{|F|} = \frac{|F^{-1} \setminus F^{-1}k^{-1}|}{|F^{-1}|} \tag{4.8}$$

since the map $g \mapsto g^{-1}$ is bijective on $G$.

Let us show that (a) implies (b). Suppose (a). Let $J$ denote the set of all pairs $(K, \varepsilon)$, where $K$ is a finite subset of $G$ and $\varepsilon > 0$. We equip $J$ with the partial ordering $\leq$ defined by

$$(K, \varepsilon) \leq (K', \varepsilon') \Leftrightarrow (K \subset K' \text{ and } \varepsilon \geq \varepsilon').$$

Note that $(J, \leq)$ is a *lattice*, that is, a partially ordered set in which any two elements admit a supremum (also called a join) and an infimum (also called a meet). Indeed, one has

$$\begin{aligned} \sup\{(K, \varepsilon), (K', \varepsilon')\} &= (K \cup K', \min(\varepsilon, \varepsilon')) \quad \text{and} \\ \inf\{(K, \varepsilon), (K', \varepsilon')\} &= (K \cap K', \max(\varepsilon, \varepsilon')). \end{aligned}$$

In particular, $J$ is a directed set. By (a), for each $j = (K, \varepsilon) \in J$, there exists a nonempty finite subset $F_j \subset G$ such that

$$\frac{|F_j \setminus kF_j|}{|F_j|} < \varepsilon \quad \text{for all } k \in K. \tag{4.9}$$

Let us fix now some element $g \in G$ and $\varepsilon_0 > 0$. Set $j_0 = (\{g\}, \varepsilon_0)$. If $j = (K, \varepsilon)$ satisfies $j \geq j_0$, then $g \in K$ and $\varepsilon \leq \varepsilon_0$. Thus we have

$$\frac{|F_j \setminus gF_j|}{|F_j|} < \varepsilon_0, \tag{4.10}$$

for all $j \geq j_0$ by (4.9). This shows that the net $(F_j)_{j\in J}$ satisfies (4.5). Consequently, (a) implies (b).

Finally, let us show (b) $\Rightarrow$ (a). Suppose (b). Let $K \subset G$ be a finite subset and let $\varepsilon > 0$. By (4.5), for every $k \in K$ there exist $j(k) \in J$ such that

$$\frac{|F_j \setminus kF_j|}{|F_j|} < \varepsilon \quad \text{for all } j \geq j(k).$$

Since $J$ is a directed set, we can find $j_0 \in J$ such that $j_0 \geq j(k)$ for all $k \in K$. Then, taking $F = F_{j_0}$ we have (4.6). This shows that (b) implies (a). □

**Definition 4.7.2.** One says that a group $G$ satisfies the *Følner conditions* if $G$ satisfies one of the equivalent conditions of Proposition 4.7.1.

Given a group $G$, a net $(F_j)_{j\in J}$ of nonempty finite subsets of $G$ which satisfies (4.5) (resp. (4.7)) is called a *left* (resp. *right*) *Følner net* for $G$.

Observe that $F_j \setminus g^{-1}F_j$ (resp. $F_j \setminus F_j g^{-1}$) is the set of elements of $F_j$ that are moved out of $F_j$ by left (resp. right) multiplication by $g$. Thus, a net $(F_j)_{j\in J}$ of nonempty finite subsets of $G$ is a left (resp. right) Følner net if and only if, for each $g \in G$, the proportion of elements of $F_j$ that are moved out of $F_j$ by left (resp. right) multiplication by $g$ converges to 0.

*Remarks 4.7.3.* (a) It immediately follows from (4.8) that $(F_j)$ is a left Følner net if and only if $(F_j^{-1})$ is a right Følner net.

(b) If $(F_j)_{j\in J}$ is a left Følner net for $G$ and $(a_j)_{j\in J}$ is a net of elements of $G$ indexed by the same directed set $J$, then the net $(F_j a_j)$ is also a left Følner net. Indeed, we have

$$\frac{|F_j a_j \setminus g F_j a_j|}{|F_j a_j|} = \frac{|F_j \setminus g F_j|}{|F_j|}$$

for all $j \in J$ and $g \in G$, since right multiplication by $a_j$ gives a bijection from $F_j$ onto $F_j a_j$ and from $F_j \setminus g F_j$ onto $F_j a_j \setminus g F_j a_j$. Similarly, if $(F_j)$ is a right Følner net, then $(a_j F_j)$ is also a right Følner net.

For commutative groups, the concepts of left and right Følner nets coincide so that one simply speaks of Følner nets in this case.

If the directed set $J$ is isomorphic to $\mathbb{N}$, one speaks of *left* and *right Følner sequences*.

*Examples 4.7.4.* (a) Let $G$ be a finite group. Then the sequence $(F_n)_{n\in\mathbb{N}}$, where $F_n = G$ for all $n \in \mathbb{N}$, is a left (and right) Følner sequence since $F_n \setminus gF_n = F_n \setminus F_n g = G \setminus G = \varnothing$ for all $g \in G$. Therefore $G$ satisfies the Følner conditions.

(b) Let $G = \mathbb{Z}$. Then the sequence $(F_n)_{n\in\mathbb{N}}$ defined by

$$F_n = \{0, 1, \ldots, n-1\}$$

is a Følner sequence. Indeed, if $g \in \mathbb{Z}$ is fixed and $n \geq |g|$, then

$$\frac{|F_n \setminus (F_n + g)|}{|F_n|} = \frac{|g|}{n}$$

which converges to 0 as $n$ tends to infinity. This shows that $\mathbb{Z}$ satisfies the Følner conditions (see Fig. 4.1).

(c) More generally, if $G = \mathbb{Z}^r$ $(r \geq 1)$, one checks that the sequence $(F_n)_{n\in\mathbb{N}}$ given by

$$F_n = \{0, 1, \ldots, n-1\}^r$$

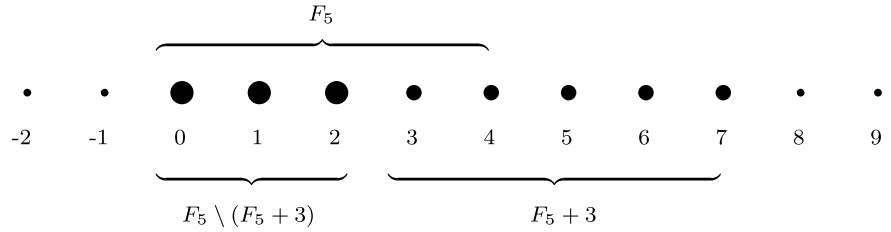



**Fig. 4.1** The Følner set $F_5$ in $\mathbb{Z}$, its translate $F_5 + 3$ and $F_5 \setminus (F_5 + 3)$

is a Følner sequence for $G$ (see Fig. 4.2). Thus, $\mathbb{Z}^r$ satisfies the Følner conditions.

**Proposition 4.7.5.** *Let $G$ be group. Then the following conditions are equivalent:*

(a) *$G$ is countable and satisfies the Følner conditions;*
(b) *$G$ admits a left Følner sequence;*
(c) *$G$ admits a right Følner sequence.*

*Proof.* (a) $\Rightarrow$ (b). Suppose (a). As $G$ is countable, we can find a sequence $(K_n)_{n\in\mathbb{N}}$ of finite subsets of $G$ such that $K_n \subset K_{n+1}$ for all $n$ and $G = \bigcup_n K_n$. Let $(\varepsilon_n)$ be a sequence of positive real numbers converging to 0. By condition (a) in Proposition 4.7.1, we can find, for each $n$, a nonempty finite subset $F_n \subset G$ such that

$$\frac{|F_n \setminus gF_n|}{|F_n|} < \varepsilon_n \quad \text{for all } g \in K_n.$$

Given an element $g \in G$, one has $g \in K_n$ for $n$ large enough and the preceding inequality implies

$$\lim_{n\to\infty} \frac{|F_n \setminus gF_n|}{|F_n|} = 0.$$

This shows that $(F_n)$ is a left Følner sequence for $G$.

(b) $\Rightarrow$ (a). Suppose that $G$ admits a left Følner sequence $(F_n)_{n\in\mathbb{N}}$. Let us show that $G$ is countable. Set $K_n = \{ab^{-1} : a, b \in F_n\}$. Consider an element $g \in G$. Then, for $n$ large enough, we have

$$\frac{|F_n \setminus gF_n|}{|F_n|} < 1.$$

This implies $F_n \cap gF_n \neq \varnothing$ and hence $g \in K_n$. It follows that $G = \bigcup_n K_n$. As $K_n$ is finite for each $n$, we deduce that $G$ is countable. This shows that (b) implies (a).

As observed above, $(F_n)$ is a left Følner sequence if and only if $(F_n^{-1})$ is a right Følner sequence. This shows that conditions (b) and (c) are equivalent. □

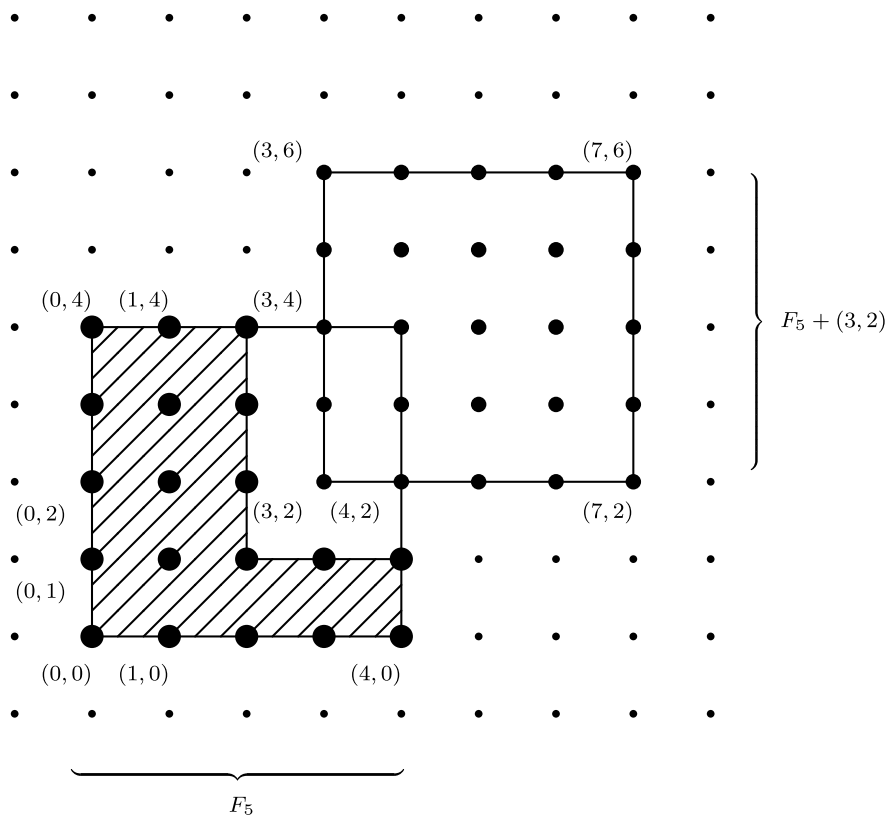



**Fig. 4.2** The Følner set $F_5$ in $\mathbb{Z}^2$, its translate $F_5 + (3, 2)$ and $F_5 \setminus (F_5 + (3, 2))$

## 4.8 Paradoxical Decompositions

**Definition 4.8.1.** Let $G$ be a group. A *left paradoxical decomposition* of $G$ is a triple $(K, (A_k)_{k\in K}, (B_k)_{k\in K})$ where $K$ is a finite subset of $G$, and $(A_k)_{k\in K}$ and $(B_k)_{k\in K}$ are families of subsets of $G$ indexed by $K$ such that

$$G = \left(\coprod_{k\in K} kA_k\right) \coprod \left(\coprod_{k\in K} kB_k\right) = \coprod_{k\in K} A_k = \coprod_{k\in K} B_k, \tag{4.11}$$

where $\coprod$ means disjoint union.

Similarly, a *right paradoxical decomposition* of $G$ is a triple $(K, (A_k)_{k\in K}$, $(B_k)_{k\in K})$ where $K$ is a finite subset of $G$, and $(A_k)_{k\in K}$ and $(B_k)_{k\in K}$ are families of subsets of $G$ indexed by $K$ such that

$$G = \left(\coprod_{k\in K} A_k k\right) \coprod \left(\coprod_{k\in K} B_k k\right) = \coprod_{k\in K} A_k = \coprod_{k\in K} B_k. \tag{4.12}$$

By taking inverses in $G$, it is clear that $(K, (A_k)_{k\in K}, (B_k)_{k\in K})$ is a left paradoxical decomposition of $G$ if and only if $(K^{-1}, (A_{k^{-1}}^{-1})_{k\in K^{-1}}, (B_{k^{-1}}^{-1})_{k\in K^{-1}})$ is a right decomposition of $G$. This shows that $G$ admits a left paradoxical decomposition if and only if it admits a right paradoxical decomposition. If this is the case, we simply say that $G$ admits a paradoxical decomposition.

*Example 4.8.2.* Consider the free group of rank two $G = F_2$ freely generated by $a$ and $b$. Let us partition $F_2$ into four subsets $A^+$, $A^-$, $B^+$ and $B^-$ as follows. We denote by $A^+$ (resp. $A^-$) the subset of $F_2$ consisting of all elements whose reduced form starts with a positive (resp. negative) power of $a$. Let also $B^+$ denote the subset of $F_2$ consisting of all elements either of the form $b^{-n}$ for $n = 0, 1, 2, \ldots$, or whose reduced form starts with a positive power of $b$. Finally, let $B^- = F_2 \setminus (A^+ \cup A^- \cup B^+)$, so that $B^-$ consists of all elements in $F_2$ whose reduced form starts with, but is not equal to, a negative power of $b$. Then we have

$$F_2 = A^+ \coprod A^- \coprod B^+ \coprod B^- = a^{-1}A^+ \coprod A^- = b^{-1}B^+ \coprod B^-. \quad (4.13)$$

For example, if $g = a^2b \in F_2$, we have $g \in A^+$ and, on the one hand, $g = a^{-1}(a^3b) \in a^{-1}A^+$ and, on the other, $g = b^{-1}(ba^2b) \in b^{-1}B^+$.

Thus, setting $A_1 = A^-$, $A_a = a^{-1}A^+$, $A_b = \varnothing$, $B_1 = B^-$, $B_a = \varnothing$ and $B_b = b^{-1}B^+$, (4.13) becomes

$$\begin{aligned} F_2 &= A_1 \coprod aA_a \coprod bA_b \coprod B_1 \coprod aB_a \coprod bB_b \\ &= A_1 \coprod A_a \coprod A_b \\ &= B_1 \coprod B_a \coprod B_b, \end{aligned} \quad (4.14)$$

which, with $K = \{1, a, b\}$, is nothing but (4.11). This shows that $F_2$ admits a paradoxical decomposition.

## 4.9 The Theorems of Tarski and Følner

The *Tarski alternative theorem* asserts that a group is either amenable or admits a paradoxical decomposition. The *theorem of Følner* says that a group is amenable if and only if it satisfies the Følner conditions. Thus, we have the following:

**Theorem 4.9.1 (Tarski-Følner).** *Let $G$ be a group. Then the following conditions are equivalent.*

(a) *$G$ is amenable;*
(b) *$G$ admits no paradoxical decompositions;*
(c) *$G$ satisfies the Følner conditions.*

Theorem 4.9.1 directly follows from the equivalence of conditions (a), (b) and (g) contained in the theorem below:

**Theorem 4.9.2.** *Let $G$ be a group. Then the following conditions are equivalent:*

(a) *$G$ is not amenable;*
(b) *$G$ does not satisfy the Følner conditions;*
(c) *there exists a finite subset $K \subset G$ such that $|KF| \geq 2|F|$ for every finite subset $F \subset G$.;*
(d) *there exists a finite subset $K \subset G$ such that $|FK| \geq 2|F|$ for every finite subset $F \subset G$;*
(e) *there exist a 2-to-one surjective map $\varphi\colon G \to G$ and a finite subset $K \subset G$ such that $g(\varphi(g))^{-1} \in K$ for all $g \in G$;*
(f) *there exist a 2-to-one surjective map $\varphi\colon G \to G$ and a finite subset $K \subset G$ such that $g^{-1}\varphi(g) \in K$ for all $g \in G$;*
(g) *$G$ admits a paradoxical decomposition.*

We recall that a surjective map $f\colon X \to Y$ from a set $X$ onto a set $Y$ is said to be *2-to-one* if every $y \in Y$ has exactly two preimages in $X$.

*Proof of Theorem 4.9.2.* By taking inverses in $G$, it is clear that (c) $\Leftrightarrow$ (d) and (e) $\Leftrightarrow$ (f). Thus, it suffices to prove (a) $\Rightarrow$ (b), (b) $\Rightarrow$ (c), (c) $\Rightarrow$ (e), (e) $\Rightarrow$ (g), and (g) $\Rightarrow$ (a).

(a) $\Rightarrow$ (b). Suppose that $G$ satisfies the Følner conditions. Then $G$ admits a left Følner net $(F_j)_{j\in J}$. Consider, for each $j \in J$, the mean with finite support $m_j\colon \ell^\infty(G) \to \mathbb{R}$ defined by

$$m_j(x) = \frac{1}{|F_j|} \sum_{h\in F_j} x(h) \quad \text{for all} \quad x \in \ell^\infty(G).$$

Let $g \in G$. For every $x \in \ell^\infty(G)$, we have

$$\begin{aligned}
(gm_j - m_j)(x) &= gm_j(x) - m_j(x) \\
&= m_j(g^{-1}x) - m_j(x) \\
&= \frac{1}{|F_j|}\left(\sum_{h\in F_j} x(gh) - \sum_{h\in F_j} x(h)\right) \\
&= \frac{1}{|F_j|}\left(\sum_{h\in gF_j} x(h) - \sum_{h\in F_j} x(h)\right) \\
&= \frac{1}{|F_j|}\left(\sum_{h\in gF_j\setminus F_j} x(h) - \sum_{h\in F_j\setminus gF_j} x(h)\right),
\end{aligned}$$

which gives, by using triangle inequality,

$$|(gm_j - m_j)(x)| \leq \frac{1}{|F_j|}\left(\sum_{h\in gF_j\setminus F_j}|x(h)| + \sum_{h\in F_j\setminus gF_j}|x(h)|\right)$$
$$\leq \frac{1}{|F_j|}(|gF_j \setminus F_j| + |F_j \setminus gF_j|)\|x\|_\infty.$$

As the sets $gF_j$ and $F_j$ have the same cardinality, we have

$$|gF_j \setminus F_j| = |F_j \setminus gF_j|. \tag{4.15}$$

Thus, we finally get

$$|(gm_j - m_j)(x)| \leq 2\frac{|F_j \setminus gF_j|}{|F_j|}\|x\|_\infty.$$

This implies that

$$\lim_j (gm_j - m_j)(x) = 0,$$

since $(F_j)$ is a left Følner net. Thus, the net $(gm_j - m_j)_{j\in J}$ converges to 0 in $(\ell^\infty(G))^*$ for the weak-$*$ topology. It follows that $G$ is amenable by Lemma 4.5.9. This shows that (a) implies (b).

(b) $\Rightarrow$ (c). Suppose that $G$ does not satisfy the Følner conditions. Then, there exist a finite subset $K_0 \subset G$ and $\varepsilon_0 > 0$ such that for every nonempty finite subset $F \subset G$ one has

$$\frac{|F \setminus k_0F|}{|F|} \geq \varepsilon_0 \tag{4.16}$$

for some $k_0 = k_0(F) \in K_0$. Consider the set $K_1 = K_0 \cup \{1_G\}$. Let $F$ be a nonempty finite subset of $G$. Observe that $K_1F \supset F$ and $K_1F\setminus F = K_0F\setminus F$. Thus, we have

$$\begin{aligned}|K_1F| - |F| &= |K_1F \setminus F|\\ &= |K_0F \setminus F|\\ &\geq |k_0F \setminus F|\\ &= |F \setminus k_0F| \quad (\text{since } |F| = |k_0F|)\\ &\geq \varepsilon_0|F| \quad (\text{by } (4.16)),\end{aligned}$$

which gives

$$|K_1F| \geq (1 + \varepsilon_0)|F|.$$

Choose $n_0 \in \mathbb{N}$ such that $(1+\varepsilon_0)^{n_0} \geq 2$ and set $K = K_1^{n_0}$. Then, we have

$$|KF| = |K_1^{n_0} F| \geq (1+\varepsilon_0)|K_1^{n_0-1} F| \geq \cdots \geq (1+\varepsilon_0)^{n_0}|F|$$

so that $|KF| \geq 2|F|$ for every finite subset $F \subset G$. This shows that (b) implies (c).

(c) $\Rightarrow$ (e). Suppose that $G$ satisfies condition (c), that is, there exists a finite subset $K \subset G$ such that

$$|KF| \geq 2|F| \quad \text{for every finite subset } F \subset G. \tag{4.17}$$

Consider the bipartite graph $\mathcal{G}_K(G) = (G, G, E)$ (see Appendix H), where the set $E \subset G \times G$ of edges consists of all the pairs $(g, h)$ such that $g \in G$ and $h \in Kg$.

We claim that $\mathcal{G}_K(G)$ satisfies the Hall 2-harem conditions. Indeed, if $F$ is a finite subset of $G$, then, using the terminology for bipartite graphs introduced in Sect. H.1, the right and left neighborhoods of $F$ in $\mathcal{G}_K(G)$ are the sets $\mathcal{N}_R(F) = KF$ and $\mathcal{N}_L(F) = K^{-1}F$ respectively. Therefore, we have

$$|\mathcal{N}_R(F)| = |KF| \geq 2|F|,$$

by applying (4.17). On the other hand, if $k \in K$, then $\mathcal{N}_L(F) \supset k^{-1}F$ so that

$$|\mathcal{N}_L(F)| \geq |k^{-1}F| = |F| \geq \frac{1}{2}|F|.$$

This proves our claim. Thus, by virtue of the Hall harem Theorem (Theorem H.4.2), we deduce the existence of a perfect $(1,2)$-matching $M$ for $\mathcal{G}_K(G)$. In other words, there exists a 2-to-one surjective map $\varphi\colon G \to G$ such that $(\varphi(g), g) \in E$, that is, $g(\varphi(g))^{-1} \in K$ for all $g \in G$. This shows that (c) implies (e).

(e) $\Rightarrow$ (g). Suppose (e), that is, there exist a 2-to-one surjective map $\varphi\colon G \to G$ and a finite set $K \subset G$ such that

$$g(\varphi(g))^{-1} \in K \text{ for all } g \in G. \tag{4.18}$$

By the axiom of choice, we can find maps $\psi_1, \psi_2\colon G \to G$ such that, for every $g \in G$, the elements $\psi_1(g)$ and $\psi_2(g)$ are the two preimages of $g$ for $\varphi$.

Observe that $\theta_1(g) = \psi_1(g)g^{-1}$ and $\theta_2(g) = \psi_2(g)g^{-1}$ belong to $K$ for every $g \in G$ by (4.18). For each $k \in K$, define $A_k$ and $B_k$ by

$$A_k = \{g \in G : \theta_1(g) = k\} \quad \text{and} \quad B_k = \{g \in G : \theta_2(g) = k\}.$$

We have

$$G = \coprod_{k \in K} A_k = \coprod_{k \in K} B_k. \tag{4.19}$$

On the other hand, observe that if $g \in A_k$ then $\psi_1(g) = kg$. Thus $\psi_1(G) = \coprod_{k\in K} kA_k$. Similarly, we have $\psi_2(G) = \coprod_{k\in K} kB_k$. As $G = \psi_1(G) \coprod \psi_2(G)$, we deduce that

$$G = \left(\coprod_{k\in K} kA_k\right) \coprod \left(\coprod_{k\in K} kB_k\right). \tag{4.20}$$

Combining together (4.19) and (4.20), ,we deduce that $(K, (A_k)_{k\in K}, (B_k)_{k\in K})$ is a left paradoxical decomposition for $G$. This shows that (e) implies (g).

(g) $\Rightarrow$ (a). Suppose that $G$ admits a left paradoxical decomposition $(K, (A_k)_{k\in K}, (B_k)_{k\in K})$. If $\mu\colon \mathcal{P}(G) \to [0,1]$ is a left invariant finitely additive probability measure on $G$, then (4.11) gives

$$\begin{aligned}
1 &= \mu(G) \\
&= \mu\left(\left(\coprod_{k\in K} kA_k\right) \coprod \left(\coprod_{k\in K} kB_k\right)\right) \\
&= \sum_{k\in K} \mu(kA_k) + \sum_{k\in K} \mu(kB_k) \\
&= \sum_{k\in K} \mu(A_k) + \sum_{k\in K} \mu(B_k) \\
&= \mu(\coprod_{k\in K} A_k) + \mu(\coprod_{k\in K} B_k) \\
&= \mu(G) + \mu(G) \\
&= 2,
\end{aligned}$$

which is clearly absurd. Therefore $G$ is not amenable. This shows that (g) implies (a). □

## 4.10 The Fixed Point Property

The following fixed point theorem is a generalization of the Markov-Kakutani theorem (Theorem G.1.1).

**Theorem 4.10.1.** *Let $G$ be an amenable group acting affinely and continuously on a nonempty convex compact subset $C$ of a Hausdorff topological vector space $X$. Then $G$ fixes at least one point in $C$.*

*Proof.* Let $(F_j)_{j\in J}$ be a left Følner net for $G$.

Choose an arbitrary point $x \in C$ and set

$$x_j = \frac{1}{|F_j|} \sum_{h\in F_j} hx.$$

for each $j \in J$. Observe that $x_j \in C$ since $C$ is convex. By compactness of $C$, we can assume, after taking a subnet, that the net $(x_j)_{j\in J}$ converges to a point $c \in C$.

Let $g \in G$. For every $j \in J$, we have

$$\begin{aligned} gx_j - x_j &= g\left(\frac{1}{|F_j|}\sum_{h\in F_j} hx\right) - \frac{1}{|F_j|}\sum_{h\in F_j} hx \\ &= \frac{1}{|F_j|}\sum_{h\in F_j} ghx - \frac{1}{|F_j|}\sum_{h\in F_j} hx \quad \text{(since the action of } G \text{ on } C \text{ is affine)} \\ &= \frac{1}{|F_j|}\sum_{h\in gF_j} hx - \frac{1}{|F_j|}\sum_{h\in F_j} hx, \end{aligned}$$

which yields, after simplification,

$$gx_j - x_j = \frac{1}{|F_j|}\sum_{h\in gF_j\setminus F_j} hx - \frac{1}{|F_j|}\sum_{h\in F_j\setminus gF_j} hx. \tag{4.21}$$

Consider the points

$$y_j = \frac{1}{|gF_j \setminus F_j|}\sum_{h\in gF_j\setminus F_j} hx, \qquad \text{and} \qquad z_j = \frac{1}{|F_j \setminus gF_j|}\sum_{h\in F_j\setminus gF_j} hx.$$

Note that $y_j$ and $z_j$ belong to $C$ by convexity of $C$.

We have $|F_j \setminus gF_j| = |gF_j \setminus F_j|$ since the sets $F_j$ and $gF_j$ have the same cardinality (cf. (4.15)). Setting

$$\lambda_j = \frac{|F_j \setminus gF_j|}{|F_j|} = \frac{|gF_j \setminus F_j|}{|F_j|},$$

equality (4.21) gives us

$$gx_j - x_j = \lambda_j y_j - \lambda_j z_j.$$

The net $(\lambda_j)$ converges to 0 since $(F_j)$ is a left Følner net. As $C$ is compact, it follows that

$$\lim_j \lambda_j y_j = \lim_j \lambda_j z_j = 0,$$

by using Lemma G.2.2. Therefore, we have

$$gc - c = \lim_j (gx_j - x_j) = 0.$$

This shows that $c$ is fixed by $G$. □

**Corollary 4.10.2.** *Let $G$ be a group. Then the following conditions are equivalent:*

(a) *the group $G$ is amenable;*
(b) *every continuous affine action of $G$ on a nonempty convex compact subset of a Hausdorff topological vector space admits a fixed point;*
(c) *every continuous affine action of $G$ on a nonempty convex compact subset of a Hausdorff locally convex topological vector space admits a fixed point.*

*Proof.* Implication (a) $\Rightarrow$ (b) follows from the preceding theorem. Implication (b) $\Rightarrow$ (c) is trivial. Finally,(c) $\Rightarrow$ (a) follows from Theorem 4.2.1, Proposition 4.3.1 and the fact that $(\ell^\infty(G))^*$ is a locally convex Hausdorff topological vector space for the weak-$*$ topology (see Sect. F.2). □

## Notes

The theory of amenable groups emerged from the study of the axiomatic properties of the Lebesgue integral and the discovery of the Banach-Tarski paradox at the beginning of the last century (see [Har3], [Pat], [Wag]). The first definition of an amenable group, by the existence of an invariant finitely additive probability measure, is due to J. von Neumann in [vNeu1]. Von Neumann proved in particular that every abelian group is amenable (Theorem 4.6.1) and that an amenable group cannot contain a subgroup isomorphic to $F_2$ (Corollary 4.5.2). The term *amenable* was introduced in the 1950s by M.M. Day, who played a central role in the development of the modern theory of amenable groups by using means and applying techniques from functional analysis. Moreover, Day extended the notion of amenability to semigroups, for which one has to distinguish between right amenability and left amenability.

The question of the existence of a non-amenable group containing no subgroup isomorphic to $F_2$, which is called by some authors the *von Neumann conjecture* or *Day's problem*, was answered in the affirmative by A.Yu. Ol'shanskii [Ols] who gave an example of a finitely generated non-amenable group all of whose proper subgroups are cyclic. Other examples of non-amenable groups with no subgroup isomorphic to $F_2$ were found by S.I. Adyan [Ady] who showed that the *free Burnside group* $B(m, n)$ is non-amenable for $m \geq 2$ and $n$ odd with $n \geq 665$. The free Burnside group $B(m, n)$ is the quotient of the free group $F_m$ by the subgroup of $F_m$ generated by all $n$-powers, that is, all elements of the form $w^n$ for some $w \in F_m$. The order of every element of $B(m, n)$ divide $n$. In particular, $B(m, n)$ is a *periodic group*, that is, a group in which every element has finite order. It is clear that a periodic group cannot contain a subgroup isomorphic to $F_2$. A geometric method for constructing finitely generated non-amenable periodic groups was described by M. Gromov in [Gro3]. Examples of finitely presented non-amenable groups

which contain no subgroup isomorphic to $F_2$ were given by A.Yu Ol'shanskii and M. Sapir [OlS].

There is also a more general notion of amenability for locally compact groups and actions of locally compact groups (see [Gre], [Pat]).

A group $G$ is called *polycyclic* if it admits a finite sequence of subgroups

$$\{1_G\} = H_0 \subset H_1 \subset H_2 \subset \cdots \subset H_n = G$$

such that $H_i$ is normal in $H_{i+1}$ and $H_{i+1}/H_i$ is a (finite or infinite) cyclic group for each $0 \leq i \leq n-1$. Clearly, every polycyclic group is solvable. It is not hard to prove that every finitely generated nilpotent group is polycyclic. It was shown by L. Auslander [Aus] that every polycyclic group is isomorphic to a subgroup of $\mathrm{SL}_n(\mathbb{Z})$ for some integer $n \geq 1$. It follows in particular that every polycyclic group is residually finite. This last result is due to K.A. Hirsch [Hir] (see also [RobD, p. 154]). P. Hall [Hall2] proved that every finitely generated metabelian group is residually finite and gave an example of a finitely generated solvable group of degree 3 which is not residually finite.

The equivalence between Følner conditions and amenability was established by E. Følner in [Føl]. The proof was later simplified by I. Namioka in [Nam]. The equivalence between amenability and the non-existence of a paradoxical decomposition is due to A. Tarski (see [Tar1], [Tar2] and [CGH1]).

Let $G$ be a group. Given a left (or right) paradoxical decomposition $\mathcal{P} = (K, (A_k)_{k\in K}, (B_k)_{k\in K})$ of $G$, the integer number $c(\mathcal{P}) = m + n$, where $m = |\{k \in K : A_k \neq \varnothing\}|$ and $n = |\{k \in K : B_k \neq \varnothing\}|$, is called the *complexity* of $\mathcal{P}$. Then the quantity $T(G) = \inf c(\mathcal{P})$, where the infimum is taken over all left (or right) paradoxical decompositions $\mathcal{P}$ of $G$, is called the *Tarski number* of $G$. One uses the convention that $T(G) = \infty$ if $G$ admits no paradoxical decompositions, that is, if $G$ is amenable (cf. Theorem 4.9.1). It was proved by B. Jonsson, a student of Tarski in the 1940s, that a group $G$ has Tarski number $T(G) = 4$ if and only if $G$ contains a subgroup isomorphic to $F_2$, the free group of rank 2. In [CGH1, CGH2] it was shown that for the free Burnside groups $B(m, n)$ with $m \geq 2$ and $n \geq 665$ odd one has $6 \leq T(B(m, n)) \leq 14$. The computations involve spectral analysis and Cheeger-Buser type isoperimetric inequalities (see Sect. 6.10 and Sect. 6.12) and Adyan's cogrowth estimates [Ady] for the free Burnside groups $B(m, n)$ (see (6.117) in the Notes for Chap. 6).

The extension of the Markov-Kakutani fixed point theorem to amenable groups (cf. Theorem 4.10.1) is due to Day [Day2].

## Exercises

**4.1.** Let $G$ be an infinite group and let $\mu\colon \mathcal{P}(G) \to [0, 1]$ be a left (or right) invariant finitely additive probability measure on $G$. Show that every finite subset $A \subset G$ satisfies $\mu(A) = 0$.

**4.2.** Let $X$ be an infinite set. Prove that the symmetric group $\mathrm{Sym}(X)$ is not amenable. Hint: Show that $\mathrm{Sym}(X)$ contains a subgroup isomorphic to the free group $F_2$.

**4.3.** Show that the finitely generated and non residually finite groups $G_1$ and $G_2$ described in Sect. 2.6 are amenable. Hint: Each of these groups is the semidirect product of a locally finite group and an infinite cyclic group.

**4.4.** Let $G$ be a group.

(a) Suppose that $H$ and $K$ are normal subgroups of $G$. Prove that $HK$ is a normal subgroup of $G$ and that the groups $HK/K$ and $H/(H \cap K)$ are isomorphic.

(b) Suppose that $H$ and $K$ are normal amenable subgroups of $G$. Prove that $HK$ is a normal amenable subgroup of $G$.

(c) Show that the set of all normal amenable subgroups of $G$ has a maximal element for inclusion. This maximal element is called the *amenable radical* of the group $G$. Hint: Use (b) and Proposition 4.5.10 to show that the union of all normal amenable subgroups of $G$ is a normal amenable subgroup of $G$.

**4.5.** Let $G$ be a group. Show that $G$ is metabelian if and only if it contains a normal subgroup $N$ such that the group $G/N$ is abelian.

**4.6.** Let $G$ be a finitely generated solvable group. Show that if all elements of $G$ have finite order then $G$ is finite. Hint: Use induction on the solvability degree of $G$.

**4.7.** Show that every subgroup of a solvable (resp. nilpotent) group is solvable (resp. nilpotent).

**4.8.** Show that every quotient of a solvable (resp. nilpotent) group is solvable (resp. nilpotent).

**4.9.** Let $G$ be a group containing a normal subgroup $H$ such that both $H$ and $G/H$ are solvable. Show that $G$ is solvable.

**4.10.** Show that the direct product of two nilpotent groups is a nilpotent group.

**4.11.** Show that a semidirect product of two nilpotent groups may fail to be nilpotent. Hint: Take for example the symmetric group $\mathrm{Sym}_3$, which is the semidirect product of two cyclic groups.

**4.12.** Let $G$ be a group. Show that $G$ is solvable if and only if there is a finite sequence

$$\{1_G\} = H_0 \subset H_1 \subset H_2 \subset \cdots \subset H_n = G$$

of subgroups of $G$ such that $H_i$ is normal in $H_{i+1}$ and $H_{i+1}/H_i$ is abelian for all $0 \leq i \leq n-1$.

**4.13.** Let $G$ be a group and $i \geq 0$. Show that the group $C^i(G)/C^{i+1}(G)$ is contained in the center of $G/C^{i+1}(G)$.

**4.14.** Show that every nontrivial nilpotent group has a nontrivial center.

**4.15.** Let $G$ be a nilpotent group of nilpotency degree $d \geq 1$. Show that the quotient group $G/C^{d-1}(G)$ is nilpotent of nilpotency degree $d-1$.

**4.16.** Let $G$ be a finite group whose order is a power of a prime number. Show that $G$ is nilpotent. Hint: Prove that the center of $G$ is nontrivial by considering the action of $G$ on itself by conjugation and then proceed by induction.

**4.17.** Let $G$ be a group. Denote by $\mathcal{N}_{nq}$ (resp. $\mathcal{N}_{sq}$, resp. $\mathcal{N}_{aq}$) the set of all normal subgroups $N \subset G$ such that the quotient group $G/N$ is nilpotent (resp. solvable, resp. amenable). The sets $\mathcal{N}_{nq}$, $\mathcal{N}_{sq}$ and $\mathcal{N}_{aq}$ are partially ordered by reverse inclusion.

(a) Show that if $N_1$ and $N_2$ are two elements of $\mathcal{N}_{nq}$ (resp. $\mathcal{N}_{sq}$, resp. $\mathcal{N}_{aq}$), then $N_1 \cap N_2$ is an element of $\mathcal{N}_{nq}$ (resp. $\mathcal{N}_{sq}$, resp. $\mathcal{N}_{aq}$). Hint: observe that if $\rho_1 \colon G \to G/N_1$ and $\rho_2 \colon G \to G/N_2$ are the canonical homomorphisms, then the map $\psi \colon G \to G/N_1 \times G/N_2$ defined by $\psi(g) = (\rho_1(g), \rho_2(g))$ is a homomorphism whose kernel is $N_1 \cap N_2$.

(b) Show that $\mathcal{N}_{nq}$ (resp. $\mathcal{N}_{sq}$, resp. $\mathcal{N}_{aq}$) gives rise to a projective system of groups in a natural way. The limit of this projective system is called the *pronilpotent completion* (resp. *prosolvable completion*, resp. *proamenable completion*) of the group $G$ and is denoted by $\widehat{G}_n$ (resp. $\widehat{G}_s$, resp. $\widehat{G}_a$).

(c) Show that there is a canonical homomorphism $G \to \widehat{G}_n$ (resp. $G \to \widehat{G}_s$, resp. $G \to \widehat{G}_a$) and that this homomorphism is injective if and only if $G$ is residually nilpotent (resp. residually solvable, resp. residually amenable).

**4.18.** Recall that a group $G$ is called *polycyclic* if it admits a finite sequence of subgroups

$$\{1_G\} = H_0 \subset H_1 \subset H_2 \subset \cdots \subset H_n = G$$

such that $H_i$ is normal in $H_{i+1}$ and $H_{i+1}/H_i$ is a (finite or infinite) cyclic group for each $0 \leq i \leq n-1$.

(a) Show that every polycyclic group is finitely generated.

(b) Show that every subgroup of a polycyclic group is polycyclic.

(c) Deduce from (a) and (b) that every subgroup of a polycyclic group is finitely generated.

**4.19.** The *lamplighter group* is the wreath product $L = (\mathbb{Z}/2\mathbb{Z}) \wr \mathbb{Z}$. Thus, $L$ is the semidirect product of the groups $H = \oplus_{n \in \mathbb{Z}} A_n$ and $\mathbb{Z}$, where $A_n = \mathbb{Z}/2\mathbb{Z}$ for all $n \in \mathbb{Z}$ and $\mathbb{Z}$ acts on $H$ by the $\mathbb{Z}$-shift.

(a) Show that the group $L$ is metabelian (and therefore solvable) and residually finite.

(b) Prove that $L$ is finitely generated. Hint: Show that $L$ is generated by the two elements $s$ and $t$ corresponding respectively to the nontrivial element of $A_0$ and to the canonical generator of $\mathbb{Z}$.

(c) Show that $L$ is not polycyclic. Hint: Observe that $H$ is not finitely generated and use Exercise 4.18(c).

**4.20.** Show that every finite solvable group is polycyclic.

**4.21.** Let $m$ be an integer such that $|m| \geq 2$. Let $G$ be the group given by the presentation $G = \langle a, b : aba^{-1} = b^m \rangle$. Use the results in Exercise 2.7 to prove that the commutator subgroup $[G, G]$ is isomorphic to the additive group $\mathbb{Z}[1/m]$ and that the quotient group $G/[G, G]$ is infinite cyclic.

**4.22.** Let $G$ be a locally finite group. Let $\mathcal{S}$ denote the directed set consisting of all finitely generated subgroups of $G$ partially ordered by inclusion. Prove that the net $(H)_{H \in \mathcal{S}}$ is a Følner net for $G$.

**4.23.** Show that the sequence $(F_n)_{n \in \mathbb{N}}$, where $\mathcal{F}_n$ consists of all rational numbers of the form $k/n!$ with $k \in \mathbb{N}$ and $k \leq (n+1)!$, is a Følner sequence for the additive group $\mathbb{Q}$.

**4.24.** Let $G = H_{\mathbb{Z}}$ denote the integral Heisenberg group (cf. Example 4.6.5). For each integer $n \geq 0$, define the subset $F_n \subset G$ by

$$F_n = \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \in G : 1 \leq x \leq n, 1 \leq y \leq n, 1 \leq z \leq n^2 \right\}.$$

Show that the sequence $(F_n)_{n \geq 0}$ is a Følner sequence for $G$.

**4.25.** Let $G$ be a group. Show that $G$ is amenable if and only if the following condition holds: for every finite subset $K \subset G$ and every $\varepsilon > 0$, there exists a finite subset $F \subset G$ such that $|KF| < (1 + \varepsilon)|F|$.

**4.26.** Let $G$ be a countable amenable group. Show that $G$ admits a left (resp. right) Følner sequence $(F_n)_{n \in \mathbb{N}}$ which satisfies $G = \bigcup_{n \in \mathbb{N}} F_n$ and $F_n \subset F_{n+1}$ for all $n \in \mathbb{N}$.

**4.27.** Let $(K, (A_k)_{k \in K}, (B_k)_{k \in K})$ be a left (or right) paradoxical decomposition of a group $G$. Show that $|K| \geq 3$.

**4.28.** Let $G$ be a group and $H \subset G$ a subgroup. Let $(K, (A_k)_{k \in K}, (B_k)_{k \in K})$ be a left paradoxical decomposition of $H$ and $T \subset G$ a set of representatives for the right cosets of $H$ in $G$. For each $k \in K$ set $A'_k = \coprod_{t \in T} A_k t$ and $B'_k = \coprod_{t \in T} B_k t$. Show that $(K, (A'_k)_{k \in K}, (B'_k)_{k \in K})$ is a left paradoxical decomposition of $G$.

**4.29.** Let $T(G) \in \mathbb{N} \cup \{\infty\}$ denote the Tarski number of a group $G$.

(a) Show that $T(G) \geq 4$ for all groups $G$.

(b) Show that $T(G) = 4$ if and only if $G$ contains a subgroup isomorphic to $F_2$. Hint: Use Example 4.8.2 and the Klein Ping-Pong theorem (Theorem D.5.1).

(c) Let $H$ be a subgroup of a group $G$. Show that $T(G) \leq T(H)$.

(d) Let $N$ be a normal subgroup of a group $G$. Show that $T(G) \leq T(G/N)$.

(e) Let $G$ be a group. Show that there exists a finitely generated subgroup $H \subset G$ such that $T(H) = T(G)$.

(f) Let $G$ be a group. Suppose that all elements of $G$ have finite order. Show that $T(G) \geq 6$.

# Chapter 5
# The Garden of Eden Theorem

The Garden of Eden Theorem gives a necessary and sufficient condition for the surjectivity of a cellular automaton with finite alphabet over an amenable group. It states that such an automaton is surjective if and only if it is pre-injective. As the name suggests it, pre-injectivity is a weaker notion than injectivity. It means that any two configurations which have the same image under the automaton must be equal if they coincide outside a finite subset of the underlying group (see Sect. 5.2). We shall establish the Garden of Eden theorem in Sect. 5.8 by showing that both the surjectivity and the pre-injectivity are equivalent to the maximality of the entropy of the image of the cellular automaton. The entropy of a set of configurations with respect to a Følner net of an amenable group is defined in Sect. 5.7. Another important tool in the proof of the Garden of Eden theorem is a notion of tiling for groups introduced in Sect. 5.6. The Garden of Eden theorem is used in Sect. 5.9 to prove that every residually amenable (and hence every amenable) group is surjunctive. In Sect. 5.10 and Sect. 5.11, we give simple examples showing that both implications in the Garden of Eden theorem become false over a free group of rank two. In Sect. 5.12 it is shown that a group $G$ is amenable if and only if every surjective cellular automaton with finite alphabet over $G$ is pre-injective. This last result gives a characterization of amenability in terms of cellular automata.

## 5.1 Garden of Eden Configurations and Garden of Eden Patterns

Let $G$ be a group and let $A$ be a set. Let $\tau\colon A^G \to A^G$ be a cellular automaton. A configuration $y \in A^G$ is called a *Garden of Eden* configuration for $\tau$ if $y$ is not in the image of $\tau$. Thus the surjectivity of $\tau$ is equivalent to the non-existence of Garden of Eden configurations.

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_5, 

The biblical terminology "Garden of Eden" (a peaceful place where we will never return) comes from the fact that one often regards a cellular automaton $\tau\colon A^G \to A^G$ from a dynamical viewpoint. This means that one thinks of a configuration as evolving with time according to $\tau$: if $x \in A^G$ is the configuration at time $t = 0, 1, 2, \ldots$, then $\tau(x)$ is the configuration at time $t + 1$. A configuration $x \in A^G \setminus \tau(A^G)$ is called a Garden of Eden configuration because it may only appear at time $t = 0$.

A pattern $p\colon \Omega \to A$ is called a *Garden of Eden* pattern for $\tau$ if there is no configuration $x \in A^G$ such that $\tau(x)|_\Omega = p$. It follows from this definition that if $p\colon \Omega \to G$ is a Garden of Eden pattern for $\tau$, then any configuration $y \in A^G$ such that $y|_\Omega = p$ is a Garden of Eden configuration for $\tau$. Thus, the existence of a Garden of Eden pattern implies the existence of Garden of Eden configurations, that is, the non-surjectivity of $\tau$. It turns out that the converse is also true when the alphabet set is finite:

**Proposition 5.1.1.** *Let $G$ be a group and let $A$ be a finite set. Let $\tau\colon A^G \to A^G$ be a cellular automaton. Suppose that $\tau$ is not surjective. Then $\tau$ admits a Garden of Eden pattern.*

*Proof.* We know that the set $\tau(A^G)$ is closed in $A^G$ for the prodiscrete topology by Lemma 3.3.2. It follows that the set $A^G \setminus \tau(A^G)$ is open in $A^G$. Therefore, if $y \in A^G$ is a Garden of Eden configuration for $\tau$, we may find a finite subset $\Omega \subset G$ such that

$$V(y, \Omega) = \{x \in A^G : x|_\Omega = y|_\Omega\} \subset A^G \setminus \tau(A^G).$$

In other words, every configuration extending $y|_\Omega$ is not in $\tau(A^G)$, that is, $y|_\Omega$ is a Garden of Eden pattern for $\tau$. □

## 5.2 Pre-injective Maps

Let $G$ be a group and let $A$ be a set.

Two configurations $x_1, x_2 \in A^G$ are called *almost equal* if the set $\{g \in G : x_1(g) \neq x_2(g)\}$ is finite. It is clear that being almost equal defines an equivalence relation on the set $A^G$.

Given a subset $X \subset A^G$ and a set $Z$, a map $f\colon X \to Z$ is called *pre-injective* if it satisfies the following condition: if two configurations $x_1, x_2 \in X$ are almost equal and such that $f(x_1) = f(x_2)$, then $x_1 = x_2$. It immediately follows from this definition that the injectivity of $f$ implies its pre-injectivity. The converse is trivially true when the group $G$ is finite. However, a pre-injective map $f\colon A^G \to Z$ may fail to be injective when $G$ is infinite.

*Examples 5.2.1.* (a) Let us take $G = \mathbb{Z}$ and $A = \mathbb{Z}/3\mathbb{Z}$. Consider the cellular automaton $\tau\colon A^G \to A^G$ defined by $\tau(x)(n) = x(n-1)+x(n)+x(n+1)$ for all

$x \in A^G$ and $n \in G$. Then $\tau$ is pre-injective. Indeed, suppose that $x_1, x_2 \in A^G$ are two configurations such that the set $\Omega = \{n \in G : x_1(n) \neq x_2(n)\}$ is a nonempty finite subset of $\mathbb{Z}$. Let $n_0$ denote the largest element in $\Omega$. Then $\tau(x_1)(n_0 + 1) \neq \tau(x_2)(n_0 + 1)$ and hence $\tau(x_1) \neq \tau(x_2)$. This sows that $\tau$ is pre-injective. However, $\tau$ is not injective since the constant configurations $c_0, c_1 \in A^G$ given by $c_0(n) = 0$ and $c_1(n) = 1$ for all $n \in \mathbb{Z}$ have the same image $c_0$ by $\tau$.

(b) Let $G = \mathbb{Z}^2$ and $A = \{0, 1\}$. Consider the cellular automaton $\tau\colon A^G \to A^G$ associated with the Game of Life (cf. Example 1.4.3(a)). Let $x_1 \in A^G$ be the configuration with no live cells ($x_1(g) = 0$ for all $g \in G$) and let $x_2 \in A^G$ be the configuration with only one live cell at the origin ($x_2(g) = 1$ if $g = (0, 0)$ and $x_2(g) = 0$ otherwise). Then $x_1$ and $x_2$ are almost equal and one has $\tau(x_1) = \tau(x_2) = x_1$. Therefore $\tau$ is not pre-injective.

(c) Let $G$ be a group and let $A = \{0, 1\}$. Let $S$ be a finite subset of $G$ having at least 3 elements. Let $\tau\colon A^G \to A^G$ be the majority action cellular automaton associated with $G$ and $S$ (cf. Example 1.4.3(c)).

Let $x_0 \in A^G$ be the configuration defined by $x_0(g) = 0$ for all $g \in G$. Let $x_1 \in A^G$ be the configuration defined by $x_1(1_G) = 1$ and $x_1(g) = 0$ if $g \neq 1_G$. One has $\tau(x_0) = \tau(x_1) = x_0$ and $\{g \in G : x_0(g) \neq x_1(g)\} = \{1_G\}$. Consequently, $\tau$ is not pre-injective.

The operations of induction and restriction of cellular automata with respect to a subgroup of the underlying group have been introduced in Sect. 1.7. It turns out that pre-injectivity, like injectivity and surjectivity (see Proposition 1.7.4), is preserved by these operations. More precisely, we have the following result (which will not be used in the proof of the Garden of Eden theorem given below):

**Proposition 5.2.2.** *Let $G$ be a group and let $A$ be a set. Let $H$ be a subgroup of $G$ and let $\tau \in \mathrm{CA}(G, H; A)$. Let $\tau_H \in \mathrm{CA}(H; A)$ denote the cellular automaton obtained by restriction of $\tau$ to $H$. Then, $\tau$ is pre-injective if and only if $\tau_H$ is pre-injective.*

*Proof.* First recall from (1.16) the factorizations

$$A^G = \prod_{c \in G/H} A^c \quad \text{and} \quad \tau = \prod_{c \in G/H} \tau_c, \tag{5.1}$$

where $\tau_c\colon A^c \to A^c$ satisfies $\tau_c(\widetilde{x}|_c) = (\tau(\widetilde{x}))|_c$ for all $\widetilde{x} \in A^c$.

Suppose that $\tau$ is pre-injective. We want to show that $\tau_H$ is pre-injective. So let $x, y \in A^H$ be almost equal configurations over $H$ such that $\tau_H(x) = \tau_H(y)$. Let us fix an arbitrary element $a_0 \in A$ and extend $x$ and $y$ to configurations $\widetilde{x}$ and $\widetilde{y}$ in $A^G$ by setting

$$\widetilde{x}(g) = \begin{cases} x(g) & \text{if } g \in H, \\ a_0 & \text{otherwise} \end{cases} \quad \text{and} \quad \widetilde{y}(g) = \begin{cases} y(g) & \text{if } g \in H, \\ a_0 & \text{otherwise} \end{cases}$$

for all $g \in G$. Note that the configurations $\widetilde{x}$ and $\widetilde{y}$ are almost equal since $\{g \in G : \widetilde{x}(g) \neq \widetilde{y}(g)\} = \{h \in H : x(h) \neq y(h)\}$. By construction, $\widetilde{x}|_c = \widetilde{y}|_c$ for all $c \in G/H \setminus \{H\}$, while $\widetilde{x}|_H = x$ and $\widetilde{y}|_H = y$. From (5.1), we deduce that $\tau(\widetilde{x}) = \tau(\widetilde{y})$. It follows that $\widetilde{x} = \widetilde{y}$, by pre-injectivity of $\tau$. This implies that $x = \widetilde{x}|_H$ equals $\widetilde{y}|_H = y$. This shows that $\tau_H$ is pre-injective.

Conversely, suppose that $\tau_H$ is pre-injective. Let $\widetilde{x}, \widetilde{y} \in A^G$ be almost equal configurations over $G$ such that $\tau(\widetilde{x}) = \tau(\widetilde{y})$. For each $g \in G$, consider the configurations $\widetilde{x}_g, \widetilde{y}_g \in A^H$ defined by $\widetilde{x}_g(h) = \widetilde{x}(gh)$ and $\widetilde{y}_g(h) = \widetilde{y}(gh)$ for all $h \in H$. Observe that the configurations $\widetilde{x}_g$ and $\widetilde{y}_g$ are almost equal since $\widetilde{x}$ and $\widetilde{y}$ are almost equal. On the other hand, we have $\widetilde{x}_g = \phi_g^*(\widetilde{x}|_c)$ and $\widetilde{y}_g = \phi_g^*(\widetilde{y}|_c)$, where $c = gH \in G/H$ and $\phi_g^* \colon A^c \to A^H$ is the bijective map defined by $\phi_g^*(u)(h) = u(gh)$ for all $u \in A^c$. As $\tau(\widetilde{x}) = \tau(\widetilde{y})$, we have $\tau_c(\widetilde{x}|_c) = \tau_c(\widetilde{y}|_c)$ by (5.1). We deduce that $\tau_H(\widetilde{x}_g) = \tau_H(\widetilde{y}_g)$ since $\phi_g^*$ conjugates $\tau_c$ and $\tau_H$ by Proposition 1.18. Consequently, we have $\widetilde{x}_g = \widetilde{y}_g$ for all $g \in G$ by the pre-injectivity of $\tau_H$. This implies that $\widetilde{x} = \widetilde{y}$. Therefore, $\tau$ is pre-injective. □

## 5.3 Statement of the Garden of Eden Theorem

The *Garden of Eden theorem* gives a necessary and sufficient condition for the surjectivity of a cellular automaton with finite alphabet over an amenable group. Its name comes from the fact that the surjectivity of a cellular automaton is equivalent to the absence of Garden of Eden configurations (see Sect. 5.1).

**Theorem 5.3.1 (The Garden of Eden theorem).**
*Let $G$ be an amenable group and let $A$ be a finite set. Let $\tau \colon A^G \to A^G$ be a cellular automaton. Then one has*

$$\tau \text{ is surjective} \iff \tau \text{ is pre-injective.}$$

The proof of Theorem 5.3.1 will be given in Sect. 5.8 (see Theorem 5.8.1). Let us first present some applications of this theorem.

*Examples 5.3.2.* (a) Let $G = \mathbb{Z}$ and $A = \mathbb{Z}/3\mathbb{Z}$. We have seen in Example 5.2.1(a) that the cellular automaton $\tau \colon A^G \to A^G$ defined by $\tau(x)(n) = x(n-1) + x(n) + x(n+1)$ is pre-injective. Since $\mathbb{Z}$ is amenable (cf. Theorem 4.6.1), it follows from the Garden of Eden theorem that $\tau$ is surjective. In fact, a direct proof of the surjectivity of $\tau$ is not difficult (see Exercise 5.5).

(b) Let $G = \mathbb{Z}^2$ and $A = \{0, 1\}$. Consider the cellular automaton $\tau \colon A^G \to A^G$ associated with the Game of Life. We have seen in Example 5.2.1(b) that $\tau$ is not pre-injective. The group $\mathbb{Z}^2$ is amenable by Theorem 4.6.1. By applying the Garden of Eden theorem, we deduce that $\tau$ is not surjective (see Sect. 5.13 for a direct proof).

(c) Let $G$ be a group and let $A = \{0, 1\}$. Let $S$ be a finite subset of $G$ having at least 3 elements. Let $\tau\colon A^G \to A^G$ be the majority action cellular automaton associated with $G$ and $S$.

We have seen in Example 5.2.1(c) that $\tau$ is not pre-injective. Thus it follows from the Garden of Eden theorem that if $G$ is amenable then $\tau$ is not surjective. We shall see in Sect. 5.10 that if $G$ is the free group $F_2$, then $\tau$ is surjective.

## 5.4 Interiors, Closures, and Boundaries

Let $G$ be a group.

Let $E$ and $\Omega$ be subsets of $G$. The *$E$-interior* $\Omega^{-E}$ and the *$E$-closure* $\Omega^{+E}$ of $\Omega$ are the subsets of $G$ defined respectively by

$$\Omega^{-E} = \{g \in G : gE \subset \Omega\} \quad \text{and}$$
$$\Omega^{+E} = \{g \in G : gE \cap \Omega \neq \varnothing\}$$

(see Figs. 5.1–5.2). Observe that

$$\Omega^{-E} = \bigcap_{e \in E} \Omega e^{-1} \tag{5.2}$$

and

$$\Omega^{+E} = \bigcup_{e \in E} \Omega e^{-1} = \Omega E^{-1}. \tag{5.3}$$

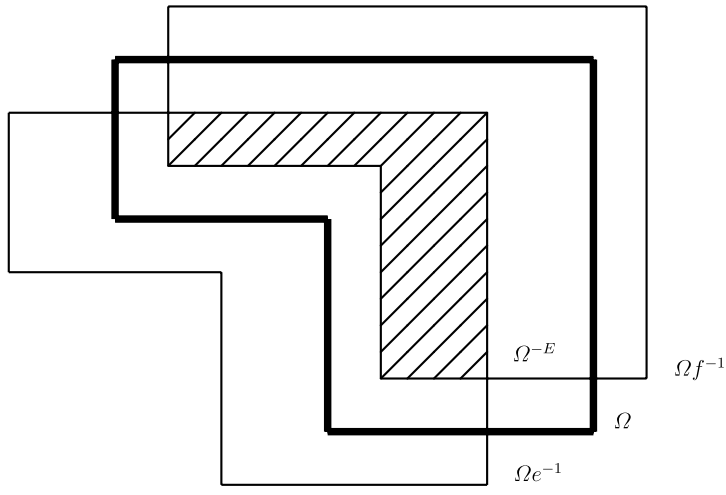



**Fig. 5.1** The $E$-interior $\Omega^{-E}$ of a set $\Omega$. Here, $\Omega \subset \mathbb{R}^2$, $E = \{e, f\}$ with $e = (2, 1)$ and $f = (-1, -1)$

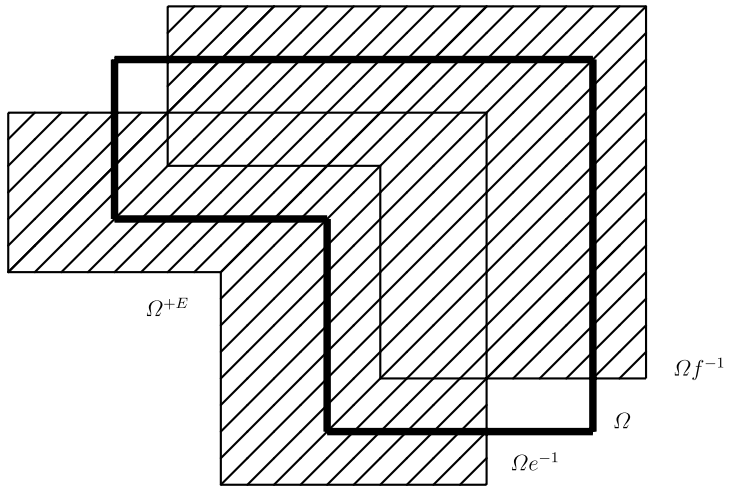



**Fig. 5.2** The $E$-closure $\Omega^{+E}$ of a set $\Omega$. Here, $\Omega \subset \mathbb{R}^2$, $E = \{e, f\}$ with $e = (2, 1)$ and $f = (-1, -1)$

The *E-boundary* of $\Omega$ is the subset $\partial_E(\Omega)$ of $G$ defined by

$$\partial_E(\Omega) = \Omega^{+E} \setminus \Omega^{-E}$$

(see Fig. 5.3).

*Examples 5.4.1.* (a) If $E = \varnothing$, then $\Omega^{-E} = G$ and $\Omega^{+E} = \partial_E(\Omega) = \varnothing$.

(b) If $E = \{1_G\}$, then $\Omega^{-E} = \Omega^{+E} = \Omega$ and $\partial_E(\Omega) = \varnothing$.

(c) If $a \in G$ and $E = \{1_G, a\}$, then $\Omega^{-E} = \Omega \cap \Omega a^{-1}$, $\Omega^{+E} = \Omega \cup \Omega a^{-1}$, so that

$$\partial_E(\Omega) = (\Omega \cup \Omega a^{-1}) \setminus (\Omega \cap \Omega a^{-1}) = (\Omega \setminus \Omega a^{-1}) \cup (\Omega a^{-1} \setminus \Omega) \qquad (5.4)$$

is the symmetric difference between the sets $\Omega$ and $\Omega a^{-1}$ (see Fig. 5.4).

(d) Let us take $G = \mathbb{Z}^2$ and $E = \{-1, 0, 1\}^2$. Let $a, b, c, d \in \mathbb{Z}$ and consider the rectangle $\Omega = [a, b] \times [c, d] = \{(x, y) \in \mathbb{Z}^2 : a \leq x \leq b, c \leq y \leq d\}$. Then one has

$$\Omega^{-E} = [a+1, b-1] \times [c+1, d-1] \quad \text{and} \quad \Omega^{+E} = [a-1, b+1] \times [c-1, d+1]$$

(see Fig. 5.5).

Here are some general properties of the sets $\Omega^{-E}$, $\Omega^{+E}$ and $\partial_E(\Omega)$ which we shall frequently use in the sequel.

**Proposition 5.4.2.** *Let $G$ be a group. Let $E, E_1, E_2$ and $\Omega$ be subsets of $G$. Then the following hold:*

(i) $(G \setminus \Omega)^{-E} = G \setminus \Omega^{+E}$;
(ii) $(G \setminus \Omega)^{+E} = G \setminus \Omega^{-E}$;

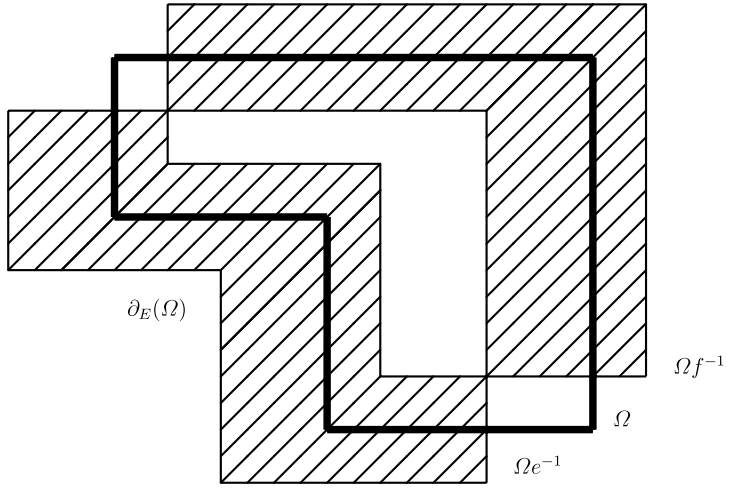



**Fig. 5.3** The $E$-boundary $\partial_E(\Omega)$ of a set $\Omega$. Here, $\Omega \subset \mathbb{R}^2$, $E = \{e, f\}$ with $e = (2, 1)$ and $f = (-1, -1)$

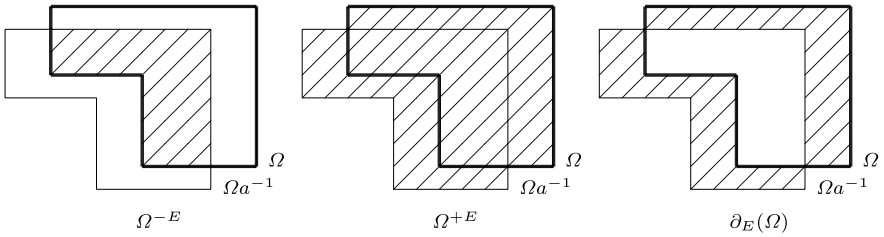



**Fig. 5.4** The $E$-interior $\Omega^{-E} = \Omega \cap \Omega a^{-1}$, the $E$-closure $\Omega^{+E} = \Omega \cup \Omega a^{-1}$, and the $E$-boundary $\partial_E(\Omega) = (\Omega \cup \Omega a^{-1}) \setminus (\Omega \cap \Omega a^{-1})$ of a set $\Omega$ when $E = \{1_G, a\}$

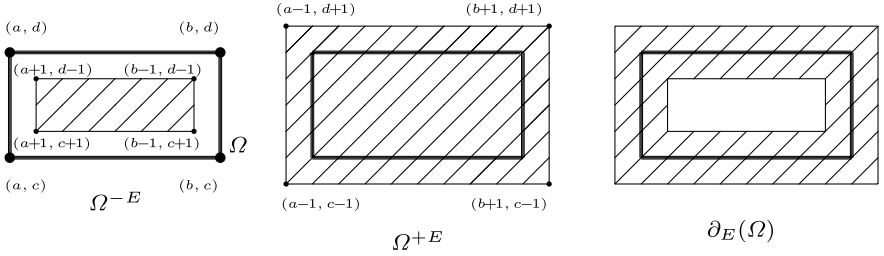



**Fig. 5.5** The $E$-interior $\Omega^{-E}$, the $E$-closure $\Omega^{+E}$, and the $E$-boundary $\partial_E(\Omega)$ of a rectangle $\Omega \subset \mathbb{Z}^2$. Here, $\Omega = [a, b] \times [b, c]$ and $E = \{-1, 0, -1\}^2$

(iii) *if* $a \in E$, *then* $\Omega^{-E} \subset \Omega a^{-1} \subset \Omega^{+E}$;
(iv) *if* $1_G \in E$, *then* $\Omega^{-E} \subset \Omega \subset \Omega^{+E}$;
(v) *if* $E$ *is nonempty and* $\Omega$ *is finite, then* $\Omega^{-E}$ *is finite;*
(vi) *if* $E$ *and* $\Omega$ *are both finite, then* $\Omega^{+E}$ *and* $\partial_E(\Omega)$ *are finite;*

(vii) *if* $E_1 \subset E_2$, *then* $\Omega^{-E_2} \subset \Omega^{-E_1}$, $\Omega^{+E_1} \subset \Omega^{+E_2}$ *and* $\partial_{E_1}(\Omega) \subset \partial_{E_2}(\Omega)$;
(viii) *if* $h \in G$, *then* $h(\Omega^{-E}) = (h\Omega)^{-E}$, $h(\Omega^{+E}) = (h\Omega)^{+E}$ *and* $h(\partial_E(\Omega)) = \partial_E(h\Omega)$.

*Proof.* (i) By definition, we have $g \in G \setminus \Omega^{+E}$ if and only if $gE$ does not meet $\Omega$, that is, if and only if $g \in (G \setminus \Omega)^{-E}$.

(ii) By replacing $\Omega$ by $G \setminus \Omega$ in (i) we get $\Omega^{-E} = G \setminus (G \setminus \Omega)^{+E}$ which gives (ii) after taking complements.

(iii) If $a \in E$, then $\Omega^{-E} \subset \Omega a^{-1}$ by (5.2) and $\Omega a^{-1} \subset \Omega^{+E}$ by (5.3).

(iv) Assertion (iii) gives (iv) by taking $a = 1_G$.

(v) If $a \in E$ and $\Omega$ is finite, then $|\Omega^{-E}| \leq |\Omega a^{-1}| = |\Omega|$ by (iii).

(vi) If $E$ and $\Omega$ are both finite, then $|\Omega^{+E}| \leq |\Omega||E|$ by (5.3) so that $\Omega^{+E}$ is finite. The set $\partial_E(\Omega)$ is then also finite since it is contained in $\Omega^{+E}$.

(vii) The first two statements follow immediately from (5.2) and (5.3), respectively, and imply that

$$\partial_{E_1}(\Omega) = \Omega^{+E_1} \setminus \Omega^{-E_1} \subset \Omega^{+E_2} \setminus \Omega^{-E_1} \subset \Omega^{+E_2} \setminus \Omega^{-E_2} = \partial_{E_2}(\Omega).$$

(viii) By using (5.2) we have

$$h(\Omega^{-E}) = h \cap_{e \in E} \Omega e^{-1} = \cap_{e \in E} h\Omega e^{-1} = (h\Omega)^{-E},$$

which gives the first statement. Similarly, from (5.3) we get

$$h(\Omega^{+E}) = h(\Omega E^{-1}) = (h\Omega)E^{-1} = (h\Omega)^{+E}.$$

Finally, we have

$$h(\partial_E(\Omega)) = h(\Omega^{+E} \setminus \Omega^{-E}) = h\Omega^{-E} - h\Omega^{-E} = \partial_E(h\Omega).$$

□

**Proposition 5.4.3.** *Let* $G$ *be a group and let* $A$ *be a set. Let* $\tau \colon A^G \to A^G$ *be a cellular automaton with memory set* $S$. *Let* $x$ *and* $x'$ *be elements of* $A^G$. *Suppose that there is a subset* $\Omega$ *of* $G$ *such that* $x$ *and* $x'$ *coincide on* $\Omega$ *(resp. on* $G \setminus \Omega$*). Then the configurations* $\tau(x)$ *and* $\tau(x')$ *coincide on* $\Omega^{-S}$ *(resp. on* $G \setminus \Omega^{+S}$*).*

*Proof.* Suppose that $x$ and $x'$ coincide on $\Omega$. If $g \in \Omega^{-S}$, then $gS \subset \Omega$ and therefore $\tau(x)(g) = \tau(x')(g)$ by Lemma 1.4.7. It follows that $\tau(x)$ and $\tau(x')$ coincide on $\Omega^{-S}$.

Suppose now $x$ and $x'$ coincide on $G \setminus \Omega$. Then $\tau(x)$ and $\tau(x')$ coincide on $(G \setminus \Omega)^{-S} = G \setminus \Omega^{+S}$ by the first part of the proof and Proposition 5.4.2(i).

□

**Proposition 5.4.4.** *Let* $G$ *be a group and let* $(F_j)_{j \in J}$ *be a net of nonempty finite subsets of* $G$. *Then the following conditions are equivalent:*

(a) *the net* $(F_j)_{j\in J}$ *is a right Følner net for* $G$*;*
(b) *one has*

$$\lim_j \frac{|\partial_E(F_j)|}{|F_j|} = 0 \quad \textit{for every finite subset } E \subset G.$$

*Proof.* (b) $\Rightarrow$ (a). Suppose (b). Let $g \in G$ and take $E = \{1_G, g^{-1}\}$. By (5.4), we have

$$F_j \setminus F_j g \subset \partial_E(F_j),$$

and hence

$$|F_j \setminus F_j g| \leq |\partial_E(F_j)|.$$

Therefore, property (b) implies

$$\lim_j \frac{|F_j \setminus F_j g|}{|F_j|} = 0.$$

This shows that $(F_j)$ is a right Følner net for $G$.
(a) $\Rightarrow$ (b). Let $E$ be a finite subset of $G$. By (5.2) and (5.3) we have

$$\begin{aligned}
\partial_E(F_j) &= \left(\bigcup_{a\in E} F_j a^{-1}\right) \setminus \left(\bigcap_{b\in E} F_j b^{-1}\right) \\
&= \left(\bigcup_{a\in E} F_j a^{-1}\right) \bigcap \left(G \setminus \bigcap_{b\in E} F_j b^{-1}\right) \\
&= \left(\bigcup_{a\in E} F_j a^{-1}\right) \bigcap \left(\bigcup_{b\in E} (G \setminus F_j b^{-1})\right) \\
&= \bigcup_{a,b\in E} (F_j a^{-1} \setminus F_j b^{-1}).
\end{aligned}$$

This implies

$$|\partial_E(F_j)| \leq \sum_{a,b\in E} |F_j a^{-1} \setminus F_j b^{-1}|. \tag{5.5}$$

Now observe that, for all $a, b \in E$, we have

$$|F_j a^{-1} \setminus F_j b^{-1}| = |F_j \setminus F_j b^{-1} a|,$$

since right multiplication by $a$ is bijective on $G$. Therefore, inequality (5.5) gives us

$$\frac{|\partial_E(F_j)|}{|F_j|} \leq |E|^2 \max_{g\in K} \frac{|F_j \setminus F_j g|}{|F_j|},$$

where $K$ is the finite subset of $G$ defined by $K = \{b^{-1}a : a, b \in E\}$. This shows that (a) implies (b). □

**Corollary 5.4.5.** *Let $G$ be a group. Then the following conditions are equivalent:*

(a) *$G$ is amenable;*
(b) *for every finite subset $E \subset G$ and every real number $\varepsilon > 0$, there exists a nonempty finite subset $F \subset G$ such that*

$$\frac{|\partial_E(F)|}{|F|} < \varepsilon. \tag{5.6}$$

*Proof.* Suppose first that $G$ is amenable. It follows from the Tarski-Følner theorem (Theorem 4.9.1) and Proposition 4.7.1 that there exists a right Følner net $(F_j)_{j\in J}$ in $G$. By Proposition 5.4.4 we have that $\lim_j \frac{|\partial_E(F_j)|}{|F_j|} = 0$ for every finite subset $E \subset G$. Thus, given $\varepsilon > 0$ and a finite subset $S \subset G$, there exists $j_0 \in J$ such that $\frac{|\partial_E(F_j)|}{|F_j|} < \varepsilon$ for all $j \geq j_0$. Taking $F = F_{j_0}$ we deduce (5.6). This shows (a) $\Rightarrow$ (b).

Conversely, suppose (b). Let $J$ denote the set of all pairs $(E, \varepsilon)$, where $E$ is a finite subset of $G$ and $\varepsilon > 0$. We equip $J$ with the partial ordering $\leq$ defined by

$$(E, \varepsilon) \leq (E', \varepsilon') \Leftrightarrow (E \subset E' \text{ and } \varepsilon' \leq \varepsilon).$$

Then $J$ is a directed set. By (b), for every $j = (E, \varepsilon) \in J$, there exists a nonempty finite subset $F_j \subset G$ such that

$$\frac{|\partial_E(F_j)|}{|F_j|} < \varepsilon. \tag{5.7}$$

Let us show that

$$\lim_j \frac{|\partial_E(F_j)|}{|F_j|} = 0 \text{ for every finite subset } E \subset G. \tag{5.8}$$

Fix a finite subset $E_0 \subset G$ and $\varepsilon_0 > 0$. Let $j \in J$ and suppose that $j \geq j_0$, where $j_0 = (E_0, \varepsilon_0) \in J$. By virtue of Proposition 5.4.2(vii) we have $\partial_{E_0}(F_j) \subset \partial_E(F_j)$ so that, from (5.7), we deduce that

$$\frac{|\partial_{E_0}(F_j)|}{|F_j|} \leq \frac{|\partial_E(F_j)|}{|F_j|} < \varepsilon \leq \varepsilon_0.$$

This shows (5.8). From Proposition 5.4.4 and Proposition 4.7.1 we deduce that $G$ satisfies the Følner conditions. Thus $G$ is amenable by virtue of the Tarski-Følner theorem (Theorem 4.9.1). □

## 5.5 Mutually Erasable Patterns

In this section, we give a characterization of pre-injective cellular automata based on the notion of mutually erasable patterns. This leads to an equivalent definition of pre-injectivity which is frequently used in the literature. However, the material contained in this section will not be used in the proof of the Garden of Eden theorem so that the reader who is only interested in this proof may go directly to the next section.

Let $G$ be a group and let $A$ be a set.

Let $Z$ be a set and let $f\colon A^G \to Z$ be a map. Two distinct patterns $p_1, p_2\colon \Omega \to Z$ with the same support $\Omega \subset G$ are called *mutually erasable* (with respect to $f$) if they satisfy the following condition: if $x_1, x_2 \in A^G$ are configurations such that $x_1|_\Omega = p_1$, $x_2|_\Omega = p_2$ and $x_1|_{G\setminus\Omega} = x_2|_{G\setminus\Omega}$, then $f(x_1) = f(x_2)$.

*Example 5.5.1.* Let $G = \mathbb{Z}^2$ and $A = \{0, 1\}$. Consider the cellular automaton $\tau\colon A^G \to A^G$ associated with the Game of Life (see Example 1.4.3(a)). Let $\Omega = \{-1, 0, 1\}^2$ be the $3 \times 3$ square in $\mathbb{Z}^2$ centered at the origin and consider the pattern $p_1$ (resp. $p_2$) with support $\Omega$ defined by $p_1(g) = 0$ for all $g \in \Omega$ (resp. $p_2(g) = 1$ if $g = (0, 0)$ and $p_2(g) = 0$ otherwise). Then it is clear that $p_1$ and $p_2$ are mutually erasable patterns for $\tau$ (see Fig. 5.6).

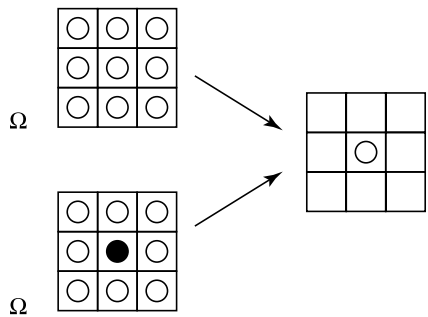



**Fig. 5.6** Two mutually erasable patterns for the Game of Life (recall that ∘ denotes a dead cell, while • denotes a live cell)

Suppose that $p_1$ and $p_2$ are mutually erasable patterns for a map $f\colon A^G \to Z$. Let $\Omega$ denote their common support and consider two configurations $x_1$ and $x_2$ which coincide outside $\Omega$ and such that $x_1|_\Omega = p_1$ and $x_2|_\Omega = p_2$. Then $x_1$ and $x_2$ are almost equal and $f(x_1) = f(x_2)$. On the other hand, $x_1 \neq x_2$ since $p_1 \neq p_2$ and therefore $f$ is not pre-injective. This shows that a pre-injective map $f\colon A^G \to Z$ admits no mutually erasable patterns. It turns out that for cellular automata, the converse is true:

**Proposition 5.5.2.** *Let $G$ be a group and let $A$ be a set. Let $\tau\colon A^G \to A^G$ be a cellular automaton. Then $\tau$ is pre-injective if and only if it does not admit mutually erasable patterns.*

*Proof.* It remains only to show that if $\tau$ is not pre-injective then it admits mutually erasable patterns. So let us assume that $\tau$ is not pre-injective. This means that there exist configurations $x_1, x_2 \in A^G$ satisfying $\tau(x_1) = \tau(x_2)$ such that $\Sigma = \{g \in G : x_1(g) \neq x_2(g)\}$ is a nonempty finite subset of $G$. Let $S$ be a memory set for $\tau$ such that $S = S^{-1}$ and $1_G \in S$. Consider the finite sets $S^2 = \{s_1 s_2 : s_1, s_2 \in S\}$ and $\Omega = \Sigma^{+S^2}$. Let us show that the patterns $p_1 = x_1|_\Omega$ and $p_2 = x_2|_\Omega$ are mutually erasable. First observe that $p_1 \neq p_2$ since $\Sigma \subset \Omega$. Suppose now that $y_1, y_2 \in A^G$ are two configurations coinciding outside $\Omega$ such that $y_1|_\Omega = p_1$ and $y_2|_\Omega = p_2$. Then $y_1$ and $y_2$ coincide outside $\Sigma$ since $p_1$ and $p_2$ coincide on $\Omega \setminus \Sigma$. This implies

$$\tau(y_1)(g) = \tau(y_2)(g) \text{ for all } g \in G \setminus \Sigma^{+S} \tag{5.9}$$

by Proposition 5.4.3. On the other hand, for $i = 1, 2$, the configurations $y_i$ and $x_i$ coincide on $\Omega$ by construction. Thus, it follows from Proposition 5.4.3 that $\tau(y_i)$ and $\tau(x_i)$ coincide on $\Omega^{-S}$. As $\tau(x_1) = \tau(x_2)$, we deduce that the configurations $\tau(y_1)$ and $\tau(y_2)$ coincide on $\Omega^{-S}$. Now observe that $\Sigma^{+S} \subset \Omega^{-S}$. Indeed, if $g \in \Sigma^{+S}$, that is, $gs_0 \in \Sigma$ for some $s_0 \in S$, then $gs \in \Omega$ for all $s \in S$ since $gss^{-1}s_0 = gs_0 \in gS^2 \cap \Sigma$. It follows that $\tau(y_1)$ and $\tau(y_2)$ coincide on $\Sigma^{+S}$. Combining this with (5.9), we conclude that $\tau(y_1) = \tau(y_2)$. This sows that $p_1$ and $p_2$ are mutually erasable patterns. □

## 5.6 Tilings

Let $G$ be a group.

Let $E$ and $E'$ be subsets of $G$. A subset $T \subset G$ is called an $(E, E')$-*tiling* of $G$ if the sets $gE$, $g \in T$ are pairwise disjoint and if the sets $gE'$, $g \in T$ cover $G$. In other words, $T \subset G$ is an $(E, E')$-tiling if and only if the following conditions are satisfied:

(T1) $g_1E \cap g_2E = \varnothing$ for all $g_1, g_2 \in T$ such that $g_1 \neq g_2$;
(T2) $G = \bigcup_{g \in T} gE'$.

*Examples 5.6.1.* (a) In the additive group $\mathbb{R}$, the set $\mathbb{Z}$ is a $([0, 1[, [0, 1[)$-tiling and the set $[0, 1]$ is a $(2\mathbb{Z}, \mathbb{Z})$-tiling.

(b) If $G$ is a group and $H$ is a subgroup of $G$, then every complete set of representatives for left cosets of $G$ modulo $H$ is an $(H, H)$-tiling.

*Remark 5.6.2.* Let $G$ be a group and let $E$ and $E'$ be subsets of $G$. If $T$ is an $(E, E')$-tiling of $G$ and if $E_1$ and $E_1'$ are subsets of $G$ such that $E_1 \subset E$ and $E' \subset E_1'$, then it is clear that $T$ is also an $(E_1, E_1')$-tiling of $G$.

The Zorn lemma may be used to prove the existence of $(E, E')$-tilings for any subset $E$ of $G$ and for $E' \subset G$ "large enough". More precisely, we have the following:

**Proposition 5.6.3.** *Let $G$ be a group and let $E$ be a nonempty subset of $G$. Let $E' = \{g_1 g_2^{-1} : g_1, g_2 \in E\}$. Then there is an $(E, E')$-tiling $T \subset G$.*

*Proof.* Consider the set $\mathcal{S}$ consisting of all subsets $S \subset G$ such that the sets $(gE)_{g\in S}$ are pairwise disjoint. Observe that $\mathcal{S}$ is not empty since $\{1_G\} \in \mathcal{S}$. On the other hand, the set $\mathcal{S}$, partially ordered by inclusion, is inductive. Indeed, if $\mathcal{S}'$ is a totally ordered subset of $\mathcal{S}$, then the set $M = \bigcup_{S\in\mathcal{S}'} S$ belongs to $\mathcal{S}$ and is an upper bound for $\mathcal{S}'$. By applying Zorn's lemma, we deduce that $\mathcal{S}$ admits a maximal element $T$. The sets $(gE)_{g\in T}$ are pairwise disjoint since $T \in \mathcal{S}$. On the other hand, consider an arbitrary element $h \in G$. By maximality of $T$, we can find $g \in T$ such that the set $hE$ meets $gE$. This implies $h \in gE'$. This shows that the sets $(gE')_{g\in T}$ cover $G$. Consequently, $T$ is an $(E, E')$-tiling of $G$. □

**Proposition 5.6.4.** *Let $G$ be an amenable group and let $(F_j)_{j\in J}$ be a right Følner net for $G$. Let $E$ and $E'$ be finite subsets of $G$ and suppose that $T \subset G$ is an $(E, E')$-tiling of $G$. Let us set, for each $j \in J$,*

$$T_j = T \cap F_j^{-E} = \{g \in T : gE \subset F_j\}.$$

*Then there exist a real number $\alpha > 0$ and an element $j_0 \in J$ such that*

$$|T_j| \geq \alpha |F_j| \quad \text{for all } j \geq j_0.$$

*Proof.* After possibly replacing $E'$ by $E \cup E'$, we can assume that $E \subset E'$. Let us set $T_j^+ = T \cap F_j^{+E'} = \{g \in T : gE' \cap F_j \neq \varnothing\}$. As the sets $gE'$, $g \in T_j^+$, cover $F_j$, we have $|F_j| \leq |T_j^+| \cdot |E'|$, which gives

$$\frac{|T_j^+|}{|F_j|} \geq \frac{1}{|E'|} \tag{5.10}$$

for all $j \in J$. Observe now that

$$\begin{aligned} T_j^+ \setminus T_j &= \left(T \cap F_j^{+E'}\right) \setminus \left(T \cap F_j^{-E}\right) \\ &= T \cap (F_j^{+E'} \setminus F_j^{-E}) \\ &\subset T \cap (F_j^{+E'} \setminus F_j^{-E'}) \\ &\subset T \cap \partial_{E'}(F_j) \\ &\subset \partial_{E'}(F_j) \end{aligned}$$

where the first inclusion follows from $E \subset E'$. Thus we have

$$|\partial_{E'}(F_j)| \geq |T_j^+| - |T_j|.$$

Using (5.10), we then deduce

$$\frac{|T_j|}{|F_j|} \geq \frac{|T_j^+|}{|F_j|} - \frac{|\partial_{E'}(F_j)|}{|F_j|} \geq \frac{1}{|E'|} - \frac{|\partial_{E'}(F_j)|}{|F_j|}.$$

Hence we have

$$\frac{|T_j|}{|F_j|} \geq \alpha = \frac{1}{2|E'|}$$

for $j$ large enough, by Proposition 5.4.4 (see Fig. 5.7). □

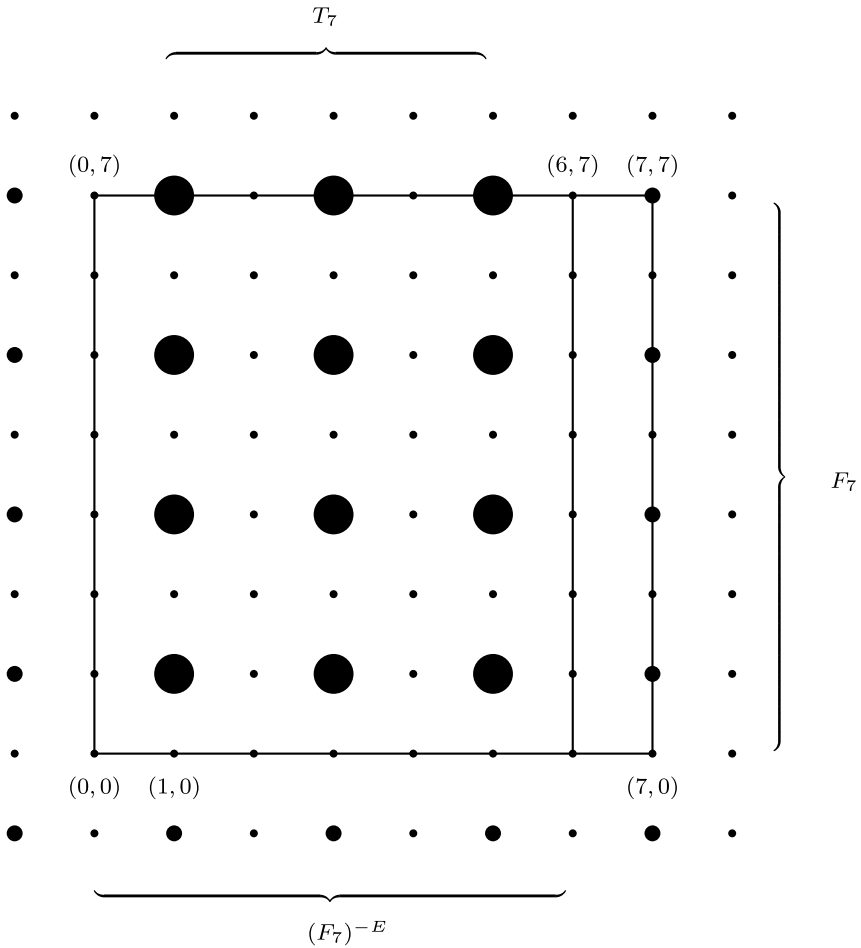



**Fig. 5.7** The set $T_7 = T \cap F_7^{-E} \subset \mathbb{Z}^2$, where $E = \{(0,0),(1,0)\}$, $E' = \{0,1\} \times \{0,1\}$ and $T = (2\mathbb{Z}+1) \times (2\mathbb{Z}+1) \subset \mathbb{Z}^2$ is an $(E,E')$-tiling in $\mathbb{Z}^2$. Note that $\frac{|T_7|}{|F_7|} = \frac{12}{64} = \frac{3}{16} \geq \frac{2}{16} = \frac{1}{8} = \frac{1}{2|E'|} = \alpha$

## 5.7 Entropy

In this section, $G$ is an amenable group, $\mathcal{F} = (F_j)_{j\in J}$ is a right Følner net for $G$, and $A$ is a finite set.

For $E \subset G$, we denote by $\pi_E \colon A^G \to A^E$ the canonical projection (restriction map). We thus have $\pi_E(x) = x|_E$ for all $x \in A^G$.

**Definition 5.7.1.** Let $X \subset A^G$. The *entropy* $\mathrm{ent}_{\mathcal{F}}(X)$ of $X$ with respect to the right Følner net $\mathcal{F} = (F_j)_{j\in J}$ is defined by

$$\mathrm{ent}_{\mathcal{F}}(X) = \limsup_j \frac{\log |\pi_{F_j}(X)|}{|F_j|}.$$

Here are some immediate properties of entropy.

**Proposition 5.7.2.** *One has*

(i) $\mathrm{ent}_{\mathcal{F}}(A^G) = \log |A|$;
(ii) $\mathrm{ent}_{\mathcal{F}}(X) \leq \mathrm{ent}_{\mathcal{F}}(Y)$ *if* $X \subset Y \subset A^G$;
(iii) $\mathrm{ent}_{\mathcal{F}}(X) \leq \log |A|$ *for all* $X \subset A^G$.

*Proof.* (i) If $X = A^G$, then, for every $j$, we have $\pi_{F_j}(X) = A^{F_j}$ and therefore

$$\frac{\log |\pi_{F_j}(X)|}{|F_j|} = \frac{\log |A|^{|F_j|}}{|F_j|} = \frac{|F_j| \log |A|}{|F_j|} = \log |A|.$$

Thus we have $\mathrm{ent}_{\mathcal{F}}(X) = \log |A|$.

(ii) If $X \subset Y$, then $\pi_{F_j}(X) \subset \pi_{F_j}(Y)$ and hence $|\pi_{F_j}(X)| \leq |\pi_{F_j}(Y)|$ for all $j$. This implies $\mathrm{ent}_{\mathcal{F}}(X) \leq \mathrm{ent}_{\mathcal{F}}(Y)$.

(iii) This follows immediately from (i) and (ii). □

An important property of cellular automata is the fact that applying a cellular automaton to a set of configurations cannot increase the entropy of the set. More precisely, we have the following:

**Proposition 5.7.3.** *Let* $\tau \colon A^G \to A^G$ *be a cellular automaton and let* $X \subset A^G$. *Then one has*

$$\mathrm{ent}_{\mathcal{F}}(\tau(X)) \leq \mathrm{ent}_{\mathcal{F}}(X).$$

*Proof.* Let $Y = \tau(X)$. Let $S \subset G$ be a memory set for $\tau$. After replacing $S$ by $S \cup \{1_G\}$, we can assume that $1_G \in S$. Let $\Omega$ be a finite subset of $G$. Observe first that $\tau$ induces a map

$$\tau_\Omega \colon \pi_\Omega(X) \to \pi_{\Omega^{-S}}(Y)$$

defined as follows. If $u \in \pi_\Omega(X)$, then

$$\tau_\Omega(u) = (\tau(x))|_{\Omega^{-S}},$$

where $x$ is an element of $X$ such that $x|_\Omega = u$. Note that the fact that $\tau_\Omega(u)$ does not depend on the choice of such an $x$ follows from Proposition 5.4.3.

Clearly $\tau_\Omega$ is surjective. Indeed, if $v \in \pi_{\Omega^{-S}}(Y)$, then there exists $x \in X$ such that $(\tau(x))|_{\Omega^{-S}} = v$. Then, setting $u = \pi_\Omega(x)$ we have, by construction, $\tau_\Omega(u) = v$. Therefore, we have

$$|\pi_{\Omega^{-S}}(Y)| \leq |\pi_\Omega(X)|. \tag{5.11}$$

Observe now that $\Omega^{-S} \subset \Omega$, since $1_G \in S$ (cf. Proposition 5.4.2(iv)). Thus $\pi_\Omega(Y) \subset \pi_{\Omega^{-S}}(Y) \times A^{\Omega \setminus \Omega^{-S}}$. This implies

$$\begin{aligned}
\log|\pi_\Omega(Y)| &\leq \log|\pi_{\Omega^{-S}}(Y) \times A^{\Omega\setminus\Omega^{-S}}| \\
&= \log|\pi_{\Omega^{-S}}(Y)| + \log|A^{\Omega\setminus\Omega^{-S}}| \\
&= \log|\pi_{\Omega^{-S}}(Y)| + |\Omega \setminus \Omega^{-S}| \log|A| \\
&\leq \log|\pi_\Omega(X)| + |\Omega \setminus \Omega^{-S}| \log|A|,
\end{aligned}$$

by (5.11). As $\Omega \setminus \Omega^{-S} \subset \partial_S(\Omega)$, we deduce that

$$\log|\pi_\Omega(Y)| \leq \log|\pi_\Omega(X)| + |\partial_S(\Omega)| \log|A|.$$

By taking $\Omega = F_j$, this gives us

$$\frac{\log|\pi_{F_j}(Y)|}{|F_j|} \leq \frac{\log|\pi_{F_j}(X)|}{|F_j|} + \frac{|\partial_S(F_j)|}{|F_j|} \log|A|.$$

Since

$$\lim_j \frac{|\partial_S(F_j)|}{|F_j|} = 0$$

by Proposition 5.4.4, we finally get

$$\mathrm{ent}_{\mathcal{F}}(Y) = \limsup_j \frac{\log|\pi_{F_j}(Y)|}{|F_j|} \leq \limsup_j \frac{\log|\pi_{F_j}(X)|}{|F_j|} = \mathrm{ent}_{\mathcal{F}}(X).$$

□

It follows from Proposition 5.7.2 that the maximal value for the entropy of a subset $X \subset A^G$ is $\log|A|$. The following result gives a sufficient condition on $X$ which implies that its entropy is strictly less than $\log|A|$.

**Proposition 5.7.4.** *Let $X \subset A^G$. Suppose that there exist finite subsets $E$ and $E'$ of $G$ and an $(E, E')$-tiling $T \subset G$ such that $\pi_{gE}(X) \subsetneq A^{gE}$ for all $g \in T$. Then one has $\mathrm{ent}_{\mathcal{F}}(X) < \log|A|$.*

*Proof.* For each $j \in J$, let us define, as in Proposition 5.6.4 (see Fig. 5.7), the subset $T_j \subset T$ by $T_j = T \cap F_j^{-E} = \{g \in T : gE \subset F_j\}$ and set

$$F_j^* = F_j \setminus \coprod_{g \in T_j} gE$$

(see Fig. 5.8). By hypothesis, we have

$$|\pi_{gE}(X)| \leq |A^{gE}| - 1 = |A|^{|gE|} - 1 \quad \text{for all } g \in T. \tag{5.12}$$

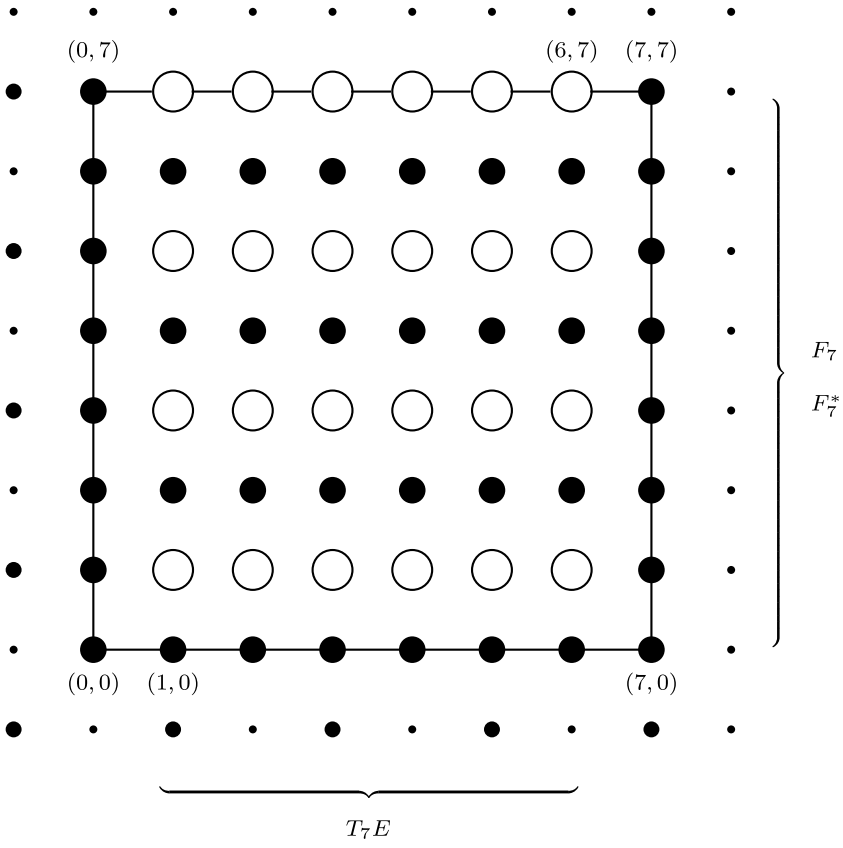



**Fig. 5.8** The set $F_7^* = F_7 \setminus \coprod_{g\in T_7} gE = F_7 \setminus T_7E \subset \mathbb{Z}^2$, where $E$ and $T$ are as in Fig. 5.7 (note that here $F_7$ is the whole square, while $F_7^* \subset F_7$ consists of its •-points)

As

$$\pi_{F_j}(X) \subset A^{F_j^*} \times \prod_{g\in T_j} \pi_{gE}(X),$$

we get

$$\begin{aligned}
\log|\pi_{F_j}(X)| &\le \log\Big|A^{F_j^*} \times \prod_{g\in T_j} \pi_{gE}(X)\Big| \\
&= |F_j^*|\log|A| + \sum_{g\in T_j} \log|\pi_{gE}(X)| \\
&\le |F_j^*|\log|A| + \sum_{g\in T_j} \log(|A|^{|gE|} - 1) \quad \text{(by (5.12))} \\
&= |F_j^*|\log|A| + \sum_{g\in T_j} |gE|\log|A| + \sum_{g\in T_j} \log(1 - |A|^{-|gE|}) \\
&= |F_j|\log|A| + |T_j|\log(1 - |A|^{-|E|}),
\end{aligned}$$

since

$$|F_j| = |F_j^*| + \sum_{g \in T_j} |gE| \quad \text{and} \quad |gE| = |E|.$$

By setting $c = -\log(1 - |A|^{-|E|})$ (note that $c > 0$), this gives us

$$\log |\pi_{F_j}(X)| \leq |F_j| \log |A| - c|T_j| \quad \text{for all } j \in J.$$

Now, by Proposition 5.6.4, there exist $\alpha > 0$ and $j_0 \in J$ such that $|T_j| \geq \alpha |F_j|$ for all $j \geq j_0$. Thus

$$\frac{\log |\pi_{F_j}(X)|}{|F_j|} \leq \log |A| - c\alpha \quad \text{for all } j \geq j_0.$$

This implies that

$$\operatorname{ent}_{\mathcal{F}} X = \limsup_j \frac{\log |\pi_{F_j}(X)|}{|F_j|} \leq \log |A| - c\alpha < \log |A|.$$

□

Recall from Sect. 1.1 that $G$ acts on the left on $A^G$ by the shift $(g, x) \mapsto gx$ defined by $gx(g') = x(g^{-1}g')$ for $g, g' \in G$ and $x \in A^G$.

**Corollary 5.7.5.** *Let $X$ be a $G$-invariant subset of $A^G$. Suppose that there exists a finite subset $E \subset G$ such that $\pi_E(X) \subsetneqq A^E$. Then one has $\operatorname{ent}_{\mathcal{F}}(X) < \log |A|$.*

*Proof.* Let $E' = \{g_1 g_2^{-1} : g_1, g_2 \in E\}$. By Proposition 5.6.3, we may find an $(E, E')$-tiling $T \subset G$. Since $\pi_E(X) \subsetneqq A^E$ and $X$ is $G$-invariant, we have $\pi_{gE}(X) \subsetneqq A^{gE}$ for all $g \in G$. This implies $\operatorname{ent}_{\mathcal{F}}(X) < \log |A|$ by Proposition 5.7.4. □

## 5.8 Proof of the Garden of Eden Theorem

The purpose of this section is to establish the following:

**Theorem 5.8.1.** *Let $G$ be an amenable group and let $A$ be a finite set. Let $\mathcal{F} = (F_j)_{j \in J}$ be a right Følner net for $G$. Let $\tau \colon A^G \to A^G$ be a cellular automaton. Then the following conditions are equivalent:*

(a) *$\tau$ is surjective;*
(b) *$\operatorname{ent}_{\mathcal{F}}(\tau(A^G)) = \log |A|$;*
(c) *$\tau$ is pre-injective.*

Note that this will prove the Garden of Eden theorem (Theorem 5.3.1) since the Garden of Eden theorem asserts the equivalence of conditions (a) and (c) in Theorem 5.8.1.

We divide the proof of Theorem 5.8.1 into several lemmas. In these lemmas, it is assumed that the hypotheses of Theorem 5.8.1 are satisfied: $G$ is an amenable group, $\mathcal{F} = (F_j)_{j\in J}$ is a right Følner net for $G$, $A$ is a finite set, and $\tau\colon A^G \to A^G$ is a cellular automaton.

**Lemma 5.8.2.** *Suppose that $\tau$ is not surjective. Then one has* $\mathrm{ent}_{\mathcal{F}}(\tau(A^G)) < \log|A|$.

*Proof.* By Proposition 5.1.1, $\tau$ admits a Garden of Eden pattern. This means that there is a finite subset $E \subset G$ such that $\pi_E(\tau(A^G)) \subsetneqq A^E$. The set $\tau(A^G)$ is $G$-invariant since $\tau$ is $G$-equivariant by Proposition 1.4.4. We deduce that $\mathrm{ent}_{\mathcal{F}}(\tau(A^G)) < \log|A|$ by applying Corollary 5.7.5. □

**Lemma 5.8.3.** *Suppose that*

$$\mathrm{ent}(\tau(A^G)) < \log|A|. \tag{5.13}$$

*Then $\tau$ is not pre-injective.*

*Proof.* Let $S$ be a memory set for $\tau$ such that $1_G \in S$. Let $Y = \tau(A^G)$. We have $F_j^{-S} \subset F_j \subset F_j^{+S}$ by Proposition 5.4.2(iv) and therefore $F_j^{+S} \setminus F_j \subset \partial_S(F_j)$. As $\pi_{F_j^{+S}}(Y) \subset \pi_{F_j}(Y) \times A^{F_j^{+S}\setminus F_j}$, it follows that

$$\begin{aligned}\log|\pi_{F_j^{+S}}(Y)| &\le \log|\pi_{F_j}(Y)| + |F_j^{+S}\setminus F_j|\log|A| \\ &\le \log|\pi_{F_j}(Y)| + |\partial_S(F_j)|\log|A|.\end{aligned}$$

This implies

$$\frac{\log|\pi_{F_j^{+S}}(Y)|}{|F_j|} \le \frac{\log|\pi_{F_j}(Y)|}{|F_j|} + \frac{|\partial_S(F_j)|}{|F_j|}\log|A|. \tag{5.14}$$

As

$$\mathrm{ent}(Y) = \limsup_j \frac{\log|\pi_{F_j}(Y)|}{|F_j|} < \log|A|$$

by hypothesis, and

$$\lim_j \frac{|\partial_S(F_j)|}{|F_j|} = 0$$

by Proposition 5.4.4, we deduce from inequality (5.14) that there exists $j_0 \in J$ such that

$$\frac{\log|\pi_{F_{j_0}^{+S}}(Y)|}{|F_{j_0}|} < \log|A|. \tag{5.15}$$

Let us fix an arbitrary element $a_0 \in A$ and denote by $Z$ the finite set of configurations $z \in A^G$ such that $z(g) = a_0$ for all $g \in G\setminus F_{j_0}$. Inequality (5.15) gives us

$$|\pi_{F_{j_0}^{+S}}(Y)| < |A|^{|F_{j_0}|} = |Z|.$$

Observe that $\tau(z_1)$ and $\tau(z_2)$ coincide outside $F_{j_0^{+S}}$ for all $z_1, z_2 \in Z$. Thus

$$|\tau(Z)| = |\pi_{F_{j_0^{+S}}}(\tau(Z))| \leq |\pi_{F_{j_0^{+S}}}(Y)| < |Z|.$$

This implies that we may find distinct configurations $z_1, z_2 \in Z$ such that $\tau(z_1)$ and $\tau(z_2)$. Since $z_1$ and $z_2$ coincide outside the finite set $F_{j_0}$, this shows that $\tau$ is not pre-injective. □

**Lemma 5.8.4.** *Suppose that $\tau$ is not pre-injective. Then one has*

$$\mathrm{ent}_{\mathcal{F}}(\tau(A^G)) < \log|A|. \tag{5.16}$$

*Proof.* Since $\tau$ is not pre-injective, we may find two configurations $x_1, x_2 \in A^G$ satisfying $\tau(x_1) = \tau(x_2)$ such that the set

$$\Omega = \{g \in G : x_1(g) \neq x_2(g)\}$$

is a nonempty finite subset of $G$. Observe that, for each $h \in G$, the configurations $hx_1$ and $hx_2$ satisfy $\tau(hx_1) = \tau(hx_2)$ (since $\tau$ is $G$-equivariant by Proposition 1.4.4) and $\{g \in G : hx_1(g) \neq hx_2(g)\} = h\Omega$.

Let $S$ be a memory set for $\tau$ such that $1_G \in S$. Then the set

$$R = \{s^{-1}s' : s, s' \in S\}$$

is finite and we have $1_G \in R$. Let $E = \Omega^{+R}$. By Proposition 5.6.3, we may find a finite subset $E' \subset G$ and an $(E, E')$-tiling $T \subset G$. Consider the subset $Z \subset A^G$ consisting of all configurations $z \in A^G$ such that

$$z|_{hE} \neq (hx_1)|_{hE} \qquad \text{for all } h \in T.$$

Observe that, for each $h \in T$, we have

$$\pi_{hE}(Z) \subsetneqq A^{hE}$$

since $(hx_1)|_{hE} \notin \pi_{hE}(Z)$. We deduce that $\mathrm{ent}_{\mathcal{F}}(Z) < \log|A|$ by applying Proposition 5.7.4. As $\mathrm{ent}_{\mathcal{F}}(\tau(Z)) \leq \mathrm{ent}_{\mathcal{F}}(Z)$ by Proposition 5.7.3, this implies

$$\mathrm{ent}_{\mathcal{F}}(\tau(Z)) < \log|A|. \tag{5.17}$$

Thus, to establish inequality (5.16), it suffices to prove that $\tau(A^G) = \tau(Z)$. To see this, consider an arbitrary configuration $x \in A^G$ and let us show that there is a configuration $z \in Z$ such that $\tau(x) = \tau(z)$. Let

$$T' = \{h \in T : x|_{hE} = (hx_1)|_{hE}\}.$$

Let $z \in A^G$ be the configuration defined by

$$z(g) = \begin{cases} hx_2(g) & \text{if there is } h \in T' \text{ such that } g \in hE, \\ x(g) & \text{otherwise.} \end{cases}$$

Notice that the configuration $z$ is obtained from $x$ by modifying the values taken by $x$ only on the subsets of the form $h\Omega$, where $h \in T'$, (since, as we have seen above, $hx_1$ and $hx_2$ coincide outside $h\Omega$).

By construction, we have $z \in Z$.

Let $g \in G$. Let us shows that $\tau(x)(g) = \tau(z)(g)$.

Suppose first that $gS$ does not meet any of the sets $h\Omega$, $h \in T'$. Then we have $z|_{gS} = x|_{gS}$. We deduce that $\tau(z)(g) = \tau(x)(g)$ by applying Lemma 1.4.7.

Suppose now that there is an element $h \in T'$ such that $gS$ meets $h\Omega$. This means that there exists an element $s_0 \in S$ such that $gs_0 \in h\Omega$. For each $s \in S$, we have $gss^{-1}s_0 = gs_0 \in h\Omega$. As $s^{-1}s_0 \in R$, this implies

$$\begin{aligned} gs \in (h\Omega)^{+R} &= h\Omega^{+R} \quad \text{(by Proposition 5.4.2(viii))} \\ &= hE. \end{aligned}$$

We deduce that $gS \subset hE$. Thus we have $\tau(x)(g) = \tau(hx_1)(g)$ since $x|_{hE} = hx_1|_{hE}$. Similarly, by applying Lemma 1.4.7, we get $\tau(z)(g) = \tau(hx_2)(g)$, since $z$ and $hx_2$ coincide on $hE$. As $\tau(hx_1) = \tau(hx_2)$, we deduce that $\tau(x)(g) = \tau(z)(g)$.

Thus $\tau(z) = \tau(x)$. This shows that $\tau(A^G) = \tau(Z)$ and completes the proof of the lemma. □

*Proof of Theorem 5.8.1.* If $\tau$ is surjective, then $\tau(A^G) = A^G$ and hence $\mathrm{ent}_{\mathcal{F}}(\tau(A^G)) = \mathrm{ent}_{\mathcal{F}}(A^G) = \log|A|$. Thus (a) implies (b). Since the converse implication follows from Lemma 5.8.2, we deduce that conditions (a) and (b) are equivalent. The fact that (c) implies (b) follows from Lemma 5.8.3 and the converse implication follows from Lemma 5.8.4. Thus, conditions (b) and (c) are also equivalent. □

## 5.9 Surjunctivity of Locally Residually Amenable Groups

The notion of a residually finite group was introduced in Chap. 2. More generally, if $\mathcal{P}$ is a property of groups, a group $G$ is called *residually* $\mathcal{P}$ if for each element $g \in G$ with $g \neq 1_G$, there exist a group $\Gamma$ satisfying $\mathcal{P}$ and an epimorphism $\phi\colon G \to \Gamma$ such that $\phi(g) \neq 1_\Gamma$. Observe that every group which satisfies $\mathcal{P}$ is residually $\mathcal{P}$.

According to the preceding definition, a group $G$ is called *residually amenable* if for each element $g \in G$ with $g \neq 1_G$, there exist an amenable group $\Gamma$ and an epimorphism $\phi\colon G \to \Gamma$ such that $\phi(g) \neq 1_\Gamma$. Note that, as every subgroup of an amenable group is amenable, it is not necessary to require that the homomorphism $\phi$ is surjective in this definition. Observe also that every subgroup of a residually amenable group is residually amenable and that the fact that every finite group is amenable (Proposition 4.4.6) implies that every residually finite group is residually amenable.

**Theorem 5.9.1.** *Every residually amenable group is surjunctive.*

Let us first establish the following:

**Lemma 5.9.2.** *Let $G$ be a residually amenable group and let $\Omega$ be a finite subset of $G$. Then there exist an amenable group $\Gamma$ and a homomorphism $\rho\colon G \to \Gamma$ such that the restriction of $\rho$ to $\Omega$ is injective.*

*Proof.* Consider the finite subset $S \subset G$ defined by

$$S = \{g^{-1}h : g, h \in \Omega \text{ and } g \neq h\}.$$

Since $G$ is residually amenable, we can find, for each $s \in S$, an amenable group $\Lambda_s$ and a homomorphism $\phi_s\colon G \to \Lambda_s$ such that $\phi_s(s) \neq 1_{\Lambda_s}$. Let us show that the group

$$\Gamma = \prod_{s \in S} \Lambda_s$$

and the homomorphism $\rho\colon G \to \Gamma$ given by

$$\rho = \prod_{s \in S} \phi_s$$

have the required properties. The fact that the group $\Gamma$ is amenable follows from Corollary 4.5.6. On the other hand, suppose that $g$ and $h$ are distinct elements of $\Omega$. Then $s = g^{-1}h \in S$ and $\phi_s(g) \neq \phi_s(h)$ since $(\phi_s(g))^{-1}\phi_s(h) = \phi_s(g^{-1}h) = \phi_s(s) \neq 1_{\Lambda_s}$. This implies $\rho(g) \neq \rho(h)$. Therefore, the restriction of $\rho$ to $\Omega$ is injective. □

*Proof of Theorem 5.9.1.* Since every injective cellular automaton is pre-injective, the Garden of Eden theorem (Theorem 5.3.1) implies that every amenable group is surjunctive. By applying Lemma 3.3.4 and Lemma 5.9.2, it follows that every residually amenable group is surjunctive. □

As an immediate consequence of Theorem 5.9.1 and Proposition 3.2.2, we obtain the following:

**Corollary 5.9.3.** *Every locally residually amenable group is surjunctive.* □

## 5.10 A Surjective but Not Pre-injective Cellular Automaton over $F_2$

The Garden of Eden theorem (Theorem 5.3.1) implies that every surjective cellular automaton with finite alphabet over an amenable group is necessarily pre-injective. In this section, we give an example of a surjective but not pre-injective cellular automaton with finite alphabet over the free group $F_2$.

Let $G = F_2$ be the free group on two generators $a$ and $b$. Let $S = \{a, b, a^{-1}, b^{-1}\}$ and $A = \{0, 1\}$. Consider the cellular automaton $\tau\colon A^G \to A^G$ defined by

$$\tau(x)(g) = \begin{cases} 1 & \text{if } \sum_{s\in S} x(gs) > 2, \\ 0 & \text{if } \sum_{s\in S} x(gs) < 2, \\ x(g) & \text{if } \sum_{s\in S} x(gs) = 2. \end{cases}$$

Thus $\tau$ is the majority action automaton associated with $G$ and $S$ (see Example 1.4.3(c)). We have already seen in Example 5.2.1(c) that $\tau$ is not pre-injective.

Let us show that $\tau$ is surjective. Let $y \in A^G$ be an arbitrary configuration. We may construct a configuration $x \in A^G$ such that $y = \tau(x)$ in the following way. Consider the map $\psi\colon G\backslash\{1_G\} \to G$ which associates to each $g \in G\backslash\{1_G\}$ the element of $G$ obtained by suppressing the last factor in the reduced form of $g$. Thus if $g = s_1 s_2 \dots s_n$ with $s_i \in S$ for $1 \le i \le n$ and $s_i s_{i+1} \neq 1_G$ for $1 \le i \le n-1$, then $\psi(g) = s_1 s_2 \dots s_{n-1}$. Let $x \in A^G$ be the configuration defined by $x(g) = y(\psi(g))$ for all $g \in G \setminus \{1_G\}$ and $x(1_G) = 0$. Then, for each $g \in G$, the configuration $x$ takes the value $y(g)$ at (at least) three of the four elements $gs$, $s \in S$. it follows that $\tau(x) = y$. Thus $\tau$ is surjective.

More generally, let $G$ be a group containing a free subgroup $H$ based on two elements $a$ and $b$. Let $\tau\colon \{0,1\}^G \to \{0,1\}^G$ be the majority action cellular automaton over $G$ associated with the set $S = \{a, b, a^{-1}, b^{-1}\}$. We have seen in Example 5.2.1(c) that $\tau$ is not pre-injective. Consider its restriction $\tau_H\colon \{0,1\}^H \to \{0,1\}^H$. Note that $\tau_H$ is the majority action cellular automaton over $H$ associated with $S$. We have seen above that $\tau_H$ is surjective. It follows from Proposition 1.7.4(ii) that $\tau$ is also surjective. Thus, every group containing a free subgroup of rank two admits a cellular automaton with finite alphabet which is surjective but not pre-injective. This will be extended to all non-amenable groups in Sect. 5.12.

## 5.11 A Pre-injective but Not Surjective Cellular Automaton over $F_2$

It follows from the Garden of Eden theorem (Theorem 5.3.1) that every pre-injective cellular automaton with finite alphabet over an amenable group is

necessarily surjective. In this section, we give examples of cellular automata with finite alphabet over the free group $F_2$ which are pre-injective but not surjective.

Let $G = F_2$ be the free group on two generators $a$ and $b$. Let $H$ be a nontrivial abelian group. Our alphabet will be the group $A = H \times H$. We shall use additive notation for the group operations on $H$ and $A$. Let $p_1, p_2 \colon A \to A$ be the group endomorphisms defined respectively by $p_1(u) = (h_1, 0)$ and $p_2(u) = (h_2, 0)$ for all $u = (h_1, h_2) \in A$. Let us equip the Cartesian product $A^G = \prod_{g \in G} A$ with its natural Abelian group structure. Consider the map $\tau \colon A^G \to A^G$ given by

$$\tau(x)(g) = p_1(x(ga)) + p_2(x(gb)) + p_1(x(ga^{-1})) + p_2(x(gb^{-1})) \tag{5.18}$$

for all $g \in G$ and $x \in A^G$.

It is clear that $\tau$ is a cellular automaton over the group $G$ and the alphabet $A$ with memory set $S = \{a, b, a^{-1}, b^{-1}\}$ and a group endomorphism of $A^G$.

**Proposition 5.11.1.** *The cellular automaton $\tau \colon A^G \to A^G$ defined by (5.18) is pre-injective but not surjective.*

*Proof.* The image of $\tau$ is contained in $(H \times \{0\})^G \subsetneqq (H \times H)^G = A^G$. Thus $\tau$ is not surjective.

Let us show that $\tau$ is pre-injective. Suppose not. Then there exist configurations $x_1, x_2 \in A^G$ satisfying $\tau(x_1) = \tau(x_2)$ such that the set $\Omega = \{g \in G : x_1(g) \neq x_2(g)\}$ is a nonempty finite subset of $G$. The configuration $x_0 = x_1 - x_2 \in A^G$ satisfies $\tau(x_0) = \tau(x_1) - \tau(x_2) = 0$ and one has $\Omega = \{g \in G : x_0(g) \neq 0\}$. Consider an element $g_0 \in \Omega$ whose reduced form has maximal length, say $n_0$. Then $x_0(g_0)$ is a nonzero element $(h_0, k_0)$ of $A$.

If $h_0 \neq 0$, take $s_0 \in \{a, a^{-1}\}$ such that the reduced form of $g_0 s_0$ has length $n_0 + 1$. Then, for each $s \in S \setminus \{s_0^{-1}\}$, the length of the reduced form of $g_0 s_0 s$ is $n_0 + 2$ and hence $x(g_0 s_0 s) = 0$. By applying 5.18, we deduce that

$$\tau(x_0)(g_0 s_0) = p_1(x_0(g_0)) = (h_0, 0) \neq 0,$$

which contradicts the fact that $\tau(x_0) = 0$.

If $h_0 = 0$, then $k_0 \neq 0$ and we proceed similarly by taking $s_0 \in \{b, b^{-1}\}$ such that the reduced form of $g_0 s_0$ has length $n_0 + 1$. This gives us

$$\tau(x_0)(g_0 s_0) = p_2(x_0(g_0)) = (k_0, 0) \neq 0,$$

which yields again a contradiction. This shows that $\tau$ is pre-injective. □

If we take for $H$ a finite abelian group of cardinality $|H| = n \geq 2$ (e.g., the group $H = \mathbb{Z}/n\mathbb{Z}$), this gives us a pre-injective but not surjective cellular automaton over $F_2$ whose alphabet is finite of cardinality $n^2$.

From Proposition 1.7.4 we deduce the following:

**Proposition 5.11.2.** *Let $G$ be a group containing a free subgroup of rank two. Then there exist a finite set $A$ and a cellular automaton $\tau\colon A^G \to A^G$ which is pre-injective but not surjective.* □

## 5.12 A Characterization of Amenability in Terms of Cellular Automata

In Sect. 5.10, we gave an example of a cellular automaton with finite alphabet over the free group $F_2$ which is surjective but not pre-injective. In fact, the existence of such an automaton holds for any non-amenable group:

**Theorem 5.12.1.** *Let $G$ be a non-amenable group. Then there exists a finite set $A$ and a cellular automaton $\tau\colon A^G \to A^G$ which is surjective but not pre-injective.*

*Proof.* Since $G$ is non-amenable, it follows from Theorem 4.9.2 that there exist a 2-to-one surjective map $\varphi\colon G \to G$ and a finite subset $S \subset G$ such that

$$(\varphi(g))^{-1}g \in S \text{ for all } g \in G. \tag{5.19}$$

Our alphabet will be the Cartesian product $A = S \times S$. Let us fix some total order $\leq$ on $S$ and an arbitrary element $s_0 \in S$. Define the map $\mu\colon A^S \to A$ by

$$\mu(y) = \begin{cases} (s', t') & \text{if there exists a unique element } (s,t) \in S \times S \text{ with } s < t \\ & \text{such that } y(s) = (s, s') \text{ and } y(t) = (t, t'), \text{ where } s', t' \in S, \\ (s_0, s_0) & \text{otherwise,} \end{cases} \tag{5.20}$$

for all $y \in A^S$.

Let us show that the cellular automaton $\tau\colon A^G \to A^G$ with memory set $S$ and local defining map $\mu$ has the required properties.

We fist observe that $S$ has at least two elements since otherwise (5.19) would imply that $\varphi$ is bijective. Let $s_1 \in S$ such that $s_1 \neq s_0$. Consider the configurations $x_0, x_1 \in A^G$, where $x_0$ is defined by $x_0(g) = (s_0, s_0)$ for all $g \in G$, and $x_1$ is defined by $x_1(g) = (s_0, s_0)$ if $g \neq 1_G$ and $x_1(1_G) = (s_0, s_1)$. The configurations $x_0$ and $x_1$ are almost equal since they differ only at $1_G$. On the other hand, it is clear that $x_0$ and $x_1$ have the same image, namely $x_0$, by $\tau$. Thus $\tau$ is not pre-injective.

We use the properties of $\varphi$ to prove that $\tau$ is surjective. Let $x \in A^G$ be an arbitrary configuration. Let us show that there is a configuration $z \in A^G$ such that $x = \tau(z)$. We construct $z$ in the following way. Let $u\colon G \to S$ and $v\colon G \to S$ be the maps defined by $x(g) = (u(g), v(g))$ for all $g \in G$. For each $g \in G$, there are exactly two elements $s_g, t_g \in S$ such that $s_g < t_g$ and $\varphi(gs_g) = \varphi(gt_g) = g$. Let us set $z(gs_g) = (s_g, u(g))$ and $z(gt_g) = (t_g, v(g))$.

Observe that $z\colon G \to A$ is well defined and that the value of $z$ at $g \in G$ is either $((\varphi(g))^{-1}g, u(\varphi(g))$ or $((\varphi(g))^{-1}g, v(\varphi(g))$. It immediately follows from the definition of $\tau$ that $x = \tau(z)$. This shows that $\tau$ is surjective. □

Combining Theorem 5.12.1 with Theorem 5.3.1, we obtain the following characterization of amenable groups in terms of cellular automata:

**Corollary 5.12.2.** *Let $G$ be a group. Then the following conditions are equivalent:*

(a) *$G$ is amenable;*
(b) *every surjective cellular automaton with finite alphabet over $G$ is pre-injective.*

□

## 5.13 Garden of Eden Patterns for Life

Let $\tau\colon \{0,1\}^{\mathbb{Z}^2} \to \{0,1\}^{\mathbb{Z}^2}$ denote the cellular automaton associated with the Game of Life (see Example 1.4.3(a)). As it was observed in Example 5.3.2(b), it follows from the Garden of Eden theorem that $\tau$ is not surjective. Thus, Proposition 5.1.1 implies that $\tau$ admits Garden of Eden patterns. The purpose of this section is to present a direct proof of the existence of such patterns. This construction will provide a concrete illustration of the ideas underlying the proof that surjectivity implies pre-injectivity in the Garden of Eden theorem.

Given an integer $n \geq 1$, one says that a subset $\Omega \subset \mathbb{Z}^2$ is a *square* of size $n \times n$ in $\mathbb{Z}^2$ if there exist $p, q \in \mathbb{Z}$ such that

$$\Omega = \{p, p+1, \ldots, p+n-1\} \times \{q, q+1, \ldots, q+n-1\}.$$

**Proposition 5.13.1.** *Let $\tau\colon \{0,1\}^{\mathbb{Z}^2} \to \{0,1\}^{\mathbb{Z}^2}$ be the cellular automaton associated with the Game of Life. Then every square $\Omega \subset \mathbb{Z}^2$ of size $n \times n$ with $n \geq 3 \times 10^9$ is the support of a Garden of Eden pattern for $\tau$.*

*Proof.* Let $n \geq 1$ be an integer. Consider a square $C_n \subset \mathbb{Z}^2$ of size $5n \times 5n$. Let $D_n \subset C_n$ be the square of size $(5n-2) \times (5n-2)$ which is the $S$-interior of $D_n$ for $S = \{-1, 0, 1\}^2 \subset \mathbb{Z}^2$. Thus $D_n$ is obtained from $C_n$ by removing the $20n - 4$ points of $\mathbb{Z}^2$ located on the (usual) boundary of $C_n$.

Let us set $X_n = \{0,1\}^{C_n}$ and $Y_n = \{0,1\}^{D_n}$. The map $\tau$ induces a map $\tau_n\colon X_n \to Y_n$ defined as follows. If $u \in X_n$, we set $\tau_n(u) = (\tau(x))|_{D_n}$, where $x \in \{0,1\}^{\mathbb{Z}^2}$ satisfies $x|_{C_n} = u$ (the fact that $(\tau(x))|_{D_n}$ does not depend of the choice of $x$ follows from Proposition 5.4.3 since $S$ is a memory set for $\tau$ and $D_n$ is the $S$-interior of $C_n$). We have

$$|Y_n| = 2^{(5n-2)^2} = 2^{25n^2-20n+4}.$$

Let us divide the square $C_n$ into $n^2$ squares of size $5 \times 5$ (see Fig. 5.10).

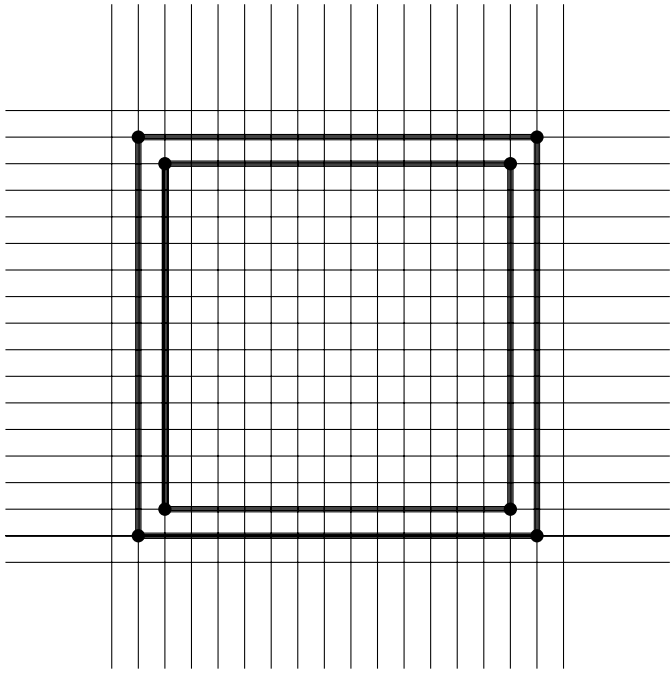


**Fig. 5.9** A square $C_n \subset \mathbb{Z}^2$ of size $5n \times 5n$ and its $S$-interior $D_n \subset C_n$, a square of size $(5n-2) \times (5n-2)$ where $S = \{-1, 0, 1\}^2 \subset \mathbb{Z}^2$; here $n = 3$

There are $2^{25}$ maps from each square of size $5 \times 5$ to the set $\{0, 1\}$. Now observe that if the restriction of an element $u \in X_n$ to one of these $5 \times 5$ squares is identically 0, then we may replace the 0 value at the center of this $5 \times 5$ square by 1 without changing $\tau_n(u)$. We deduce that

$$|\tau_n(X_n)| \leq (2^{25} - 1)^{n^2} = (2^{\log_2(2^{25}-1)})^{n^2} = 2^{(25+\log_2(1-2^{-25}))n^2}.$$

Therefore we have $|\tau_n(X_n)| < |Y_n|$ if

$$(25 + \log_2(1 - 2^{-25}))n^2 < 25n^2 - 20n + 4.$$

This inequality is equivalent to

$$-n^2 \log_2(1 - 2^{-25}) - 20n + 4 > 0,$$

which is verified if and only if

$$n > \frac{2}{-\log_2(1 - 2^{-25})}(5 + \sqrt{25 + \log_2(1 - 2^{-25})}),$$

that is, if and only if

$$n \geq 465\,163\,744.$$

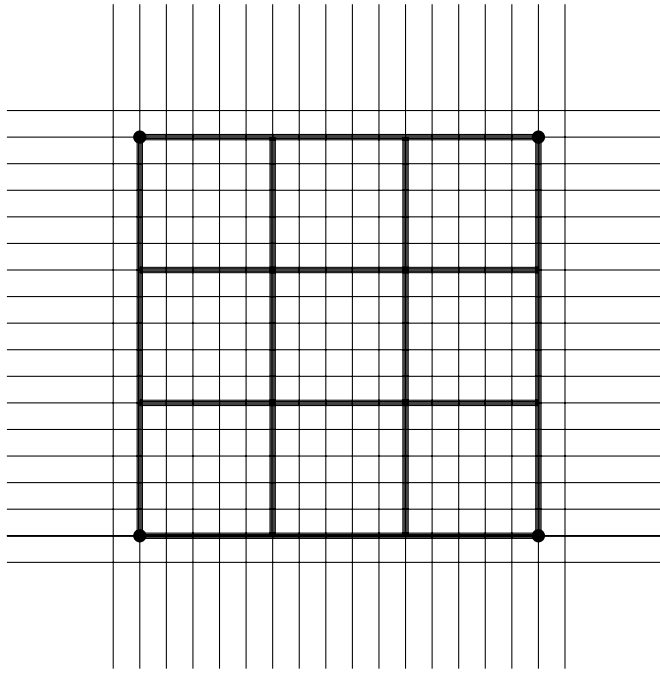


**Fig. 5.10** The square $C_n \subset \mathbb{Z}^2$ of size $5n \times 5n$ is divided into $n^2$ squares of size $5 \times 5$; here, as in Fig. 5.9, $n = 3$

For such values of $n$, the map $\tau_n$ is not surjective, i.e., there exists a Garden of Eden pattern whose support is $D_n$. This shows the existence of a Garden of Eden pattern of size

$$2\,325\,818\,718 \times 2\,325\,818\,718.$$

□

## Notes

According to M. Gardner (see [Gar-2, p.230]), it was J. Tuckey who introduced the term "Garden of Eden" in the theory of cellular automata. The first contribution to the Garden of Eden theorem goes back to E.F. Moore [Moo] who proved that every surjective cellular automaton with finite alphabet over $\mathbb{Z}^2$ is pre-injective. The converse implication was established shortly after by J. Myhill [Myh]. This is the reason why the Garden of Eden theorem for $\mathbb{Z}^2$ is often referred to as the Moore-Myhill theorem. The next step in the proof of the Garden of Eden theorem was done by A. Machì and F. Mignosi [MaM] who extended it to finitely generated groups of subexponential growth (see Chap. 6 for the definition of finitely generated groups of subexponential growth). Then, the Garden of Eden theorem was proved for

all finitely generated amenable groups by A. Machì, F. Scarabotti and the first author [CMS1]. The general case may be reduced to the case of finitely generated groups by considering the restriction of the cellular automaton to the subgroup generated by a memory set and applying Proposition 1.7.4(ii) and Proposition 5.2.2 (see [CeC8]). The proof based on Følner nets which is presented in this chapter is more direct.

The term "pre-injective" was introduced by M. Gromov in the appendix of [Gro5].

It follows from a result due to D.S. Ornstein and B. Weiss (see [OrW], [Gro6], [Kri]) that the lim sup appearing in the definition of entropy (Definition 5.7.1) is in fact a true limit and is independent of the particular choice of the right Følner net $\mathcal{F}$ in the group.

Versions of the Garden of Eden theorem for cellular automata over certain classes of subshifts (closed invariant subsets of the full shift) may be found in the appendix of [Gro5] and in two papers of F. Fiorenzi [Fio1], [Fio2]. See also [CFS].

The first examples of pre-injective (resp. surjective) cellular automata which are not surjective (resp. not pre-injective) where described by D.E. Muller (unpublished class notes). The underlying group was the *modular group* $G = PSL(2,\mathbb{Z}) = (\mathbb{Z}/2\mathbb{Z}) * (\mathbb{Z}/3\mathbb{Z})$ (note that $G$ contains the free group $F_2$).

Theorem 5.12.1 is due to L. Bartholdi [Bar].

The computation in Sect. 5.13 is taken from [BCG, Page 828]. The size of the corresponding Garden of Eden pattern is far from being optimal. The first explicit example of a Garden of Eden pattern for the cellular automaton associated with Conway's Game of Life was found by R. Banks in 1971. Banks' pattern is supported by a rectangle of size $33 \times 9$ and has 226 alive cells. The smallest known Garden of Eden pattern for Life, found by N. Beluchenko on September 2009, has as support a square $11 \times 11$ and bears 69 alive cells. It was proved that there exist no Garden of Eden patterns for Life with support contained in a rectangle of size $6 \times 5$.

## Exercises

**5.1.** Let $G$ be a group and let $A$ be a finite set. Let $\tau\colon A^G \to A^G$ be a cellular automaton admitting a memory set $M \subset G$ such that $|M| = 1$. Show that the following conditions are equivalent:

(i) $\tau$ is pre-injective;
(ii) $\tau$ is injective;
(iii) $\tau$ is surjective.

**5.2.** Let $G$ be a group and let $A$ be a finite set. Let $\tau\colon A^G \to A^G$ be a non-surjective cellular automaton. Show that there are uncountably many Garden of Eden configurations in $A^G$.

**5.3.** *Life on* $\mathbb{Z}$. Let $A = \{0, 1\}$ and consider the cellular automaton $\tau\colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ with memory set $S = \{-1, 0, 1\}$ and local defining map $\mu\colon A^S \to A$ given by

$$\mu(y) = \begin{cases} 1 & \text{if } \sum_{s\in S} y(s) = 2 \\ 0 & \text{otherwise} \end{cases}$$

for all $y \in A^S$.

(a) Show that $\tau$ is not pre-injective.

(b) Deduce from (a) that $\tau$ is not surjective either. Hint: Use Theorem 4.6.1 and the Garden of Eden Theorem (cf. Exemple 5.3.2(c)).

(c) It follows from Proposition 5.1.1 that $\tau$ admits a Garden of Eden pattern. Check that the map $p\colon \{0, 1, 2, 3, 4, 5, 6, 7, 8\} \to A$ defined by $p(0) = p(1) = p(3) = p(5) = p(8) = 1$ and $p(2) = p(4) = p(6) = p(7) = 0$ is a Garden of Eden pattern for $\tau$.

**5.4.** Let $G = \mathbb{Z}$, $A = \{0, 1\}$ and let $\tau\colon A^G \to A^G$ be the majority action cellular automaton (cf. Example 1.4.3(c)). We have seen in Example 5.2.1(c) that $\tau$ is not pre-injective so that, by amenability of the group $G$ (cf. Theorem 4.6.1) and the Garden of Eden Thoerem, $\tau$ is not surjective either (cf. Exemple 5.3.2(c)). It follows from Proposition 5.1.1 that $\tau$ admits a Garden of Eden pattern. Check that the map $p\colon \{1, 2, 3, 4, 5\} \to A$ defined by $p(1) = p(3) = p(4) = 0$ and $p(2) = p(5) = 1$ is a Garden of Eden pattern for $\tau$.

**5.5.** Let $(A, +)$ be an abelian group (not necessarily finite). Let $\tau\colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ be the cellular automaton defined by $\tau(x)(n) = x(n-1) + x(n) + x(n+1)$ for all $x \in A^{\mathbb{Z}}$ and $n \in \mathbb{Z}$. Show that $\tau$ is surjective. Hint: Given an arbitrary configuration $y \in A^{\mathbb{Z}}$, construct a configuration $x \in A^{\mathbb{Z}}$ such that $x(0) = x(1) = 0_A$ and $\tau(x) = y$.

**5.6.** Take $G = \mathbb{Z}$ and $A = \mathbb{Z}/2\mathbb{Z} = \{0, 1\}$. Let $x_1 \in A^G$ be the configuration defined by $x_1(n) = 1$ for all $n \in \mathbb{Z}$. Let $x_2 \in A^G$ be the configuration defined by $x_2(0) = 1$ and $x_2(n) = 0$ for all $n \in \mathbb{Z} \setminus \{0\}$. Let $f\colon A^G \to A^G$ be the map defined by $f(x_2) = x_1$ and $f(x) = x$ for all $x \in A^G \setminus \{x_2\}$. Let $g\colon A^G \to A^G$ be the map defined by $g(x)(n) = x(n+1) + x(n)$ for all $n \in \mathbb{Z}$ and $x \in A^G$. Verify that the maps $f$ and $g$ are both pre-injective but that the map $g \circ f$ is not pre-injective.

**5.7.** Let $G$ be group and let $A$ be a set. Suppose that $\sigma\colon A^G \to A^G$ and $\tau\colon A^G \to A^G$ are pre-injective cellular automata. Show that the cellular automaton $\tau \circ \sigma$ is pre-injective.

**5.8.** Let $G$ be a locally finite group and let $A$ be a set. Show that every pre-injective cellular automaton $\tau\colon A^G \to A^G$ is injective.

**5.9.** Give a direct proof of the Garden of Eden theorem (Theorem 5.3.1) in the case when the group $G$ is locally finite.

**5.10.** *Life in a tree.* Let $G = F_4$ denote the free group based on four generators $a, b, c, d$. Let $A = \{0, 1\}$. Consider the cellular automaton $\tau\colon A^G \to A^G$ with memory set

$$S = \{1_G, a, b, c, d, a^{-1}, b^{-1}, c^{-1}, d^{-1}\}$$

and local defining map $\mu\colon A^S \to A$ given by

$$\mu(y) = \begin{cases} 1 & \text{if } \begin{cases} \sum_{s\in S} y(s) = 3 \\ \text{or} \\ \sum_{s\in S} y(s) = 4 \text{ and } y((0,0)) = 1, \end{cases} \\ 0 & \text{otherwise} \end{cases}$$

for all $y \in A^S$ (cf. Example 1.4.3(a)). Show that $\tau$ is surjective but not pre-injective.

**5.11.** Let $G$ be a group. Let $E, F$ and $\Omega$ be subsets of $G$.

(a) Show that $(\Omega^{-E})^{-F} = \Omega^{-FE}$ and $(\Omega^{+E})^{+F} = \Omega^{+FE}$.

(b) Let $G = \mathbb{Z}$, $E = \{-2, -1, 0, 1, 2\}$, $F = \{-1, 0, 1\}$ and $\Omega = \{-10, -9, \ldots, -1, 0, 1, \ldots, 9, 10\}$. Check that $\partial_F(\partial_E(\Omega)) = \{-13, -12, -9, -8, 8, 9, 12, 13\}$ and $\partial_{FE}(\Omega) = \{-13, -12, -11, -10, -9, -8, 8, 9, 10, 11, 12, 13\}$. Deduce that, in general, one has $\partial_F(\partial_E(\Omega)) \neq \partial_{EF}(\Omega)$.

**5.12.** Let $G$ be a group and let $E \subset G$. Determine the sets $\partial_E(\varnothing)$ and $\partial_E(G)$.

**5.13.** Let $G$ be a group. Let $E$ and $\Omega$ be subsets of $G$. Show that $\partial_E(G \setminus \Omega) = \partial_E(\Omega)$.

**5.14.** Let $G$ be a group. Let $E$, $\Omega_1$ and $\Omega_2$ be subsets of $G$. Show that $\partial_E(\Omega_1 \cup \Omega_2) \subset \partial_E(\Omega_1) \cup \partial_E(\Omega_2)$. Give an example showing that one may have $\partial_E(\Omega_1 \cup \Omega_2) \neq \partial_E(\Omega_1) \cup \partial_E(\Omega_2)$.

**5.15.** Let $G$ be a group. Let $E$, $\Omega_1$ and $\Omega_2$ be subsets of $G$. Show that $\partial_E(\Omega_1 \setminus \Omega_2) \subset \partial_E(\Omega_1) \cup \partial_E(\Omega_2)$. Give an example showing that one may have $\partial_E(\Omega_1 \setminus \Omega_2) \neq \partial_E(\Omega_1) \cup \partial_E(\Omega_2)$.

**5.16.** Let $G$ be a group. Let $E$ and $\Omega$ be subsets of $G$. Show that $\Omega^{-gE} = \Omega^{-E} g^{-1}$, $\Omega^{+gE} = \Omega^{+E} g^{-1}$, and $\partial_{gE}(\Omega) = \partial_E(\Omega) g^{-1}$ for all $g \in G$.

**5.17.** Let $G$ be a group. Show that a subset $R \subset G$ is syndetic (cf. Exercise 3.39) if and only if there exists a finite subset $S \subset G$ such that the set $\Omega = R^{-1}$ satisfies $\Omega^{+S} = G$.

**5.18.** Let $G$ be a group. Let $E$ and $\Omega$ be subsets of $G$. Show that $\Omega \subset (\Omega^{+E^{-1}})^{-E}$.

**5.19.** Let $G$ be a group and let $A$ be a set. Let $\tau\colon A^G \to A^G$ be a cellular automaton with memory set $M$. Let $\Omega$ be a subset of $G$. Show that if two configurations $x_1, x_2 \in A^G$ coincide on $\Omega^{+M^{-1}}$ then the configurations $\tau(x_1)$ and $\tau(x_2)$ coincide on $\Omega$.

**5.20.** Let $(A, +)$ be a finite abelian group and let $\varphi\colon A \times A \to A$ be a map. Let $\tau\colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ be the map defined by $\tau(x)(n) = x(n)+\varphi(x(n+1), x(n+2))$ for all $n \in \mathbb{Z}$ and $x \in A^{\mathbb{Z}}$.

(a) Show that $\tau$ is a cellular automaton over the group $\mathbb{Z}$ and the alphabet $A$.

(b) Show that $\tau$ is pre-injective.

(c) Show that $\tau$ is surjective.

**5.21.** Let $G = \mathbb{Z}$ and $A = \{0, 1\}$. Verify that, among the 16 cellular automata $\tau\colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ with memory set $S = \{0, 1\} \subset \mathbb{Z}$, there are exactly 6 of them which are surjective, 4 of them which are injective, 4 of them which are bijective, and 10 of them which are neither surjective nor injective. Hint: The Garden of Eden Theorem may help.

**5.22.** *Topological entropy.* Let $G$ be an amenable group and let $\mathcal{F} = (F_j)_{j\in J}$ be a right Følner net for $G$. If $X$ is a compact topological space equipped with a continuous action of $G$, the *topological entropy* $h_{\mathcal{F}}(X, G)$, $0 \leq h_{\mathcal{F}}(X, G) \leq \infty$, of $X$ is defined as follows. Suppose that $\mathcal{U} = (U_i)_{i\in I}$ is an open cover of $X$. We denote by $N(\mathcal{U})$ the smallest integer $n \geq 0$ such that there is a finite subset $I_0 \subset I$ of cardinality $n$ such that $\bigcup_{i\in I_0} U_i = X$ (note that there exists such a finite subset $I_0 \subset I$ by compactness of $X$). Given a nonempty subset $F \subset G$, we define the open cover $\mathcal{U}_F = (W_\alpha)_{\alpha\in I^F}$ indexed by the set $I^F = \{\alpha\colon F \to I\}$ by setting

$$W_\alpha = \bigcap_{g\in F} gU_{\alpha(g)}$$

for all $\alpha \in I^F$. Finally, we set

$$h_{\mathcal{F}}(\mathcal{U}) = \limsup_j \frac{\log N(\mathcal{U}_{F_j})}{|F_j|}$$

and

$$h_{\mathcal{F}}(X, G) = \sup_{\mathcal{U}} h_{\mathcal{F}}(\mathcal{U}),$$

where $\mathcal{U}$ ranges over all open covers of $X$.

Observe that it is clear from this definition that if $X$ and $Y$ are two compact topological spaces, each equipped with a continuous action of $G$, such that there exists a $G$-equivariant homeomorphism $f\colon X \to Y$, then one has $h_{\mathcal{F}}(X, G) = h_{\mathcal{F}}(Y, G)$.

(a) Let $X$ be a compact space and let $\mathcal{U} = (U_i)_{i\in I}$ and $\mathcal{V} = (V_k)_{k\in K}$ be two open covers of $X$. One says that the open cover $\mathcal{V}$ is *finer* than $\mathcal{U}$ if, for each $k \in K$, there exists $i \in I$ such that $V_k \subset U_i$. Show that if the open cover $\mathcal{V}$ is finer than the open cover $\mathcal{U}$ then one has $N(\mathcal{U}) \leq N(\mathcal{V})$.

(b) Let $X$ be a compact space and let $\mathcal{U} = (U_i)_{i\in I}$ be an open cover of $X$ which forms a partition of $X$ (i.e., such that $U_{i_1} \cap U_{i_2} = \varnothing$ for all $i_1, i_2 \in I$ with $i_1 \neq i_2$). Show that $N(\mathcal{U}) = |\{i \in I : U_i \neq \varnothing\}|$.

(c) Let $A$ be a finite set and let $X \subset A^G$ be a subshift. Show that if $X$ is equipped with the action of $G$ induced by the $G$-shift on $A^G$, then one has

$$h_{\mathcal{F}}(X, G) = \mathrm{ent}_{\mathcal{F}}(X).$$

Hint: Consider the open cover $\mathcal{T} = (T_a)_{a \in A}$ of $X$ defined by $T_a = \{x \in X : x(1_G) = a\}$ and deduce from (b) that $N(\mathcal{T}_{F_j}) = |\pi_{F_j}(X)|$, where $\pi_{F_j} \colon A^G \to A^{F_j}$ denotes the projection map. This gives $h_{\mathcal{F}}(X, G) \geq h_{\mathcal{F}}(\mathcal{T}) = \mathrm{ent}_{\mathcal{F}}(X)$. To prove $h_{\mathcal{F}}(X, G) \leq \mathrm{ent}_{\mathcal{F}}(X)$, observe that if $\mathcal{U} = (U_i)_{i \in I}$ is an arbitrary open cover of $X$, then there exists a finite subset $\Omega \subset G$ such that the open cover $\mathcal{T}_\Omega$ is finer than $\mathcal{U}$. Apply (a) to get $N(\mathcal{U}_{F_j}) \leq |\pi_{F_j \Omega}(X)|$ and finally use the fact that $\mathcal{F}$ is a right Følner net to conclude.

Note: It can be shown by using a result due to Ornstein and Weiss (cf. the notes above) that the topological entropy $h_{\mathcal{F}}(X, G)$ of a compact space $X$ equipped with a continuous action of an amenable group $G$ is in fact independent of the choice of the right Følner net $\mathcal{F}$.

**5.23.** Let $G$ be an amenable group and let $\mathcal{F} = (F_j)_{j \in J}$ be a right Følner net for $G$. Let $A$ and $B$ be finite sets. Suppose that $X \subset A^G$ and $Y \subset B^G$ are two subshifts such that there exists a bijective continuous $G$-equivariant map $\varphi \colon X \to Y$. Show that $\mathrm{ent}_{\mathcal{F}}(X) = \mathrm{ent}_{\mathcal{F}}(Y)$. Hint: Use Exercise 5.22.

**5.24.** Let $G$ be an amenable group, $\mathcal{F}$ a right Følner net for $G$, and $A$ a finite set. Let $X$ be a subset of $A^G$ and let $\overline{X}$ denote the closure of $X$ in $A^G$ for the prodiscrete topology. Show that one has $\mathrm{ent}_{\mathcal{F}}(\overline{X}) = \mathrm{ent}_{\mathcal{F}}(X)$.

**5.25.** Let $G$ be an amenable group and let $(F_j)_{j \in J}$ be a right Følner net for $G$. Let $A$ be a finite set. Suppose that $X$ (resp. $Y$) is a subset of $A^G$ having at least two distinct elements. Show that $\mathrm{ent}_{\mathcal{F}}(X \cup Y) \leq \mathrm{ent}_{\mathcal{F}}(X) + \mathrm{ent}_{\mathcal{F}}(Y)$.

**5.26.** Let $G$ be an amenable group, $A$ a finite set, $\mathcal{F}$ be a right Følner net for $G$, and $X \subset A^G$ a subshift. Let $F$ be a nonempty finite subset of $G$ and consider the subshift $X^{[F]} \subset B^G$, where $B = A^F$ (cf. Exercise 1.34). Show that $\mathrm{ent}_{\mathcal{F}}(X^{[F]}) = \mathrm{ent}_{\mathcal{F}}(X)$.

**5.27.** Let $G$ be a group, $A$ be a set and $X \subset A^G$ a subshift. Let also $H \subset G$ be a subgroup of $G$ and $T \subset G$ a complete set of representatives for the right cosets of $H$ in $G$, and consider the subshift $X^{(H,T)} \subset B^H$, where $B = A^{H \backslash G}$ (cf. Exercise 1.35). Suppose that $G$ and $H$ are isomorphic and denote by $\phi \colon G \to H$ an isomorphism. Let $\mathcal{F} = (F_j)_{j \in J}$ be a right Følner net in $G$ so that $\mathcal{F}' = (\phi(F_j))_{j \in J}$ is a right Følner net in $G$. Show that $\mathrm{ent}_{\mathcal{F}}(X^{(H,T)}) = \mathrm{ent}_{\mathcal{F}'}(X)$.

**5.28.** Let $G = \mathbb{Z}$ and $A = \{0, 1\}$. Set $F_n = \{0, 1, \ldots, n-1\}$ and recall from Example 4.7.4(b) that $\mathcal{F} = (F_n)_{n \in \mathbb{N}}$ is a Følner sequence in $\mathbb{Z}$. Show that if $X \subset A^{\mathbb{Z}}$ is either the even subshift considered in Exercise 1.38 or the golden mean subshift considered in Exercise 1.39 then $\mathrm{ent}_{\mathcal{F}}(X) = \log \varphi$, where $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden number.

**5.29.** *Entropy of the Morse subshift.* Let $A = \{0,1\}$ and let $x \in A^{\mathbb{N}}$ and $X \subset A^{\mathbb{Z}}$ denote the Thue-Morse sequence (cf. Exercises 3.41) and the Morse subshift (cf. Exercise 3.42) respectively.

(a) Show that $|L_n(X)| \leq 8n$ for all $n \geq 1$. Hint: Let $n \geq 1$ and denote by $k$ the unique integer such that $2^{k-1} \leq n < 2^k$. Consider the words $u = x(0)x(1)\cdots x(2^k-1)$ and $v = \iota(u)$ (cf. Exercise 3.41(e)). Observe that any word $w \in L_n(X)$ is necessarily a subword of one of the words $uv$, $vv$, $vu$, or $uu$. Altogether, this gives at most $4 \cdot 2^k = 8 \cdot 2^{k-1} \leq 8n$ distinct possibilities for $w$.

(b) Let $\mathcal{F} = (F_n)_{n\geq 1}$ denote the Følner sequence of $\mathbb{Z}$ where $F_n = \{0,1,\ldots,n-1\}$. Deduce from (a) that $\mathrm{ent}_{\mathcal{F}}(X) = 0$.

**5.30.** Let $A = \{0,1,2\}$ and set $Y = \{y \in A^{\mathbb{Z}} : y(g) \in \{1,2\} \text{ for all } g \in \mathbb{Z}\}$ and $X = Y \cup \{x_0\}$, where $x_0 \in A^{\mathbb{Z}}$ denotes the constant configuration defined by $x_0(g) = 0$ for all $g \in \mathbb{Z}$.

(a) Show that $X$ and $Y$ are subshifts of finite type of $A^{\mathbb{Z}}$ and that one has the strict inclusion $Y \subsetneq X$.

(b) Show that $X$ is not irreducible.

(c) Check that $X$ and $Y$ have the same entropy $\mathrm{ent}_{\mathcal{F}}(X) = \mathrm{ent}_{\mathcal{F}}(Y) = \log 2$ with respect to the Følner sequence $\mathcal{F} = (F_n)_{n\in\mathbb{N}}$ of $\mathbb{Z}$ given by $F_n = \{0,1,\ldots,n-1\}$.

**5.31.** Let $G$ be an amenable group, $A$ a finite set, $X \subset A^G$ a strongly irreducible subshift, and $Y \subset A^G$ a nonempty subshift such that $Y \subsetneq X$. Let also $\mathcal{F} = (F_j)_{j\in J}$ be a right Følner net in $G$. For a subset $\Omega \subset G$ we denote by $\pi_\Omega \colon A^G \to A^\Omega$ the projection (restriction) map.

(a) Show that there exists a finite subset $\Omega_0 \subset G$ such that $\pi_{\Omega_0}(Y) \subsetneq \pi_{\Omega_0}(X)$.

(b) Let $\Delta \subset G$ be a finite subset such that $1_G \in \Delta$ and $X$ is $\Delta$-irreducible. Let $E = \Omega_0^{+\Delta}$ and set $\xi = |\pi_E(X)|^{-1}$ and $E' = EE^{-1} = \{ab^{-1} : a,b \in E\}$. By virtue of Proposition 5.6.3 we can find an $(E,E')$-tiling $T \subset G$. For $j \in J$ set $T_j = \{g \in T : gE \subset F_j\}$. Show that $\xi|\pi_{F_j}(X)| \leq |\pi_{F_j \setminus gE}(X)|$ for all $g \in T_j$.

(c) Let $p \in \pi_{\Omega_0}(X) \setminus \pi_{\Omega_0}(Y)$ (cf. (a)) and let $x \in X$ such that $x|_{\Omega_0} = p$. For $g \in G$ set $p_g = (gx)|_{g\Omega_0} \in A^{g\Omega_0}$. Show that $p_g \in \pi_{g\Omega_0}(X) \setminus \pi_{g\Omega_0}(Y)$.

(d) For $j \in J$ and $g \in T_j$ denote by $\pi_{g\Omega_0}^{F_j} \colon \pi_{F_j}(X) \to \pi_{g\Omega_0}(X)$ the projection (restriction) map and let $p_g \in \pi_{g\Omega_0}(X) \setminus \pi_{g\Omega_0}(Y)$ (cf. (c)). Using (b) and the $\Delta$-irreducibility of $X$ show that $|\pi_{F_j}(X) \setminus (\pi_{g\Omega_0}^{F_j})^{-1}(p_g)| \leq (1-\xi)|\pi_{F_j}(X)|$ for all $j \in J$ and $g \in T_j$.

(e) For $j \in J$ denote by $\pi_{F_j}(X)^*$ the set of patterns $p \in \pi_{F_j}(X)$ such that $\pi_{g\Omega_0}^{F_j}(p) \neq p_g$ for all $g \in T_j$. Observe that $\pi_{F_j}(X)^* = \pi_{F_j}(X) \setminus \bigcup_{g\in T_j} (\pi_{g\Omega_0}^{F_j})^{-1}(p_g)$ and using the $\Delta$-irreducibility of $X$ and an inductive argument based on (d), show that

$$|\pi_{F_j}(X)^*| \leq (1-\xi)^{|T_j|}|\pi_{F_j}(X)|.$$

(f) Observe that $\pi_{F_j}(Y) \subset \pi_{F_j}(X)^*$ and deduce from (e) that $\log|\pi_{F_j}(Y)| \leq |T_j|\log(1-\xi) + \log|\pi_{F_j}(X)|$.

(g) By Proposition 5.6.4 there exists a real number $\alpha > 0$ and an element $j_0 \in J$ such that $|T_j| \geq \alpha|F_j|$ for all $j \geq j_0$. Deduce from (f) that $\mathrm{ent}_{\mathcal{F}}(Y) \leq \alpha\log(1-\xi) + \mathrm{ent}_{\mathcal{F}}(X)$.

(h) Deduce from (g) that $\mathrm{ent}_{\mathcal{F}}(Y) < \mathrm{ent}_{\mathcal{F}}(X)$.

**5.32.** Let $G$ be an amenable group and let $A$ be a finite set. Let $\mathcal{F} = (F_j)_{j\in J}$ be a right Følner net in $G$. Let $X \subset A^G$ be a nonempty strongly irreducible subshift. Show that if $X$ is not minimal then one has $\mathrm{ent}_{\mathcal{F}}(X) > 0$. Hint: Show that, with the notation introduced in Exercise 5.31, one has $\mathrm{ent}_{\mathcal{F}}(X) \geq \alpha\log 2$.

**5.33.** Let $G$ be a group and let $A$ be a finite set. Let $\Delta$ be a finite subset of $G$ and let $X \subset A^G$ be a $\Delta$-irreducible subshift. Suppose that $(\Omega_i)_{i\in I}$ is a (possibly infinite) family of (possibly infinite) subsets of $G$ such that

$$\Omega_i^{+\Delta} \cap \left( \bigcup_{k\in I\setminus\{i\}} \Omega_k \right) = \varnothing \quad \text{for all } i \in I.$$

Let also $(x_i)_{i\in I}$ be a family of configurations in $X$.

Denote by $\mathcal{P}_f(G)$ the set of all finite subsets of $G$. For each $\Lambda \in \mathcal{P}_f(G)$ let $X(\Lambda) \subset X$ denote the set consisting of all configurations in $X$ which coincide with $x_i$ on $\Lambda \cap \Omega_i$ for all $i \in I$.

(a) Show that $X(\Lambda)$ is closed in $X$ for each $\Lambda \in \mathcal{P}_f(G)$.

(b) Show that if we fix $\Lambda \in \mathcal{P}_f(G)$, then the subsets $\Psi_i = \Lambda \cap \Omega_i$ are all contained in $\Lambda$ and satisfy

$$\Psi_i^{+\Delta} \cap \left( \bigcup_{k\in I\setminus\{i\}} \Psi_k \right) = \varnothing \quad \text{for all } i \in I.$$

(c) Fix $\Lambda \in \mathcal{P}_f(G)$ and set $I_\Lambda = \{i \in I : \Lambda \cap \Omega_i \neq \varnothing\}$. Show that $I_\Lambda$ is finite. Then, applying induction on the cardinality of $I_\Lambda$, show that, by $\Delta$-irreducibility of $X$, one has $X(\Lambda) \neq \varnothing$.

(d) Show that $X(\Lambda_1) \cap X(\Lambda_2) \cap \cdots \cap X(\Lambda_n) = X(\Lambda_1 \cup \Lambda_2 \cup \cdots \cup \Lambda_n)$ and deduce that $X(\Lambda_1) \cap X(\Lambda_2) \cap \cdots \cap X(\Lambda_n) \neq \varnothing$ for all $\Lambda_1, \Lambda_2, \ldots, \Lambda_n \in \mathcal{P}_f(G)$.

(e) Deduce from (d) that the family $(X(\Lambda))_{\Lambda\in\mathcal{P}_f(G)}$ has a nonempty intersection.

(f) Let $x \in X$ be such that $x \in X(\Lambda)$ for each finite subset $\Lambda \subset G$ (whose existence is guaranteed by (e)). Show that $x$ coincides with $x_i$ on $\Omega_i$ for all $i \in I$.

**5.34.** Let $G$ be an amenable group and let $A$ be a finite set. Let $\mathcal{F} = (F_j)_{j\in J}$ be a right Følner net in $G$. Let $X \subset A^G$ be a (possibly minimal) strongly irreducible subshift containing at least two distinct configurations $x_0$ and $x_1$.

(a) Show that there exists a finite subset $D \subset G$ such that $(gx_0)|_{gD} \neq (gx_1)|_{gD}$ for all $g \in G$.

(b) Let $\Delta$ be a finite subset of $G$ such that $X$ is $\Delta$-irreducible and $1_G \in \Delta$. Set $E = D^{+\Delta}$. By Proposition 5.6.3 there exists an $(E, E')$-tiling $T \subset G$ for some finite subset $E' \subset G$. Consider the subset $Z \subset X$ consisting of all the configurations $z \in X$ such that, for all $g \in T$, one has either $z|_{gD} = (gx_0)|_{gD}$ or $z|_{gD} = (gx_1)|_{gD}$. Applying Exercise 5.33 to the family $(gD)_{g\in T}$ show that, given any map $\iota\colon T \to \{0, 1\}$, there exists a configuration $x \in X$ such that $x|_{gD} = (gx_{\iota(g)})|_{gD}$ for all $g \in T$.

(c) Deduce that $|Z_{F_j}| \geq 2^{|T_j|}$ for all $j \in J$, where, $T_j = \{g \in T : gE \subset F_j\}$.

(d) By Proposition 5.6.4 there exists a real number $\alpha > 0$ and an element $j_0 \in J$ such that $|T_j| \geq \alpha |F_j|$ for all $j \geq j_0$. Deduce from (c) that $\mathrm{ent}_{\mathcal{F}}(Z) \geq \alpha \log 2$.

(e) Deduce that $\mathrm{ent}_{\mathcal{F}}(X) > 0$.

**5.35.** Let $G$ be a group and let $A$ be a finite set. Let $X \subset A^G$ be a strongly irreducible subshift. Suppose that $\Delta$ is a finite subset of $G$ such that $X$ is $\Delta$-irreducible. Show that if $\Omega_1$ and $\Omega_2$ are (possibly infinite) subsets of $G$ such that $\Omega_1^{+\Delta} \cap \Omega_2 = \varnothing$, then, given any two configurations $x_1$ and $x_2$ in $X$, there exists a configuration $x \in X$ which coincides with $x_1$ on $\Omega_1$ and with $x_2$ on $\Omega_2$. Hint: For each finite subset $\Omega \subset G$, consider the subset $X(\Omega) \subset X$ consisting of all configurations in $X$ which coincide with $x_1$ on $\Omega \cap \Omega_1$ and with $x_2$ on $\Omega \cap \Omega_2$. Observe that the family formed by the sets $X(\Omega)$, where $\Omega$ runs over all finite subsets of $G$ is a family of closed subsets of $X$ having the finite intersection property and use the compactness of $X$.

**5.36.** (cf. Lemma 5.8.3) Let $G$ be an amenable group, $\mathcal{F} = (F_j)_{j\in J}$ a right Følner net for $G$, and $A$ a finite set. Let $\tau\colon X \to Y$ be a cellular automaton, where $X, Y \subset A^G$ are subshifts such that $X$ is strongly irreducible and $\mathrm{ent}_{\mathcal{F}}(Y) < \mathrm{ent}_{\mathcal{F}}(X)$.

(a) Let $S \subset G$ be a memory set for $\tau$. Up to enlarging the subset $S$ if necessary, we can also suppose that $1_G \in S$ and that $X$ is $S$-irreducible. Deduce from the hypothesis $\mathrm{ent}_{\mathcal{F}}(Y) < \mathrm{ent}_{\mathcal{F}}(X)$ that there exists $j_0 \in J$ such that $|\pi_{F_{j_0}^{+S^2}}(Y)| < |\pi_{F_{j_0}}(X)|$.

(b) Fix an arbitrary configuration $x_0 \in X$ and consider the finite subset $Z \subset X$ consisting of all configurations $z \in X$ which coincide with $x_0$ on $G \setminus F_{j_0}^{+S}$. Show that $\pi_{F_{j_0}}(Z) = \pi_{F_{j_0}}(X)$. Hint: Use the result of Exercise 5.35(a) by taking $\Omega_1 = F_{j_0}$ and $\Omega_2 = G \setminus F_{j_0}^{+S}$.

(c) Use Proposition 5.4.3 to show that $\tau(z)$ coincide with $\tau(x_0)$ on $G \setminus F_{j_0}^{+S^2}$ for all $z \in Z$.

(d) Deduce from (c) that $|\tau(Z)| = |\pi_{F_{j_0}^{+S^2}}(\tau(Z))|$.

(e) Using the fact that $\tau(Z) \subset Y$, deduce from (c), (a) and (b) that there exist configurations $z_1 \neq z_2$ in $Z$ such that $\tau(z_1) = \tau(z_2)$.

(f) Deduce from (e) that $\tau$ is not pre-injective.

**5.37.** Let $G$ be an amenable group, $\mathcal{F} = (F_j)_{j\in J}$ a right Følner net for $G$, and $A$ a finite set. Let $X, Y \subset A^G$ be two strongly irreducible subshifts such that $\mathrm{ent}_{\mathcal{F}}(X) = \mathrm{ent}_{\mathcal{F}}(Y)$. Show that every pre-injective cellular automaton $\tau\colon X \to Y$ is surjective. Hint: Use the results of Exercise 5.31 and Exercise 5.36.

**5.38.** Let $G$ be an amenable group and let $A$ be a finite set. Let $X \subset A^G$ be a strongly irreducible subshift. Show that every pre-injective cellular automaton $\tau\colon X \to X$ is surjective. Hint: Use Exercise 5.37 with $Y = X$.

**5.39.** Let $G$ be an amenable group and let $A$ be a finite set. Show that every strongly irreducible subshift $X \subset A^G$ is surjunctive. Hint: This immediately follows from Exercise 5.38.

**5.40.** (cf. Lemma 5.8.4) Let $G$ be an amenable group, $\mathcal{F} = (F_j)_{j\in J}$ a right Følner net for $G$, and $A$ a finite set. Let $X \subset A^G$ be a strongly irreducible subshift of finite type and let $\tau\colon X \to A^G$ be a cellular automaton which is not pre-injective. Let $x_1, x_2 \in X$ such that $\tau(x_1) = \tau(x_2)$ and $\Omega = \{g \in G : x_1(g) \neq x_2(g)\}$ is a nonempty finite subset of $G$. Let $S$ be a memory set for both $\tau$ and $X$. Up to enlarging the subset $S$ if necessary, we can also suppose that $1_G \in S$ and that $X$ is $S$-irreducible. Set $E = \Omega^{+(S^{-1}S)} = \{\omega s^{-1} s' : \omega \in \Omega, s, s' \in S\}$ and $E' = EE^{-1} = \{ab^{-1} : a, b, \in E\}$. By Proposition 5.6.3 there exists an $(E, E')$-tiling $T \subset G$. Let $Z \subset X$ be the set consisting of all configurations $z$ in $X$ such that $z|_{tE} \neq (tx_1)|_{tE}$ for all $t \in T$.

(a) Using the $S$-irreducibility of $X$, show that $\mathrm{ent}_{\mathcal{F}}(Z) < \mathrm{ent}_{\mathcal{F}}(X)$. Hint: Use the arguments in Exercise 5.31.

(b) From (a) and Proposition 5.7.3 deduce that $\mathrm{ent}_{\mathcal{F}}(\tau(Z)) \leq \mathrm{ent}_{\mathcal{F}}(X)$.

(c) Using the fact that $S$ is a memory set for $X$, show that for every $x \in X$ the configuration $z$ defined by

$$z(g) = \begin{cases} tx_2(g) & \text{if there is } t \in T \text{ such that } g \in tE \text{ and } x|_{tE} = tx_1|_{tE}, \\ x(g) & \text{otherwise.} \end{cases}$$

satisfies $z \in Z$ and $\tau(z) = \tau(x)$ (cf. the end of the proof of Lemma 5.8.4).

(d) Deduce from (c) that $\tau(Z) = \tau(X)$.

(e) Deduce from (b) and (d) that $\mathrm{ent}_{\mathcal{F}}(\tau(X)) < \mathrm{ent}_{\mathcal{F}}(X)$.

**5.41.** Let $G$ be an amenable group, $\mathcal{F} = (F_j)_{j\in J}$ a right Følner net for $G$, and $A$ a finite set. Let $X, Y \subset A^G$ be two subshifts such that $X$ is strongly irreducible of finite type and $\mathrm{ent}_{\mathcal{F}}(X) = \mathrm{ent}_{\mathcal{F}}(Y)$. Show that every surjective cellular automaton $\tau\colon X \to Y$ is pre-injective. Hint: Use the result of Exercise 5.40.

**5.42.** *The Garden of Eden theorem for strongly irreducible subshifts of finite type.* (cf. [Fio2, Theorem 4.7]). Let $G$ be an amenable group, $A$ a finite set, $X \subset A^G$ a strongly irreducible subshift of finite type, and $\tau\colon X \to X$ a cellular automaton. Show that $\tau$ is pre-injective if and only if it is surjective.

Hint: Combine together the results from Exercise 5.41 with $Y = X$ and Exercise 5.38.

**5.43.** Let $G$ be a group. Let $F$ be a nontrivial finite abelian group and set $A = F \times F$. Suppose that $G$ contains an element $h_0 \in G$ of infinite order (e.g. $G = \mathbb{Z}$). Denote by $H \subset G$ the infinite cyclic subgroup generated by $h_0$. For each $x \in A^G$, let $x_1, x_2 \colon G \to F$ be the maps defined by $x(g) = (x_1(g), x_2(g))$ for all $g \in G$. Consider the subset $X \subset A^G$ consisting of all configurations $x \in A^G$ which satisfy $x_2(gh_0) = x_2(g)$ for any $g \in G$. In other words, $X$ is formed by all the configurations $x \in A^G$ such that $x_2$ is constant on each left coset of $H$ in $G$.

(a) Show that $X$ is a subshift of finite type.

(b) Consider the map $\tau \colon X \to X$ defined by $\tau(x) = (x_1 + x_2, 0)$ for all $x = (x_1, x_2) \in X$ (here 0 denotes the zero configuration, i.e., the constant configuration whose value at each element of $G$ is equal to the identity element of $F$). Show that $\tau$ is a cellular automaton.

(c) Show that the cellular automaton $\tau$ defined in (b) is pre-injective but not surjective.

(d) Show that if $H$ is of infinite index in $G$ (this is the case, for example, when $G = \mathbb{Z}^2$) then the subshift $X$ is irreducible.

**5.44.** Let $G$ be a group and let $A$ be a set. Let $H$ be a subgroup of $G$. Suppose that $\sigma \colon A^H \to A^H$ is a cellular automaton and let $\sigma^G \colon A^G \to A^G$ be the induced cellular automaton (cf. Sect. 1.7). Let $X \subset A^H$ be a subshift and denote by $X^{(G)} \subset A^G$ the associated subshift defined in Exercise 1.33. Suppose that $\sigma(X) \subset X$. Show that the cellular automaton $\sigma^G|_{X^{(G)}} \colon X^{(G)} \to X^{(G)}$ is pre-injective if and only if the cellular automaton $\sigma|_X \colon X \to X$ is pre-injective.

**5.45.** Let $A = \{0, 1\}$. Let $x_0, x_1 \in A^{\mathbb{Z}}$ denote the two constant configurations defined by $x_0(n) = 0$ and $x_1(n) = 1$ for all $n \in \mathbb{Z}$.

(a) Show that $X = \{x_0, x_1\}$ is a subshift of finite type.

(b) Show that the map $\tau \colon X \to X$ given by $\tau(x_0) = \tau(x_1) = x_0$ is a cellular automaton which is pre-injective but neither surjective nor injective.

**5.46.** Let $A = \{0, 1\}$, $G = \mathbb{Z}^2$, and $H = \mathbb{Z} \times \{0\} \subset G$. Consider the subset $X = \mathrm{Fix}(H) \subset A^G$ consisting of all the configurations $x \in A^G$ which are fixed by each element of $H$.

(a) Show that $X$ is an irreducible subshift of finite type of $A^G$.

(b) Consider the map $\tau \colon X \to X$ defined by $\tau(x)(g) = 0$ for all $x \in X$ and $g \in G$. Show that $\tau$ is a cellular automaton which is pre-injective but neither surjective nor injective.

**5.47.** ([Fio1, Counterexample 4.27]) Let $A = \{0, 1, 2\}$ and let $X \subset A^{\mathbb{Z}}$ be the subshift of finite type with defining set of forbidden words $\{01, 02\}$.

(a) Show that $X$ is not irreducible.

(b) Consider the cellular automaton $\sigma\colon A^{\mathbb{Z}} \to A^{Z}$ with memory set $S = \{0, 1\}$ and local defining map

$$\mu(y) = \begin{cases} y(0) & \text{if } y(1) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Show that $\sigma(X) \subset X$.

(c) Show that the cellular automaton $\sigma|_X\colon X \to X$ is surjective but not pre-injective.

**5.48.** Let $A$, $X \subset A^{\mathbb{Z}}$ and $\sigma\colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ as in Exercise 5.47. Let $H = \mathbb{Z}\times\{0\} \subset \mathbb{Z}^2$ and let $X_{(\mathbb{Z}^2)}$ be the associated irreducible subshift of finite type in $A^{\mathbb{Z}^2}$ defined as in Exercise 1.33 and let $\sigma^{\mathbb{Z}^2}\colon A^{\mathbb{Z}^2} \to A^{\mathbb{Z}^2}$ be the induced cellular automaton. Show that the cellular automaton $\sigma^{\mathbb{Z}^2}|_{X^{(\mathbb{Z}^2)}}\colon X^{(\mathbb{Z}^2)} \to X^{(\mathbb{Z}^2)}$ is surjective but not pre-injective.

**5.49.** ([Fio1, Section 3]) Let $A = \{0, 1\}$ and let $X \subset A^{\mathbb{Z}}$ be the even subshift (cf. Exercise 1.38). Let also $\sigma\colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ be the cellular automaton with memory set $S = \{0, 1, 2, 3, 4\}$ and local defining map $\mu\colon A^S \to A$ given by

$$\mu(y) = \begin{cases} 1 & \text{if } y(0)y(1)y(2) \in \{000, 111\} \text{ or } y(0)y(1)y(2)y(3)y(4) = 00100 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Show that $\sigma(X) \subset X$.

(b) Set $\tau = \sigma|_X\colon X \to X$. Show that $\tau$ is surjective.

(c) Show that $\tau$ is not pre-injective. Hint: Show that the configurations

$$x_1 = \cdots 0000(100100)0000\cdots$$

and

$$x_2 = \cdots 0000(011100)0000\cdots$$

satisfy

$$\tau(x_1) = \tau(x_2) = \cdots 1111(100100)1111\cdots$$

# Chapter 6
# Finitely Generated Amenable Groups

This chapter is devoted to the growth and amenability of finitely generated groups. The choice of a finite symmetric generating subset for a finitely generated group defines a word metric on the group and a labelled graph, which is called a Cayley graph. The associated growth function counts the number of group elements in a ball of radius $n$ with respect to the word metric. We define a notion of equivalence for such growth functions and observe that the growth functions associated with different finite symmetric generating subsets are in the same equivalence class (Corollary 6.4.5). This equivalence class is called the growth type of the group. The notions of polynomial, subexponential and exponential growth are introduced in Sect. 6.5. In Sect. 6.7 we give an example of a finitely generated metabelian group with exponential growth. We prove that finitely generated nilpotent groups have polynomial growth (Theorem 6.8.1). In Sect. 6.9 we consider the Grigorchuk group. It is shown that it is an infinite periodic (Theorem 6.9.8), residually finite (Corollary 6.9.5) finitely generated group of intermediate growth (Theorem 6.9.17). In the subsequent section, we show that every finitely generated group of subexponential growth is amenable (Theorem 6.11.2). In Sect. 6.12 we prove the Kesten-Day characterization of amenability (Theorem 6.12.9) which asserts that a group with a finite (not necessarily symmetric) generating subset is amenable if and only if 0 is in the $\ell^2$-spectrum of the associated Laplacian. Finally, in Sect. 6.13 we consider the notion of quasi-isometry for not necessarily countable groups and we show that amenability is a quasi-isometry invariant (Theorem 6.13.23).

## 6.1 The Word Metric

Let $G$ be a group.

One says that a subset $S \subset G$ *generates* $G$, or that $S$ is a *generating subset* of $G$, if every element $g \in G$ can be expressed as a product of elements in

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_6, 

$S \cup S^{-1}$, that is, for each $g \in G$ there exist $n \geq 0$, $s_1, s_2, \ldots, s_n \in S$ and $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n \in \{1, -1\}$ such that

$$g = s_1^{\varepsilon_1} s_2^{\varepsilon_2} \cdots s_n^{\varepsilon_n}. \tag{6.1}$$

A subset $S \subset G$ is called *symmetric* if $S = S^{-1}$, that is, $s^{-1} \in S$ whenever $s \in S$. Observe that if $S \subset G$ is a generating subset of $G$, then $S \cup S^{-1}$ is a symmetric generating subset of $G$.

Recall that $G$ is said to be finitely generated if it admits a finite generating subset. If $S \subset G$ is a finite generating subset of $G$, then $S \cup S^{-1}$ is a finite symmetric generating subset of $G$. Thus any finitely generated group admits a finite symmetric generating subset.

Let $G$ be a finitely generated group and let $S$ be a finite symmetric generating subset of $G$.

The *$S$-word-length* $\ell_S(g) = \ell_S^G(g)$ of an element $g \in G$ is the minimal integer $n \geq 0$ such that $g$ can be expressed as a product of $n$ elements in $S$, that is,

$$\ell_S(g) = \min\{n \geq 0 : g = s_1 s_2 \cdots s_n,\ s_i \in S, 1 \leq i \leq n\}. \tag{6.2}$$

It immediately follows from the definition that for $g \in G$ one has

$$\ell_S(g) = 0 \text{ if an only if } g = 1_G. \tag{6.3}$$

**Proposition 6.1.1.** *One has*

$$\ell_S(g^{-1}) = \ell_S(g) \tag{6.4}$$

*and*

$$\ell_S(gh) \leq \ell_S(g) + \ell_S(h) \tag{6.5}$$

*for all $g, h \in G$.*

*Proof.* Let $g, h \in G$. Set $m = \ell_S(g)$ and $n = \ell_S(h)$. Then there exist $s_1, s_2, \ldots, s_m$ and $t_1, t_2, \ldots, t_n$ in $S$ such that $g = s_1 s_2 \cdots s_m$ and $h = t_1 t_2 \cdots t_n$. We have $g^{-1} = s_m^{-1} \cdots s_2^{-1} s_1^{-1}$ so that $\ell_S(g^{-1}) \leq m = \ell_S(g)$. Exchanging the roles of $g$ and $g^{-1}$ we get $\ell_S(g) \leq \ell_S(g^{-1})$ and therefore $\ell_S(g^{-1}) = \ell_S(g)$.

On the other hand, we have $gh = s_1 s_2 \cdots s_m t_1 t_2 \cdots t_n$ so that $\ell_S(gh) \leq m + n = \ell_S(g) + \ell_S(h)$. □

Consider the map $d_S = d_S^G \colon G \times G \to \mathbb{N}$ defined by

$$d_S(g, h) = \ell_S(g^{-1}h) \tag{6.6}$$

for all $g, h \in G$.

**Proposition 6.1.2.** *The map $d_S$ is a metric on $G$.*

*Proof.* Let $g, h, k \in G$. It follows from (6.3) that $d_S(g,h) = 0$ if and only if $g = h$. Moreover, from (6.4) we deduce that $d_S(g,h) = d_S(h,g)$. Finally, using (6.5) we get

$$\begin{aligned} d_S(g,k) + d_S(k,h) &= \ell_S(g^{-1}k) + \ell_S(k^{-1}h) \\ &\geq \ell_S((g^{-1}k)(k^{-1}h)) \\ &= \ell_S(g^{-1}h) \\ &= d_S(g,h). \end{aligned}$$

□

The metric $d_S$ is called the *word metric* on $G$ associated with the finite symmetric generating subset $S$.

**Proposition 6.1.3.** *The metric $d_S$ is invariant by left multiplication, that is,*

$$d_S(gg_1, gg_2) = d_S(g_1, g_2)$$

*for all $g, g_1, g_2 \in G$.*

*Proof.* We have

$$d_S(gg_1, gg_2) = \ell_S((gg_1)^{-1}gg_2) = \ell_S(g_1^{-1}g^{-1}gg_2) = \ell_S(g_1^{-1}g_2) = d_S(g_1, g_2).$$

□

We may rephrase the previous result by saying that the action of $G$ on itself by left multiplication is an isometric action with respect to the word metric.

For $g \in G$ and $n \in \mathbb{N}$, we denote by

$$B_S^G(g,n) = \{h \in G : d_S(g,h) \leq n\}$$

the *ball of radius* $n$ in $G$ centered at the element $g \in G$. When $g = 1_G$ we have $B_S^G(1_G, n) = \{h \in G : \ell_S(h) \leq n\}$ and we simply write $B_S^G(n)$ instead of $B_S^G(1_G, n)$. Also, when there is no ambiguity on the group $G$, we omit the supscript "$G$" and we simply write $B_S(g,n)$ and $B_S(n)$ instead of $B_S^G(g,n)$ and $B_S^G(n)$.

## 6.2 Labeled Graphs

Let $S$ be a set. An *$S$-labeled graph* is a pair $\mathcal{Q} = (Q, E)$, where $Q$ is a set, called the set of *vertices*, and $E$ is a subset of $Q \times S \times Q$, called the set of *edges*. The projection map $\lambda\colon E \to S$, defined by $\lambda(e) = s$ for all $e = (q, s, q') \in E$, is called the *labelling map*. Also consider the projection maps $\alpha, \omega\colon E \to Q$ defined by $\alpha(e) = q$ and $\omega(e) = q'$ for all $e = (q, s, q') \in E$. Then, $\alpha(e)$ and $\omega(e)$ are called the *initial* and *terminal* vertices of the edge $e \in E$.

Let $\mathcal{Q}_1 = (Q_1, E_1)$ and $\mathcal{Q}_2 = (Q_2, E_2)$ be two $S$-labeled graphs. An *$S$-labeled graph homomorphism* from $\mathcal{Q}_1$ to $\mathcal{Q}_2$ is a map $\phi\colon Q_1 \to Q_2$ such that $(\phi(q), s, \phi(q')) \in E_2$ for all $(q, s, q') \in E_1$. An $S$-labeled graph homomorphism $\phi\colon Q_1 \to Q_2$ which is bijective and such that the inverse map $\phi^{-1}\colon Q_2 \to Q_1$ is also an $S$-labeled graph homomorphism is called a *$S$-labeled graph isomorphism*. If such an $S$-labeled graph isomorphism exists one says that the $S$-labeled graphs $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are isomorphic.

An $S$-labeled graph $\mathcal{Q} = (Q, E)$ is said to be *finite* if the sets $Q$ and $E$ are both finite.

Let $\mathcal{Q} = (Q, E)$ be an $S$-labeled graph.

An *$S$-labeled subgraph* of $\mathcal{Q}$ is an $S$-labeled graph $\mathcal{Q}' = (Q', E')$ such that $Q' \subset Q$ and $E' \subset E$.

Let $Q' \subset Q$ and set $E' = E \cap (Q' \times S \times Q')$. Then the $S$-labeled graph $\mathcal{Q}' = (Q', E')$ is called the *$S$-labeled subgraph* of $\mathcal{Q}$ *induced* on $Q'$. Sometimes, by abuse of language, we shall simply denote by $Q'$ the $S$-labeled graph $\mathcal{Q}' = (Q', E')$.

If there exist vertices $q, q' \in Q$ and labels $s_1 \neq s_2$ in $S$ such that $(q, s_1, q'), (q, s_2, q') \in E$, in other words, if there exist two edges with the same initial and terminal vertices but with different labels, then one says that $\mathcal{Q}$ has *multiple edges*.

An edge of the form $(q, s, q)$, $q \in Q$, $s \in S$, is called a *loop* at $q$.

A *path* in $\mathcal{Q}$ is a finite sequence of edges $\pi = (e_1, e_2, \ldots, e_n)$, $e_1, e_2, \ldots, e_n \in E$, such that $\omega(e_i) = \alpha(e_{i+1})$ for all $i = 1, \ldots, n-1$. The integer $n$ is called the *length* of the path $\pi$ and it is denoted by $\ell(\pi)$. The vertices $\pi^- = \alpha(e_1)$ and $\pi^+ = \omega(e_n)$ are called the *initial* and *terminal* vertices of $\pi$ and one says that $\pi$ *connects* $\pi^-$ to $\pi^+$. The *label* of a path $\pi = (e_1, e_2, \ldots, e_n)$ is defined by $\lambda(\pi) = (\lambda(e_1), \lambda(e_2), \ldots, \lambda(e_n)) \in S^n$. For $q \in Q$, we also admit the *empty path* starting and ending at $q$. It has length 0 and its label is the empty word $\epsilon$. Let $\pi_1 = (e_1, e_2, \ldots, e_n)$ and $\pi_2 = (e'_1, e'_2, \ldots, e'_m)$ be two paths with $\pi_1^+ = \pi_2^-$. Then the path

$$\pi_1\pi_2 = (e_1, e_2, \ldots, e_n, e'_1, e'_2, \ldots, e'_m)$$

is called the *composition* of the paths $\pi_1$ and $\pi_2$. Note that $\ell(\pi_1\pi_2) = \ell(\pi_1) + \ell(\pi_2)$ and that $\lambda(\pi_1\pi_2) = \lambda(\pi_1)\lambda(\pi_2)$.

Let $q, q' \in Q$. If there is no path connecting $q$ to $q'$ we set $d_{\mathcal{Q}}(q, q') = \infty$, otherwise, we set

$$d_{\mathcal{Q}}(q, q') = \min\{\ell(\pi) : \pi \text{ a path connecting } q \text{ to } q'\}. \tag{6.7}$$

A path $\pi$ connecting $q$ to $q'$ with minimal length, that is, such that $\ell(\pi) = d_{\mathcal{Q}}(q, q')$, is called a *geodesic path* from $q$ to $q'$.

**Proposition 6.2.1.** *Let $\mathcal{Q} = (Q, E)$ be a labeled graph. Then, for all $q, q', q'' \in Q$ one has:*

(i) $d_{\mathcal{Q}}(q,q') \in \mathbb{N} \cup \{\infty\}$;
(ii) $d_{\mathcal{Q}}(q,q') = 0$ *if and only if* $q = q'$;
(iii) $d_{\mathcal{Q}}(q,q') < \infty$ *if and only if there exists a path which connects* $q$ *to* $q'$;
(iv) $d_{\mathcal{Q}}(q,q'') \leq d_{\mathcal{Q}}(q,q') + d_{\mathcal{Q}}(q',q'')$ *(triangular inequality).*

*Proof.* The statements (i), (ii), and (iii) are trivial. Let us prove (iv). If $d_{\mathcal{Q}}(q,q') = \infty$ or $d_{\mathcal{Q}}(q',q'') = \infty$ there is nothing to prove. Otherwise, let $\pi_1$ be a geodesic path connecting $q$ to $q'$ and $\pi_2$ a geodesic path connecting $q'$ to $q''$. Then $\pi_1\pi_2$ connects $q$ to $q''$ and therefore

$$d_{\mathcal{Q}}(q,q'') \leq \ell(\pi_1\pi_2) = \ell(\pi_1) + \ell(\pi_2) = d_{\mathcal{Q}}(q,q') + d_{\mathcal{Q}}(q',q'').$$

□

Note that, in general, for $q, q' \in Q$ one has $d_{\mathcal{Q}}(q,q') \neq d_{\mathcal{Q}}(q',q)$. For $q \in Q$ and $r \in \mathbb{N}$ we denote by $B(q,r) = \{q' \in Q : d_{\mathcal{Q}}(q,q') \leq r\}$ the *ball of radius* $r$ centered at $q$.

The labeled graph $\mathcal{Q}$ is said to be *connected* if given any two vertices $q, q' \in Q$ there exists a path which connects $q$ to $q'$. Equivalently, $\mathcal{Q}$ is connected if and only if $d_{\mathcal{Q}}(q,q') < \infty$ for all $q, q' \in Q$.

A path $\pi$ in $\mathcal{Q}$ such that $\pi^- = \pi^+$ is said to be *closed*.

The *valence* (or *degree*) $\delta(q)$ of a vertex $q \in Q$ is the cardinality of the set $\{e \in E : \alpha(e) = q\}$. If $\delta(q) < \infty$ for all $q \in Q$, one says that $\mathcal{Q}$ is *locally finite*. If all vertices have the same finite valence $k$ then one says that $\mathcal{Q}$ is *regular* of *degree* $k$.

Suppose that $S$ is equipped with an involution $s \mapsto \overline{s}$. Then the labeled graph $\mathcal{Q}$ is said to be *edge-symmetric* if for all $e = (q,s,q') \in E$ one has that the *inverse edge* $e^{-1} = (q', \overline{s}, q)$ also belongs to $E$.

If $\mathcal{Q}$ is edge-symmetric and $(q,s,q') \in E$ one says that the vertices $q$ and $q'$ are *neighbors*. Suppose that $\mathcal{Q}$ is edge-symmetric. If $\pi = (e_1, e_2, \ldots, e_n)$ is a path in $\mathcal{Q}$ connecting a vertex $q$ to a vertex $q'$, then $\pi^{-1} = ({e_n}^{-1}, {e_{n-1}}^{-1}, \ldots, {e_2}^{-1}, {e_1}^{-1})$ is a path in $\mathcal{Q}$ connecting $q'$ to $q$. The path $\pi^{-1}$ is called the *inverse* of $\pi$. Note that $\ell(\pi^{-1}) = \ell(\pi)$.

**Corollary 6.2.2.** *Suppose that* $\mathcal{Q}$ *is edge-symmetric and connected. Then the map* $d_{\mathcal{Q}} \colon Q \times Q \to \mathbb{N}$ *defined in* (6.7) *is a metric on the set* $Q$ *of vertices of* $\mathcal{Q}$.

*Proof.* This follows immediately from Proposition 6.2.1 and the fact that $d_{\mathcal{Q}}(q,q') = d_{\mathcal{Q}}(q',q)$ for all $q, q' \in Q$. The latter immediately follows from the fact that the map $\pi \mapsto \pi^{-1}$ yields a length-preserving bijection from the set of paths connecting $q$ to $q'$ onto the set of paths connecting $q'$ to $q$. □

The metric $d_{\mathcal{Q}}$ is called the *graph metric* on $\mathcal{Q}$.

Let $\pi = (e_1, e_2, \ldots, e_n)$ be a path in $\mathcal{Q}$. We associate with $\pi$ the sequence $\pi_Q = (q_0, q_1, \ldots, q_n) \in Q^{n+1}$ defined by $q_i = \alpha(e_{i+1})$ for all $i = 0, 1, \ldots, n$ and $q_n = \omega(e_n)$. The vertices $q_0, q_1, \ldots, q_n$ are called the vertices *visited* by $\pi$.

One says that $\pi$ is *simple* if $q_i \neq q_j$ for all $0 \leq i \neq j \leq n$. If $\pi$ is closed, one says that $\pi$ is a *closed simple* path if $q_i \neq q_j$ for all $0 \leq i, j \leq n$ such that $(i, j) \neq (0, n)$. One says that $\pi$ is *proper*, or with no *back–tracking*, if $e_{i+1} \neq e_i^{-1}$ for $i = 1, 2, \ldots, n-1$.

An edge-symmetric, connected labeled graph with no loops, no multiple edges and with no closed proper paths is called a *tree*.

## 6.3 Cayley Graphs

Let $G$ be a finitely generated group and $S$ a finite symmetric generating subset of $G$. The *Cayley graph* of $G$ with respect to $S$ is the $S$-labeled graph $\mathcal{C}_S(G) = (Q, E)$ whose vertices are the group elements, that is, $Q = G$ and the edge set is $E = \{(g, s, gs) : g \in G \text{ and } s \in S\}$.

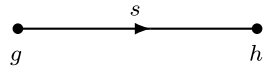



**Fig. 6.1** An edge in the Cayley graph $\mathcal{C}_S(G)$ is a triple $(g, s, h)$, where $g \in G$, $s \in S$, and $h = gs$

Note that, as $S$ is symmetric, the inverse map $s \mapsto s^{-1}$ is an involution on $S$. Since we have $(h, s^{-1}, g) \in E$ for all $(g, s, h) \in E$, it follows that the Cayley graph $\mathcal{C}_S(G)$ is edge-symmetric with respect to the inverse map on $S$. Moreover, the Cayley graph is connected. For, given $g, h \in G$, as $S$ generates $G$, we can find a nonnegative integer $n$ and $s_1, s_2, \ldots, s_n \in S$ such that $g^{-1}h = s_1 s_2 \cdots s_n$. Then the path $\pi = (e_1, e_2, \ldots, e_n)$, where $e_i = (gs_1 s_2 \cdots s_{i-1}, s_i, gs_1 s_2 \cdots s_{i-1} s_i)$, $i = 1, 2, \ldots, n$, connects $g$ to $h = gs_1 s_2 \cdots s_n$. Also observe that if $1_G \in S$ then $\mathcal{C}_S(G)$ has a loop at each vertex and that, on the contrary, $\mathcal{C}_S(G)$ has no loops if $1_G \notin S$. On the other hand, $\mathcal{C}_S(G)$ has no multiple edges, that is, for all $g, h \in G$, there exists at most one $s \in S$ such that $(g, s, h) \in E$, namely $s = g^{-1}h$ if $g^{-1}h \in S$.

**Proposition 6.3.1.** *Let $G$ be a finitely generated group. Let $S \subset G$ be a finite symmetric generating subset. Then, the $S$-distance of two group elements equals the graph distance of the same elements viewed as vertices in the associated Cayley graph $\mathcal{C}_S(G)$. In other words,*

$$d_S(g, h) = d_{\mathcal{C}_S(G)}(g, h) \tag{6.8}$$

*for all $g, h \in G$.*

*Proof.* Let $g, h \in G$ and suppose that $\pi = (e_1, e_2, \ldots, e_n)$ is a geodesic path connecting $g$ and $h$. Let $\lambda(\pi) = (s_1, s_2, \ldots, s_n)$ be the label of $\pi$. It then follows that $d_S(g, h) = \ell_S(g^{-1}h) = \ell_S(s_1 s_2 \cdots s_n) \leq n = \ell(\pi) = d_{\mathcal{C}_S(G)}(g, h)$.

Conversely, suppose that $d_S(g,h) = m$. Then we can find $s'_1, s'_2, \ldots, s'_m \in S$ such that $g^{-1}h = s'_1 s'_2 \cdots s'_m$. Then the unique path $\pi' = (e'_1, e'_2, \ldots, e'_m)$ with $(\pi')^- = g$ and label $\lambda(\pi') = s'_1 s'_2 \cdots s'_m$ clearly satisfies $(\pi')^+ = h$, that is, it connects $g$ to $h$. We deduce that $d_{\mathcal{C}_S(G)}(g,h) \leq \ell(\pi') = m = d_S(g,h)$. Then (6.8) follows. □

Note that, in particular, two distinct elements $g$ and $h$ in $G$ are neighbors in the graph $\mathcal{C}_S(G)$ if and only if $d_S(g,h) = 1$. For all $g \in G$, the map $s \mapsto gs$ is a bijection from $S$ onto the set $gS$ of all neighbors of $g$ in $\mathcal{C}_S(G)$. In particular, all vertices $g$ in $\mathcal{C}_S(G)$ have the same degree $\delta(g) = |S|$.

Summarizing, we have that a Cayley graph is a connected, edge-symmetric and regular labeled graph.

*Examples 6.3.2.* We graphically represent Cayley graphs by connecting two neighboring vertices by a single directed labeled arc $e$ (see Fig. 6.1). Thus one should think of the inverse edge $e^{-1}$ as the oppositely directed arc with label $\lambda(e)^{-1}$.

(a) Let $G = \mathbb{Z}$ and take $S = \{1, -1\}$ as a finite symmetric generating subset of $G$. Then the Cayley graph $\mathcal{C}_S(\mathbb{Z})$ is represented in Fig. 6.2.

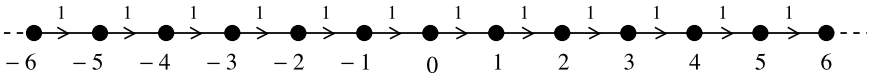



**Fig. 6.2** The Cayley graph of $G = \mathbb{Z}$ for $S = \{1, -1\}$

(b) Let $G = \mathbb{Z}$ and take $S = \{1, 0, -1\}$ as a finite symmetric generating subset of $G$. Then the Cayley graph $\mathcal{C}_S(\mathbb{Z})$ is represented in Fig. 6.3. Note that as $0 = 1_{\mathbb{Z}} \in S$, we have a loop at each vertex in $\mathcal{C}_S(\mathbb{Z})$.

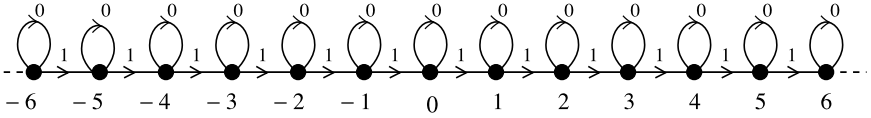



**Fig. 6.3** The Cayley graph of $G = \mathbb{Z}$ for $S = \{1, 0, -1\}$

(c) Let $G = \mathbb{Z}$ and $S = \{2, -2, 3, -3\}$. Then, the corresponding Cayley graph $\mathcal{C}_S(\mathbb{Z})$ is represented in Fig. 6.4.

(d) Let $G = \mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z})$, where $\mathbb{Z}/2\mathbb{Z} = \{\overline{0}, \overline{1}\}$ and take $S = \{(1, \overline{0}), (-1, \overline{0}), (0, \overline{1})\}$. Then the Cayley graph $\mathcal{C}_S(G)$ is represented by the bi-infinite ladder as in Fig. 6.5.

(e) Let $G = D_\infty$ be the infinite dihedral group, that is, the group of isometries of the real line $\mathbb{R}$ generated by the reflections $r \colon \mathbb{R} \to \mathbb{R}$ and

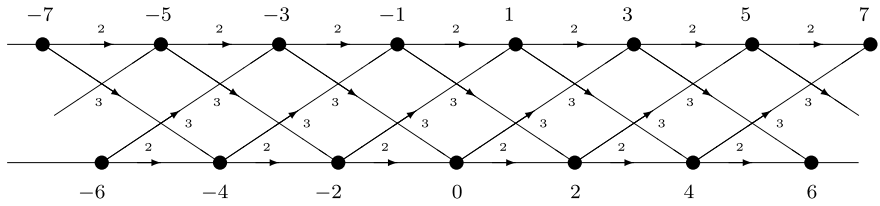



**Fig. 6.4** The Cayley graph of $G = \mathbb{Z}$ for $S = \{2, -2, 3, -3\}$

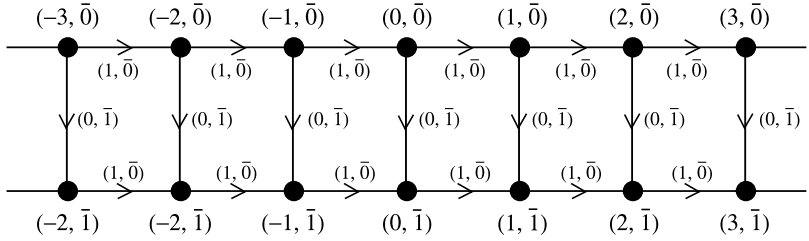



**Fig. 6.5** The Cayley graph of $G = \mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z})$ for $S = \{(1, \overline{0}), (-1, \overline{0}), (0, \overline{1})\}$

$s\colon \mathbb{R} \to \mathbb{R}$ defined by

$$\begin{aligned} r(x) &= -x && \text{(symmetry with respect to 0)} \\ s(x) &= 1 - x && \text{(symmetry with respect to 1/2)} \end{aligned}$$

for all $x \in \mathbb{R}$. Note that $r^2 = s^2 = 1_G$. Taking $S = \{r, s\}$, the Cayley graph $\mathcal{C}_S(G)$ is as in Fig. 6.6.

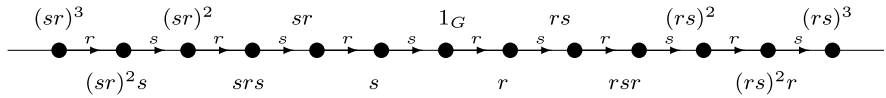



**Fig. 6.6** The Cayley graph of $G = D_\infty$ for $S = \{r, s\}$

(f) Let $G$ be the infinite dihedral group, as in the previous example. Denote by $t\colon \mathbb{R} \to \mathbb{R}$ the map defined by $t(x) = x + 1$ for all $x \in \mathbb{R}$ (translation). Then we have $t = sr$ and hence $s = tr$. It follows that the set $S = \{r, t, t^{-1}\}$ is also a symmetric generating subset of $G$ and the corresponding Cayley graph $\mathcal{C}_S(G)$ is as in Fig. 6.7.

(g) Let $G = \mathbb{Z}^2$ and take $S = \{(1, 0), (-1, 0), (0, 1), (0, -1)\}$ as a finite and symmetric generating subset of $G$. Then, the corresponding Cayley graph $\mathcal{C}_S(\mathbb{Z}^2)$ is given in Fig. 6.8.

(h) Let $G = \mathbb{Z}^2$ and take

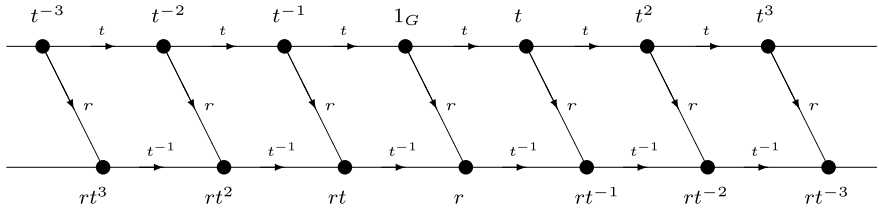



**Fig. 6.7** The Cayley graph of $G = D_\infty$ for $S = \{r, t, t^{-1}\}$

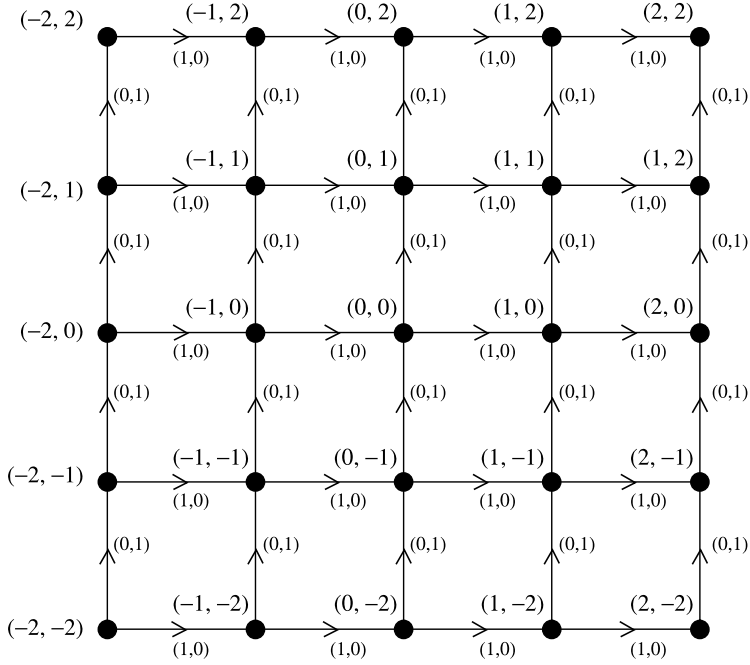



**Fig. 6.8** The Cayley graph of $G = \mathbb{Z}^2$ for $S = \{(1,0), (-1,0), (0,1), (0,-1)\}$

$$S = \{(1,0), (-1,0), (0,1), (0,-1), (1,1), (-1,1), (1,-1), (-1,-1)\}.$$

Then, the corresponding Cayley graph $\mathcal{C}_S(\mathbb{Z}^2)$ is as in Fig. 6.9.

(i) Let $G = F_X$ be the free group based on a nonempty finite set $X$ and take $S = X \cup X^{-1}$. Then, the Cayley graph $\mathcal{C}_S(F_X)$ is a regular tree of degree $|S| = 2|X|$ (see Fig. 6.10). Indeed, since $1_{F_X} \notin S$, the Cayley graph $\mathcal{C}_S(F_X)$ does not have loops. Moreover, if $\pi$ is a nontrivial proper path in $\mathcal{C}_S(F_X)$ with label $\lambda(\pi) = (s_1, s_2, \ldots, s_n) \in S^*$, then, the word $w = s_1 s_2 \cdots s_n \in S^*$ is nonempty and, by definition of properness, it is reduced. Therefore $1_{F_X}$ and $h = s_1 s_2 \cdots s_n \in F_X$ are distinct and we have $\pi^+ = \pi^- h \neq \pi^-$, that is, $\pi$ is not closed. This shows that $\mathcal{C}_S(F_X)$ is a tree.

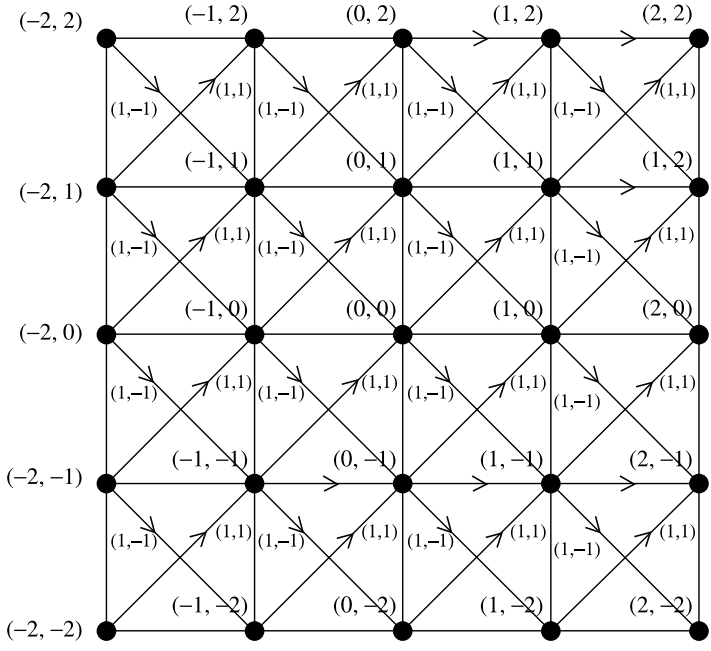



**Fig. 6.9** The Cayley graph of $G = \mathbb{Z}^2$ for $S = \{(1,0), (-1,0), (0,1), (0,-1), (1,1), (-1,1), (1,-1), (-1,-1)\}$

## 6.4 Growth Functions and Growth Types

Let $G$ be a finitely generated group and $S \subset G$ a finite and symmetric generating subset of $G$. The *growth function* of $G$ relative to $S$ is the function $\gamma_S^G : \mathbb{N} \to \mathbb{N}$ defined by

$$\gamma_S^G(n) = |B_S^G(n)| = |\{g \in G : \ell_S(g) \leq n\}| \tag{6.9}$$

for all $n \in \mathbb{N}$. When there is no ambiguity we omit the supscript "$G$" and simply denote it by $\gamma_S$.

Note that $\gamma_S(0) = |B_S(0)| = |\{1_G\}| = 1$ and that $\gamma_S(n) \leq \gamma_S(n+1)$ for all $n \in \mathbb{N}$. Also, as the map $(s_1, s_2, \ldots, s_n) \mapsto s_1 s_2 \cdots s_n$ is a surjection from $(S \cup \{1_G\})^n$ to $B_S(n)$, one has

$$\gamma_S(n) \leq |S \cup \{1_G\}|^n \tag{6.10}$$

for all $n \in \mathbb{N}$.

**Proposition 6.4.1.** *Let $G$ be a finitely generated group and let $S$ and $S'$ be two finite symmetric generating subsets of $G$. Let $c = \max\{\ell_{S'}(s) : s \in S\}$.*

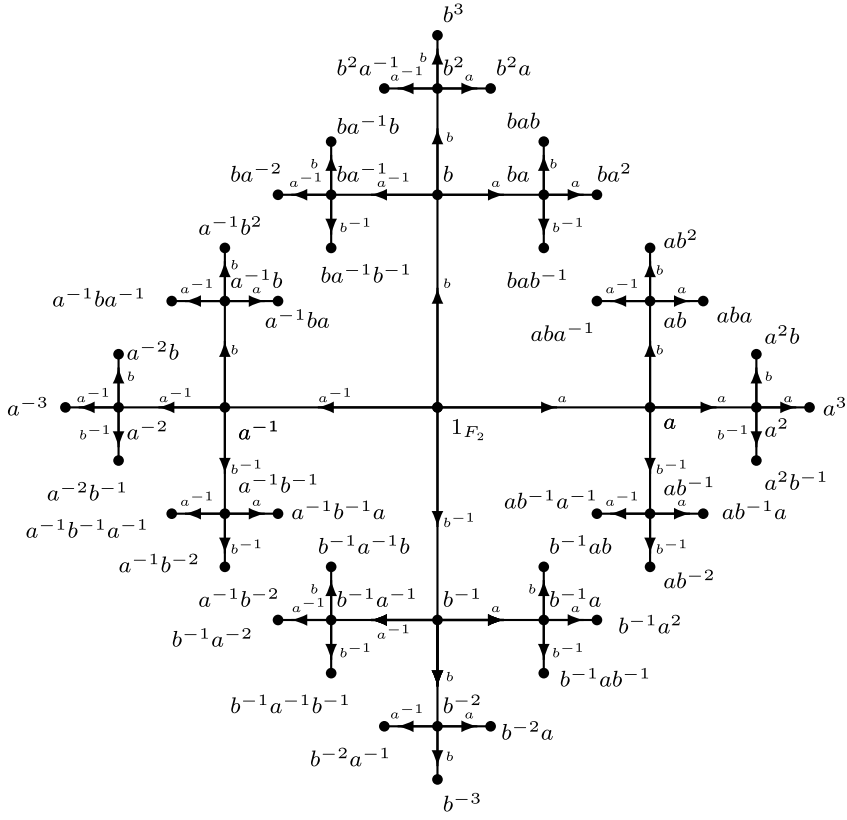



**Fig. 6.10** The Cayley graph of $G = F_2$ for $S = \{a, a^{-1}, b, b^{-1}\}$

*Then the following holds.*

(i) $\ell_{S'}(g) \leq c\ell_S(g)$ *for all* $g \in G$;
(ii) $d_{S'}(g,h) \leq cd_S(g,h)$ *for all* $g, h \in G$;
(iii) $B_S(n) \subset B_{S'}(cn)$ *for all* $n \in \mathbb{N}$;
(iv) $\gamma_S(n) \leq \gamma_{S'}(cn)$ *for all* $n \in \mathbb{N}$.

*Proof.* (i) Let $g \in G$. Suppose that $\ell_S(g) = n$. Then there exist $s_1, s_2, \ldots, s_n \in S$ such that $g = s_1 s_2 \cdots s_n$. Using (6.5) we get

$$\ell_{S'}(g) = \ell_{S'}(s_1 s_2 \cdots s_n) \leq \sum_{i=1}^{n} \ell_{S'}(s_i) \leq cn.$$

(ii) By applying (i) we have, for all $g, h \in G$,

$$d_{S'}(g,h) = \ell_{S'}(g^{-1}h) \leq c\ell_S(g^{-1}h) = cd_S(g,h).$$

Finally, from (ii), we have that $d_{S'}(g, 1_G) \leq cn$ if $d_S(g, 1_G) \leq n$ for all $g \in G$. This gives (iii). Thus

$$\gamma_S(n) = |B_S(n)| \leq |B_{S'}(cn)| = \gamma_{S'}(cn)$$

for all $n \in \mathbb{N}$. □

Two metrics $d$ and $d'$ on a set $X$ are said to be *Lipschitz-equivalent* if there exist constants $c_1, c_2 > 0$ such that

$$c_1 d(x, y) \leq d'(x, y) \leq c_2 d(x, y)$$

for all $x, y \in X$.

**Corollary 6.4.2.** *Let $G$ be a finitely generated group and let $S$ and $S'$ be two finite symmetric generating subsets of $G$. Then, the word metrics $d_S$ and $d_{S'}$ are Lipschitz-equivalent.* □

A non-decreasing function $\gamma\colon \mathbb{N} \to [0, +\infty)$ is called a *growth function*. Let $\gamma, \gamma'\colon \mathbb{N} \to [0, +\infty)$ be two growth functions. One says that $\gamma'$ *dominates* $\gamma$, and one writes $\gamma \preceq \gamma'$, if there exists an integer $c \geq 1$ such that

$$\gamma(n) \leq c\gamma'(cn) \text{ for all } n \geq 1.$$

One says that $\gamma$ and $\gamma'$ are *equivalent* and one writes $\gamma \sim \gamma'$ if $\gamma \preceq \gamma'$ and $\gamma' \preceq \gamma$.

**Proposition 6.4.3.** *We have the following:*

(i) $\preceq$ *is reflexive and transitive;*
(ii) $\sim$ *is an equivalence relation;*
(iii) *let $\gamma_1, \gamma_2, \gamma_1', \gamma_2'\colon \mathbb{N} \to [0, +\infty)$ be growth functions. Suppose that $\gamma_1 \sim \gamma_1'$, $\gamma_2 \sim \gamma_2'$ and that $\gamma_1 \preceq \gamma_2$. Then $\gamma_1' \preceq \gamma_2'$.*

*Proof.* It is clear that $\preceq$ is reflexive. Let $\gamma_1, \gamma_2, \gamma_3\colon \mathbb{N} \to [0, +\infty)$ be growth functions. Suppose that $\gamma_1 \preceq \gamma_2$ and that $\gamma_2 \preceq \gamma_3$. Let $c_1$ and $c_2$ be positive integers such that $\gamma_1(n) \leq c_1\gamma_2(c_1 n)$ and $\gamma_2(n) \leq c_2\gamma_3(c_2 n)$ for all $n \geq 1$. Then, taking $c = c_1c_2$ one has

$$\gamma_1(n) \leq c_1\gamma_2(c_1 n) \leq c_1c_2\gamma_3(c_2c_1 n) = c\gamma_3(cn)$$

for all $n \geq 1$. Thus $\preceq$ is also transitive. This shows (i).

Property (ii) immediately follows from (i) and the definition of $\sim$.

Finally, suppose that $\gamma_1, \gamma_2, \gamma_1', \gamma_2'$ satisfy the hypotheses of (iii). Then we have in particular, $\gamma_1' \preceq \gamma_1$, $\gamma_1 \preceq \gamma_2$ and $\gamma_2 \preceq \gamma_2'$. From (i) we deduce that $\gamma_1' \preceq \gamma_2'$. □

Let $\gamma\colon \mathbb{N} \to [0, +\infty)$ be a growth function. We denote by $[\gamma]$ the $\sim$-equivalence class of $\gamma$. By abuse of notation, we shall also write $[\gamma] \sim \gamma(n)$.

If $\gamma_1$ and $\gamma_2$ are two growth functions, we write $[\gamma_1] \preceq [\gamma_2]$ if $\gamma_1 \preceq \gamma_2$. This definition makes sense by virtue of Proposition 6.4.3(iii). Note that, this way, $\preceq$ becomes a partial ordering on the set of equivalence classes of growth functions.

*Examples 6.4.4.* (a) Let $\alpha$ and $\beta$ be nonnegative real numbers. Then $n^\alpha \preceq n^\beta$ if and only if $\alpha \leq \beta$, and $n^\alpha \sim n^\beta$ if and only if $\alpha = \beta$.

(b) Let $\gamma\colon \mathbb{N} \to [0, +\infty)$ be a growth function. Suppose that $\gamma$ is a polynomial of degree $d$ for some $d \geq 0$. Then one has $\gamma(n) \sim n^d$.

(c) Let $a, b \in (1, +\infty)$. Then

$$a^n \sim b^n. \tag{6.11}$$

Indeed, suppose for instance that $a \leq b$. On the one hand we have $a^n \leq b^n$ for all $n \geq 1$, so that trivially $a^n \preceq b^n$. On the other, setting $c = [\log_a b] + 1 > 1$ (here $[\,\cdot\,]$ denotes the integer part), one has

$$b^n = (a^{\log_a b})^n = a^{(\log_a b)n} \leq a^{cn} \leq ca^{cn}$$

for all $n \geq 1$, so that $b^n \preceq a^n$. We deduce that $a^n \sim b^n$. In particular, we have $a^n \sim \exp(n)$ for all $a \in (1, +\infty)$.

(d) Let $d \geq 0$ be an integer. Then $n^d \preceq \exp(n)$ and $n^d \not\sim \exp(n)$. As a consequence if $\gamma\colon \mathbb{N} \to [0, +\infty)$ is a growth function such that $\gamma(n) \preceq n^d$, then $\gamma \preceq \exp(n)$ and $\gamma \not\sim \exp(n)$. Indeed, since $\lim_{n\to\infty} \frac{n^d}{\exp(n)} = 0$, the sequence $(\frac{n^d}{\exp(n)})_{n\geq 1}$ is bounded and we can find an integer $c \geq 1$ such that $\frac{n^d}{\exp(n)} \leq c$ for all $n \geq 1$. It follows that $n^d \leq c\exp(n) \leq c\exp(cn)$ for all $n \geq 1$ and therefore $n^d \preceq \exp(n)$.

On the other hand, suppose by contradiction that $\exp(n) \preceq n^d$. Then we can find an integer $c > 0$ such that $\exp(n) \leq c(cn)^d$ for all $n \geq 1$. But then $\frac{\exp(n)}{n^d} \leq c^{d+1}$ for all $n \geq 1$ contradicting the fact that $\lim_{n\to\infty} \frac{\exp(n)}{n^d} = +\infty$. Thus $n^d \not\sim \exp(n)$.

The remaining statements concerning the growth function $\gamma$ immediately follow from transitivity of $\preceq$ (cf. Proposition 6.4.3(i)) and symmetry of $\sim$ (cf. Proposition 6.4.3(ii)).

From Proposition 6.4.1(iii) and (6.10) we deduce the following:

**Corollary 6.4.5.** *Let $G$ be a finitely generated group and let $S$ and $S'$ be two finite symmetric generating subsets of $G$. Then, the growth functions associated with $S$ and $S'$ are equivalent, that is, $\gamma_S \sim \gamma_{S'}$. Moreover, $\gamma_S(n) \preceq \exp(n)$.* □

Let $G$ be a finitely generated group. The *equivalence class* $[\gamma_S]$ of the growth functions associated with the finite symmetric generating subsets $S$ of $G$ is called the *growth type* of $G$ and we denote it by $\gamma(G)$.

**Proposition 6.4.6.** *Let $\gamma\colon \mathbb{N} \to [0, +\infty)$ be a growth function with $\gamma(0) > 0$. Then $\gamma \sim 1$ if and only if $\gamma$ is bounded.*

*Proof.* Suppose first that $\gamma$ is bounded. Then we can find an integer $c \geq 1$ such that $\gamma(n) \leq c$ for all $n \geq 1$. It follows that $\gamma(n) \leq c1(n) \leq c1(cn)$ for all $n \geq 1$. Thus $\gamma \preceq 1$. On the other hand, setting $c = [\frac{1}{\gamma(0)}] + 1$, we have $1(n) = 1 \leq c\gamma(0) \leq c\gamma(n) \leq c\gamma(cn)$ for all $n \geq 1$ so that $1 \preceq \gamma$. This shows that $\gamma \sim 1$.

Conversely, suppose that $\gamma \sim 1$. Then $\gamma \preceq 1$ and we can find an integer $c \geq 1$ such that $\gamma(n) \leq c1(cn) = c$ for all $n \geq 1$. This shows that $\gamma$ is bounded. □

**Corollary 6.4.7.** *Let $G$ be a finitely generated group. Then $\gamma(G) \sim 1$ if and only if $G$ is finite. As a consequence, all finite groups have the same growth type.*

*Proof.* Let $S$ be a finite and symmetric generating subset of $G$. Suppose that $\gamma(G) \sim \gamma_S(n) \sim 1$. Then, by Proposition 6.4.6 we deduce that $\gamma_S$ is bounded, say by an integer $c \geq 1$. This shows that $|G| \leq c$ and therefore $G$ is finite. Conversely, if $G$ is finite, we have $\gamma_S(n) = |B_S(n)| \leq |G|$ for all $n \geq 1$, that is, $\gamma_S$ is bounded. From Proposition 6.4.6 we deduce that $\gamma(G) \sim \gamma_S(n) \sim 1$. □

**Proposition 6.4.8.** *Let $G$ be an infinite finitely generated group. Then $n \preceq \gamma(G)$.*

*Proof.* Let $S$ be a finite symmetric generating subset of $G$. Consider the inclusions

$$\{1_G\} = B_S(0) \subset B_S(1) \subset B_S(2) \subset \cdots \subset B_S(n) \subset B_S(n+1) \subset \cdots \quad (6.12)$$

Let us show that if $B_S(n) = B_S(n+1)$ for some $n \in N$, then $B_S(n) = B_S(m)$ for all $m \geq n$. We proceed by induction on $m$. Suppose that $B_S(n) = B_S(m)$ for some $m \geq n+1$. For all $g \in B_S(m+1)$ there exist $g' \in B_S(m)$ and $s \in S$ such that $g = g's$. By the inductive hypothesis, $g' \in B_S(m-1)$ so that $g = g's \in B_S(m-1)S \subset B_S(m)$. Since $B_S(m) \subset B_S(m+1)$, it follows that $B_S(m+1) = B_S(m) = B_S(n)$.

As a consequence, if $B_S(n) = B_S(n+1)$ for some $n \in \mathbb{N}$, then we have $G = B_S(n)$. Since, by our assumptions, $G$ is infinite, we deduce that all the inclusions in (6.12) are strict. It follows that for all $n \in \mathbb{N}$ we have $n \leq |B_S(n)| = \gamma_S(n)$. This shows that $n \preceq \gamma_S(n)$ and therefore $n \preceq \gamma(G)$. □

**Definition 6.4.9.** Let $G$ be a finitely generated group.

One says that $G$ has *exponential* (resp. *subexponential*) *growth* if $\gamma(G) \sim \exp(n)$ (resp. $\gamma(G) \not\sim \exp(n)$).

One says that $G$ has *polynomial growth* if there exists an integer $d \geq 0$ such that $\gamma(G) \preceq n^d$.

**Proposition 6.4.10.** *Every finitely generated group of polynomial growth has subexponential growth.*

*Proof.* Let $G$ be a finitely generated group. It follows from Example 6.4.4(d) that if $\gamma(G) \preceq n^d$ for some integer $d \geq 0$, then $\gamma(G) \not\sim \exp(n)$. □

*Examples 6.4.11.* (a) Let $G = \mathbb{Z}$. With $S = \{1, -1\}$ one has that the ball of radius $r$ centered at the element $g \in G$ is the interval $[g - r, g + r] = \{n \in \mathbb{Z} : g - r \leq n \leq g + r\}$, see Fig. 6.11. We have $\gamma_S(n) = 2n + 1$. It follows that $\gamma(\mathbb{Z}) \sim \gamma_S(n) \sim n$. In particular, $\mathbb{Z}$ has polynomial growth.

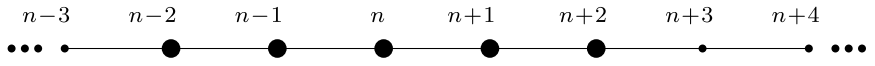



**Fig. 6.11** The ball $B_S(n, r) \subset \mathbb{Z}$ with $S = \{1, -1\}$; here $r = 2$

(b) Let $G = \mathbb{Z}^2$ and $S = \{(1, 0), (-1, 0), (0, 1), (0, -1)\}$. Then the ball of radius $r$ centered at the element $g = (n, m) \in G$ is the diagonal square with vertices $(n + r, m), (n - r, m), (n, m - r), (n, m + r)$, see Fig. 6.12. In the language of cellular automata, the ball $B_S^{\mathbb{Z}^2}(g, 1)$ is commonly called the *von Neumann neighborhood* of $g$. We have $\gamma_S(n) = 1 + \sum_{k=1}^{n} 4k = 2n^2 + 2n + 1$. It follows that $\gamma(\mathbb{Z}^2) \sim \gamma_S(n) \sim n^2$. In particular, $\mathbb{Z}^2$ has polynomial growth.

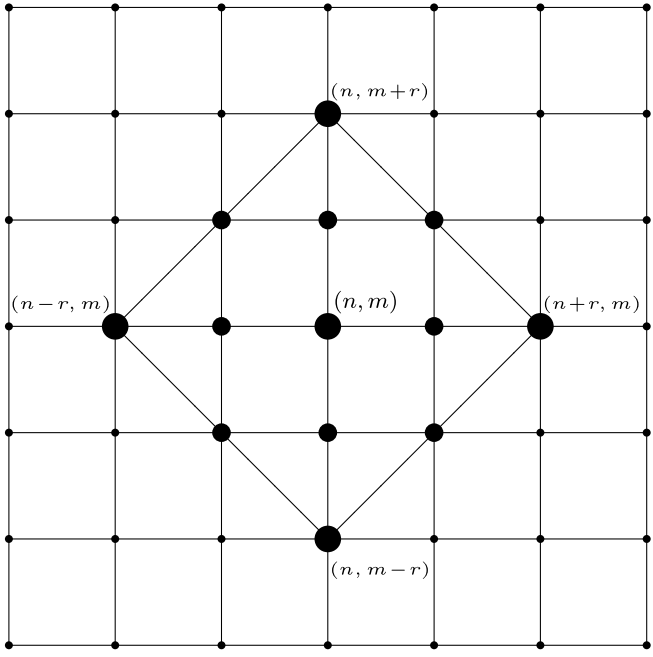



**Fig. 6.12** The ball $B_S(n, r) \subset \mathbb{Z}^2$ with $S = \{(1, 0), (-1, 0), (0, 1), (0, -1)\}$ and $r = 2$

(c) Let $G = \mathbb{Z}^2$ and

$$S = \{(1,0), (-1,0), (0,1), (0,-1), (1,1), (-1,1), (1,-1), (-1,-1)\}.$$

Then the ball of radius $r$ centered at the element $g = (n,m) \in G$ is the square $[n-r, n+r] \times [m-r, m+r]$, see Fig. 6.13. In the language of cellular automata, the ball $B_S^{\mathbb{Z}^2}(g,1)$ is commonly called the *Moore neighborhood* of $g$. We have $\gamma_S^{\mathbb{Z}^2}(n) = 4n^2 + 4n + 1 = (2n+1)^2 = (\gamma_S^{\mathbb{Z}}(n))^2$. Note that this yields again $\gamma(\mathbb{Z}^2) \sim \gamma_S(n) \sim n^2$ and the polynomial growth of $\mathbb{Z}^2$ (cf. Example 6.4.11(b) above).

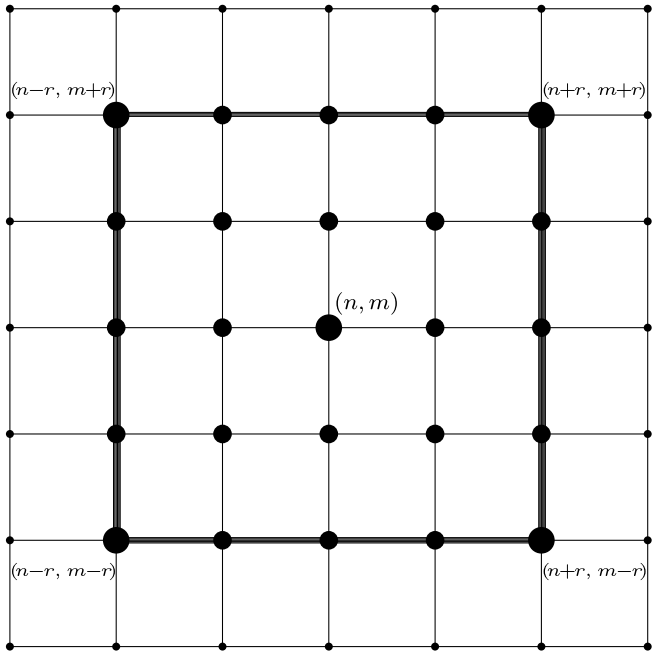



**Fig. 6.13** The ball $B_S(n,r) \subset \mathbb{Z}^2$ with $S = \{\pm(1,0), \pm(0,1), \pm(1,1), \pm(1,-1)\}$ and $r = 2$

(d) Let $G = \mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z})$ and $S = \{(1,\bar{0}), (-1,\bar{0}), (1,\bar{1}), (-1,\bar{1})\}$. Then the ball of radius $r$ centered at the element $(n,\bar{0})$ is represented in Fig. 6.14. We deduce that $\gamma_S(n) = (2n+1) + 2(n-1) + 1 = 4n$. Thus $\gamma(\mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z})) \sim \gamma_S(n) \sim n$. In particular, $\mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z})$ has polynomial growth.

(e) Let $G = D_\infty$ be the infinite dihedral group and $S = \{r, s\}$. Then the ball of radius $n$ centered at the element $g \in G$ is represented in Fig. 6.15. It follows that $\gamma_S(n) = 2n + 1$. Thus $\gamma(D_\infty) \sim \gamma_S(n) \sim n$. In particular, $D_\infty$ has polynomial growth.

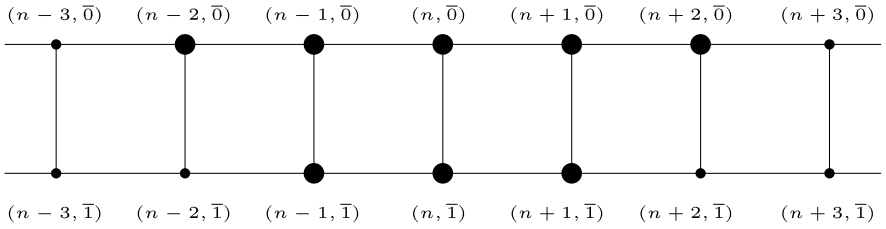



**Fig. 6.14** The ball $B_S((n,\overline{0}),r) \subset \mathbb{Z}\times(\mathbb{Z}/2\mathbb{Z})$ with $S=\{(1,\overline{0}),(-1,\overline{0}),(0,\overline{1})\}$; here $r=2$

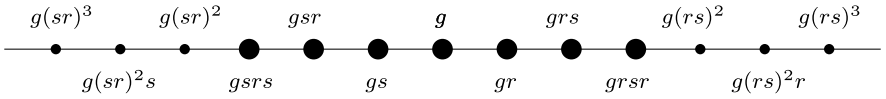



**Fig. 6.15** The ball $B_S(g,n) \subset D_\infty$ with $S=\{r,s\}$; here $n=3$

(f) Let $G=D_\infty$ be the infinite dihedral group and $S=\{r,t,t^{-1}\}$. Then the ball of radius $r$ centered at the element $g\in G$ is represented in Fig. 6.16. It follows that $\gamma_S(n)=(2n+1)+2(n-1)+1=4n$. Note that this yields again $\gamma(D_\infty)\sim\gamma_S(n)\sim n$ and the polynomial growth of $D_\infty$.

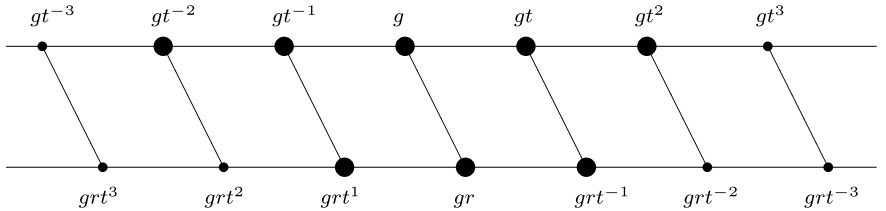



**Fig. 6.16** The ball $B_S(g,n) \subset D_\infty$ with $S=\{r,t,t^{-1}\}$; here $n=2$

(g) Let $G=F_k$ be the free group of rank $k\geq 2$. Let $\{a_1,a_2,\ldots,a_k\}$ be a free basis and set $S=\{a_1,a_1^{-1},a_2,a_2^{-1},\ldots,a_k,a_k^{-1}\}$. Then the ball of radius $r$ centered at the element $g\in G$ is the finite tree rooted at $g$ of depth $r$ (see Fig. 6.17). We have

$$\gamma_S^{F_k}(n)=1+2k\sum_{j=0}^{n-1}(2k-1)^j=\frac{k(2k-1)^n-1}{k-1}.$$

It follows that $\gamma(F_k)\sim\gamma_S(n)\sim(2k-1)^n\sim\exp(n)$. In particular, $F_k$ has exponential growth.

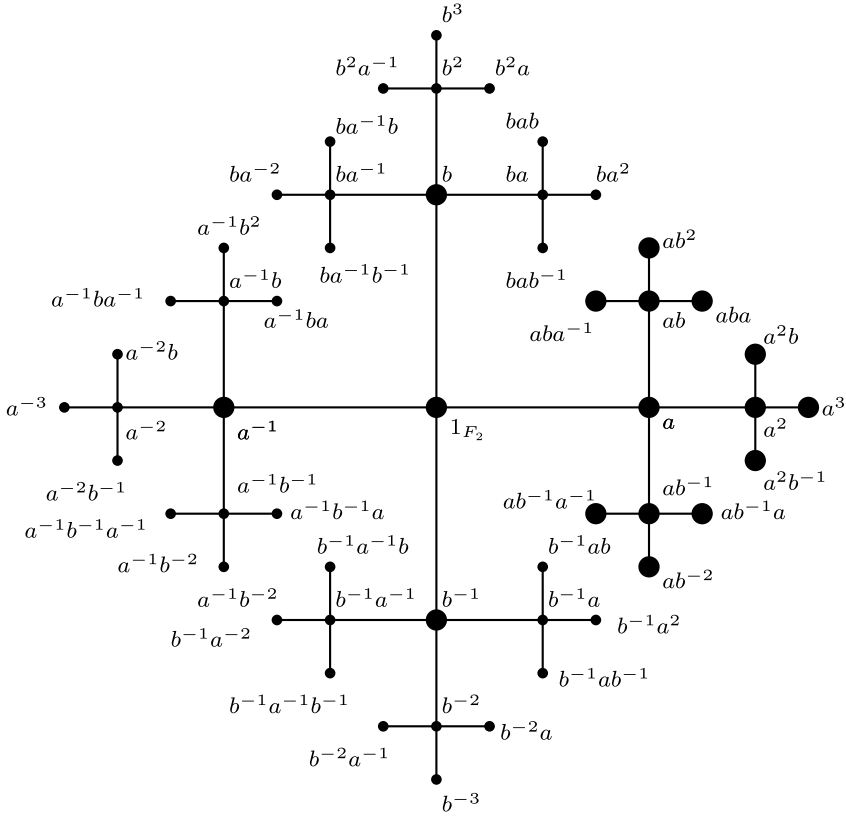



**Fig. 6.17** The ball $B_S(a,2) \subset F_2$ with $S = \{a, a^{-1}, b, b^{-1}\}$

## 6.5 The Growth Rate

**Lemma 6.5.1.** *Let $(a_n)_{n\geq 1}$ be a sequence of positive real numbers such that $a_{n+m} \leq a_n a_m$ for all $n, m \geq 1$. Then the limit*

$$\lim_{n\to\infty} \sqrt[n]{a_n}$$

*exists and equals* $\inf_{n\geq 1} \sqrt[n]{a_n}$.

*Proof.* Fix an integer $t \geq 1$ and, for all $n \geq 1$ write $n = qt + r$, with $0 \leq r < t$. Then $a_n \leq a_{qt} a_r \leq a_t^q a_r$ and $\sqrt[n]{a_n} \leq a_t^{q/n} \sqrt[n]{a_r}$. As $0 \leq r/n < t/n$ and $\lim_{n\to\infty} t/n = 0$, we have that $\lim_{n\to\infty} r/n = 0$ and $\lim_{n\to\infty} q/n = 1/t$. In particular $\lim_{n\to\infty} a_t^{q/n} a_r^{1/n} = a_t^{1/t} = \sqrt[t]{a_t}$. Thus, $\limsup_{n\to\infty} \sqrt[n]{a_n} \leq \sqrt[t]{a_t}$. This gives $\limsup_{n\to\infty} \sqrt[n]{a_n} \leq \inf_{t\geq 1} \sqrt[t]{a_t} \leq \liminf_{t\to\infty} \sqrt[t]{a_t}$ completing the proof. □

**Proposition 6.5.2.** *Let $G$ be a finitely generated group and let $S$ be a finite symmetric generating subset of $G$. Then, the limit*

$$\lambda_S = \lim_{n\to\infty} \sqrt[n]{\gamma_S(n)} \tag{6.13}$$

*exists and $\lambda_S \in [1, +\infty)$.*

*Proof.* From (6.5) we deduce that $B_S(n+m) \subseteq B_S(n)B_S(m)$ and therefore

$$\begin{aligned}\gamma_S(n+m) &= |B_S(n+m)| \\ &\leq |B_S(n)B_S(m)| \\ &\leq |B_S(n)| \cdot |B_S(m)| \\ &= \gamma_S(n)\gamma_S(m).\end{aligned}$$

Thus, the sequence $\{\gamma_S(n)\}_{n\geq 1}$ satisfies the hypotheses of the previous lemma and $\lambda_S$ exists and is finite. As $\gamma_S(n) \geq 1$ for all $n \in \mathbb{N}$ we also have $\lambda_S \geq 1$. □

**Definition 6.5.3.** The number $\lambda_S = \lambda_S^G$ in (6.13) is called the *growth rate* of $G$ with respect to $S$.

**Proposition 6.5.4.** *Let $G$ be a finitely generated group and let $S$ be a finite symmetric generating subset of $G$. Then $\lambda_S > 1$ if and only if $G$ has exponential growth.*

*Proof.* Suppose that $\gamma(G) \sim \exp(n)$. We then have $\exp(n) \preceq \gamma_S$ so that there exists an integer $c \geq 1$ such that $e^n \leq c\gamma_S(cn)$ for all $n \geq 1$. This implies

$$1 < \sqrt[c]{e} = \lim_{n\to\infty} \sqrt[cn]{e^n} \leq \lim_{n\to\infty} \sqrt[cn]{c\gamma_S(cn)} = \left(\lim_{n\to\infty} \sqrt[cn]{c}\right) \cdot \left(\lim_{n\to\infty} \sqrt[cn]{\gamma_S(cn)}\right) = \lambda_S.$$

Conversely, suppose that $\lambda_S > 1$. By Lemma 6.5.1 we have $\sqrt[n]{\gamma_S(n)} \geq \lambda_S$ so that

$$\gamma_S(n) \geq \lambda_S^n.$$

This shows that $\exp(n) \sim \lambda_S^n \preceq \gamma_S(n)$. By Corollary 6.4.5 we have $\gamma_S(n) \preceq \exp(n)$ and therefore $\gamma(G) \sim \gamma_S \sim \exp(n)$. □

**Corollary 6.5.5.** *Let $G$ be a finitely generated group and let $S$ be a finite symmetric generating subset of $G$. Then $\lambda_S = 1$ if and only if $G$ has subexponential growth.* □

As an immediate consequence of Proposition 6.5.4 and Proposition 6.4.3(ii) we deduce the following.

**Corollary 6.5.6.** *Let $G$ be a finitely generated group and let $S$ and $S'$ be two finite symmetric generating subsets of $G$. Then $\lambda_S = 1$ (resp. $\lambda_S > 1$) if and only if $\lambda_{S'} = 1$ (resp. $\lambda_{S'} > 1$).* □

## 6.6 Growth of Subgroups and Quotients

**Proposition 6.6.1.** *Let $G$ be a group and let $N$ be a normal subgroup of $G$. Suppose that $N$ and the quotient group $G/N$ are both finitely generated. Then $G$ is finitely generated.*

*Proof.* Denote by $\pi\colon G \to G/N$ the quotient homomorphism. Let $U \subset N$ and $T \subset G/N$ be two finite symmetric generating subsets of $N$ and $G/N$ respectively. Let $S \subset G$ be a finite symmetric set such that $\pi(S) \supset T$ and $S \supset U$. Let us prove that $S$ generates $G$. Let $g \in G$. Then there exist $h \geq 0$ and $t_1, t_2, \ldots, t_h \in T$ such that $\pi(g) = t_1 t_2 \cdots t_h$. Let $s_1, s_2, \ldots, s_h \in S$ be such that $\pi(s_i) = t_i$ for $i = 1, 2, \ldots, h$. Setting $g' = s_1 s_2 \cdots s_h$ we then have $\pi(g') = \pi(g)$. It follows that $n = (g')^{-1} g \in \ker(\pi) = N$. Therefore, there exist $k \geq 0$ and $s_{h+1}, s_{h+2}, \ldots, s_{h+k} \in S$ such that $n = s_{h+1} s_{h+2} \cdots s_{h+k}$. It follows that $g = g'n = s_1 s_2 \cdots s_h s_{h+1}, s_{h+2} \cdots s_{h+k}$. This shows that $S$ generates $G$. □

**Proposition 6.6.2.** *Let $G$ be a group and let $H$ be a subgroup of finite index in $G$. Then, $G$ is finitely generated if and only if $H$ is finitely generated.*

*Proof.* Let $R \subset G$ be a complete set of representatives of the right cosets of $H$ in $G$ such that $1_G \in R$. Note that $|R| = [G : H] < \infty$. Suppose first that $H$ is finitely generated. Let $S \subset H$ be a finite symmetric generating subset of $H$. Given $g \in G$, there exist $r \in R$ and $h \in H$ such that $g = hr$. Let $s_1, s_2, \ldots, s_n \in S$ be such that $h = s_1 s_2 \cdots s_n$. Then $g = s_1 s_2 \cdots s_n r$. This shows that the set $S \cup R$ is a finite generating subset of $G$.

Suppose now that $G$ is finitely generated and let $S$ be a finite symmetric generating subset of $G$. Let us show that the finite set

$$S' = RSR^{-1} \cap H \tag{6.14}$$

generates $H$. Let $h \in H$ and write $h = s_1 s_2 \cdots s_n$ where $s_i \in S$. Then, there exists $h_1 \in H$ and $r_1 \in R$ such that $s_1 = h_1 r_1$. Note that $h_1 = 1_G s_1 r_1^{-1} \in S'$. By induction, for all $i = 2, 3, \ldots, n-1$ there exist $h_i \in H$ and $r_i \in R$ such that $r_{i-1} s_i = h_i r_i$. Moreover, $h_i = r_{i-1} s_i r_i^{-1} \in S'$. Thus, setting $h_n = r_{n-1} s_n$ we have

$$\begin{aligned} h &= s_1 s_2 \cdots s_n \\ &= (1_G s_1 r_1^{-1})(r_1 s_2 r_2^{-1}) \cdots (r_{n-2} s_{n-1} r_{n-1}^{-1})(r_{n-1} s_n) \\ &= h_1 h_2 \cdots h_{n-1} h_n. \end{aligned}$$

Note that $h_n = h_{n-1}^{-1} \cdots h_2^{-1} h_1^{-1} h \in H$ and, on the other hand, $h_n = r_{n-1} s_n = r_{n-1} s_n 1_G \in RSR^{-1}$. Thus also $h_n \in S'$. This shows that $S'$ generates $H$. □

**Proposition 6.6.3.** *Let $G$ be a finitely generated group and let $H$ be a finitely generated subgroup of $G$. Then $\gamma(H) \preceq \gamma(G)$.*

*Proof.* Let $S_G$ (resp. $S_H$) be a finite symmetric generating subset of $G$ (resp. $H$). Then the set $S = S_H \cup S_G$ is a finite symmetric generating subset of $G$. As $S_H \subset S$ we have $B^H_{S_H}(n) \subset B^G_S(n)$ and therefore $\gamma^H_{S_H}(n) \leq \gamma^G_S(n)$ for all $n \in \mathbb{N}$. Thus, $\gamma(H) \preceq \gamma(G)$. □

**Corollary 6.6.4.** *Every finitely generated group which contains a finitely generated subgroup of exponential growth has exponential growth.* □

From Example 6.4.11(d) and Corollary 6.6.4 we deduce the following.

**Corollary 6.6.5.** *Every finitely generated group which contains a subgroup isomorphic to the free group $F_2$ has exponential growth.* □

In the following proposition we show that finitely generated groups and their finite index subgroups (which are finitely generated as well, by Proposition 6.6.2) have the same growth type.

**Proposition 6.6.6.** *Let $G$ be a finitely generated group and let $H$ be a finite index subgroup of $G$. Then $H$ is finitely generated and $\gamma(H) = \gamma(G)$.*

*Proof.* Since $[G : H] < \infty$, we deduce from Proposition 6.6.2 that $H$ is finitely generated. It follows from Proposition 6.6.3 that $\gamma(H) \preceq \gamma(G)$.

Let now $S$ be a finite symmetric generating subset of $G$ and consider the finite symmetric set $S' = RSR^{-1} \cap H$. It follows from the proof of Proposition 6.6.2 that $S'$ generates $H$. Let $g \in B^G_S(n)$ and write $g = s_1 s_2 \cdots s_n$, where $s_1, s_2, \ldots, s_n \in S$. As in the proof of Proposition 6.6.2 we may find $r_0 = 1_G, r_1, \ldots, r_n \in R$ such that

$$\begin{aligned} g &= s_1 s_2 \cdots s_n \\ &= (1_G s_1 r_1^{-1})(r_1 s_2 r_2^{-1}) \cdots (r_{n-2} s_{n-1} r_{n-1}^{-1})(r_{n-1} s_n r_n^{-1}) r_n \\ &= h_1 h_2 \cdots h_{n-1} h_n r_n \end{aligned}$$

where $h_i = r_{i-1} s_i r_i^{-1} \in S'$. It follows that $B^G_S(n) \subset B^H_{S'}(n)R$ so that, taking cardinalities,

$$\gamma^G_S(n) = |B^G_S(n)| \leq |B^H_{S'}(n)||R| = [G : H]\gamma^H_{S'}(n) \leq [G : H]\gamma^H_{S'}([G : H]n).$$

This shows that $\gamma(G) \preceq \gamma(H)$. It follows that $\gamma(G) = \gamma(H)$. □

Two groups $G_1$ and $G_2$ are called *commensurable* if there exist finite index subgroups $H_1 \subset G_1$ and $H_2 \subset G_2$ such that $H_1$ and $H_2$ are isomorphic. From the previous proposition we immediately deduce:

**Corollary 6.6.7.** *If $G_1$ and $G_2$ are commensurable groups and $G_2$ is finitely generated, then $G_1$ is finitely generated and one has $\gamma(G_1) = \gamma(G_2)$.* □

**Proposition 6.6.8.** *Let $G$ be a finitely generated group and let $N$ be a normal subgroup of $G$. Then the quotient group $G/N$ is finitely generated and one has $\gamma(G/N) \preceq \gamma(G)$. If in addition $N$ is finite, then $\gamma(G/N) = \gamma(G)$.*

*Proof.* Let $S$ be a finite symmetric generating subset of $G$ and let $\pi\colon G \to G/N$ denote the canonical quotient homomorphism. Then $S' = \pi(S)$ is a finite symmetric generating subset of $G/N$. Thus, for all $n \in \mathbb{N}$ one has

$$B_{S'}^{G/N}(n) = \pi(B_S^G(n)) \tag{6.15}$$

and therefore

$$\gamma_{S'}^{G/N}(n) = |B_{S'}^{G/N}(n)| \leq |B_S^G(n)| = \gamma_S^G(n).$$

This shows that $\gamma(G/N) \preceq \gamma(G)$.

Suppose now that $N$ is finite. From (6.15) we deduce that $B_S^G(n) \subset \pi^{-1}(B_{S'}^{G/N}(n))$ and therefore

$$\gamma_S^G(n) = |B_S^G(n)| \leq |N||B_{S'}^{G/N}(n)| = |N|\gamma_{S'}^{G/N}(n) \leq |N|\gamma_{S'}^{G/N}(|N|n).$$

This shows that $\gamma(G) \preceq \gamma(G/N)$. It follows that $\gamma(G) = \gamma(G/N)$. □

**Lemma 6.6.9.** *Let $\gamma_1, \gamma_2, \gamma_1', \gamma_2'\colon \mathbb{N} \to [0, +\infty)$ be growth functions. Suppose that $\gamma_1 \preceq \gamma_1'$, $\gamma_2 \preceq \gamma_2'$. Then the products $\gamma_1\gamma_2, \gamma_1'\gamma_2'\colon \mathbb{N} \to [0, \infty)$ are also growth functions and one has $\gamma_1\gamma_2 \preceq \gamma_1'\gamma_2'$.*

*Proof.* Since the product of non-decreasing functions is also non-decreasing, it is clear that $\gamma_1\gamma_2$ and $\gamma_1'\gamma_2'$ are also growth functions. Let $c_1$ and $c_2$ be positive integers such that $\gamma_1(n) \leq c_1\gamma_1'(c_1 n)$ and $\gamma_2(n) \leq c_2\gamma_2'(c_2 n)$ for all $n \geq 1$. Taking $c = c_1c_2$ we have

$$\begin{aligned}(\gamma_1\gamma_2)(n) &= \gamma_1(n)\gamma_2(n)\\ &\leq c_1c_2\gamma_1'(c_1 n)\gamma_2'(c_2 n)\\ &\leq c_1c_2\gamma_1'(c_1c_2 n)\gamma_2'(c_1c_2 n)\\ &= c(\gamma_1'\gamma_2')(cn)\end{aligned}$$

for all $n \geq 1$. This shows $\gamma_1\gamma_2 \preceq \gamma_1'\gamma_2'$. □

Given two growth functions $\gamma_1$ and $\gamma_2$ we set $[\gamma_1] \cdot [\gamma_2] = [\gamma_1\gamma_2]$. This definition makes sense since if $\gamma_1 \sim \gamma_1'$ and $\gamma_2 \sim \gamma_2'$ then $\gamma_1\gamma_2 \sim \gamma_1'\gamma_2'$ as it immediately follows from Lemma 6.6.9.

**Proposition 6.6.10.** *Let $G_1$ and $G_2$ be two finitely generated groups. Then the direct product $G_1 \times G_2$ is also finitely generated and $\gamma(G_1 \times G_2) = \gamma(G_1)\gamma(G_2)$.*

*Proof.* Let $S_1$ and $S_2$ be finite symmetric generating subsets of $G_1$ and $G_2$. Then the set

$$S = (S_1 \times \{1_{G_2}\}) \cup (\{1_{G_1}\} \times S_2)$$

is a finite symmetric generating subset of $G_1 \times G_2$. Let $(g_1, g_2) \in B_S^{G_1 \times G_2}(n)$. Then there exist $s_{1,1}, s_{2,1}, \ldots, s_{k,1} \in S_1$ and $s_{1,2}, s_{2,2}, \ldots, s_{h,2} \in S_2$, where $h + k \leq n$, such that

$$\begin{aligned}(g_1, g_2) &= (s_{1,1}, 1_{G_2})(s_{2,1}, 1_{G_2}) \cdots (s_{k,1}, 1_{G_2}) \cdot (1_{G_1}, s_{1,2})(1_{G_1}, s_{2,2}) \cdots (1_{G_1}, s_{h,2}) \\ &= (s_{1,1} s_{2,1} \cdots s_{k,1}, s_{1,2} s_{2,2} \cdots s_{h,s}).\end{aligned}$$

Thus, $B_S^{G_1 \times G_2}(n) \subset B_{S_1}^{G_1}(n) \times B_{S_2}^{G_2}(n)$ and $\gamma_S^{G_1 \times G_2}(n) \leq \gamma_{S_1}^{G_1}(n) \gamma_{S_2}^{G_2}(n)$. This shows that $\gamma(G_1 \times G_2) \preceq \gamma(G_1)\gamma(G_2)$.

On the other hand, if $g_1 \in B_{S_1}^{G_1}(n)$ and $g_2 \in B_{S_2}^{G_2}(n)$, then $(g_1, g_2) \in B_S^{G_1 \times G_2}(2n)$ and one has $\gamma_{S_1}^{G_1}(n)\gamma_{S_2}^{G_2}(n) \leq \gamma_S^{G_1 \times G_2}(2n) \leq 2\gamma_S^{G_1 \times G_2}(2n)$. This shows that $\gamma(G_1)\gamma(G_2) \preceq \gamma(G_1 \times G_2)$. It follows that $\gamma(G_1 \times G_2) = \gamma(G_1)\gamma(G_2)$. □

From Example 6.4.11(a) and the previous proposition one immediately deduces the following.

**Corollary 6.6.11.** *Let $d$ be a positive integer. Then $\gamma(\mathbb{Z}^d) \sim n^d$.* □

**Corollary 6.6.12.** *Every finitely generated abelian group has polynomial growth.*

*Proof.* Let $G$ be a finitely generated abelian group. Then there exist an integer $d \geq 0$ and a finite group $F$ such that $G$ is isomorphic to the cartesian product $\mathbb{Z}^d \times F$. It follows from Proposition 6.6.10, Corollary 6.6.11 and Corollary 6.4.7 that $\gamma(G) = \gamma(\mathbb{Z}^d)\gamma(F) \sim n^d$. This shows that $G$ has polynomial growth. □

## 6.7 A Finitely Generated Metabelian Group with Exponential Growth

We have seen in Corollary 6.6.12 that every finitely generated abelian group has polynomial growth. The purpose of the present section is to give an example of a finitely generated metabelian group with exponential growth. As every metabelian group is solvable and hence amenable (Theorem 4.6.3), this will show in particular that there exist finitely generated amenable groups, and even finitely generated solvable groups, whose growth is exponential.

**Proposition 6.7.1.** *Let $G$ denote the subgroup of $\mathrm{GL}_2(\mathbb{Q})$ generated by the two matrices*

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

*Then $G$ is a finitely generated metabelian group with exponential growth.*

*Proof.* The group $G$ is finitely generated by definition. We have

$$G = \left\{ \begin{pmatrix} 2^k & r \\ 0 & 1 \end{pmatrix} : k \in \mathbb{Z}, r \in \mathbb{Z}[1/2] \right\}, \tag{6.16}$$

where $\mathbb{Z}[1/2]$ is the subring of $\mathbb{Q}$ consisting of all dyadic rationals. Indeed, if we denote by $H$ the right-hand side of (6.16), we first observe that $G \subset H$ since $H$ is clearly a subgroup of $\mathrm{GL}_2(\mathbb{Q})$ containing $A$ and $B$. On the other hand, we also have the inclusion $H \subset G$ since any $r \in \mathbb{Z}[1/2]$ can be written in the form $r = 2^m n$ for some $m, n \in \mathbb{Z}$, so that

$$\begin{pmatrix} 2^k & r \\ 0 & 1 \end{pmatrix} = A^m B^n A^{k-m} \in G.$$

From (6.16), we deduce that the determinant map yields a surjective homomorphism from $G$ onto an infinite cyclic group whose kernel is isomorphic to the additive group $\mathbb{Z}[1/2]$ and is therefore abelian. Since any element in $[G, G]$ has determinant 1, this shows that $[G, G]$ is abelian, that is, that $G$ is metabelian.

It remains to show that $G$ has exponential growth. To see this, let us estimate from below the cardinality of the ball $B_S(3n-2)$, where $S$ is the finite symmetric generating subset of $G$ defined by $S = \{A, B, A^{-1}, B^{-1}\}$ and $n \geq 1$ is a fixed integer. Consider the subset $E \subset G$ consisting of all matrices of the form

$$M(q) = \begin{pmatrix} 1 & q \\ 0 & 1 \end{pmatrix},$$

where $q$ is an integer such that $0 \leq q \leq 2^n - 1$. Developing $q$ in base 2, we get

$$q = \sum_{i=1}^{n} u_i 2^{i-1},$$

where $u_i \in \{0, 1\}$ for $1 \leq i \leq n$. This gives us

$$M(q) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{u_1} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}^{u_2} \begin{pmatrix} 1 & 2^2 \\ 0 & 1 \end{pmatrix}^{u_3} \cdots \begin{pmatrix} 1 & 2^{n-1} \\ 0 & 1 \end{pmatrix}^{u_n}.$$

As we have

$$\begin{pmatrix} 1 & 2^{i-1} \\ 0 & 1 \end{pmatrix} = A^{i-1} B A^{-(i-1)}$$

for all $1 \leq i \leq n$, it follows that we can write $M(q)$ in the form

$$\begin{aligned} M(q) &= B^{u_1} A B^{u_2} A^{-1} A^2 B^{u_3} A^{-2} \cdots A^{n-2} B^{u_{n-1}} A^{-(n-2)} A^{n-1} B^{u_n} A^{-(n-1)} \\ &= B^{u_1} A B^{u_2} A B^{u_3} A \cdots A B^{u_{n-1}} A B^{u_n} A^{-(n-1)}. \end{aligned}$$

This shows that

$$\ell_S(M(q)) \leq (2n-1) + (n-1) = 3n-2.$$

We deduce that $|B_S(3n-2)| \geq |E| = 2^n$. It follows that the growth rate of $G$ with respect to $S$ satisfies

$$\lambda_S = \lim_{n\to\infty} \sqrt[3n-2]{|B_S(3n-2)|} \geq \lim_{n\to\infty} \sqrt[3n-2]{2^n} = \sqrt[3]{2} > 1.$$

Thus $G$ has exponential growth. □

## 6.8 Growth of Finitely Generated Nilpotent Groups

In this section we prove the following

**Theorem 6.8.1.** *Every finitely generated nilpotent group has polynomial growth.*

In order to prove this result we need some preliminaries.

**Lemma 6.8.2.** *Let $G$ be any group. Let $H$ and $K$ be two normal subgroups of $G$. Suppose that $S \subseteq H$ and $T \subseteq K$ generate $H$ and $K$ respectively. Then $[H,K]$ is equal to the normal closure in $G$ of the set $\{[s,t] : s \in S, t \in T\}$.*

*Proof.* Denote by $N$ the normal closure in $G$ of the set $\{[s,t] : s \in S, t \in T\}$ and let us show that $N = [H,K]$. Since $\{[s,t] : s \in S, t \in T\} \subset [H,K]$ and $[H,K]$ is a normal subgroup, we have

$$N \subset [H,K]. \tag{6.17}$$

Consider the quotient homomorphism $\pi\colon G \to G/N$. For all $s \in S$ and $t \in T$ we have $[\pi(s),\pi(t)] = \pi([s,t]) = 1_{G/N}$, that is, $\pi(s)$ and $\pi(t)$ commute. It follows that all elements of $\pi(H)$ commute with all elements of $\pi(K)$, since $S$ generates $H$ and $T$ generates $K$. In other words, $\pi([h,k]) = [\pi(h),\pi(k)] = 1_{G/N}$, for all $h \in H$ and $k \in K$. This gives $[h,k] \in N$ for all $h \in H$ and $k \in K$ and therefore $[H,K] \subset N$. From (6.17) we deduce that $[H,K] = N$. □

Let $G$ be a group. Recall that the lower central series of $G$ is the sequence $(C^i(G))_{i\geq 0}$ of normal subgroups of $G$ defined by $C^0(G) = G$ and $C^{i+1}(G) = [C^i(G), G]$ for all $i \geq 0$. Given elements $g_1, g_2, \ldots, g_i \in G$, $i \geq 3$, we inductively set

$$[g_1, g_2, \ldots, g_i] = [[g_1, g_2, \ldots, g_{i-1}], g_i] \in C^{i-1}(G).$$

If $S \subset G$ and $i \geq 2$ we then denote by $S_G^{(i)}$ the set consisting of all elements of the form

$$[s_1, s_2, \dots, s_i], \quad s_1, s_2, \dots, s_i \in S. \tag{6.18}$$

The elements (6.18) are called *simple $S$-commutators of weight $i$*. Note that $S_G^{(i)} \subset C^{i-1}(G)$.

**Lemma 6.8.3.** *Let $G$ be any group. Let $S_G \subset G$ be a generating subset of $G$. Then for all $i \geq 1$, the subgroup $C^i(G)$ is the normal closure in $G$ of the set $S_G^{(i+1)}$.*

*Proof.* Let us prove the statement by induction on $i$. We have that $C^1(G) = [G, G]$ is the normal closure in $G$ of $S_G^{(2)} = \{[s_1, s_2] : s_1, s_2 \in S_G\}$, as it follows from Lemma 6.8.2 by taking $H = K = G$ and $S = T = S_G$. Thus, the statement holds for $i = 1$. Suppose by induction that $C^{i-1}(G)$ is the normal closure in $G$ of $S_G^{(i)}$, denote by $N$ the normal closure in $G$ of $S_G^{(i+1)}$, and let us show that $C^i(G) = N$. Since $S_G^{(i+1)} \subset C^i(G)$ and $C^i(G)$ is a normal subgroup, we deduce that

$$N \subset C^i(G). \tag{6.19}$$

Denote by $\pi\colon G \to G/N$ the quotient map. Let $w \in S_G^{(i)}$ and $s \in S_G$. Then, by definition, $[w, s] \in S_G^{(i+1)}$ and therefore $[w, s] \in N$. It follows that $[\pi(w), \pi(s)] = \pi([w, s]) = 1_{G/N}$, that is, $\pi(w)$ and $\pi(s)$ commute. As $S_G$ generates $G$, we have that $\pi(S_G)$ generates $G/N$ and therefore $\pi(w) \in Z(G/N)$. It follows that $\pi(hwh^{-1}) = \pi(h)\pi(w)\pi(h)^{-1} = \pi(w)$ for all $h \in G$, so that $\pi([hwh^{-1}, s]) = [\pi(hwh^{-1}), \pi(s)] = [\pi(w), \pi(s)] = 1_{G/N}$. It follows that

$$[hwh^{-1}, s] \in N. \tag{6.20}$$

By the inductive hypothesis, every element in $C^{i-1}(G)$ can be expressed as a product $(h_1 w_1 h_1^{-1})(h_2 w_2 h_2^{-1}) \cdots (h_m w_m h_m^{-1})$ with $w_k \in S_G^{(i)}$ and $h_k \in G$, $k = 1, 2, \dots, m$, for some $m \in \mathbb{N}$. By taking $H = C^{i-1}(G)$, $K = G$, $S = \{w^h : w \in S_G^{(i)}, h \in G\}$ and $T = S_G$ in Lemma 6.8.2, we deduce that $C^i(G) = [C^{i-1}(G), G]$ is the normal closure in $G$ of the set $\{[hwh^{-1}, s] : w \in S_G^{(i)}, h \in G, s \in S_G\}$. Thus from (6.20) it follows that $C^i(G) \subset N$. By using (6.19) this shows that $C^i(G) = N$. □

**Lemma 6.8.4.** *Let $G$ be a finitely generated nilpotent group of nilpotency degree $d \geq 1$. Then the subgroups $C^i(G)$, $i = 1, 2, \dots, d-1$ are finitely generated.*

*Proof.* Let $S_G \subset G$ be a finite generating subset of $G$. We first show that the quotient groups $C^i(G)/C^{i+1}(G)$ are finitely generated, $i = 0, 1, \dots, d-1$. From Lemma 6.8.3 we have that $C^i(G)$ is the normal closure in $G$ of the finite set $S_G^{(i+1)}$. Therefore, if $\pi\colon G \to G/C^{i+1}(G)$ denotes the quotient homomorphism, we have that $C^i(G)/C^{i+1}(G) = \pi(C^i(G))$ is the normal closure in $G/C^{i+1}(G)$ of the set $\pi(S_G^{(i+1)})$. But $\pi(S_G^{(i+1)}) = (\pi(S_G))^{(i+1)} \subset Z(C^i(G)/C^{i+1}(G))$ and therefore $\pi(S_G^{(i+1)})$ generates $C^i(G)/C^{i+1}(G)$.

Let us now prove the statement by reverse induction starting from $i = d-1$. It follows from the first part of the proof that the subgroup $C^{d-1}(G) \cong C^{d-1}(G)/\{1_G\} = C^{d-1}(G)/C^d(G)$ is finitely generated. Suppose by induction that the subgroup $C^{i+1}(G)$ is finitely generated for some $i \leq d-2$. From the first part of the proof we have that $C^i(G)/C^{i+1}(G)$ is also finitely generated. Therefore from Proposition 6.6.1 we deduce that $C^i(G)$ is finitely generated. □

We are now in position to prove the main result of this section.

*Proof of Theorem 6.8.1.* Let $G$ be a finitely generated nilpotent group of nilpotency degree $d$. Let us prove the statement by induction on $d$. If $d = 0$ then $G = \{1_G\}$ and therefore $G$ has polynomial growth. Suppose now that $d \geq 1$ and that all finitely generated nilpotent groups of nilpotency degree $\leq d-1$ have polynomial growth.

We first observe that the subgroup $H = C^1(G)$ is nilpotent of nilpotency degree $\leq d-1$. Indeed, as one immediately checks by induction, we have $C^i(H) \subset C^{i+1}(G)$ for all $i = 0, 1, \ldots, d-1$, so that $C^{d-1}(H) \subset C^d(G) = \{1_G\}$. Moreover, by Lemma 6.8.4, we have that $H$ is finitely generated. Thus, by the inductive hypothesis we have that $H$ has polynomial growth. Let $T \subset H$ be a finite symmetric generating subset of $H$. Then there exist integers $c_1 > 0$ and $q \geq 0$ such that

$$\gamma_T^H(n) \leq c_1(c_1 n)^q \tag{6.21}$$

for all $n \geq 1$.

Let $S = \{s_1, s_2, \ldots, s_k\} \subset G$ be a finite symmetric generating subset of $G$. Let $g \in G$ and suppose that $m = \ell_S^G(g) \leq n$. Then there exist $1 \leq i_1, i_2, \ldots, i_m \leq k$ such that

$$g = s_{i_1} s_{i_2} \cdots s_{i_m}. \tag{6.22}$$

Since $g_2 g_1 = g_1 g_2 [g_2^{-1}, g_1^{-1}]$ and $[g_2^{-1}, g_1^{-1}] \in H$ for all $g_1, g_2 \in G$, we can permute the generators in (6.22) and express $g$ in the form

$$g = s_{j_1} s_{j_2} \cdots s_{j_m} h \tag{6.23}$$

where $1 \leq j_1 \leq j_2 \leq \cdots \leq j_m \leq k$ and $h \in H$. Let us estimate $\ell_T^H(h)$. We set

$$L = \max\{\ell_T^H(w) : w \in S^{(i)}, 2 \leq i \leq d\}. \tag{6.24}$$

As we observed above, exchanging a pair of consecutive generators produces a simple $S$-commutator of weight two on their right. It is clear that one needs at most $n$ such exchanges to bring $s_{j_1}$, where $j_1 = i_{p_1} = \min\{i_p : p = 1, 2, \ldots, m\}$, to the leftmost place. Analogously, one needs at most $n$ such exchanges to bring $s_{j_2}$, where $j_2 = \min\{i_p : p = 1, 2, \ldots, m; p \neq p_1\}$, to the leftmost but one place (in fact on the right of $s_{j_1}$). And so on. Altogether,

there are at most $mn \leq n^2$ such exchanges. This produces at most $n^2$ simple $S$-commutators of weight two. But at each step, when moving a generator $s_{j_t}$ to the left, we also have to exchange it with all the simple $S$-commutators on its left that were produced before (namely by $s_{j_1}, s_{j_2}, \ldots, s_{j_{t-1}}$). As one easily checks by induction, that this produces, altogether, at most $n^3$ simple $S$-commutators of weight three, $n^4$ simple $S$-commutators of weight four, and so on. Continuing this way, since $G$ is nilpotent of nilpotency degree $d$, all simple $S$-commutators of weight $d+1$ are equal to $1_G$. Thus, the total number of simple $S$-commutators that are eventually produced in this process is at most $n^2 + n^3 + \cdots + n^d \leq dn^d$. It follows from (6.24) that

$$\ell_T^H(h) \leq Ldn^d. \tag{6.25}$$

On the other hand, we can bound the number of elements in $G$ which are of the form $s_{j_1} s_{j_2} \cdots s_{j_m}$, where $1 \leq j_1 \leq j_2 \leq \cdots \leq j_m \leq k$, by $c_2 n^k$, where $c_2 > 0$ is a constant independent of $n$. Indeed, each such group element can be written in the form $s_1^{n_1} s_2^{n_2} \cdots s_k^{n_k}$, where $0 \leq n_i \leq n$ for all $i = 1, 2, \ldots, k$.

We deduce from (6.21) and (6.25) that

$$\gamma_S^G(n) \leq c_2 n^k c_1 (Ldn^d)^q = Cn^\delta \leq C(Cn)^\delta$$

for all $n \geq 1$, where $\delta = k + qd$ and $C = c_1 c_2 (Ld)^q$. Note that $C > 0$ is a constant independent of $n$. It follows that $G$ has polynomial growth. □

From Theorem 6.8.1 and Proposition 6.6.6 we deduce the following:

**Corollary 6.8.5.** *Every finitely generated virtually nilpotent group has polynomial growth.* □

## 6.9 The Grigorchuk Group and Its Growth

In this section we present the Grigorchuk group and some of its main properties, namely being infinite, periodic, residually finite and of intermediate growth.

Let $\Sigma = \{0, 1\}$. We denote by $\Sigma^* = \cup_{n \in \mathbb{N}} \Sigma^n$ the set of all words on the alphabet $\Sigma$. Recall that $\Sigma^*$ is a monoid for the word concatenation whose identity element is the empty word $\epsilon$ (cf. Sect. D.1). Every word $w \in \Sigma^*$ may be uniquely written in the form $w = \sigma_1 \sigma_2 \cdots \sigma_n$, where $\sigma_1, \sigma_2, \ldots, \sigma_n \in \Sigma$ and $n = \ell(w) \in \mathbb{N}$ is the *length* of $w$.

Denote by $\mathrm{Sym}(\Sigma^*)$ the symmetric group on $\Sigma^*$ (cf. Appendix C). We introduce a partial order $\preceq$ in $\Sigma^*$ by setting $u \preceq v$, $u, v \in \Sigma^*$, if there exists $w \in \Sigma^*$ such that $uw = v$. We then set

$$\mathrm{Sym}(\Sigma^*, \preceq) = \{g \in \mathrm{Sym}(\Sigma^*) : g(u) \preceq g(v) \text{ for all } u, v \in \Sigma^* \text{ such that } u \preceq v\}. \tag{6.26}$$

Note that $\mathrm{Sym}(\Sigma^*, \preceq)$ is a subgroup of $\mathrm{Sym}(\Sigma^*)$. Moreover, for $w \in \Sigma^*$ the length $\ell(w)$ equals the maximum of $n \in \mathbb{N}$ such that there exists a sequence $(w_k)_{0\le k\le n}$ of distinct words in $\Sigma^*$ such that $\epsilon = w_0 \preceq w_1 \preceq \cdots \preceq w_n = w$. It follows that if $g \in \mathrm{Sym}(\Sigma^*, \preceq)$ then $\ell(g(w)) = \ell(w)$ for all $w \in \Sigma^*$.

Consider the elements $a, b, c, d \in \mathrm{Sym}(\Sigma^*)$ defined as follows. We first define $a$ by setting

$$a(\epsilon) = \epsilon \tag{6.27}$$

and

$$a(0w) = 1w, \quad a(1w) = 0w \tag{6.28}$$

for all $w \in \Sigma^*$. Then, for all $w \in \Sigma^*$, we define $b(w), c(w), d(w)$ by induction on $\ell(w)$. We start by setting

$$b(\epsilon) = c(\epsilon) = d(\epsilon) = \epsilon. \tag{6.29}$$

Then we set

$$\begin{array}{l} b(0w) = 0a(w),\ b(1w) = 1c(w), \\ c(0w) = 0a(w),\ c(1w) = 1d(w), \\ d(0w) = 0w, \quad d(1w) = 1b(w) \end{array} \tag{6.30}$$

for all $w \in \Sigma^*$. By an obvious induction we have

$$a, b, c, d \in \mathrm{Sym}(\Sigma^*, \preceq). \tag{6.31}$$

*Example 6.9.1.* Let $w = 10110 \in \Sigma^*$. Then we have

$$\begin{aligned} a(w) &= a(10110) = 00110, \\ b(w) &= b(10110) = 1c(0110) = 10a(110) = 10010, \\ c(w) &= c(10110) = 1d(0110) = 10110, \\ d(w) &= d(10110) = 1b(0110) = 10a(110) = 10010. \end{aligned}$$

**Definition 6.9.2.** The *Grigorchuk group* is the subgroup $G$ of $\mathrm{Sym}(\Sigma^*)$ generated by the elements $a, b, c, d$.

Observe that by (6.31) we have

$$G \subset \mathrm{Sym}(\Sigma^*, \preceq). \tag{6.32}$$

**Proposition 6.9.3.** *The following relations hold in $G$.*

$$a^2 = b^2 = c^2 = d^2 = 1_G \tag{6.33}$$

*and*

$$bc = cb = d,\ dc = cd = b,\ db = bd = c. \tag{6.34}$$

*Proof.* To prove (6.33), we have to show that

$$g^2(w) = w \tag{6.35}$$

for all $w \in \Sigma^*$ and $g \in \{a, b, c, d\}$. We have $a(\varepsilon) = \varepsilon$ by (6.27). Moreover from (6.28) we have

$$a^2(0w) = a(1w) = 0w \quad \text{and} \quad a^2(1w) = a(0w) = 1w, \tag{6.36}$$

for all $w \in \Sigma^*$. This shows that $a^2 = 1_G$. To prove (6.35) for $g = b, c$ and $d$, we use induction on $\ell(w)$. By virtue of (6.29) this holds when $\ell(w) = 0$. Suppose that (6.35) holds when $\ell(w) = n$. From (6.30) and the fact that $a^2 = 1_G$ we deduce, for all $w \in \Sigma^*$ with $\ell(w) = n$,

$$\begin{aligned} b^2(0w) &= b(0a(w)) = 0a^2(w) = 0w, & b^2(1w) &= b(1c(w)) = 1c^2(w) = 1w, \\ c^2(0w) &= c(0a(w)) = 0a^2(w) = 0w, & c^2(1w) &= c(1d(w)) = 1d^2(w) = 1w, \\ d^2(0w) &= d(0w) = 0w, & d^2(1w) &= d(1b(w)) = 1b^2(w) = 1w. \end{aligned}$$

This shows that (6.35) holds for all $w \in \Sigma^*$ with $\ell(w) = n + 1$. This proves (6.35) for $g = b, c, d$ and (6.33) follows.

To prove (6.34), we have to show that

$$ij(w) = k(w) \tag{6.37}$$

for all $w \in \Sigma^*$ and all distinct $i, j, k \in \{b, c, d\}$. Again, we can use induction on $\ell(w)$. If $\ell(w) = 0$ then (6.37) follows from (6.29). Suppose by induction that (6.37) holds when $\ell(w) = n$. From (6.28),(6.30) and (6.33) we deduce, for all $w \in \Sigma^*$ with $\ell(w) = n$,

$$\begin{aligned} bc(0w) &= b(0a(w)) = 0a^2(w) = 0w = d(0w), \\ bc(1w) &= b(1d(w)) = 1cd(w) = 1b(w) = d(1w), \\ cd(0w) &= c(0w) = 0a(w) = b(0w), \\ cd(1w) &= c(1b(w)) = 1db(w) = 1c(w) = b(1w), \\ db(0w) &= d(0a(w)) = 0a(w) = c(0w), \\ db(1w) &= d(1c(w)) = 1bc(w) = 1d(w) = c(1w). \end{aligned}$$

It follows that (6.37) holds for all $w \in \Sigma^*$ with $\ell(w) = n + 1$ whenever $(i, j, k) = (b, c, d), (c, d, b), (d, b, c)$. Thus, by induction $bc = d$, $cd = b$ and $db = c$. Finally, using (6.33) we deduce $cb = cd^2b = (cd)(db) = bc$, $dc = db^2c = (db)(bc) = cd$ and $bd = bc^2d = (bc)(cd) = db$. This completes the proof of (6.34). □

It follows from (6.33) that the set $S = \{a, b, c, d\}$ is a symmetric generating subset of $G$. We denote by $\ell_S \colon G \to \mathbb{N}$ the corresponding word-length function. Every group element $g \in G$ can be expressed in the form

$$g = s_1 s_2 \cdots s_n \tag{6.38}$$

with $s_1, s_2, \ldots, s_n \in S$. We say that the expression (6.38) is a *reduced form* of $g$ provided that for all $i, j = 1, 2, \ldots, n-1$ one has that if $s_i = a$ then $s_{i+1} \in \{b, c, d\}$, and if $s_j \in \{b, c, d\}$ then $s_{j+1} = a$. It immediately follows from (6.33) and (6.34) that every group element $g \in G$ can be expressed (not necessarily in a unique way) as in (6.38) in reduced form.

For all $n \in \mathbb{N}$ we set

$$H_n = \{g \in G : g(w) = w \text{ for all } w \in \Sigma^n\}. \tag{6.39}$$

**Proposition 6.9.4.** *For all $n \in \mathbb{N}$, the set $H_n$ is a normal subgroup of $G$ and*

$$[G : H_n] < \infty. \tag{6.40}$$

*Moreover,*

$$G = H_0 \supset H_1 \supset H_2 \supset \cdots \supset H_n \supset H_{n+1} \supset \cdots. \tag{6.41}$$

*Proof.* Since $g(\Sigma^n) \subset \Sigma^n$ for all $g \in G$ (cf. (6.32)) we may consider the (restriction) map $\theta_n \colon G \to \mathrm{Sym}(\Sigma^n)$ defined by $\theta_n(g)(w) = g(w)$ for all $g \in G$ and $w \in \Sigma^n$. Clearly, $\theta_n$ is a homomorphism and $\ker(\theta_n) = H_n$. This shows that $H_n$ is a normal subgroup and that $[G : H_n] = |G/H_n| = |\theta_n(G)| \leq |\mathrm{Sym}(\Sigma^n)| < \infty$.

Finally, let $n \in \mathbb{N}$, $u \in \Sigma^n$ and $g \in H_{n+1}$. Let $\sigma \in \Sigma$ and set $w = u\sigma \in \Sigma^{n+1}$ so that $u \preceq w$. From (6.32) we deduce that $g(u) \preceq g(w) = u\sigma$ and therefore $g(u) = u$. This shows that $g \in H_n$. Thus $H_n \supset H_{n+1}$ and (6.41) follows. □

**Corollary 6.9.5.** *The Grigorchuk group $G$ is residually finite.*

*Proof.* Let $h \in \cap_{n\in\mathbb{N}} H_n$. We have $h(w) = w$ for all $w \in \Sigma^*$ and therefore $h = 1_G$. This shows that $\cap_{n\in\mathbb{N}} H_n = \{1_G\}$. As $[G : H_n] < \infty$ for all $n \in \mathbb{N}$ we deduce from Proposition 2.1.11 that $G$ is residually finite. □

**Proposition 6.9.6.** *We have:*

(i) *the subgroup $H_1$ consists of all group elements $g \in G$ which can be expressed in the form (6.38) (not necessarily reduced) with an even number of occurrences of the generator $a$;*
(ii) *$[G : H_1] = 2$;*
(iii) *the group $H_1$ is generated by the elements $b, c, d, aba, aca, ada$;*
(iv) *the group $H_1$ equals the normal closure in $G$ of the elements $b, c$ and $d$.*

*Proof.* It follows from (6.28) and (6.30) that a generator $s \in S$ satisfies $s(\sigma) = \sigma$ for all $\sigma \in \Sigma$ if and only if $s \in \{b, c, d\}$. Thus, $g = s_1 s_2 \cdots s_n$, $s_i \in S$, belongs to $H_1$ if and only if $|\{i : s_i = a\}|$ is an even number. This also shows that $[G : H_1] = 2$. Finally, let $g \in H_1$. Then it can be expressed in one of the following reduced forms:

$$g = t_0 a t_1 a t_2 a t_3 a t_4 \cdots a t_{2k-1} a t_{2k} = t_0(a t_1 a) t_2 (a t_3 a) t_4 \cdots (a t_{2k-1} a) t_{2k},$$
$$g = a t_1 a t_2 a t_3 a_t 4 \cdots a t_{2k-1} a t_{2k} = (a t_1 a) t_2 (a t_3 a) t_4 \cdots (a t_{2k-1} a) t_{2k},$$
$$g = t_0 a t_1 a t_2 a t_3 a t_4 \cdots a t_{2k-1} a = t_0(a t_1 a) t_2 (a t_3 a) t_4 \cdots (a t_{2k-1} a),$$
$$g = a t_1 a t_2 a t_3 a t_4 \cdots a t_{2k-1} a = (a t_1 a) t_2 (a t_3 a) t_4 \cdots (a t_{2k-1} a),$$

where $t_0, t_1, \ldots, t_{2k+1} \in \{b, c, d\}$ and $k \in \mathbb{N}$ (observe that the occurrences of the $a$'s is $2k$). This shows that the set $\{b, c, d, aba, aca, ada\}$ generates $H_1$. Now, the normal closure in $G$ of the generators $b, c, d$ is the subgroup $H \subset G$ generated by all the conjugates $gtg^{-1}$ with $g \in G$ and $t \in \{b, c, d\}$. Thus taking $g = 1_G, a$ we deduce from (iii) that $H_1 \subset H$. On the other hand, since $b, c, d \in H_1$ we also have $H \subset H_1$. Thus $H = H_1$. □

Let $h \in H_1$. For all $w \in \Sigma^*$ there exist $w_0, w_1 \in \Sigma^*$, $\ell(w_0) = \ell(w_1) = \ell(w)$ such that $h(0w) = 0w_0$ and $h(1w) = 1w_1$. Denote by $h_0, h_1 \in \mathrm{Sym}(\Sigma^*, \preceq)$ the maps defined by $h_0(w) = w_0$ and $h_1(w) = w_1$. We thus have

$$h(0w) = 0h_0(w) \quad \text{and} \quad h(1w) = 1h_1(w)$$

for all $w \in \Sigma^*$. We denote by $\phi_0 \colon H_1 \to \mathrm{Sym}(\Sigma^*, \preceq)$ (resp. $\phi_1 \colon H \to \mathrm{Sym}(\Sigma^*, \preceq)$) the map defined by $\phi_0(h) = h_0$ (resp. $\phi_1(h) = h_1$) and by $\phi \colon H_1 \to \mathrm{Sym}(\Sigma^*, \preceq) \times \mathrm{Sym}(\Sigma^*, \preceq)$ the product map $\phi(h) = (h_0, h_1)$. From (6.30) we immediately deduce

$$\begin{array}{ll} \phi(b) = (a, c), & \phi(aba) = (c, a), \\ \phi(c) = (a, d), & \phi(aca) = (d, a), \\ \phi(d) = (1_G, b), & \phi(ada) = (b, 1_G). \end{array} \tag{6.42}$$

**Proposition 6.9.7.** *We have:*

(i) *the maps $\phi_0, \phi_1 \colon H_1 \to G$ are surjective homomorphisms;*
(ii) *the map $\phi \colon H_1 \to G \times G$ is an injective homomorphism.*

*Proof.* Let $\sigma \in \{0, 1\}$. For all $h, h' \in H$ and $w \in \Sigma^*$ we have

$$\sigma \phi_\sigma(hh')(w) = hh'(\sigma w) = h(\sigma \phi_\sigma(h')(w)) = \sigma \phi_\sigma(h) \phi_\sigma(h')(w).$$

This shows that $\phi_\sigma(hh') = \phi_\sigma(h)\phi_\sigma(h')$. It follows that $\phi_0$ and $\phi_1$ are homomorphisms. From (6.42) and Proposition 6.9.6(iii) we immediately deduce that $\phi_0(H_1) = \phi_1(H_1) = G$. This shows (i).

To show the injectivity of $\phi$, let $h \in H_1$ and suppose that $\phi(h) = 1_{G \times G} = (1_G, 1_G)$. We have $h(\epsilon) = \epsilon$ and $h(\sigma w) = \sigma h_\sigma(w) = \sigma w$ for all $\sigma \in \{0, 1\}$ and $w \in \Sigma^*$. It follows that $h = 1_G = 1_H$. This shows (ii). □

As a consequence of Proposition 6.9.7(ii), we can identify each element $h \in H_1$ with its image $\phi(h) \in G \times G$. Thus, we shall write $h = (h_0, h_1)$ if $h \in H_1$ and $\phi(h) = (h_0, h_1)$.

Recall that a group is called periodic if it contains no elements of infinite order.

**Theorem 6.9.8.** *The Grigorchuk group $G$ is an infinite finitely generated periodic group.*

*Proof.* Since $H_1$ is a proper subset of $G$ and $\phi_0 \colon H_1 \to G$ is surjective (cf. Proposition 6.9.7(i)), we deduce that $G$ is infinite.

Let now show that $G$ is a 2–group, that is, that the order of every element $g \in G$ is a power of 2. The proof is by induction on $\ell_S(g)$. By (6.33) the statement is true for $\ell_S(g) = 1$.

Let us first show that every element $g \in G \setminus \{1_G\}$ is conjugate either to an element in $S$ or to an element $g'$ which can be expressed in a reduced form

$$g' = at_1at_2 \cdots at_k \tag{6.43}$$

with $t_1, t_2, \ldots, t_k \in \{b, c, d\}$, $k \geq 1$, and such that $\ell_S(g') \leq \ell_S(g)$. Indeed, if $g$ is not conjugate to an element in $S$, then any element of minimal $S$-length in the conjugacy class of $g$ necessarily admits one of the following two reduced forms, besides (6.43): $t_1at_2 \cdots at_ka$ and $t_0at_1 \cdots at_k$ with $t_0, t_1, \ldots, t_k \in \{b, c, d\}$, $t_0 \neq t_k$, $k \geq 1$. Conjugating by $a$ and $t_0$ respectively and replacing $t_kt_0$ with $t \in \{b, c, d\}$ according to (6.34), we transform the above two expressions into one of the form (6.43). It is clear that such a process does not increase the word lengths so that $\ell_S(g') \leq \ell_S(g)$.

Suppose first that $g \in H_1$ and $\ell(g) > 1$. Since the order of any element equals the order of all its conjugates, we may suppose, up to conjugacy, that $g$ admits a reduced expression of the form (6.43) (with $k \geq 2$ even, since $g \in H_1$). Now, the image of each quadruple $at_{2i-1}at_{2i}$, $i = 1, 2, \ldots, k/2$, via $\phi_0$ or $\phi_1$ has word-length $\leq 2$. Thus, $\ell_S(g_0), \ell_S(g_1) \leq \ell_S(g)/2 < \ell_S(g)$. By induction, $g_0$ e $g_1$ are 2−elements, and thus $g = (g_0, g_1)$ is a 2−element as well.

Suppose now that $g \notin H_1$. As before, we may suppose, up to conjugacy, that $g$ admits a reduced expression of the form (6.43) (with $k$ odd, since $g \notin H_1$). We distinguish three cases.

**Case 1.** The generator $d$ appears in the expression of $g$, say $d = t_i$ for some $1 \leq i \leq k$. Then, up to conjugating by the element $t_{i-1}at_{i-2}a \cdots at_1a$, we can suppose that $d = t_1$.

Now $g^2 = (adat_2)(at_3at_4) \cdots (at_kad)(at_2at_3) \cdots (at_{k-1}at_k) \in H$. The image of each quadruple via $\phi_0$ (resp. $\phi_1$) has length 2, with the only exception for those of the form $at_jad = (at_ja)d$ (resp. $adat_j = (ada)t_j$) since $\phi_0(d) = 1_G$ (resp. $\phi_1(ada) = 1_G$). But at least one of these quadruples occurs in $g^2$, for instance for $j = k$ (resp. $j = 1$), so that $\ell_S(\phi_0(g^2)) \leq 2k - 1 < 2k = \ell_S(g)$ (resp. $\ell_S(\phi_1(g^2)) < \ell_S(g)$). By induction, $\phi_0(g^2)$ and $\phi_1(g^2)$ are 2-elements so that $g^2 = (\phi_0(g^2), \phi_1(g^2))$ and therefore $g$ are 2-elements as well.

**Case 2.** Suppose now that $d$ doesn't occur in the expression of $g$ but $c$ does. As before, up to conjugacy, we may suppose that $t_1 = c$. We then have $\phi_0(abat') = ca, \phi_0(acat') = da$ and $\phi_1(abat') = ac, \phi_1(acat') = ad$ for all $t' \in \{b, c\}$. It follows that $\ell_S(\phi_0(g^2)) = \ell_S(\phi_1(g^2)) = 2k = \ell(g)$. Observe that $\phi_0(g^2) = da \cdots a$ (resp. $\phi_1(g^2) = a \cdots ad$) so that, by conjugating by $a$ (resp.

$ad$) we fall into Case 1. This shows that $\phi_0(g^2)$ and $\phi_1(g^2)$ are 2-elements. Thus $g^2 = (\phi_0(g^2), \phi_1(g^2))$ and therefore $g$ are 2-elements.

**Case 3.** Finally, suppose that neither $d$ nor $c$ appear in (6.43) so that necessarily $g = (ab)^{2h+1}$ for some $h \geq 0$. We have $g^2 = (ab)^{4h+2} = ((aba)b)^{2h+1} \in H$, so that $\phi_0(g^2) = (ca)^{2h+1}$ and $\phi_1(g^2) = (ac)^{2h+1}$. Since $\ell_S(\phi_0(g^2)) = \ell_S(\phi_1(g^2)) = 4h + 2 = \ell_S(g)$, we are in Case 2 and we deduce that $\phi_0(g^2)$ and $\phi_1(g^2)$ are 2-elements. Thus $g^2 = (\phi_0(g^2), \phi_1(g^2))$ and therefore $g$ itself are 2-elements. □

**Theorem 6.9.9.** *The Grigorchuk group $G$ does not have polynomial growth.*

In order to prove this theorem, let us prove some preliminary results.

**Lemma 6.9.10.** *The subgroup of $G$ generated by $a$ and $d$ is dihedral of order 8.*

*Proof.* The dihedral group $D_8$ of order 8 has presentation $D_8 = \langle x, y : x^2, y^2, (xy)^4 \rangle$. Since $a^2 = d^2 = 1_G$ by (6.33), we are only left to verify that the order of the element $ad$ is 4. We have $(ad)^2 = (ada)d = (b, 1)(1, b) = (b, b) \neq 1_G$, while $(ad)^4 = ((ad)^2)^2 = (b, b)^2 = (b^2, b^2) = (1_G, 1_G) = 1_G$. □

Recall that two groups $G_1$ and $G_2$ are commensurable if there exist two subgroups of finite index $K_1 \subset G_1$ and $K_2 \subset G_2$ such that $K_1$ and $K_2$ are isomorphic.

**Lemma 6.9.11.** *The Grigorchuk group $G$ is commensurable with its own square $G \times G$.*

*Proof.* Let us start by showing that the index of $\phi(H_1)$ inside $G \times G$ is finite.

Since $b, c, d, aba, aca$ and $ada$ generate $H_1$ (cf. Proposition 6.9.4), we deduce that the elements in (6.42) generate $\phi(H_1)$.

Denote by $B \subset G$ the normal closure in $G$ of $b$. The quotient $G/B$ is generated by the images of the generators of $G$ and since $cd = b \in B$, it is generated by the images of $a$ and $d$. From Lemma 6.9.10 we deduce

$$[G : B] = |G/B| \leq 8. \tag{6.44}$$

Let $g \in G$ and consider the element $gbg^{-1}$. Since $\phi_0$ (resp. $\phi_1$) is surjective, there exists an element $h \in H$ (resp. $h' \in H$) such that $\phi_0(h) = g$ (resp. $\phi_1(h') = g$). It follows that $(gbg^{-1}, 1_G) = \phi(hadah^{-1}) \in \phi(H_1)$ (resp. $(1_G, gbg^{-1}) = \phi(h'd(h')^{-1}) \in \phi(H_1)$). As $g$ varies in $G$ the elements $(gbg^{-1}, 1)$ (resp. $(1, gbg^{-1})$) generate the subgroup $B_0 = B \times \{1_G\} \subset \phi(H_1)$ (resp. $B_1 = \{1_G\} \times B \subset \phi(H_1)$). Observe that $B_0 \simeq B_1 \simeq B$ and that $B_0$ and $B_1$ are normal subgroups of $G \times G$. Moreover, $B_0 \cap B_1 = \{1_{G\times G}\}$, so that $(G \times G)/(B_0 B_1) \simeq G/B \times G/B$. From (6.44) and the fact that $B_0 B_1 \subset \phi(H_1)$ we deduce that

$$[G \times G : \phi(H_1)] \leq [G \times G : B_0 B_1] = [G : B]^2 = 64. \tag{6.45}$$

This shows that $\phi(H_1)$ has finite index in $G \times G$.

On the other hand $[G : H_1] = 2$ (Proposition 6.9.6(ii)) and since $\phi$ is an injective homomorphism (Proposition 6.9.7(ii)), we have that $H_1$ and $\phi(H_1)$ are isomorphic. It follows that $G$ and $G \times G$ are commensurable. □

*Proof of Theorem 6.9.9.* Since $G$ is infinite (Theorem 6.9.8) we have $n \preceq \gamma(G)$ (cf. Proposition 6.4.8). Since $G$ and $G \times G$ are commensurable (Lemma 6.9.11) we deduce from Corollary 6.6.7 that $\gamma(G) = \gamma(G \times G)$. On the other hand, by Proposition 6.6.10 we have $\gamma(G \times G) = \gamma(G)^2$. Using Lemma 6.6.9 it follows that $n^2 \preceq \gamma(G)^2 = \gamma(G)$. By induction, we have that $n^{2h} \preceq \gamma(G)$ for all $h \in \mathbb{N}$. Thus $G$ cannot have polynomial growth. □

The remaining of this section is devoted to showing the following:

**Theorem 6.9.12.** *The Grigorchuk group $G$ has subexponential growth.*

To prove this theorem we need some preliminaries. We start with a useful criterion for detecting that certain finitely generated groups have subexponential growth.

**Lemma 6.9.13.** *Let $G$ be a finitely generated group. Let $S \subset G$ be a finite symmetric generating subset and suppose that there exist an integer $M \geq 2$, two constants $0 < k < 1$ and $K \geq 0$, and an injective homomorphism*

$$\begin{aligned} \psi \colon H &\to G^M \\ g &\mapsto (g_i)_{i=1}^M \end{aligned}$$

*where $H \subset G$ is a finite index subgroup of $G$, such that*

$$\sum_{i=1}^{M} \ell_S(g_i) \leq k\ell_S(g) + K \tag{6.46}$$

*for all $g \in H$. Then $G$ has subexponential growth.*

*Proof.* Let us show that $\lambda_S = \lim_{n\to\infty} \gamma_S^G(n)^{\frac{1}{n}}$ equals 1.

Fix $\varepsilon > 0$. Then there exists an integer $n_0 \geq 1$ such that $\gamma_S^G(n) < (\lambda_S + \varepsilon)^n$ for all $n \geq n_0$. Since $\lambda_S \geq 1$, it follows that

$$\gamma_S^G(n) \leq \gamma_S^G(n_0)(\lambda_S + \varepsilon)^n \tag{6.47}$$

for all $n \in \mathbb{N}$.

Let

$$\gamma_S^H(n) = |\{h \in H : \ell_S(h) \leq n\}|$$

and fix a system $T$ of left coset representatives of $H$ in $G$. Set $C = \max_{t\in T} \ell_S(t)$. Then, given $g \in G$, there exist unique $h \in H$ and $t \in T$ such that $g = th$. Therefore $\ell_S(h) \leq \ell_S(t) + \ell_S(g) \leq C + \ell(g)$, so that $B_S^G(n) \subset TB_S^H(n + C)$. We deduce that

$$\gamma_S^G(n) \leq [G : H]\gamma_S^H(n + C). \tag{6.48}$$

On the other hand, by (6.46) we have

$$\gamma_S^H(n) \leq \sum \gamma_S^G(n_1)\gamma_S^G(n_2)\cdots\gamma_S^G(n_M)$$

where the sum runs over all $M-$tuples $n_1, n_2, \ldots, n_M$ such that $\sum n_i \leq kn + K$.

From (6.47) we then deduce:

$$\begin{aligned}\gamma_S^H(n) &\leq \gamma_S^G(n_0)^M \sum (\lambda_S+\epsilon)^{n_1}(\lambda_S+\epsilon)^{n_2}\cdots(\lambda_S+\epsilon)^{n_M} \\ &= \gamma_S^G(n_0)^M \sum (\lambda_S+\epsilon)^{n_1+n_2+\cdots+n_M} \\ &\leq \left(\gamma_S^G(n_0)(kn+K)\right)^M (\lambda_S+\epsilon)^{kn+K}.\end{aligned}$$

Using (6.48) we then obtain

$$\gamma_S^G(n) \leq [G:H]\gamma_S^H(n+C) \leq [G:H]\left(\gamma_S^G(n_0)(kn+K')\right)^M (\lambda_S+\varepsilon)^{kn+K'} \tag{6.49}$$

where $K' = K + kC$. Taking the $n$th roots and passing to the limit for $n \to \infty$ in (6.49), the first term on the left tends to $\lambda_S$, while the last term on the right tends to $(\lambda_S+\varepsilon)^k$. Thus $\lambda_S \leq (\lambda_S+\varepsilon)^k$. Since $\varepsilon$ was arbitrary we deduce that $\lambda_S \leq \lambda_S^k$. Since by hypothesis $k < 1$, we have that $\lambda_S = 1$.

Finally, from Corollary 6.5.5 we deduce that $G$ has subexponential growth. □

**Lemma 6.9.14.** *Let $n \geq 1$. Then $\phi_\sigma(H_{n+1}) \subset H_n$ for $\sigma = 0, 1$.*

*Proof.* We proceed by induction on $n$. We have already noticed that $\phi_\sigma(H_1) \subset G = H_0$ (Proposition 6.9.7(i)). Suppose that $\phi_\sigma(H_n) \subset H_{n-1}$ and let us show that $\phi_\sigma(H_{n+1}) \subset H_n$. Let $g \in H_{n+1}$ and $w = \sigma u$, with $u \in \Sigma^n$ and $\sigma \in \{0,1\}$. Then

$$\sigma u = w = g(w) = \sigma\phi_\sigma(g)(u),$$

so that $\phi_\sigma(g)(u) = u$. This shows that $\phi_\sigma(g) \in H_n$. It follows that $\phi_\sigma(H_{n+1}) \subset H_n$. □

As a consequence of the previous lemma, for all $n \geq 1$ and $w = \sigma_1\sigma_2\cdots\sigma_n \in \Sigma^n$ the homomorphisms $\phi_w \colon H_n \to G$ defined by

$$\phi_w = \phi_{\sigma_1} \circ \phi_{\sigma_2} \circ \cdots \circ \phi_{\sigma_n}$$

are well defined. We then define the homomorphism $\psi_n \colon H_n \to \prod_{w\in\Sigma^n} G$ by setting $\psi_n(g) = (\phi_w(g))_{w\in\Sigma^n}$ for all $g \in H_n$. Note that by Proposition 6.9.7(ii) $\psi_n$ is injective. Thus, identifying $H_n$ with its image $\psi_n(H_n) \subset \prod_{w\in\Sigma^n} G$, we simply write $g = (g_w)_{w\in\Sigma^n}$ for all $g \in H_n$.

In the following, we consider the alphabet $\Lambda = \{\alpha, \beta, \gamma, \delta\}$ and the monoid $\Lambda^*$. The map $\Lambda \to S$ given by $\alpha \mapsto a$, $\beta \mapsto b$, $\gamma \mapsto c$ and $\delta \mapsto d$ uniquely extends to a surjective monoid homomorphism $\pi \colon \Lambda^* \to G$. We say that a

word $w = \lambda_1\lambda_2\cdots\lambda_n \in \Lambda^*$ is *reduced* if for all $i, j = 1, 2, \ldots, n-1$ one has $\lambda_i = \alpha$ implies $\lambda_{i+1} \in \{\beta, \gamma, \delta\}$ and $\lambda_j \in \{\beta, \gamma, \delta\}$ implies $\lambda_{j+1} = \alpha$.

Consider the following transformation $p\colon \Lambda^2 \to \Lambda^*$ defined by setting

$$p(\lambda\lambda') = \begin{cases} \epsilon & \text{if } \lambda = \lambda' \\ \delta & \text{if } \lambda = \beta, \lambda' = \gamma \text{ or } \lambda = \gamma, \lambda' = \beta \\ \gamma & \text{if } \lambda = \beta, \lambda' = \delta \text{ or } \lambda = \delta, \lambda' = \beta \\ \beta & \text{if } \lambda = \gamma, \lambda' = \delta \text{ or } \lambda = \delta, \lambda' = \gamma \\ \lambda\lambda' & \text{otherwise.} \end{cases} \tag{6.50}$$

Given a word $u = \lambda_1\lambda_2\cdots\lambda_n \in \Lambda^*$, and $1 \le i \le n$, we set

$$u_{(i)} = \lambda_1\lambda_2\cdots\lambda_{i-1}p(\lambda_i\lambda_{i+1})\lambda_{i+2}\cdots\lambda_n \in \Lambda^*.$$

Note that $u$ is reduced if and only if $u = u_{(i)}$ for all $i = 1, 2, \ldots, n-1$. By induction we define $u_{(i_1,i_2,\ldots,i_k)} = (u_{(i_1,i_2,\ldots,i_{k-1})})_{i_k}$ for $1 \le i_k \le \ell(u_{(i_1,i_2,\ldots,i_{k-1})}) - 1$.

Let now $w = \lambda_1\lambda_2\cdots\lambda_n \in \Lambda^*$. If $w$ is reduced we set $\overline{w} = w$. Otherwise, let $i_1$ be the minimal integer such that $w \ne w_{(i_1)}$. If $w_{(i_1)}$ is reduced we set $\overline{w} = w_{(i_1)}$. Otherwise, let $i_2$ be the minimal integer such that $w_{(i_1)} \ne w_{(i_1,i_2)}$. If $w_{(i_1,i_2)}$ is reduced we set $\overline{w} = w_{(i_1,i_2)}$. And so on. Since $\ell(w) > \ell(w_{(i_1)}) > \ell(w_{(i_1,i_2)}) > \cdots > \ell(w_{(i_1,i_2,\ldots,i_{k-1})}) > \ell(w_{(i_1,i_2,\ldots,i_k)}) > \cdots$ there exists an integer $k$ such that $w_{(i_1,i_2,\ldots,i_k)}$ is reduced. We then set $\overline{w} = w_{(i_1,i_2,\ldots,i_k)}$. A transformation $w_{(i_1,i_2,\ldots,i_{j-1})} \mapsto w_{(i_1,i_2,\ldots,i_j)}$ is called a *leftmost cancellation*. It follows that every word in $\Lambda^*$ can be transformed into a reduced word by a finite sequence of leftmost cancellations.

We denote by $\Theta_1 \subset \Lambda^*$ the subset consisting of all reduced words $w$ on $\Lambda$ which contain an even number $\ell_\alpha(w)$ of occurrences of the letter $\alpha$. Thus, a word $w \in \Theta_1$ is an alternate product of terms of the form $(\alpha\lambda\alpha)$ and $\lambda'$, with $\lambda, \lambda' \in \{\beta, \gamma, \delta\}$. Consider the maps $\Phi_0\colon \Theta_1 \to \Lambda^*$ and $\Phi_1\colon \Theta_1 \to \Lambda^*$ defined by setting $\Phi_\sigma((\alpha\lambda_1\alpha)\lambda_2\cdots) = \Phi_\sigma(\alpha\lambda_1\alpha)\Phi_\sigma(\lambda_2)\cdots$ (resp. $\Phi_\sigma(\lambda_1(\alpha\lambda_2\alpha)\cdots) = \Phi_\sigma(\lambda_1)\Phi_\sigma(\alpha\lambda_2\alpha)\cdots$ for all $\lambda_1, \lambda_2, \ldots \in \{\beta, \gamma, \delta\}$ where

$$\begin{array}{llll} \Phi_0(\beta) = \alpha, & \Phi_0(\alpha\beta\alpha) = \gamma, & \Phi_1(\beta) = \gamma, & \Phi_1(\alpha\beta\alpha) = \alpha, \\ \Phi_0(\gamma) = \alpha, & \Phi_0(\alpha\gamma\alpha) = \delta, & \Phi_1(\gamma) = \delta, & \Phi_1(\alpha\beta\alpha) = \alpha, \\ \Phi_0(\delta) = \epsilon, & \Phi_0(\alpha\delta\alpha) = \beta, & \Phi_1(\beta) = \delta, & \Phi_1(\alpha\beta\alpha) = \epsilon. \end{array}$$

For $n \ge 1$ we inductively define $\Theta_{n+1} \subset \Lambda^*$ as the subset consisting of all reduced words $w$ in $\Theta_n$ such that the reduced words $\overline{\Phi_0(w)}$ and $\overline{\Phi_1(w)}$ belong to $\Theta_n$. Also, given $w \in \Theta_n$ we recursively set $w_{\sigma_0\sigma_1\cdots\sigma_{n-1}\sigma} = \overline{\Phi_\sigma(w_{\sigma_0\sigma_1\cdots\sigma_{n-1}})}$ for all $\sigma, \sigma_1, \sigma_2, \ldots, \sigma_{n-1} \in \{0, 1\}$.

**Lemma 6.9.15.** *For all $w \in \Theta_1$ one has*

$$\ell(w_0) + \ell(w_1) \le \ell(w) + 1. \tag{6.51}$$

*Proof.* Let $w \in \Theta_1$. We distinguish a few cases.

**Case 1.** Suppose that $w$ starts and ends with $\alpha$ so that $w = (\alpha\lambda_1\alpha)\lambda_2 \cdots \lambda_{k-1}(\alpha\lambda_k\alpha)$ with $\lambda_i \in \{\beta, \gamma, \delta\}$. Note that $\ell(w) = 2k+1$. Then $w_\sigma = \overline{\Phi_\sigma(w)}$ and $\Phi_\sigma(w) = \Phi_\sigma(\alpha\lambda_1\alpha)\Phi_\sigma(\lambda_2)\cdots\Phi_\sigma(\lambda_{k-1})\Phi_\sigma(\alpha\lambda_k\alpha)$ and therefore

$$\begin{aligned}\ell(w_\sigma) &= \ell(\overline{\Phi_\sigma(w)}) \le \ell(\Phi_\sigma(w)) \\ &\le \ell(\Phi_\sigma(\alpha\lambda_1\alpha)) + \ell(\Phi_\sigma(\lambda_2)) + \cdots + \ell(\Phi_\sigma(\lambda_{k-1})) + \ell(\Phi_\sigma(\alpha\lambda_k\alpha)) \\ &\le k = \frac{(2k+1)-1}{2} \\ &= \frac{\ell(w)-1}{2}.\end{aligned}$$

**Case 2.** Suppose that $w$ starts and ends with letters in $\{\beta, \gamma, \delta\}$, that is, $w = \lambda_1(\alpha\lambda_2\alpha)\lambda_3\cdots(\alpha\lambda_{k-1}\alpha)\lambda_k$ with $\lambda_i \in \{\beta, \gamma, \delta\}$. Note that $\ell(w) = 2k-1$. Then $w_\sigma = \overline{\Phi_\sigma(w)}$ and $\Phi_\sigma(w) = \Phi_\sigma(\lambda_1)\Phi_\sigma(\alpha\lambda_2\alpha)\cdots\Phi_\sigma(\alpha\lambda_{k-1}\alpha)\Phi_\sigma(\lambda_k)$ and therefore

$$\begin{aligned}\ell(w_\sigma) &= \ell(\overline{\Phi_\sigma(w)}) \le \ell(\Phi_\sigma(w)) \\ &\le \ell(\Phi_\sigma(\lambda_1)) + \ell(\Phi_\sigma(\alpha\lambda_2\alpha)) + \cdots + \ell(\Phi_\sigma(\alpha\lambda_{k-1}\alpha)) + \ell(\Phi_\sigma(\lambda_k)) \\ &\le k = \frac{(2k-1)+1}{2} \\ &= \frac{\ell(w)+1}{2}.\end{aligned}$$

**Case 3.** Finally, suppose that $w$ starts with $\alpha$ and ends with $\lambda \in \{\beta, \gamma, \delta\}$, or viceversa. To fix ideas, suppose that $w = (\alpha\lambda_1\alpha)\lambda_2\cdots(\alpha\lambda_{k-1}\alpha)\lambda_k$ with $\lambda_i \in \{\beta, \gamma, \delta\}$. Passing to the inverse word $w^{-1}$ one handles the other possibility. Note that $\ell(w) = 2k$. Then $w_\sigma = \overline{\Phi_\sigma(w)}$ and $\Phi_\sigma(w) = \Phi_\sigma(\alpha\lambda_1\alpha)\Phi_\sigma(\lambda_2)\cdots\Phi_\sigma(\alpha\lambda_{k-1}\alpha)\Phi_\sigma(\lambda_k)$ and therefore

$$\begin{aligned}\ell(w_\sigma) &= \ell(\overline{\Phi_\sigma(w)}) \le \ell(\Phi_\sigma(w)) \\ &\le \ell(\Phi_\sigma(\alpha\lambda_1\alpha)) + \ell(\Phi_\sigma(\lambda_2)) + \cdots + \ell(\Phi_\sigma(\alpha\lambda_{k-1}\alpha)) + \ell(\Phi_\sigma(\lambda_k)) \\ &\le k = \frac{2k}{2} \\ &= \frac{\ell(w)}{2}.\end{aligned}$$

This shows (6.51). □

**Lemma 6.9.16.** *For all $g \in H_3$ one has*

$$\sum_{i,j,k=0}^{1} \ell_S(g_{ijk}) \le \frac{5}{6}\ell_S(g) + 8. \tag{6.52}$$

*Proof.* Let $g \in H_3$ and let $w \in \Theta_3$ be such that $g = \pi(w)$ and $\ell_S(g) = \ell(w)$. As a consequence of Lemma 6.9.15 we have

$$\begin{aligned} \sum_{i,j=0}^{1} \ell(w_{ij}) &\leq \sum_{i=0}^{1} (\ell(w_i) + 1) \\ &= \sum_{i=0}^{1} \ell(w_i) + 2 \\ &\leq \ell(w) + 3 \end{aligned} \tag{6.53}$$

and therefore

$$\begin{aligned} \sum_{i,j,k=0}^{1} \ell(w_{ijk}) &\leq \sum_{i,j=0}^{1} (\ell(w_{ij}) + 1) \\ &\leq \sum_{i,j=0}^{1} \ell(w_{ij}) + 4 \\ \text{(by (6.53))} \quad &\leq \left( \sum_{i=0}^{1} \ell(w_i) + 2 \right) + 4 = \sum_{i=0}^{1} \ell(w_i) + 6 \\ \text{(by (6.51))} \quad &\leq \ell(w) + 1 + 6 = \ell(w) + 7. \end{aligned} \tag{6.54}$$

Now we observe that by definition of the maps $\Phi_0$ and $\Phi_1$, the inequalities in (6.51), (6.53) and (6.54) can be sharpened as follows.

$$\ell(w_0) + \ell(w_1) \leq \ell(w) + 1 - \ell_\delta(w), \tag{6.55}$$

where $\ell_\delta(w)$ denotes the number of occurrences of the letter $\delta$ in the word $w$. Indeed, every such $\delta$, may appear in $w$ either isolated or in a triplet $(\alpha\delta\alpha)$. Then, in the first case we have $\Phi_0(\delta) = \epsilon$, while in the second one, $\Phi_1(\alpha\delta\alpha) = \epsilon$.

Similarly, since every letter $\gamma$ in the word $w$ will give rise via $\Phi_0$ or $\Phi_1$ to a letter $\delta$ in one of $w_0$ and $w_1$, the above argument shows that, even if some cancellation occurs,

$$\ell(w_{00}) + \ell(w_{01}) + \ell(w_{10}) + \ell(w_{11}) \leq \ell(w) + 3 - \ell_\gamma(w), \tag{6.56}$$

where $\ell_\gamma(w)$ denotes the number of occurrences of the letter $\gamma$ in the word $w$.

Finally, since every letter $\beta$ in the word $w$ will give rise via $\Phi_0$ or $\Phi_1$ to a letter $\gamma$ in one of $w_0$ and $w_1$, and therefore to a letter $\delta$ in one of $w_{00}, w_{01}, w_{10}$ and $w_{11}$, the above arguments show that indeed, even if some cancellation occurs,

$$\sum_{i,j,k=0}^{1} \ell(w_{ijk}) \leq \ell(w) + 7 - \ell_\beta(w), \tag{6.57}$$

where $\ell_\beta(w)$ denotes the number of occurrences of the letter $\beta$ in the word $w$.

Since $\ell(w) = \ell_\alpha(w)+\ell_\beta(w)+\ell_\gamma(w)+\ell_\delta(w)$ and $w$ is reduced, we necessarily have $\ell_\beta(w) + \ell_\gamma(w) + \ell_\delta(w) \geq \frac{\ell(w)-1}{2}$ and therefore

$$\max_{\lambda\in\{\beta,\gamma,\delta\}} \ell_\lambda(w) > \frac{\ell(w)}{6} - 1. \tag{6.58}$$

Taking into account the inequalities in (6.54) we obtain

$$\sum_{i,j,k=0}^{1} \ell(w_{ijk}) \leq \min\left\{ \sum_{i=0}^{1} \ell(w_i) + 6, \sum_{i,j=0}^{1} \ell(w_{ij}) + 4 \right\}. \tag{6.59}$$

Observe that $\pi(w_{ijk}) = g_{ijk}$ so that $\ell_S(g_{ijk}) \leq \ell(w_{ijk})$ for all $i, j, k = 0, 1$. Thus, using first (6.59), (6.55), (6.56) and (6.57), and then (6.58) we deduce

$$\begin{aligned}
\sum_{i,j,k=0}^{1} \ell_S(g_{ijk}) &\leq \sum_{i,j,k=0}^{1} \ell(w_{ijk}) \\
&\leq \min\{\ell(w) + 7 - \ell_\beta(w), \ell(w) + 7 - \ell_\gamma(w), \ell(w) + 7 - \ell_\delta(w)\} \\
&\leq \ell(w) + 7 - \max\{\ell_\beta(w), \ell_\gamma(w), \ell_\delta(w)\} \\
&\leq \ell(w) + 7 - \left(\frac{\ell(w)}{6} - 1\right) \\
&\leq \frac{5}{6}\ell(w) + 8 \\
&= \frac{5}{6}\ell(g) + 8.
\end{aligned}$$

This shows (6.52). □

*Proof of Theorem 6.9.12.* Let $H = H_3$. By Proposition 6.9.4 we have $[G : H] < \infty$. Moreover, by Lemma 6.9.16 we can take $\psi = \psi_3$, $M = 8$, $k = 5/6$ and $K = 8$, and apply Lemma 6.9.13 to obtain that $G$ has subexponential growth. □

A finitely generated group $G$ is said to have *intermediate growth* if $G$ has subexponential growth but does not have polynomial growth. Note that every finitely generated group of intermediate growth contains no subgroups isomorphic to $F_2$, by Corollary 6.6.5. From Theorem 6.9.9 and Theorem 6.9.12 we then get:

**Theorem 6.9.17.** *The Grigorchuk group $G$ has intermediate growth.* □

## 6.10 The Følner Condition for Finitely Generated Groups

Let $G$ be a group. As right multiplication by elements in $G$ is bijective we have, for all $A, B, F \subset G$, $F$ finite, and $g \in G$,

$$(A \setminus B)g = Ag \setminus Bg, \tag{6.60}$$

$$|F \setminus Fg| = |Fg \setminus F| = |F \setminus Fg^{-1}| = |Fg^{-1} \setminus F|. \tag{6.61}$$

**Lemma 6.10.1.** *Let $A, B, C \subset G$ be three finite sets. Then we have:*

$$|A \setminus B| \leq |A \setminus C| + |C \setminus B|. \tag{6.62}$$

*Proof.* Suppose that $x \in A \setminus B$, that is, (i) $x \in A$ and (ii) $x \notin B$. We distinguish two cases. If $x \notin C$, then, by (i), $x \in A \setminus C$. If $x \in C$, then, by (ii), $x \in C \setminus B$. In both cases, $x \in (A \setminus C) \cup (C \setminus B)$. It follows that $A \setminus B \subset (A \setminus C) \cup (C \setminus B)$ and (6.62) follows. □

Let now $S \subset G$ be a finite subset. The *isoperimetric constant* of $G$ with respect to $S$ is the nonnegative number

$$\iota_S(G) = \inf_F \frac{|FS \setminus F|}{|F|} \tag{6.63}$$

where $F$ runs over all non empty finite subsets of $G$. Observe that $\iota_S(G) = \iota_{S \cup \{1_G\}}(G)$.

**Proposition 6.10.2.** *Let $G$ be a finitely generated group and let $S \subset G$ be a finite generating subset. Then the following conditions are equivalent:*

(a) *$G$ is amenable;*
(b) *for all $\varepsilon > 0$ there is a finite subset $F = F(\varepsilon) \subset G$ such that*

$$|F \setminus Fs| < \varepsilon |F| \qquad \text{for all } s \in S; \tag{6.64}$$

(c) $\iota_S(G) = 0$.

*Proof.* Suppose (a) and fix $\varepsilon > 0$. By Theorem 4.9.1, $G$ satisfies the Følner conditions. For our convenience, we express the Følner conditions as follows. Given any finite subset $K \subset G$ and $\varepsilon' > 0$ there exists a finite subset $F = F(K, \varepsilon') \subset G$ such that

$$|F \setminus Fk| < \varepsilon' |F| \qquad \text{for all } k \in K. \tag{6.65}$$

Taking $\varepsilon' = \varepsilon$, $K = S$ and $F = F(S, \varepsilon)$ in 6.65, we then immediately obtain (6.64), and the implication (a) ⇒ (b) follows.

Suppose (b). Taking $g = s$ in (6.61) we immediately deduce

$$|F\backslash Fs| < \varepsilon|F| \qquad \text{for all } s \in S \cup S^{-1}. \tag{6.66}$$

Fix a finite set $K \subset G$ and $\varepsilon' > 0$. Since $S$ generates $G$ we can find $n \in \mathbb{N}$ such that $K \subset (S \cup S^{-1})^n$. Set $\varepsilon = \varepsilon'/n$. Let $F = F(\varepsilon) \subset G$ be a finite set such that (6.66) holds. Given $k \in K$ we can find $a_1, a_2, \dots, a_n \in S \cup S^{-1}$ such that $k = a_1 a_2 \cdots a_n$. Then, recalling (6.62) and (6.60), we have

$$\begin{aligned}
|F \setminus Fk| &= |F \setminus Fa_1a_2\cdots a_n| \\
&\le |F \setminus Fa_n| + |Fa_n \setminus Fa_{n-1}a_n| + |Fa_{n-1}a_n \setminus Fa_{n-2}a_{n-1}a_n| \\
&\qquad + \cdots + |Fa_2a_3\cdots a_n \setminus Fa_1a_2\cdots a_n| \\
&= |F \setminus Fa_n| + |F \setminus Fa_{n-1}| + |F \setminus Fa_{n-2}| + \cdots + |F \setminus Fa_1| \\
&< n\varepsilon|F| \\
&\le \varepsilon'|F|
\end{aligned}$$

for all $k \in K$. This shows that $G$ satisfies the Følner conditions (6.65) and therefore $G$ is amenable by Theorem 4.9.1. Thus (b) $\Rightarrow$ (a).

Finally, the equivalence (b) $\Leftrightarrow$ (c) immediately follows from (6.61) and the inequalities

$$|F \setminus Fs| = |Fs \setminus F| \le |FS \setminus F| \le \sum_{s' \in S} |Fs' \setminus F| = \sum_{s' \in S} |F \setminus Fs'|.$$

for any $s \in S$ and any finite subset $F \subset G$. □

## 6.11 Amenability of Groups of Subexponential Growth

In this section we show that every finitely generated group of subexponential growth is amenable.

We first prove the following:

**Lemma 6.11.1.** *Let $(a_n)_{n\ge 1}$ be a sequence of positive real numbers. Then*

$$\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} \le \liminf_{n\to\infty} \sqrt[n]{a_n}. \tag{6.67}$$

*Proof.* Set $\alpha = \liminf_{n\to\infty} \frac{a_{n+1}}{a_n}$. If $\alpha = 0$ there is nothing to prove. Otherwise, let $0 < \beta < \alpha$. Then, there exists an integer $N \ge 1$ such that $\frac{a_{n+1}}{a_n} \ge \beta$ for all $n \ge N$. Thus, for all $p = 1, 2 \dots$ one has

$$\frac{a_{N+p}}{a_N} = \frac{a_{N+p}}{a_{N+p-1}} \cdot \frac{a_{N+p-1}}{a_{N+p-2}} \cdots \frac{a_{N+1}}{a_N} \ge \beta^p.$$

This gives $a_{N+p} \ge \beta^p a_N$. It follows that setting $n = N + p$, one has, for all $n \ge N$

$$a_n \geq \beta^{n-N} a_N = \beta^n (\beta^{-N} a_N).$$

Thus, taking the $n$th roots we obtain $\sqrt[n]{a_n} \geq \beta \sqrt[n]{\beta^{-N} a_N}$ and therefore,

$$\liminf_{n\to\infty} \sqrt[n]{a_n} \geq \beta. \tag{6.68}$$

As (6.68) holds for all $\beta < \alpha$, we have $\liminf_{n\to\infty} \sqrt[n]{a_n} \geq \alpha = \liminf_{n\to\infty} \frac{a_{n+1}}{a_n}$. □

We are now in position to prove the main result of this section.

**Theorem 6.11.2.** *Every finitely generated group of subexponential growth is amenable.*

*Proof.* Let $G$ be a finitely generated group of subexponential growth. Let $S \subset G$ be a finite symmetric generating subset of $G$. Then, by virtue of the previous lemma we have $1 \leq \liminf_{n\to\infty} \frac{\gamma_S(n+1)}{\gamma_S(n)} \leq \lim_{n\to\infty} \sqrt[n]{\gamma_S(n)} = 1$, so that $\liminf_{n\to\infty} \frac{\gamma_S(n+1)}{\gamma_S(n)} = 1$. Fix $\varepsilon > 0$ and let $n_0 \in \mathbb{N}$ be such that

$$\frac{\gamma_S(n_0+1)}{\gamma_S(n_0)} < 1 + \varepsilon. \tag{6.69}$$

Set $F = B_S(n_0)$ and let us show that $|F \setminus Fs| < \varepsilon |F|$ for all $s \in S$. Let $s \in S$. Then, as $B_S(n_0)s \subset B_S(n_0+1)$ for all $s \in S$, we have

$$\begin{aligned} |F \setminus Fs| &= |Fs \setminus F| \\ &= |B_S(n_0)s \setminus B_S(n_0)| \\ &\leq |B_S(n_0+1) \setminus B_S(n_0)| \\ &= \gamma_S(n_0+1) - \gamma_S(n_0) \\ &< \varepsilon \gamma_S(n_0) \\ &= \varepsilon |F|, \end{aligned}$$

where the last inequality follows from (6.69). From Proposition 6.10.2 we deduce that $G$ is amenable. □

From Theorem 6.9.12 and Theorem 6.11.2 we deduce the following:

**Corollary 6.11.3.** *The Grigorchuk group $G$ is amenable.* □

## 6.12 The Theorems of Kesten and Day

Let $G$ be a group.

Let $p \in [1, +\infty)$. We consider the real Banach space

$$\ell^p(G) = \left\{ x \in \mathbb{R}^G : \sum_{g\in G} |x(g)|^p < \infty \right\}$$

consisting of all $p$-summable real functions on $G$. For $x \in \ell^p(G)$, the nonnegative number $\|x\|_p = \left(\sum_{g\in G} |x(g)|^p\right)^{\frac{1}{p}}$ is called the $\ell^p$-*norm* of $x$.

The support of a configuration $x \in \mathbb{R}^G$ is the set $\{g \in G : x(g) \neq 0\}$. We denote by $\mathbb{R}[G] \subset \mathbb{R}^G$ the vector subspace consisting of all finitely supported configurations in $\mathbb{R}^G$. Note that $\mathbb{R}[G]$ is a dense subspace in $\ell^p(G)$.

When $p = 2$, it is possible to endow $\ell^2(G)$ with a scalar product $\langle \cdot, \cdot \rangle$ defined by setting

$$\langle x, y \rangle = \sum_{g\in G} x(g)y(g)$$

for all $x, y \in \ell^2(G)$. Then, the $\ell^2$-norm of an element $x \in \ell^2(G)$ is given by $\|x\|_2 = \sqrt{\langle x, x \rangle}$. The space $(\ell^2(G), \langle \cdot, \cdot \rangle)$ is a real Hilbert space.

The $\ell^p$-*norm* of a linear map $T \colon \ell^p(G) \to \ell^p(G)$ is defined by

$$\|T\|_{p\to p} = \sup_{\substack{x\in\ell^p(G)\\ \|x\|_p\leq 1}} \|Tx\|_p = \sup_{\substack{x\in\ell^p(G)\\ x\neq 0}} \frac{\|Tx\|_p}{\|x\|_p}.$$

Then, $T$ is continuous if and only if $\|T\|_{p\to p} < \infty$. We denote by $\mathcal{L}(\ell^p(G))$ the space of all continuous linear maps $T \colon \ell^p(G) \to \ell^p(G)$ and by $I \colon \ell^p(G) \to \ell^p(G)$ the identity map.

Note that, by density of $\mathbb{R}[G]$ in $\ell^p(G)$, we have

$$\|T\|_{p\to p} = \sup_{\substack{x\in\mathbb{R}[G]\\ \|x\|_p\leq 1}} \|Tx\|_p = \sup_{\substack{x\in\mathbb{R}[G]\\ x\neq 0}} \frac{\|Tx\|_p}{\|x\|_p} \tag{6.70}$$

for all $T \in \mathcal{L}(\ell^p(G))$.

Let $p \in [1, +\infty)$ and $s \in G$. For all $x \in \ell^p(G)$ and $g \in G$ we set

$$(T_s^{(p)}x)(g) = x(gs).$$

We then have

$$\begin{aligned}
\|T_s^{(p)}x\|_p^p &= \sum_{g\in G} |T_s^{(p)}x(g)|^p \\
&= \sum_{g\in G} |x(gs)|^p \\
&= \sum_{h\in G} |x(h)|^p \text{ (by setting } h = gs) \\
&= \|x\|_p^p
\end{aligned}$$

for all $x \in \ell^p(G)$. We deduce that $T_s^{(p)}x \in \ell^p(G)$ and that the linear map $T_s^{(p)} \colon \ell^p(G) \to \ell^p(G)$ has $\ell^p$-norm

$$\|T_s^{(p)}\|_{p\to p} = 1. \tag{6.71}$$

In particular, $T_s^{(p)} \in \mathcal{L}(\ell^p(G))$.

Let now $S \subset G$ be a non-empty finite set. We denote by $M_S^{(p)} \colon \ell^p(G) \to \ell^p(G)$ the map defined by $M_S^{(p)} = \frac{1}{|S|}\sum_{s\in S} T_s^{(p)}$. In other words,

$$(M_S^{(p)}x)(g) = \frac{1}{|S|}\sum_{s\in S} x(gs)$$

for all $x \in \ell^p(G)$ and $g \in G$. The map $M_S^{(p)}$ is called the *$\ell^p$-Markov operator* on $\ell^p(G)$ associated with $S$.

**Proposition 6.12.1.** *Let $G$ be a group and $S \subset G$ a non-empty finite set. Then the $\ell^p$-Markov operator $M_S^{(p)} \colon \ell^p(G) \to \ell^p(G)$ is linear and continuous. Moreover,*

$$\|M_S^{(p)}\|_{p\to p} \leq 1. \tag{6.72}$$

*Proof.* Since $T_s^{(p)} \in \mathcal{L}(\ell^p(G))$, we deduce that $M_S^{(p)} \in \mathcal{L}(\ell^p(G))$. Finally, we have

$$\|M_S^{(p)}\|_{p\to p} = \|\frac{1}{|S|}\sum_{s\in S} T_s^{(p)}\|_{p\to p} \leq \frac{1}{|S|}\sum_{s\in S}\|T_s^{(p)}\|_{p\to p} = 1$$

where the last equality follows from (6.71). □

**Proposition 6.12.2.** *Let $G$ be a group. Let $S \subset G$ be a finite subset containing $1_G$. Then, the following conditions are equivalent:*

(a) $\|M_S^{(2)}\|_{2\to 2} = 1$;
(b) *given $\varepsilon > 0$ there exists $x \in \mathbb{R}[G]$ such that $\|x\|_2 = 1$ and $\|M_S^{(2)}x\|_2 \geq 1-\varepsilon$;*
(c) *given $\varepsilon > 0$ there exists $x \in \mathbb{R}[G]$ such that $x \geq 0$, $\|x\|_2 = 1$ and $\|M_S^{(2)}x\|_2 \geq 1-\varepsilon$;*
(d) *given $\varepsilon > 0$ there exists $x \in \mathbb{R}[G]$ such that $x \geq 0$, $\|x\|_2 = 1$ and*

$$\|x - T_s^{(2)}x\|_2 \leq \varepsilon \quad \text{for all } s \in S; \tag{6.73}$$

(e) *1 belongs to the real spectrum $\sigma(M_S^{(2)})$ of $M_S^{(2)}$.*

*Proof.* The implication (a) ⇒ (b) follows from the definition of $\ell^2$-norm and the density of $\mathbb{R}[G]$ in $\ell^2(G)$ (cf. (6.70)).

Let now $x \in \mathbb{R}[G]$. We have

$$\|M_S^{(2)}x\|_2^2 \leq \left\|M_S^{(2)}|x|\right\|_2^2. \tag{6.74}$$

Indeed,

$$\begin{aligned}\|M_S^{(2)}x\|_2^2 &= \left\|\frac{1}{|S|}\sum_{s\in S}T_s^{(2)}x\right\|_2^2\\ &= \sum_{g\in G}\left(\frac{1}{|S|}\sum_{s\in S}x(gs)\right)^2\\ &\le \sum_{g\in G}\left(\frac{1}{|S|}\sum_{s\in S}|x(gs)|\right)^2\\ &= \left\|\frac{1}{|S|}\sum_{s\in S}T_s^{(2)}|x|\right\|_2^2\\ &= \left\|M_S^{(2)}|x|\right\|_2^2.\end{aligned}$$

Thus, replacing $x$ by $|x|$ gives the implication (b) $\Rightarrow$ (c).

Suppose that (d) fails to hold, that is, there exists $\varepsilon_0 > 0$ such that for all $x \in \mathbb{R}[G]$, $x \ge 0$, $\|x\|_2 = 1$, there exists $s_0 \in S$ such that $\|x - T_{s_0}^{(2)}x\|_2 \ge \varepsilon_0$. Since $\ell^2(G)$ is uniformly convex (Lemma I.4.2) there exists $\delta_0 > 0$ such that

$$\left\|\frac{x + T_{s_0}^{(2)}x}{2}\right\|_2 \le 1 - \delta_0 \tag{6.75}$$

for all $x \in \mathbb{R}[G]$ such that $\|x\|_2 = 1$. It then follows

$$\begin{aligned}\|M_S^{(2)}x\|_2 &= \left\|\frac{1}{|S|}\sum_{s\in S}T_s^{(2)}x\right\|_2\\ &= \left\|\frac{2}{|S|}\left(\frac{x + T_{s_0}^{(2)}x}{2}\right) + \frac{1}{|S|}\sum_{s\in S\setminus\{1_G,s_0\}}T_s^{(2)}x\right\|_2\\ &\le \frac{2}{|S|}\left\|\frac{x + T_{s_0}^{(2)}x}{2}\right\|_2 + \frac{1}{|S|}\sum_{s\in S\setminus\{1_G,s_0\}}\|T_s^{(2)}x\|_2\\ \text{(by (6.75))}\quad &\le \frac{2}{|S|}(1-\delta_0) + \frac{1}{|S|}(|S|-2)\\ &= 1 - \frac{2\delta_0}{|S|}\end{aligned}$$

for all $x \in \mathbb{R}[G]$, such that $\|x\|_2 = 1$ and $x \ge 0$. This clearly contradicts (c). We have shown (c) $\Rightarrow$ (d).

Let us show (d) $\Rightarrow$ (e). Suppose that (6.73) holds. Then, for every $\varepsilon > 0$ there exists $x \in \mathbb{R}[G]$ such that

$$\begin{aligned}\|x - M_S^{(2)}x\|_2 &= \left\|x - \frac{1}{|S|}\sum_{s\in S} T_s^{(2)}x\right\|_2 \\ &= \left\|\frac{1}{|S|}\sum_{s\in S}(x - T_s^{(2)}x)\right\|_2 \\ &\leq \frac{1}{|S|}\sum_{s\in S}\|x - T_s^{(2)}x\|_2 \\ &\leq \varepsilon.\end{aligned}$$

Therefore, by Corollary I.2.3, the linear map $I - M_S^{(2)} \in \mathcal{L}(\ell^2(G))$ is not bijective, that is, $1 \in \sigma(M_S^{(2)})$.

Finally, the implication (e) $\Rightarrow$ (a) follow from (I.14) and the fact that $\|M_S^{(2)}\|_{2\to 2} \leq 1$ (Proposition 6.12.1). □

**Lemma 6.12.3.** *Let $S \subset G$ be a non-empty finite subset. The following conditions are equivalent:*

(a) $1 \in \sigma(M_S^{(2)})$;
(b) $1 \in \sigma(M_{S\cup\{1_G\}}^{(2)})$.

*Proof.* If $S$ contains $1_G$ there is nothing to prove.

Suppose that $1_G \notin S$ and set $\alpha = \frac{|S|}{|S|+1}$. Note that $0 < \alpha < 1$ and that

$$M_{S\cup\{1_G\}}^{(2)} = (1-\alpha)I + \alpha M_S^{(2)}. \tag{6.76}$$

We then have $I - M_{S\cup\{1_G\}}^{(2)} = (1-\alpha)(I - M_S^{(2)})$ so that $I - M_{S\cup\{1_G\}}^{(2)}$ is bijective if and only if $I - M_S^{(2)}$ is bijective. In other words, $1 \in \sigma(M_{S\cup\{1_G\}}^{(2)})$ if and only if $1 \in \sigma(M_S^{(2)})$. □

Before stating and proving the main result of this section we need a little more work. More precisely, we want to show the equivalence between the existence of an almost $S$-invariant positive element $x \in \mathbb{R}[G]$ of norm $\|x\|_2 = 1$ (cf. condition (d) in Proposition 6.12.2) and the existence of an almost $S$-invariant non-empty finite set $F \subset G$ (cf. condition (b) in Proposition 6.10.2).

Let $F \subset G$ be a finite set. Denote, as usual, by $\chi_F$ the characteristic map of $F$ and observe that $\chi_F \in \mathbb{R}[G]$.

**Proposition 6.12.4.** *Let $A, B \subset G$ be two finite sets. Then*

$$\|\chi_A - \chi_B\|_1 = \|\chi_A - \chi_B\|_2^2 = |A \setminus B| + |B \setminus A|. \tag{6.77}$$

*Proof.* First note that the map $\chi_A - \chi_B$ takes value 1 at each point of $A \setminus B$, value $-1$ at each point of $B \setminus A$, and value 0 everywhere else, that is, outside of the set $(A \setminus B) \cup (B \setminus A) = (G \setminus (A \cup B)) \cup (A \cap B)$. We then have

$$\begin{aligned}
\|\chi_A - \chi_B\|_2^2 &= \sum_{g \in G} |\chi_A(g) - \chi_B(g)|^2 \\
&= \sum_{g \in G} |\chi_A(g) - \chi_B(g)| \\
&= \sum_{\substack{g \in G: \\ \chi_A(g) - \chi_B(g) = 1}} |\chi_A(g) - \chi_B(g)| \\
&\quad + \sum_{\substack{g \in G: \\ \chi_A(g) - \chi_B(g) = -1}} |\chi_A(g) - \chi_B(g)| \\
&= \sum_{g \in G} \chi_{A \setminus B}(g) + \sum_{g \in G} \chi_{B \setminus A}(g) \\
&= |A \setminus B| + |B \setminus A|.
\end{aligned}$$

□

**Lemma 6.12.5.** *Let $x, y \in \ell^2(G)$ such that $\|x\|_2 = \|y\|_2 = 1$. Then, the following holds:*

(i) $x^2 \in \ell^1(G)$ *and* $\|x^2\|_1 = 1$;
(ii) $\|x^2 - y^2\|_1 \leq 2\|x - y\|_2$;
(iii) *suppose that $x, y \geq 0$. Then,* $\|x - y\|_2 \leq \left(\|x^2 - y^2\|_1\right)^{\frac{1}{2}}$.

*Proof.* (i) We have $\|x^2\|_1 = \sum_{g \in G} |x^2(g)| = \|x\|_2^2 = 1$.
(ii) We have

$$\begin{aligned}
\|x^2 - y^2\|_1 &= \sum_{g \in G} |x^2(g) - y^2(g)| \\
&= \sum_{g \in G} |x(g) - y(g)| \cdot |x(g) + y(g)| \\
&= \langle |x - y|, |x + y| \rangle \\
&\leq \|x - y\|_2 \cdot \|x + y\|_2 \\
&\leq \|x - y\|_2 \cdot (\|x\|_2 + \|y\|_2) \\
&= 2\|x - y\|_2,
\end{aligned}$$

where the first inequality follows from Cauchy-Schwarz.

(iii) We have

$$\begin{aligned}\|x-y\|_2^2 &= \sum_{g\in G}(x(g)-y(g))^2\\ &= \sum_{g\in G}|x(g)-y(g)|\cdot|x(g)-y(g)|\\ &\le \sum_{g\in G}|x(g)-y(g)|\cdot|x(g)+y(g)|\\ &= \sum_{g\in G}|x^2(g)-y^2(g)|\\ &= \|x^2-y^2\|_1,\end{aligned}$$

where the inequality follows from the fact that $x, y \ge 0$. □

**Lemma 6.12.6.** *Let $F \subset G$ be a non-empty finite set and $s \in G$. Then the following holds:*

(i) $T_s^{(p)}\chi_F = \chi_{Fs^{-1}}$, *for* $p = 1, 2$;
(ii) $\|\chi_F\|_1 = \|\chi_F\|_2^2 = |F|$;
(iii) $\|\chi_F - T_s^{(1)}\chi_F\|_1 = \|\chi_F - T_s^{(2)}\chi_F\|_2^2 = 2|F \setminus Fs|$.

*Proof.* For $g \in G$ and $p = 1, 2$, one has that $(T_p^{(s)}\chi_F)(g) = \chi_F(gs)$ equals 1 if and only if $gs \in F$, that is, if and only if $g \in Fs^{-1}$. This shows (i).

Moreover, recalling that $\chi_F(g) \in \{0, 1\}$, for all $g \in G$, on has

$$\|\chi_F\|_2^2 = \sum_{g\in G}\chi_F(g)^2 = \sum_{g\in G}\chi_F(g) = \|\chi_F\|_1 = |\{g \in G : \chi_F(g) = 1\}| = |F|.$$

This shows (ii).

Finally, (iii) follows from (i), (6.77) (with $A = F$ and $B = Fs^{-1}$) and (6.61). □

**Lemma 6.12.7.** *Let $x \in \mathbb{R}[G]$ such that $x \ge 0$ and $\|x\|_1 = 1$. Then there exist an integer $n \ge 1$, nonempty finite subsets $A_i \subset G$ and real numbers $\lambda_i > 0$, $1 \le i \le n$, satisfying $A_1 \supset A_2 \supset \cdots \supset A_n$ and $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$ such that*

$$x = \sum_{i=1}^{n}\lambda_i\frac{\chi_{A_i}}{|A_i|}. \tag{6.78}$$

*Proof.* Let $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_n$ be the values taken by $x$. For each $1 \le i \le n$, let us set

$$A_i = \{g \in G : x(g) \ge \alpha_i\}.$$

Clearly the sets $A_i$ are nonempty finite subsets of $G$ such that $A_1 \supset A_2 \supset \cdots \supset A_n$. On the other hand, we have

$$\begin{aligned} x &= \alpha_1 \chi_{A_1} + (\alpha_2 - \alpha_1)\chi_{A_2} + \cdots + (\alpha_n - \alpha_{n-1})\chi_{A_n} \\ &= \lambda_1 \frac{\chi_{A_1}}{|A_1|} + \lambda_2 \frac{\chi_{A_2}}{|A_2|} + \cdots + \lambda_n \frac{\chi_{A_n}}{|A_n|}, \end{aligned}$$

by setting $\lambda_1 = \alpha_1 |A_1|$ and $\lambda_i = (\alpha_i - \alpha_{i-1})|A_i|$ for $2 \le i \le n$. Thus $\lambda_i > 0$ for $1 \le i \le n$ and

$$\begin{aligned} \sum_{i=1}^{n} \lambda_i &= \alpha_1 |A_1| + (\alpha_2 - \alpha_1)|A_2| + \cdots + (\alpha_n - \alpha_{n-1})|A_n| \\ &= \alpha_1(|A_1| - |A_2|) + \alpha_2(|A_2| - |A_3|) + \cdots + \alpha_n |A_n| \\ &= \sum_{g \in G} x(g) = 1. \end{aligned}$$

□

**Lemma 6.12.8.** *With the same notation and hypotheses as in Lemma 6.12.7, we have*

$$\|x - T_g^{(1)} x\|_1 = \sum_{i=1}^{n} \lambda_i \frac{2|A_i \setminus A_i g|}{|A_i|} \tag{6.79}$$

*for every $g \in G$.*

*Proof.* Equality (6.78) gives us

$$\begin{aligned} x - T_g^{(1)} x &= \sum_{i=1}^{n} \lambda_i \frac{\chi_{A_i} - T_g^{(1)} \chi_{A_i}}{|A_i|} \\ \text{(by Lemma 6.12.6(a))} \quad &= \sum_{i=1}^{n} \lambda_i \frac{\chi_{A_i} - \chi_{A_i g^{-1}}}{|A_i|}. \end{aligned}$$

As we observed before (cf. the proof of Proposition 6.12.4), the map $\chi_{A_i} - \chi_{A_i g^{-1}}$ takes value 1 at each point of $A_i \setminus A_i g^{-1}$, value $-1$ at each point of $A_i g^{-1} \setminus A_i$, and value 0 everywhere else. Let us set

$$B = \bigcup_{1 \le i \le n} (A_i \setminus A_i g^{-1}) \quad \text{and} \quad C = \bigcup_{1 \le i \le n} (A_i g^{-1} \setminus A_i).$$

Note that the sets $B$ and $C$ are disjoint. Indeed, for all $1 \le i, j \le n$, we have $(A_i \setminus A_i g^{-1}) \cap (A_j \setminus A_j g^{-1}) = \varnothing$ since either $A_i \subset A_j$ or $A_j \subset A_i$ (which implies $A_j g^{-1} \subset A_i g^{-1}$).

It follows that

$$\begin{aligned}
\|x - T_g^{(1)}x\|_1 &= \sum_{a\in G} |(x - T_g^{(1)}x)(a)| \\
&= \sum_{a\in G} \left| \sum_{i=1}^n \lambda_i \frac{(\chi_{A_i} - \chi_{A_i g^{-1}})(a)}{|A_i|} \right| \\
&= \sum_{a\in B} \left| \sum_{i=1}^n \lambda_i \frac{(\chi_{A_i} - \chi_{A_i g^{-1}})(a)}{|A_i|} \right| + \sum_{a\in C} \left| \sum_{i=1}^n \lambda_i \frac{(\chi_{A_i} - \chi_{A_i g^{-1}})(a)}{|A_i|} \right| \\
&= \sum_{i=1}^n \lambda_i \frac{|A_i \setminus A_i g^{-1}|}{|A_i|} + \sum_{i=1}^n \lambda_i \frac{|A_i g^{-1} \setminus A_i|}{|A_i|} \\
\text{(by (6.61))} \quad &= \sum_{i=1}^n \lambda_i \frac{2|A_i \setminus A_i g|}{|A_i|}.
\end{aligned}$$

□

Let $G$ be a finitely generated group and let $S$ be a finite (not necessarily symmetric) generating subset of $G$. Consider the combinatorial Laplacian $\Delta_S \colon \mathbb{R}^G \to \mathbb{R}^G$ (cf. Example 1.4.3(b)). For all $x \in \ell^2(G) \subset \mathbb{R}^G$ and $g \in G$ we have

$$\begin{aligned}
(\Delta_S x)(g) &= |S|x(g) - \sum_{s\in S} x(gs) \\
&= |S|x(g) - \sum_{s\in S} (T_s^{(2)}x)(g) \\
&= |S|(I - M_S^{(2)})(x)(g).
\end{aligned}$$

This shows that for the map $\Delta_S^{(2)} = \Delta_S|_{\ell^2(G)}$ one has

$$\Delta_S^{(2)} = |S| \left( I - M_S^{(2)} \right) \tag{6.80}$$

and therefore $\Delta_S^{(2)} \in \mathcal{L}(\ell^2(G))$.

Let $\lambda \in \mathbb{R}$. From (6.80) we deduce that

$$\lambda I - \Delta_S^{(2)} = \lambda I - |S|(I - M_S^{(2)}) = -|S| \left( \left( 1 - \frac{\lambda}{|S|} \right) I - M_S^{(2)} \right).$$

It follows that

$$\lambda \in \sigma(\Delta_S^{(2)}) \Leftrightarrow \left( 1 - \frac{\lambda}{|S|} \right) \in \sigma(M_S^{(2)})$$

and therefore

$$0 \in \sigma(\Delta_S^{(2)}) \Leftrightarrow 1 \in \sigma(M_S^{(2)}). \tag{6.81}$$

**Theorem 6.12.9 (Kesten-Day).** *Let $G$ be a finitely generated group. Let $S \subset G$ be a finite (not necessarily symmetric) generating subset of $G$. Then*

*the following conditions are equivalent:*

(a) $G$ *is amenable;*
(b) $0 \in \sigma(\Delta_S^{(2)})$.

*Proof.* Suppose (a). By Theorem 4.9.1 we have that $G$ satisfies the Følner conditions. Fix $\varepsilon > 0$. By Proposition 6.10.2 there exists a finite subset $F \subset G$ such that $|F \setminus Fs| < \frac{\varepsilon^2}{2}|F|$ for all $s \in S$. Set $x = \frac{1}{\sqrt{|F|}}\chi_F$ and note that $\|x\|_2 = 1$. We have

$$\begin{aligned}
\|(I - M_S^{(2)})x\|_2 &= \|x - \frac{1}{|S|}\sum_{s\in S} T_s^{(2)}x\|_2 \\
&= \frac{1}{|S|}\|\sum_{s\in S}(x - T_s^{(2)}x)\|_2 \\
&\leq \frac{1}{|S|}\sum_{s\in S}\|x - T_s^{(2)}x\|_2 \\
&= \frac{1}{|S|\cdot\sqrt{|F|}}\sum_{s\in S}\|\chi_F - T_s^{(2)}\chi_F\|_2 \\
\text{(by Lemma 6.12.6(iii))} \quad &= \frac{1}{|S|\cdot\sqrt{|F|}}\sum_{s\in S} 2|F \setminus Fs|^{\frac{1}{2}} \\
&= \frac{1}{|S|}\sum_{s\in S} 2\left(\frac{|F \setminus Fs|}{|F|}\right)^{\frac{1}{2}} \\
&= \frac{2}{|F|^{\frac{1}{2}}}\varepsilon^{\frac{1}{2}} \\
&< \varepsilon.
\end{aligned}$$

From Corollary I.2.3 we deduce that $I - M_S^{(2)}$ is not bijective, that is, $1 \in \sigma(M_S^{(2)})$. From (6.81) we deduce that $0 \in \sigma(\Delta_S^{(2)})$. Thus (a) implies (b).

To prove the converse implication, first observe that, by virtue of Proposition 6.10.2, in order to prove (a), it suffices to show that given any $\varepsilon > 0$ there exists a finite subset $F \subset G$ such that

$$|F \setminus Fs| < \varepsilon|F| \qquad \text{for all } s \in S. \tag{6.82}$$

Fix $\varepsilon > 0$. Suppose (b) and observe that by (6.81) we have $1 \in \sigma(M_S^{(2)})$. Also note that, by Lemma 6.12.3, we can suppose that $S \ni 1_G$. Thus, by virtue of the implication (e) $\Rightarrow$ (c) in Proposition 6.12.2, we can find a non-negative function $x_\varepsilon \in \mathbb{R}[G]$ such that $\|x_\varepsilon\|_2 = 1$ and $\|x_\varepsilon - T_s^{(2)}x_\varepsilon\|_2 \leq \frac{\varepsilon}{2|S|}$ for all $s \in S$. Setting $x = x_\varepsilon^2$, from Lemma 6.12.5 we deduce that $\|x\|_1 = 1$ and

$$\|x - T_s^{(1)}x\|_1 \leq \frac{\varepsilon}{|S|} \qquad \text{for all } s \in S. \tag{6.83}$$

By Lemma 6.12.7, there exist an integer $n \geq 1$, nonempty finite subsets $A_i \subset G$ and real numbers $\lambda_i > 0$, $1 \leq i \leq n$, satisfying $A_1 \supset A_2 \supset \cdots \supset A_n$ and $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$, such that $x = \sum_{i=1}^{n} \lambda_i \frac{\chi_{A_i}}{|A_i|}$.

Set $\Omega = \{1, 2, \ldots, n\}$ and consider the unique probability measure $\mu$ on $\Omega$ such that $\mu(\{i\}) = \lambda_i$ for every $i \in \Omega$. Finally, for each $g \in G$, let $\Omega_g$ denote the subset of $\Omega$ defined by

$$\Omega_g = \left\{ i \in \Omega : \frac{|A_i \setminus A_i g|}{|A_i|} \geq \varepsilon \right\}.$$

It follows from Lemma 6.12.8 that

$$\begin{aligned} \|x - T_g^{(1)} x\|_1 &= \sum_{i \in \Omega} \lambda_i \frac{2|A_i \setminus A_i g|}{|A_i|} \\ &\geq \sum_{i \in \Omega_g} \lambda_i \frac{2|A_i \setminus A_i g|}{|A_i|} \\ &\geq 2\varepsilon \sum_{i \in \Omega_g} \lambda_i \\ &= 2\varepsilon \mu(\Omega_g). \end{aligned}$$

Therefore, we have

$$\mu(\Omega_g) \leq \frac{\|x - T_g^{(1)} x\|_1}{2\varepsilon} \quad \text{for all } g \in G.$$

By using (6.83), we deduce

$$\mu(\Omega_s) < \frac{1}{|S|} \quad \text{for all } s \in S,$$

which implies

$$\mu\left( \bigcup_{s \in S} \Omega_s \right) \leq \sum_{s \in S} \mu(\Omega_s) < 1.$$

Thus

$$\bigcup_{s \in S} \Omega_s \neq \Omega.$$

This means that there is some $i_0 \in \Omega$ such that

$$\frac{|A_{i_0} \setminus A_{i_0} s|}{|A_{i_0}|} < \varepsilon \quad \text{for all } s \in S.$$

Thus, in order to satisfy (6.82), we can take $F = A_{i_0}$. □

## 6.13 Quasi-Isometries

Let $G$ and $H$ be two groups.

**Definition 6.13.1.** A map $\varphi\colon G \to H$ is called a *quasi-isometric embedding* of $G$ into $H$ if the following two conditions hold:

(i) for every finite subset $K \subset G$ there exists a finite subset $F \subset H$ such that

$$g_1^{-1}g_2 \in K \implies \varphi(g_1)^{-1}\varphi(g_2) \in F \text{ for all } g_1, g_2 \in G; \tag{6.84}$$

(ii) for every finite subset $F \subset H$ there exists a finite subset $K \subset G$ such that

$$\varphi(g_1)^{-1}\varphi(g_2) \in F \implies g_1^{-1}g_2 \in K \text{ for all } g_1, g_2 \in G. \tag{6.85}$$

**Definition 6.13.2.** A *quasi-isometry* from $G$ to $H$ is a quasi-isometric embedding $\varphi$ of $G$ into $H$ for which there exists a finite subset $C \subset H$ such that

$$\varphi(G)C = H, \tag{6.86}$$

that is, for all $h \in H$ there exists $g \in G$ and $c \in C$ such that

$$h = \varphi(g)c. \tag{6.87}$$

If such a quasi-isometry from $G$ to $H$ exists then one says that $G$ is *quasi-isometric* to $H$.

*Example 6.13.3.* Let $G$ and $H$ be two groups and let $\varphi\colon G \to H$ be an injective homomorphism. Then $\varphi$ is a quasi-isometric embedding. Indeed, given a finite subset $K \subset G$, the set $F = \varphi(K)$ is a finite subset of $H$. On the other hand, if $g_1, g_2 \in G$ satisfy $g_1^{-1}g_2 \in K$, then we have $\varphi(g_1)^{-1}\varphi(g_2) = \varphi(g_1^{-1}g_2) \in F$. Conversely, suppose that $F'$ is a finite subset of $H$ and consider the set $K' = \varphi^{-1}(F') \subset G$. Note that $|K'| \leq |F'| < \infty$, since $\varphi$ is injective. On the other hand, if $g_1, g_2 \in G$ satisfy $\varphi(g_1)^{-1}\varphi(g_2) \in F'$, then we have $\varphi(g_1^{-1}g_2) = \varphi(g_1)^{-1}\varphi(g_2) \in F'$ so that $g_1^{-1}g_2 \in K'$. This shows that $\varphi$ is a quasi-isometric embedding.

Moreover, $\varphi$ is a quasi-isometry if and only if the index of the image subgroup $\varphi(G)$ is finite in $H$. Indeed, suppose first that $[H : \varphi(G)] < \infty$ and let $R \subset H$ be a set of representatives for the right cosets of $\varphi(G)$ in $H$. Note that $R$ is finite and

$$H = \coprod_{h \in R} \varphi(G)h = \varphi(G)R. \tag{6.88}$$

Thus, $\varphi$ is a quasi-isometry. Conversely, if $C \subset H$ is a finite set such that $H = \varphi(G)C$ one clearly has $[H : \varphi(G)] = |H/\varphi(G)| \leq |C| < \infty$.

*Remark 6.13.4.* Let $G$ and $H$ be two groups. Observe that if $\psi\colon G \to H$ is a quasi-isometric embedding (resp. a quasi-isometry) and $h_0 \in H$, then the map $\varphi\colon G \to H$ defined by $\varphi(g) = h_0\psi(g)$ for all $g \in G$ is a quasi-isometric embedding (resp. a quasi-isometry). Indeed, $\varphi(g_1)^{-1}\varphi(g_2) = \psi(g_1)^{-1}\psi(g_2)$ for all $g_1, g_2 \in G$ and, if $C \subset H$ is a finite set such that $\psi(G)C = H$, then $\varphi(G)C = h_0\psi(G)C = h_0H = H$. Note that taking $h_0 = \psi(1_G)^{-1}$ we have $\varphi(1_G) = 1_H$. It follows that if there exists a quasi-isometric embedding of $G$ into $H$, then there is a quasi-isometric embedding $\varphi\colon G \to H$ such that $\varphi(1_G) = 1_H$. Similarly, if $G$ is quasi-isometric to $H$, one can find a quasi-isometry $\varphi\colon G \to H$ such that $\varphi(1_G) = 1_H$.

**Proposition 6.13.5.** *Let $G$, $H$ and $L$ be three groups and let $\varphi\colon G \to H$ and $\psi\colon H \to L$ be two quasi-isometric embeddings (resp. quasi-isometries). Then the composition map $\psi \circ \varphi\colon G \to L$ is a quasi-isometric embedding (resp. quasi-isometry).*

*Proof.* Since $\varphi$ is a quasi-isometric embedding, given a finite subset $K \subset G$, we can find a finite subset $E \subset H$ such that if $g_1^{-1}g_2 \in K$, $g_1, g_2 \in G$, then $\varphi(g_1)^{-1}\varphi(g_2) \in E$. Similarly, as $\psi$ is a quasi-isometric embedding, we can find a finite subset $F \subset L$ such that if $h_1^{-1}h_2 \in E$, $h_1, h_2 \in H$, then $\psi(h_1)^{-1}\psi(h_2) \in F$. It follows that if $g_1^{-1}g_2 \in K$, $g_1, g_2 \in G$, then $\psi(\varphi(g_1))^{-1}\psi(\varphi(g_2)) \in F$. In the same way one proves that given a finite subset $F \subset L$ one can find a finite set $K \subset G$ such that if $\psi(\varphi(g_1))^{-1}\psi(\varphi(g_2)) \in F$, $g_1, g_2 \in G$, then $g_1^{-1}g_2 \in K$. This shows that $\psi \circ \varphi$ is a quasi-isometric embedding.

Let now $C \subset H$ and $D \subset L$ be two finite sets such that $\varphi(G)C = H$ and $\psi(H)D = L$. As $\psi$ is a quasi-isometric embedding there exists a finite set $F \subset L$ such that $h_1^{-1}h_2 \in C$ implies $\psi(h_1)^{-1}\psi(h_2) \in F$ for all $h_1, h_2 \in H$. Let us show that

$$\psi(\varphi(G))FD = L. \tag{6.89}$$

Given $\ell \in L$ there exists $h \in H$ such that

$$\ell \in \psi(h)D. \tag{6.90}$$

Let also $g \in G$ and $c \in C$ be such that $h = \varphi(g)c$. Set $h' = \varphi(g) \in H$ and observe that $(h')^{-1}h = c \in C$. It follows that $\psi(\varphi(g))^{-1}\psi(h) = \psi(h')^{-1}\psi(h) \in F$, that is,

$$\psi(h) \in \psi(\varphi(g))F. \tag{6.91}$$

From (6.90) and (6.91) we deduce $\ell \in \psi(\varphi(g))FD$. This shows (6.89). It follows that $\psi \circ \varphi\colon G \to L$ is a quasi-isometry. □

**Corollary 6.13.6.** *Let $G$ and $H$ be two groups and let $\varphi\colon G \to H$ be a quasi-isometric embedding. Let $L \subset G$ be a subgroup of $G$. Then the restriction map $\varphi|_L\colon L \to H$ is a quasi-isometric embedding.*

*Proof.* From Example 6.13.3 we deduce that the inclusion map $\iota\colon L \to G$ is a quasi-isometric embedding. Since $\varphi|_L = \varphi \circ \iota$, the statement follows from Proposition 6.13.5. □

**Proposition 6.13.7.** *Quasi-isometry is an equivalence relation in the class of groups.*

*Proof.* Every group is quasi-isometric to itself. Indeed, for any group $G$, the identity map $\mathrm{Id}_G\colon G \to G$ is clearly a quasi-isometry (just take $F = K$ in (6.84) and (6.84), and $C = \{1_G\}$ in (6.86)).

Suppose that a group $G$ is quasi-isometric to a group $H$. Let $\varphi\colon G \to H$ be a quasi-isometry and let $C \subset H$ be a finite subset satisfying (6.86). We define a map $\varphi'\colon H \to G$ as follows. For every $h \in H$ we can find $c \in C$ and $g \in G$ such that (6.87) holds. We then set $\varphi'(h) = g$. Let us show that $\varphi'$ is a quasi-isometry.

Let $h_1, h_2 \in H$. We set $g_i = \varphi'(h_i) \in G$ and $c_i = \varphi(g_i)^{-1}h_i \in C$, $i = 1, 2$. Note that $h_i = \varphi(g_i)c_i$, equivalently $\varphi(g_i) = h_i c_i^{-1}$, $i = 1, 2$.

Let now $F \subset H$ be a finite subset and suppose that $h_1^{-1}h_2 \in F$. Consider the finite set $F' = CFC^{-1} \subset H$. As $\varphi$ is a quasi-isometric embedding, we can find a finite subset $K \subset G$ such that if $\varphi(g_1)^{-1}\varphi(g_2) \in F'$ then $g_1^{-1}g_2 \in K$. But $\varphi(g_1)^{-1}\varphi(g_2) = c_1 h_1^{-1}h_2 c_2^{-1} \in CFC^{-1} = F'$ so that $\varphi'(h_1)^{-1}\varphi'(h_2) = g_1^{-1}g_2 \in K$. Conversely, let $K \subset G$ be a finite subset and suppose that $\varphi'(h_1)^{-1}\varphi'(h_2) = g_1^{-1}g_2$ belongs to $K$. As $\varphi$ is a quasi-isometric embedding, we can find a finite set $F' \subset H$ such that if $g_1^{-1}g_2 \in K$ then $\varphi(g_1)^{-1}\varphi(g_2) \in F'$. Set $F = C^{-1}F'C$. It follows that

$$h_1^{-1}h_2 = c_1^{-1}\varphi(g_1)^{-1}\varphi(g_2)c_2 \in c_1^{-1}F'c_2 \subset F.$$

This shows that $\varphi'$ is a quasi-isometric embedding of $H$ into $G$.

Now, as $\varphi$ is a quasi-isometric embedding, we can find a finite set $K \subset G$ such that if $\varphi(g_1)^{-1}\varphi(g_2) \in C$ then $g_1^{-1}g_2 \in K$. Let us show that

$$G = \varphi'(H)K. \tag{6.92}$$

Let $g \in G$. Set $h = \varphi(g) \in H$ and $g' = \varphi'(h) \in G$. Note that by (6.86) there exists $c \in C$ such that $h = \varphi(g')c$. Now $\varphi(g')^{-1}\varphi(g) = (ch^{-1})h = c \in C$ and therefore $(g')^{-1}g \in K$, that is, $g \in g'K = \varphi'(h)K \subset \varphi'(H)K$. This shows (6.92). We deduce that $\varphi'$ is a quasi-isometry. It follows that the symmetric property of quasi-isometry holds true.

Finally, the transitivity property of quasi-isometry was proved in Proposition 6.13.5. □

As a consequence, if a group $G$ is quasi-isometric to a group $H$ then $H$ is quasi-isometric to $G$ and we simply say that $G$ and $H$ are quasi-isometric.

**Proposition 6.13.8.** *Let $G$ be a group and let $N$ be a finite normal subgroup of $G$. Then $G$ and $G/N$ are quasi-isometric.*

*Proof.* Let $\varphi\colon G \to G/N$ denote the quotient homomorphism. Let us show that $\varphi$ is a quasi-isometry. Given a finite subset $K \subset G$, the set $F = \varphi(K)$ is a finite subset of $G/N$. On the other hand, if $g_1, g_2 \in G$ satisfy $g_1^{-1}g_2 \in K$, then we have $\varphi(g_1)^{-1}\varphi(g_2) = \varphi(g_1^{-1}g_2) \in F$. Conversely, suppose that $F'$ is a finite subset of $G/N$ and consider the set $K' = \varphi^{-1}(F') \subset G$. Note that $|K'| = |N| \cdot |F'| < \infty$. On the other hand, if $g_1, g_2 \in G$ satisfy $\varphi(g_1)^{-1}\varphi(g_2) \in F'$, then we have $\varphi(g_1^{-1}g_2) = \varphi(g_1)^{-1}\varphi(g_2) \in F'$ so that $g_1^{-1}g_2 \in K'$. This shows that $\varphi$ is a quasi-isometric embedding. Since $\varphi$ is surjective, we deduce that $\varphi$ is indeed a quasi-isometry. □

**Proposition 6.13.9.** *Let $G$ be a group and let $H$ be a subgroup of finite index of $G$. Then $G$ and $H$ are quasi-isometric.*

*Proof.* It follows from Example 6.13.3 that the inclusion map $\iota\colon H \to G$ is a quasi-isometry. □

**Corollary 6.13.10.** *If two groups are commensurable then they are quasi-isometric.* □

**Proposition 6.13.11.** *Let $G$ and $H$ be two groups. Suppose that $G$ and $H$ are quasi-isometric and that $G$ is finitely generated. Then $H$ is finitely generated.*

*Proof.* Let $\varphi\colon G \to H$ by a quasi-isometry and let $C \subset H$ be a finite subset such that $H = \varphi(G)C$. Let $S \subset G$ be a finite symmetric generating subset of $G$. Since $\varphi$ is a quasi-isometric embedding, we can find a finite subset $F \subset H$ such that $\varphi(g_1)^{-1}\varphi(g_2) \in F$ whenever $g_1, g_2 \in G$ satisfy $g_1^{-1}g_2 \in S$. Let us show that the set $T \subset H$ defined by $T = \{\varphi(1_G)\} \cup F \cup C$ is a generating subset for $H$. Let $h \in H$. As $H = \varphi(G)C$, we can find $g \in G$ and $c \in C$ such that $h = \varphi(g)c$. Since $S$ is a symmetric generating subset of $G$, there exist an integer $n \geq 0$ and elements $s_1, s_2, \ldots, s_n \in S$ such that $g = s_1 s_2 \cdots s_n$. Consider the elements $g_0, g_1, \ldots, g_n \in G$ defined by $g_0 = 1_G$ and $g_i = g_{i-1}s_i$ for all $1 \leq i \leq n$. Observe that $g_n = g$ and that $\varphi(g_{i-1})^{-1}\varphi(g_i) \in F$, for all $1 \leq i \leq n$, since $g_{i-1}^{-1}g_i = s_i \in S$. Writing

$$h = \varphi(g)c = \varphi(1_G)(\varphi(g_0)^{-1}\varphi(g_1)(\varphi(g_1)^{-1}\varphi(g_2)\cdots(\varphi(g_{n-1})^{-1}\varphi(g_n))c,$$

we deduce that $T$ generates $H$. As $T$ is finite, this shows that $H$ is finitely generated. □

In the two following propositions we characterize quasi-isometric embeddings and quasi-isometries for finitely generated groups in terms of the word metrics associated with finite symmetric generating subsets of the groups.

**Proposition 6.13.12.** *Let $G$ and $H$ be two finitely generated groups and denote by $d_G$ and $d_H$ the word metric associated with two finite symmetric generating subsets $S_G$ and $S_H$ for $G$ and $H$ respectively. Let $\varphi\colon G \to H$ be a map. Then the following conditions are equivalent:*

(a) $\varphi$ *is an isometric embedding;*
(b) *there exist constants* $\alpha \geq 1$ *and* $\beta \geq 0$ *such that, for all* $g, g' \in G$,

$$\frac{1}{\alpha} d_G(g, g') - \beta \leq d_H(\varphi(g), \varphi(g')) \leq \alpha d_G(g, g') + \beta. \tag{6.93}$$

*Proof.* Suppose (a). Let $K \subset G$ (resp. $F \subset H$) be a finite set such that $g_1^{-1} g_2 \in K$ (resp. $\varphi(g_1)^{-1}\varphi(g) \in F$) whenever $\varphi(g_1)^{-1}\varphi(g) \in S_H$ (resp. $g_1^{-1} g_2 \in S_G$) for all $g_1, g_2 \in G$. Also let $K' \subset G$ be a finite set such that $g_1^{-1} g_2 \in K'$ whenever $\varphi(g_1)^{-1}\varphi(g) \in \{1_H\}$, equivalently, $\varphi(g_1) = \varphi(g_2)$, for all $g_1, g_2 \in G$. We set

$$\alpha = \max\{\mathrm{diam}_{d_G}(K), \mathrm{diam}_{d_H}(F)\} \tag{6.94}$$

and

$$\beta = \frac{1}{\alpha} \mathrm{diam}_{d_G}(K'), \tag{6.95}$$

where we denoted by $\mathrm{diam}_{d_G}(K) = \max\{d_G(k_1, k_2) : k_1, k_2 \in K\}$ (resp. $\mathrm{diam}_{d_H}(F) = \max\{d_H(f_1, f_2) : f_1, f_2 \in F\}$) the diameter of $K \subset G$ (resp. $F \subset H$).

Let now $g, g' \in G$. Note that (6.93) trivially holds if $g = g'$. Suppose first that $d_G(g, g') = 1$. Thus $g^{-1}g' \in S_G$ and therefore $\varphi(g)^{-1}\varphi(g') \in F$ so that $d_H(\varphi(g), \varphi(g')) \leq \alpha$. Suppose now that $d_G(g, g') = n \geq 1$ and let $g_0, g_1, \ldots, g_n \in G$ be such that $g_0 = g$, $g_n = g'$ and $d_G(g_i, g_{i+1}) = 1$ for all $i = 0, 1, \ldots, n-1$. Using the triangular inequality we have:

$$\begin{aligned} d_H(\varphi(g), \varphi(g')) &= d_H(\varphi(g_0), \varphi(g_n)) \\ &\leq \sum_{i=0}^{n-1} d_H(\varphi(g_i), \varphi(g_{i+1})) \\ &\leq n\alpha \\ &= \alpha d_G(g, g'). \end{aligned}$$

This shows that

$$d_H(\varphi(g), \varphi(g')) \leq \alpha d_G(g, g') \tag{6.96}$$

for all $g, g' \in G$ and the second inequality in (6.93) follows.

Arguing as above, we deduce that for all $n \geq 1$ the condition $d_H(\varphi(g), \varphi(g')) = n$ implies $d_G(g, g') \leq n\alpha$. Thus, provided that $d_H(\varphi(g), \varphi(g')) \neq 0$, we have $d_G(g, g') \leq \alpha d_H(\varphi(g), \varphi(g'))$, equivalently

$$\frac{1}{\alpha} d_G(g, g') \leq d_H(\varphi(g), \varphi(g')). \tag{6.97}$$

On the other hand, if $\varphi(g) = \varphi(g')$, equivalently, $\varphi(g)^{-1}\varphi(g') \in \{1_H\}$, then $g^{-1}g' \in K'$ and from (6.95) we deduce that $\frac{1}{\alpha} d_G(g, g') \leq \beta$. It follows that

$$\frac{1}{\alpha} d_G(g, g') - \beta \leq 0 = d_H(\varphi(g), \varphi(g')). \tag{6.98}$$

From (6.97) (for $\varphi(g) \neq \varphi(g')$) and (6.98) (for $\varphi(g) = \varphi(g')$) we deduce that (6.98) holds for all $g, g' \in G$. Thus also the first inequality in (6.93) is proved. This shows (a) $\Rightarrow$ (b).

Conversely, suppose (b). Let $K \subset G$ be a finite subset. Consider the finite set $F = B^H_{S_H}(1_H, \delta) \subset H$ where $\delta = \alpha \max\{d_G(1_G, k) : k \in K\} + \beta$. Let $g_1, g_2 \in G$ and suppose that $g_1^{-1} g_2 \in K$. Using (6.93) we deduce that $d_H(1_H, \varphi(g_1)^{-1}\varphi(g_2)) = d_H(\varphi(g_1), \varphi(g_2)) \leq \alpha d_G(g_1, g_2) + \beta = \alpha d_G(1_G, g_1^{-1} g_2) + \beta \leq \delta$, so that $\varphi(g_1)^{-1}\varphi(g_2) \in B^H_{S_H}(1_H, \delta) = F$. On the other hand, let $F \subset H$ be a finite set. Consider the finite set $K = B^G_{S_G}(1_G, \delta')$ where $\delta' = \alpha \max\{d_H(1_H, f) : f \in F\} + \alpha\beta$. Let $g_1, g_2 \in G$ and suppose that $\varphi(g_1)^{-1}\varphi(g_2) \in F$. From (6.93) it follows that $d_G(1_G, g_1^{-1} g_2) = d_G(g_1, g_2) \leq \alpha d_H(\varphi(g_1), \varphi(g_2)) + \alpha\beta = \alpha d_H(1_H, \varphi(g_1)^{-1}\varphi(g_2)) + \alpha\beta \leq \delta'$, so that $g_1^{-1} g_2 \in B^G_{S_G}(1_G, \delta') = K$. This shows that $\varphi$ is a quasi-isometric embedding and the implication (b) $\Rightarrow$ (a) follows. □

**Proposition 6.13.13.** *Let $G$ and $H$ be two finitely generated groups and denote by $d_G$ and $d_H$ the word metric associated with two finite symmetric generating subsets $S_G$ and $S_H$ for $G$ and $H$ respectively. Let $\varphi\colon G \to H$ be an isometric embedding. Then the following conditions are equivalent:*

(a) *$\varphi$ is a quasi-isometry;*
(b) *there exists $\delta > 0$ such that for every $h \in H$ there exists $g \in G$ such that*

$$d_H(\varphi(g), h) \leq \delta. \tag{6.99}$$

*Proof.* Suppose that $\varphi$ is a quasi-isometric embedding. Let $C \subset H$ be a finite subset such that $H = \varphi(H)C$ and set $\delta = \max\{d_H(1_H, c) : c \in C\}$. Let $h \in H$, then there exist $c \in C$ and $g \in G$ such that $h = \varphi(g)c$, that is, $\varphi(g)^{-1} h = c \in C$. We deduce that $d_H(\varphi(g), h) \leq \delta$ and (6.99) follows. This shows (a) $\Rightarrow$ (b).

Conversely, suppose (b) and set $C = B_{S_G}(1_G, \delta) \subset G$. Let $h \in H$. By (6.99) there exists $g \in G$ such that $d_H(\varphi(g), h) \leq \delta$. It follows that $\varphi(g)^{-1} h \in C$, equivalently $h \in \varphi(g)C$. This shows that $H = \varphi(G)C$. Therefore $\varphi$ is a quasi-isometry and (a) follows. □

**Proposition 6.13.14.** *Let $G$ and $H$ be two finitely generated groups. Suppose that there is a quasi-isometric embedding $\varphi\colon G \to H$. Then one has $\gamma(G) \preceq \gamma(H)$.*

*Proof.* By Remark 6.13.4, we can assume $\varphi(1_G) = 1_H$. Denote by $d_G$ and $d_H$ the word metric associated with two finite symmetric generating subsets $S_G$ and $S_H$ for $G$ and $H$ respectively. By Proposition 6.13.12, there exist integers $\alpha \geq 1$ and $\beta \geq 0$ such that

$$\frac{1}{\alpha} d_G(g, g') - \beta \leq d_H(\varphi(g), \varphi(g')) \leq \alpha d_G(g, g') + \beta \tag{6.100}$$

for all $g, g' \in G$. Note that the left inequality in (6.100) implies that $\varphi(g) \neq \varphi(g')$ whenever $g, g' \in G$ satisfy $d_G(g, g') \geq \alpha\beta+1$. For $n \in \mathbb{N}$, let $B^G_{S_G}(n) \subset G$ (resp. $B^H_{S_H}(n) \subset H$) denote the ball of radius $n$ centered at $1_G$ (resp. $1_H$). Choose a subset $E_n \subset B^G_{S_G}(n)$ of maximal cardinality such that $d_G(g, g') \geq \alpha\beta + 1$ for all $g, g' \in E_n$. Setting $C = |B^G_{S_G}(\alpha\beta + 1)|$, we have

$$|B^G_{S_G}(n)| \leq C|E_n| \tag{6.101}$$

for all $n \in N$. Indeed, the balls of radius $C$ centered at the elements of $E_n$ cover $B^G_{S_G}(n)$ by the maximality of $E_n$. Now observe that the images by $\varphi$ of the elements of $E_n$ are all distinct and belong to $B^H_{S_H}(\alpha n + \beta)$ by the right inequality in (6.100). This implies that

$$|E_n| \leq |B^H_{S_H}(\alpha n + \beta)|.$$

By using (6.101), we then deduce that

$$|B^G_{S_G}(n)| \leq C|B^H_{S_H}(\alpha n + \beta)| \leq C|B^H_{S_H}(\beta)||B^H_{S_H}(\alpha n)| \leq C'|B^H_{S_H}(C'n)|$$

for all $n \geq 1$, where $C' = C\alpha|B^H_{S_H}(\beta)|$. It follows that $\gamma(G) \preceq \gamma(H)$. □

**Corollary 6.13.15.** *If two finitely generated groups are quasi-isometric, then they have the same growth type.*

*Proof.* If $G$ and $H$ are quasi-isometric finitely generated groups, then there exist a quasi-isometric embedding from $G$ into $H$ and a quasi-isometric embedding from $H$ into $G$. It follows that $\gamma(G) \preceq \gamma(H)$ and $\gamma(H) \preceq \gamma(G)$. Thus we have $\gamma(G) = \gamma(H)$. □

**Corollary 6.13.16.** *Let $G$ and $H$ be two quasi-isometric finitely generated groups. Then $G$ has exponential (resp. subexponential, resp. polynomial, resp. intermediate) growth if and only if $H$ has exponential (resp. subexponential, resp. polynomial, resp. intermediate) growth.* □

An important property of quasi-isometric embeddings is the fact that every quasi-isometric embedding is uniformly finite-to-one:

**Proposition 6.13.17.** *Let $G$ and $H$ be two groups. Let $\varphi\colon G \to H$ be a quasi-isometric embedding. Then there exists an integer $M \geq 1$ such that $|\varphi^{-1}(h)| \leq M$ for all $h \in H$.*

*Proof.* Since $\varphi$ is a quasi-isometric embedding, we can find a finite set $K \subset G$ such that $g_1^{-1}g_2 \in K$ whenever $g_1, g_2 \in G$ satisfy $\varphi(g_1)^{-1}\varphi(g_2) \in \{1_H\}$. Let us set $M = |K|$. Let $h \in H$. Suppose that $g_0 \in \varphi^{-1}(h)$. Then, every $g \in \varphi^{-1}(h)$ satisfies $\varphi(g_0)^{-1}\varphi(g) = h^{-1}h = 1_H$ and therefore $g_0^{-1}g \in K$. Thus, we have $\varphi^{-1}(h) \subset g_0K$ and hence $|\varphi^{-1}(h)| \leq |g_0K| = |K| = M$. □

**Corollary 6.13.18.** *Let $G$ and $H$ be two groups. Suppose that there exists a quasi-isometric embedding $\varphi\colon G \to H$ and that $H$ is finite. Then $G$ is finite.*

*Proof.* Taking $M$ as in the preceding proposition, we have $|G| \leq M|H|$. □

**Corollary 6.13.19.** *Let $G$ and $H$ be two groups. Suppose that $H$ is finite. Then $G$ is quasi-isometric to $H$ if and only if $G$ is finite.*

*Proof.* If $G$ is finite then it is clear from the definition that any map from $G$ to $H$ is a quasi-isometry. On the other hand, if $G$ is quasi-isometric to $H$ then $G$ is finite by Corollary 6.13.18. □

**Proposition 6.13.20.** *Let $G$ and $H$ be two groups. Suppose that there exists a quasi-isometric embedding $\varphi\colon G \to H$ and that $H$ is locally finite. Then $G$ is locally finite.*

*Proof.* Let $K$ be a finitely generated subgroup of $G$ and let $S \subset K$ be a finite symmetric generating subset of $K$. Since $\varphi$ is a quasi-isometric embedding, we can find a finite subset $F \subset H$ such that $\varphi(g_1)^{-1}\varphi(g_2) \in F$ whenever $g_1, g_2 \in G$ satisfy $g_1^{-1}g_2 \in S$. Let $L$ denote the subgroup of $H$ generated by $\{\varphi(1_G)\} \cup F$. Suppose that $k \in K$. Since $S$ is a symmetric generating subset of $K$, there exist an integer $n \geq 0$ and elements $s_1, s_2, \ldots, s_n \in S$ such that $k = s_1 s_2 \cdots s_n$. Consider the elements $k_0, k_1, \ldots, k_n \in K$ defined by $k_0 = 1_G$ and $k_i = k_{i-1}s_i$ for all $1 \leq i \leq n$. Observe that $k_n = k$ and that $\varphi(k_{i-1})^{-1}\varphi(k_i) \in F$, for all $1 \leq i \leq n$, since $k_{i-1}^{-1}k_i = s_i \in S$. Thus, we have

$$\varphi(k) = \varphi(1_G)(\varphi(k_0)^{-1}\varphi(k_1))(\varphi(k_1)^{-1}\varphi(k_2)) \cdots (\varphi(k_{n-1})^{-1}\varphi(k_n)) \in L.$$

It follows that $\varphi(K) \subset L$. As $H$ is locally finite, the subgroup $L$ is finite. On the other hand, it follows from Proposition 6.13.17 that there exists an integer $M \geq 1$ such that $|\varphi^{-1}(h)| \leq M$ for all $h \in H$. We deduce that $|K| \leq M|L| < \infty$. This shows that $G$ is locally finite. □

**Corollary 6.13.21.** *Let $G$ and $H$ be two groups. Suppose that $G$ and $H$ are quasi-isometric and that the group $G$ is locally finite. Then $H$ is locally finite.* □

**Lemma 6.13.22.** *Let $G$ be a group. Let $E, \Omega$ and $C$ be three subsets of $G$. Then, for every $c_0 \in C$ one has*

$$\partial_{Ec_0}(\Omega) = \partial_E(\Omega c_0^{-1}). \tag{6.102}$$

*and*

$$\partial_{EC}(\Omega) \supset \partial_{Ec_0}(\Omega). \tag{6.103}$$

*Proof.* By (5.3) we have $\Omega^{+Ec_0} = \bigcup_{e\in E} \Omega(ec_0)^{-1} = \bigcup_{e\in E} \Omega c_0^{-1}e^{-1} = (\Omega c_0^{-1})^{+E}$ and from (5.2) we deduce that $\Omega^{-Ec_0} = \bigcap_{e\in E} \Omega(ec_0)^{-1} = \bigcap_{e\in E} \Omega c_0^{-1}e^{-1} = (\Omega c_0^{-1})^{-E}$. It follows that $\partial_{Ec_0}(\Omega) = \Omega^{+Ec_0} \setminus \Omega^{-Ec_0} = (\Omega c_0^{-1})^{+E} \setminus (\Omega c_0^{-1})^{-E} = \partial_E(\Omega c_0^{-1})$. Similarly, we have

$$\Omega^{+EC} = \bigcup_{\substack{e\in E\\ c\in C}} \Omega(ec)^{-1} = \bigcup_{\substack{e\in E\\ c\in C}} \Omega c^{-1}e^{-1} \supset \bigcup_{e\in E} \Omega c_0^{-1}e^{-1} = (\Omega c_0^{-1})^{+E}$$

and

$$\Omega^{-EC} = \bigcap_{\substack{e\in E\\ c\in C}} \Omega(ec)^{-1} = \bigcap_{\substack{e\in E\\ c\in C}} \Omega c^{-1}e^{-1} \subset \bigcap_{e\in E} \Omega c_0^{-1}e^{-1} = (\Omega c_0^{-1})^{+E}.$$

We deduce that $\partial_{EC}(\Omega) = \Omega^{+EC} \setminus \Omega^{-EC} \supset (\Omega c_0^{-1})^{+E} \setminus (\Omega c_0^{-1})^{-E} = \partial_E(\Omega c_0^{-1}) = \partial_{Ec_0}(\Omega)$, where the last equality follows from (6.102). □

**Theorem 6.13.23.** *Let $G$ and $H$ be two quasi-isometric groups. Suppose that $H$ is amenable. Then $G$ is amenable.*

*Proof.* Let $\varphi\colon G \to H$ be a quasi-isometry and let $C \subset H$ be a finite set such that

$$H = \varphi(G)C = \bigcup_{c\in C} \varphi(G)c. \tag{6.104}$$

Let $E_G \subset G$ be a finite set and $\varepsilon > 0$. Let us show that there exists a finite set $F_G \subset G$ such that

$$|\partial_{E_G}(F_G)| < \varepsilon |F_G|. \tag{6.105}$$

Since $\varphi$ is a quasi-isometric embedding, we can find a finite set $E'_H \subset H$ such that

$$g_1^{-1}g_2 \in E_G \Rightarrow \varphi(g_1)^{-1}\varphi(g_2) \in E'_H \tag{6.106}$$

for all $g_1, g_2 \in G$. We then set $E_H = E'_H C$. Also, by Proposition 6.13.17 we can find an integer $M \geq 1$ such that

$$|\varphi^{-1}(F)| \leq M|F| \tag{6.107}$$

for all finite sets $F \subset H$. Since $H$ is amenable, it follows from Corollary 5.4.5 that we can find a finite subset $F'_H \subset H$ such that

$$|\partial_{E_H}(F'_H)| < \frac{\varepsilon}{M|C|}|F'_H|. \tag{6.108}$$

By (6.104) we can find $c_0 \in C$ such that

$$F'_H c_0^{-1} \cap \varphi(G) \neq \varnothing \tag{6.109}$$

and

$$|F'_H c^{-1} \cap \varphi(G)| \leq |F'_H c_0^{-1} \cap \varphi(G)| \tag{6.110}$$

for all $c \in C$. Set $F_H = F'_H c_0^{-1} \subset H$ and $F_G = \varphi^{-1}(F_H) \subset G$. Note that $F_G \neq \varnothing$ by (6.109). Then we have

$$\varphi(F_G) = F_H \cap \varphi(G) \subset F_H \tag{6.111}$$

and

$$\varphi(G \setminus F_G) \subset H \setminus F_H. \tag{6.112}$$

Moreover,

$$|F'_H| \leq |C| \cdot |F_G|. \tag{6.113}$$

Indeed, from (6.104) we deduce that $F'_H = \bigcup_{c\in C}(F'_H \cap \varphi(G)c)$ so that

$$\begin{aligned}
|F'_H| &\leq \sum_{c\in C} |F'_H \cap \varphi(G)c| \\
&= \sum_{c\in C} |F'_H c^{-1} \cap \varphi(G)| \\
&\leq \sum_{c\in C} |F'_H c_0^{-1} \cap \varphi(G)| \quad \text{(by (6.110))} \\
&= |C| \cdot |F_H \cap \varphi(G)| \\
&\leq |C| \cdot |F_G| \quad \text{(by the equality in (6.111))}.
\end{aligned}$$

Let us show that the set $F_G \subset G$ has the required property. Suppose that $g \in \partial_{E_G}(F_G)$. This means that the set $gE_G$ meets both $F_G$ and $G \setminus F_G$. Thus, there exist $g_1 \in F_G$ and $g_2 \in G \setminus F_G$ such that $g^{-1}g_1 \in E_G$ and $g^{-1}g_2 \in E_G$. This implies $\varphi(g)^{-1}\varphi(g_1) \in E'_H$ and $\varphi(g)^{-1}\varphi(g_2) \in E'_H$ by applying (6.106). As $\varphi(g_1) \in F_H$ (by (6.111)) and $\varphi(g_2) \in H \setminus F_H$ (by (6.112)), we deduce that the set $\varphi(g)E'_H$ meets both $F_H$ and $H \setminus F_H$. In other words, we have $\varphi(g) \in \partial_{E'_H}(F_H)$. This shows that

$$\partial_{E_G}(F_G) \subset \varphi^{-1}(\partial_{E'_H}(F_H)). \tag{6.114}$$

By taking cardinalities, we finally get

$$\begin{aligned}
|\partial_{E_G}(F_G)| &\leq |\varphi^{-1}(\partial_{E'_H}(F_H))| \quad \text{(by (6.114))} \\
&\leq M|\partial_{E'_H}(F_H)| \quad \text{(by (6.107))} \\
&= M|\partial_{E'_H c_0}(F'_H)| \quad \text{(by (6.102))} \\
&\leq M|\partial_{E_H}(F'_H)| \quad \text{(by (6.103))} \\
&\leq \frac{\varepsilon}{|C|}|F'_H| \quad \text{(by (6.108))} \\
&\leq \varepsilon|F_G| \quad \text{(by (6.113))}.
\end{aligned}$$

This shows that $F_G$ satisfies (6.105). From Corollary 5.4.5 we deduce that $G$ is amenable. □

## Notes

The idea of looking at a finitely generated group with a geometer's eye by investigating the properties of the family consisting of all its word metrics is due to M. Gromov ([Gro1, Gro3, Gro4]) and gave birth to the flourishing branch of mathematics which is commonly known as *geometric group theory* in the late 1970s.

In the 1950s, the notion of growth of a finitely generated group arose in group theory in relation to volume growth in Riemannian manifolds. This line of study was initiated by V.A. Efremovich [Efr] and A.S. Švarc [Sva] in the USSR and, slightly later and completely independently, by J. Milnor [Mil1] and J.A. Wolf [Wol] in the USA. In [Mil1] Milnor proved that fundamental groups of closed Riemannian manifolds with negative sectional curvature have exponential growth. Wolf [Wol] proved that a polycyclic group has polynomial growth if it contains a nilpotent subgroup of finite index and has exponential growth otherwise. Then, Milnor [Mil2] proved that every finitely generated non-polycyclic solvable group has exponential growth. Finally, in 1972 H. Bass [Bas] showed that the growth of a nilpotent group $G$ with a finite symmetric generating subset $S$ is exactly polynomial in the sense that there are positive constants $C_1$ and $C_2$ such that $C_1 n^d \leq \gamma_S(n) \leq C_2 n^d$, for all $n \geq 1$, where $d = d(G) \geq 0$ is an integer which can be computed explicitly from the lower central series of $G$ (see [Har1, page 201] for more information on the history and prehistory of these results). Note that the growth estimates of Milnor and Bass imply that a finitely generated solvable group has either polynomial or exponential growth. It was shown by J. Tits [Tits] (see also [Har1]) that every finitely generated linear group either is virtually nilpotent or contains a free subgroup of rank two. This last result, which is known as the *Tits alternative* for linear groups, implies that every finitely generated linear group has either polynomial growth or exponential growth.

The problem of the characterization of finitely generated groups with polynomial growth remained open until Gromov proved in [Gro2] that a finitely generated group with polynomial growth contains a nilpotent subgroup of finite index. It follows from the above mentioned result of Bass and Proposition 6.6.6 that a group of polynomial growth has in fact exactly polynomial growth. Thus, for finitely generated groups, the notions of polynomial and exactly polynomial growth coincide.

The *(general) Burnside problem*, posed by W. Burnside in 1902, asked whether a finitely generated periodic group is necessarily finite. It was answered in the negative in 1964 by E.S. Golod and I.R. Shafarevich [GolS], who gave an example of a finitely generated infinite $p$-group. The Grigorchuk group, also known as the *first Grigorchuk group*, was originally constructed by R. I. Grigorchuk in 1980 [Gri2] as a new example of a finitely generated infinite periodic group, thus providing another counterexample to the general Burnside problem. In 1984 Grigorchuk [Gri4] proved that this group has intermediate growth (this was announced by Grigorchuk in 1983 [Gri3]),

thus providing a positive answer to the *Milnor problem*, posed by Milnor in 1968, about the existence of finitely generated groups of intermediate growth. More precisely, in [Gri4] Grigorchuk proved, among other things, that $\exp(\sqrt{n}) \preceq \gamma(G) \preceq \exp(n^s)$, where $s = \log_{32}(31) \approx 0.991$. The Grigorchuk group also provides the first example of an amenable but not elementary amenable group (the class of *elementary amenable* groups is the smallest class of groups containing all finite and all abelian groups that is closed under taking subgroups, quotients, extensions, and directed unions), thus answering a question posed by Day in 1957 [Day1]. Among other interesting properties of the Grigorchuk group $G$, we mention the following (see [CMS2], [Har2], [Gri5], [GriP]): (a) $G$ is not finitely presented (a recursive set of defining relations for $G$ was found by I.G. Lysënok [Lys]), (b) $G$ is *just infinite* (it is infinite but every proper quotient is finite), (c) $G$ has *solvable word problem*, that is, there exists an algorithm that establishes whether, given $s_1, s_2, \ldots, s_n \in \{a, b, c, d\}$, one has $s_1 s_2 \cdots s_n = 1_G$ or not. Originally, the Grigorchuk group was defined as a group of Lebesgue measure-preserving transformations of the unit interval. By representing the elements of $\Sigma^*$ as the vertices of an infinite binary rooted tree, the Grigorchik group may be also realized as a subgroup of the full automorphism group of the tree. Another description of this group was provided by regarding it as a *group generated by a finite automaton* (see [GriNS]).

Simple random walks on groups were first considered by H. Kesten in [Kes1]. Given a finitely generated group $G$ and a finite symmetric generating subset $S \subset G$, the *simple random walk* on $G$ relative to $S$ is the $G$-invariant Markov chain with state space $G$ and transition probabilities given by

$$p(g,h) = \begin{cases} \frac{1}{|S|} & \text{if } g^{-1}h \in S \\ 0 & \text{otherwise.} \end{cases}$$

This can be interpreted as follows: a "random walker" on $G$ moves from a group element $g$ with equal probability to one of its $|S|$ neighbors $gs$, where $s \in S$. For $g, h \in G$, denote by $p^{(n)}(g,h)$ the probability of reaching $h$ from $g$ after exactly $n$ steps. We then have $p^{(0)}(g,h) = \delta_{g,h}$ where $\delta_{g,h}$ is the Kronecker symbol, $p^{(1)}(g,h) = p(g,h)$ and, more generally,

$$p^{(n)}(g,h) = \sum_{k \in G} p^{(n-1)}(g,k) p(k,h).$$

The quantity $p^{(n)}(g,g)$ does not depend on $g \in G$ and is called the return probability after $n$ steps. The number

$$\rho(G,S) = \limsup_{n \to \infty} \sqrt[n]{p^{(n)}(g,g)} \tag{6.115}$$

is called the *spectral radius* of the simple random walk on $G$ relative to $S$. Kesten [Kes1, Kes2] proved that one always has

$$\frac{2\sqrt{|S|-1}}{|S|} \leq \rho(G,S) \leq 1$$

with equality on the right if and only if $G$ is amenable. Moreover, if $S$ contains no involutions, equality on the left holds if and only if $G$ is a free group and $S$ is the symmetrization of a free base. Theorem 6.12.9 is just an $\ell^2$-reformulation of the amenability criterion of Kesten's theorem. Kesten also proved that if $S$ is a finite symmetric generating subset of a group $G$ and $N \subset G$ is a normal subgroup, then, denoting by $\overline{G} = G/N$ (resp $\overline{S} \subset \overline{G}$) the corresponding quotient group (resp. generating subset of $\overline{G}$) then $\rho(G,S) \leq \rho(\overline{G},\overline{S})$ with equality if and only if $N$ is amenable.

Day [Day3] extended Kesten's amenability criterion to non symmetric random walks. The associated Markov chain is then determined by a probability density whose support generates the group. In this setting, the associated Markov operator on $\ell^2(G)$ is no more self-adjoint, in general. The key ingredient of this new proof is the uniform convexity (*uniform rotundity* in Day's terminology) of Hilbert spaces (cf. Lemma I.4.2) and more generally of $\ell^p$-spaces with $p > 1$ (note that in fact Day considers, more generally, Markov operators on the Banach spaces $\ell^p(G)$, for $p > 1$). For more on this we refer to the paper [KaV] by V.A. Kaimanovich and A.M. Vershik and to W. Woess' review [Woe1] and monograph [Woe2].

Another important criterion for amenability of finitely generated groups has been obtained by R.I. Grigorchuk [Gri1]. Let $G$ be a group with $m$ generators. Then $G$ is isomorphic to $F/N$ where $F$ is the free group on $m$ generators and $N \subset F$ is a normal subgroup. Let $\alpha = \alpha(G;F,N)$ be defined by setting

$$\alpha = \limsup_{n\to\infty} \sqrt[n]{w(n)},$$

where $w(n)$ equals the number of elements in $N$ at distance at most $n$ from the identity element $1_F$ in the free group $F$. The non-negative number $\alpha$ is called the *cogrowth* of $G$ relative to the presentation $G = \langle F; N\rangle$. Grigorchuk [Gri1] proved that either $\alpha = 1$ (this holds if and only if $N = \{1_F\}$) or

$$\sqrt{2m-1} \leq \alpha \leq 2m-1. \tag{6.116}$$

He also showed that if $\rho = \rho(G,S)$ is the spectral radius of the simple random walk on $G$ relative to $S$ (the symmetrization of the image of a free base of $F$ under the canonical quotient homomorphism $F \to G = F/N$), then the following relation holds:

$$\rho = \begin{cases} \frac{\sqrt{2m-1}}{m} & \text{if } 1 \leq \alpha \leq \sqrt{2m-1} \\ \frac{\sqrt{2m-1}}{2m}\left(\frac{\sqrt{2m-1}}{\alpha} + \frac{\alpha}{\sqrt{2m-1}}\right) & \text{if } \sqrt{2m-1} \leq \alpha \leq 2m-1. \end{cases} \tag{6.117}$$

Then, from (6.117) and Kesten's criterion he deduced that in (6.116) equality holds on the right if and only if $G$ is amenable. This is called the *Grigorchuk*

*criterion* (or the *cogrowth criterion*) of amenability. Grigorchuk's criterion was used by Ol'shanskii [Ols] to show the existence of non-amenable groups without non-abelian free subgroups and by Adyan [Ady] to show that the free Burnside groups $B(m,n)$, with $m \geq 2$ generators and exponent $n \geq 665$ odd, are non-amenable.

The program of the classification of all finitely generated groups up to quasi-isometries was posed and initiated by Gromov [Gro4]. The definition of quasi-isometry presented here is modeled after Y. Shalom [Sha].

## Exercises

**6.1.** Let $S = \{s_1, s_2, \ldots, s_n\}$ be a finite subset of $\mathbb{Z}$. Show that $S$ generates $\mathbb{Z}$ if and only if $\gcd(s_1, s_2, \ldots, s_n) = 1$.

**6.2.** Let $G$ be a finitely generated group and let $S$ be a finite symmetric generating subset of $G$. Let $x \in G$ and denote by $L_x, R_x \colon G \to G$ the maps defined by $L_x(g) = xg$ and $R_x(g) = gx$ for all $g \in G$.

(a) Show that the map $g \mapsto d_S(g, R_x(g))$ is constant on $G$.

(b) Show that if $x$ is in the center of $G$, then $R_x$ is an isometry of $(G, d_S)$.

(c) Show that $\sup_{g \in G} d_S(g, L_x(g)) < \infty$ if and only if the conjugacy class of $x$ in $G$ is finite.

**6.3.** Suppose that $S$ and $S'$ are two finite symmetric generating subsets of a group $G$ with $S \subset S'$. Show that one has: (i) $\ell_S(g) \geq \ell_{S'}(g)$ and $d_S(g,h) \geq d_{S'}(g,h)$ for all $g, h \in G$; (ii) $B_S(n) \subset B_{S'}(n)$ and $\gamma_S(n) \leq \gamma_{S'}(n)$ for all $n \in \mathbb{N}$; (iii) $\lambda_S \leq \lambda_{S'}$.

**6.4.** Let $G_1$ and $G_2$ be two finitely generated groups and let $G = G_1 \times G_2$. Let $S_1$ (resp. $S_2$) be a finite symmetric generating subset of $G_1$ (resp. $G_2$). Show that $S = (S_1 \times \{1_{G_2}\}) \cup (\{1_{G_1}\} \times S_2)$ is a finite symmetric generating subset of $G$ and that one has $\ell_S^G(g) = \ell_{S_1}^{G_1}(g_1) + \ell_{S_2}^{G_2}(g_2)$ for all $g = (g_1, g_2) \in G$.

**6.5.** Let $S_1$ and $S_2$ be two sets. For $i = 1, 2$, let $\mathcal{Q}_i = (Q_i, E_i)$ be an $S_i$-labeled graph. We define their *direct product* $\mathcal{Q}_1 \times \mathcal{Q}_2$ as the $S$-labeled graph $\mathcal{Q} = (Q, E)$ with:

(1) $S = S_1 \coprod S_2$, where $\coprod$ denotes the disjoint union;
(2) $Q = Q_1 \times Q_2$;
(3) $E = \{((q_1, q_2), s, (q_1', q_2')) :$ either $q_1 = q_1'$ and $(q_2, s, q_2') \in E_2$, or $q_2 = q_2'$ and $(q_1, s, q_1') \in E_1\}$.

Suppose that $S_1$ (resp. $S_2$) is endowed with an involution $\iota_1 \colon S_1 \to S_1$ (resp. $\iota_2 \colon S_2 \to S_2$) and that $\mathcal{Q}_1$ (resp. $\mathcal{Q}_2$) is edge-symmetric with respect to such involution. Denote by $\iota \colon S \to S$ the map defined by $\iota(s) = \iota_i(s)$ if $s \in S_i$, $i = 1, 2$, and observe that $\iota$ is an involution. Show that $\mathcal{Q}$ is edge-symmetric with respect to $\iota$.

**6.6.** Let $G_1$ and $G_2$ be two finitely generated groups and let $S_1 \subset G_1$ and $S_2 \subset G_2$ be two finite and symmetric generating subsets such that $1_{G_1} \notin S_1$ and $1_{G_2} \notin S_2$. Consider the direct product group $G = G_1 \times G_2$ together with the (finite and symmetric) generating subset $S = (S_1 \times \{1_{G_2}\}) \cup (\{1_{G_1}\} \times S_2)$. Denote by $\mathcal{C}_{S_1}(G_1)$, $\mathcal{C}_{S_2}(G_2)$ and $\mathcal{C}_S(G)$ the corresponding Cayley graphs. If we identify $S_1$ with $S_1 \times \{1_{G_2}\}$ (resp. $S_2$ with $\{1_{G_1}\} \times S_2$), we may regard $\mathcal{C}_{S_1}(G_1)$ (resp. $\mathcal{C}_{S_2}(G_2)$) as an $(S_1 \times \{1_{G_2}\})$-labeled graph (resp. $(\{1_{G_1}\} \times S_2)$-labeled graph). Show that $\mathcal{C}_S(G) = \mathcal{C}_{S_1}(G_1) \times \mathcal{C}_{S_2}(G_2)$.

**6.7.** Let $G = \mathbb{Z}$, $S = \{1, -1\}$ and $S' = \{2, -2, 3, -3\}$. Find the best possible positive constants $C_1$ and $C_2$ such that $C_1 \ell_S(g) \leq \ell_{S'}(g) \leq C_2 \ell_S(g)$ for all $g \in G$. Hint: Check that $C_1 = 1/3$ and $C_2 = 2$.

**6.8.** Let $G = \mathbb{Z}^m$, where $m \geq 1$ is an integer. Consider the finite and symmetric generating subset

$$S = \{\pm(1, 0, 0, \ldots, 0), \pm(0, 1, 0, \ldots, 0), \ldots, \pm(0, 0, \ldots, 0, 1)\} \subset \mathbb{Z}^m.$$

(a) Show that if $g = (a_1, a_2, \ldots, a_m) \in \mathbb{Z}^m$ then $\ell_S(g) = |a_1| + |a_2| + \cdots + |a_m|$.

(b) Let $n \in \mathbb{N}$. Set $P_0(n) = 1$ and, for all integers $t \geq 1$ denote by $P_t(n)$ the number of distinct $t$-tuples $(a_1, a_2, \ldots, a_t)$ of positive integers such that $a_1 + a_2 + \cdots + a_t \leq n$. Show that $P_t(n) = \binom{n}{t}$ for $1 \leq t \leq n$. Hint: The map $(a_1, a_2, \ldots, a_t) \mapsto \{a_1, a_1 + a_2, \ldots, a_1 + a_2 + \cdots + a_t\}$ establishes a bijection between the set $\{(a_1, a_2, \ldots, a_t) \in \mathbb{N}^t : a_i \geq 1 \text{ and } a_1 + a_2 + \cdots + a_t \leq n\}$ and the set of all subsets of cardinality $t$ of the set $\{1, 2, \ldots, n\}$.

(c) For $n, t \in \mathbb{N}$ and $t \geq 1$ denote by $N_t(n)$ the number of all $m$-tuples $(a_1, a_2, \ldots, a_m) \in \mathbb{Z}^m$ with $\sum_{i=1}^m |a_i| \leq n$ and exactly $t$ many of the $a_i$'s nonzero. Show that $\gamma_S^{\mathbb{Z}^m}(n) = \sum_{t=0}^m N_t(n)$.

(d) Let $0 \leq t \leq n$ and let $I$ be a subset of $\{1, 2, \ldots, n\}$ such that $|I| = t$. Show that there are precisely $P_t(n)$ distinct elements $g = (a_1, a_2, \ldots, a_m) \in \mathbb{N}^m$ with $I = \{i : a_i > 0\}$ and such that $\ell_S(g) \leq n$.

(e) Deduce from (d) that there are exactly $\binom{m}{t}\binom{n}{t}$ elements $g = (a_1, a_2, \ldots, a_m) \in \mathbb{N}^m$ with $|\{i : a_i > 0\}| = t$ such that $\ell_S(g) \leq n$.

(f) Deduce from (e) that $N_t(n) = 2^t \binom{m}{t}\binom{n}{t}$.

(g) Deduce from (f) and (c) that $\gamma_S^{\mathbb{Z}^m}(n) = \sum_{t=0}^m 2^t \binom{m}{t}\binom{n}{t}$.

**6.9.** Suppose that $\gamma, \gamma' \colon \mathbb{N} \to [0, +\infty)$ are two growth functions such that $\gamma \preceq \gamma'$. Show that $\limsup_{n\to\infty} \sqrt[n]{\gamma(n)} \leq \limsup_{n\to\infty} \sqrt[n]{\gamma'(n)}$.

**6.10.** Let $\alpha = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. Consider the metabelian group $G = \mathbb{Z}^2 \rtimes_\alpha \mathbb{Z}$, that is, the semidirect product of $\mathbb{Z}^2$ by the infinite cyclic subgroup of $\mathrm{SL}_2(\mathbb{Z})$ generated by $\alpha$. Recall that

$$G = \left\{ \left( \begin{pmatrix} x \\ y \end{pmatrix}, z \right) : x, y, z \in \mathbb{Z} \right\}$$

with the multiplication defined by

$$\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, z_1\right)\left(\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, z_2\right) = \left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \alpha^{z_1}\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, z_1 + z_2\right)$$

for all $x_1, x_2, y_1, y_2, z_1, z_2 \in \mathbb{Z}$.

(a) Show that $\alpha^n\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} f_{2n+1} \\ f_{2n} \end{smallmatrix}\right)$ for all $n \in \mathbb{N}$, where $(f_k)_{k\in\mathbb{N}}$ is the *Fibonacci sequence* which is inductively defined by $f_0 = 0$, $f_1 = 1$, and $f_k = f_{k-2} + f_{k-1}$ for all $k \geq 2$.

(b) Deduce from (a) that, for any integer $n \geq 1$, the set

$$A(n) = \left\{ \left( \sum_{i=1}^{n} u_i \alpha^{i-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, 0 \right) : u_i \in \{0, 1\} \text{ for } 1 \leq i \leq n \right\} \subset G$$

has cardinality $|A(n)| = 2^n$.

(c) Consider the subset $S \subset G$ defined by $S = \{a, b, c, a^{-1}, b^{-1}, c^{-1}\}$, where

$$a = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, 0\right), b = \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, 0\right) \text{ and } c = \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, 1\right).$$

Show that $S$ is a finite symmetric generating subset of $G$ and that one has $A(n) \subset B_S^G(3n-2)$ for all $n \geq 1$.

(d) Deduce from (b) and (c) that $G$ has exponential growth.

**6.11.** Let $n \geq 2$. Show that $\mathrm{GL}_n(\mathbb{Z})$ is a finitely generated group of exponential growth. Hint: Use Exercise 2.18, Lemma 2.3.2 and Corollary 6.6.5.

**6.12.** Let $F_2$ denote the free group of rank two. Show that the groups $\mathrm{GL}_2(\mathbb{Z})$ and $F_2$ are commensurable. Hint: Use elementary row operations to show that the matrices $\left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)$ generate a finite index subgroup of $\mathrm{GL}_2(\mathbb{Z})$ and apply Lemma 2.3.2.

**6.13.** *Growth of the Baumslag-Solitar group $BS(1, m)$.* Let $m$ be an integer such that $|m| \geq 2$. Prove that the metabelian group $G = \langle a, b : aba^{-1} = b^m \rangle$ studied in Exercises 2.7 and 4.21 has exponential growth. Hint: Use an argument similar to the one used for the case $m = 2$ in the proof of Proposition 6.7.1. More precisely, take $S = \{a, b, a^{-1}, b^{-1}\}$ and prove that every element of the form $g = b^k$, where $0 \leq k \leq |m|^n - 1$ and $n \geq 1$, has word length $\ell_S(g) \leq |m|n + n - 2$ by developing $k$ in base $|m|$.

**6.14.** *Affine representation of the Baumslag-Solitar group $BS(1, m)$.* Let $m$ be an integer such that $|m| \geq 2$. Consider the group $G$ given by the presentation $G = \langle a, b : aba^{-1} = b^m \rangle$ (cf. Exercises 2.7, 4.21 and 6.13).

(a) Let $\alpha, \beta \colon \mathbb{R} \to \mathbb{R}$ be the maps respectively defined by $\alpha(x) = mx$ and $\beta(x) = x + 1$ fro all $x \in \mathbb{R}$. Show that there is a unique homomorphism $\varphi \colon G \to \mathrm{Sym}(\mathbb{R})$ satisfying $\varphi(a) = \alpha$ and $\varphi(b) = \beta$.

(b) Show that $\varphi$ is injective. Hint: Use Exercise 2.7(a).

(c) Let $n_0$ be an integer such that $|m|^{n_0} \geq 3$. Consider the elements $\lambda, \mu \in \mathrm{Sym}(\mathbb{R})$ respectively defined by $\lambda = \alpha^{-n_0}$ and $\mu = \beta\lambda\beta^{-1}$. Check that the open intervals $i = (-1/2, 1/2)$ and $J = (1/2, 3/2)$ satisfy $\lambda(J) \subset I$ and $\mu(I) \subset J$.

(d) Let $n \geq 1$ be an integer and let $\sigma_1, \sigma_2, \ldots, \sigma_n \in \{\lambda, \mu\}$. Prove that

$$\sigma_1\sigma_2\cdots\sigma_n \neq \mathrm{Id}_{\mathbb{R}}\,.$$

Hint: Use (c) and play ping-pong as in the proof of Theorem D.5.1.

(e) Use (a), (b) and (d) to get another proof of the fact that $G$ has exponential growth.

**6.15.** *Growth of the lamplighter group.* Let $L = (\mathbb{Z}/2Z) \wr \mathbb{Z}$ denote the lamplighter group (cf. Exercise 4.19). Recall that $L$ is the semidirect product of a normal subgroup $H = \oplus_{n\in\mathbb{Z}} A_n$, where each $A_n$ is a subgroup of order 2, with an infinite cyclic subgroup $N$ generated by an element $t$ which satisfies $tat^{-1} = (a_{n-1})_{n\in\mathbb{Z}}$ for all $a = (a_n)_{n\in\mathbb{Z}} \in H$. Let $s$ denote the nontrivial element of $A_0$.

(a) Show that $S = \{s, t, t^{-1}\}$ is a symmetric generating subset of $L$.

(b) For $n \geq 1$, let $B_n = \oplus_{k=0}^{n-1} A_k$. Prove that $B_n$ is a subgroup of $H$ generated by the elements $s, tst^{-1}, t^2st^{-2}, \ldots, t^{n-1}st^{-n+1}$ and that $|B_n| = 2^n$.

(c) Deduce from (b) that $2^n \leq \gamma_S^L(3n-2)$ for all $n \geq 1$.

(d) Deduce from (c) that $L$ has exponential growth.

**6.16.** *Growth of the integral Heisenberg group.* Let $G = H_{\mathbb{Z}}$ denote the Heisenberg group over the ring of integers (cf. Example 4.6.5). Recall that $G$ is the subgroup of $\mathrm{SL}_3(\mathbb{Z})$ consisting of all matrices of the form

$$M(x,y,z) = \begin{pmatrix} 1 & y & z \\ 0 & 1 & x \\ 0 & 0 & 1 \end{pmatrix} \quad (x,y,z \in \mathbb{Z}).$$

Let us set $A = M(1,0,0)$, $B = M(0,1,0)$, $C = M(0,0,1)$, and $S = \{A, A^{-1}, B, B^{-1}, C, C^{-1}\}$.

(a) Verify that $M(x,y,z) = A^xB^yC^z$ for all $x, y, z \in \mathbb{Z}$.

(b) Show that $S$ is a finite symmetric generating subset of $G$.

(c) Show that $C^xA^y = A^yC^x$, $C^xB^y = B^yC^x$, and $B^xA^y = A^yB^xC^{xy}$ for all $x, y, z \in \mathbb{Z}$.

(d) Deduce from (c) that if $P \in G$ satisfies $\ell_S(P) \leq n$, then there exist $x, y, z \in \mathbb{Z}$ with $|x| \leq n$, $|y| \leq n$, and $|z| \leq n^2 + n$, such that $P = A^xB^yC^z$.

(e) Deduce from (d) that there exists a constant $C_1 > 0$ such that $\gamma_S(n) \leq C_1n^4$ for all $n \geq 1$.

(f) Let $n, x, y, z$ be integers such that $0 \leq x \leq n$, $0 \leq y \leq n$, and $0 \leq z \leq n^2$. Show that there exist integers $q, r$ with $0 \leq q \leq n$ and $0 \leq r \leq n-1$ such that

$$A^x B^y C^z = A^{x-q} B^n A^q B^{y-n} C^r.$$

Hint: Use (c) and Euclidean division of $z$ by $n$.

(g) Deduce from (f) that $\gamma_S(5n) \geq (n+1)^2(n^2+1)$ for all $n \geq 0$.

(h) Deduce from (g) that there exists a constant $C_2 > 0$ such that $\gamma_S(n) \geq C_2 n^4$ for all $n \geq 0$.

(i) Deduce from (e) and (h) that $\gamma(G) \sim n^4$.

**6.17.** Let $G$ be the Grigorchuk group.

(a) Show that $G$ acts transitively on $\Sigma^n$ for all $n \in \mathbb{N}$. Hint: Use induction on $n$. More precisely, let $w_1, w_2 \in \Sigma^n$ and suppose first that $w_1$ and $w_2$ start with the same letter, that is, there exist $x \in \Sigma$ and $u_1, u_2 \in \Sigma^{n-1}$ such that $w_1 = xu_1$ and $w_2 = xu_2$. Use induction and Proposition 6.9.7(i). Otherwise, if $w_1$ and $w_2$ do not start with the same letter, observe that $w_1$ and $a(w_2)$ do start with the same letter and reduce to the previous case.

(b) Use (a) to recover the fact that $G$ is infinite (cf. Theorem 6.9.8).

**6.18.** Let $K_1 = \mathbb{Z}/2\mathbb{Z} = \{\overline{0}, \overline{1}\}$ and, for $n \geq 2$, define by induction $K_n = K_{n-1} \wr (\mathbb{Z}/2\mathbb{Z})$. Recall that $K_{n-1} \wr (\mathbb{Z}/2\mathbb{Z}) = (K_{n-1})^{\mathbb{Z}/2\mathbb{Z}} \rtimes (\mathbb{Z}/2\mathbb{Z})$, so that $K_n$ consists of the elements $(f, a) \in (K_{n-1})^{\mathbb{Z}/2\mathbb{Z}} \times (\mathbb{Z}/2\mathbb{Z})$ with the multiplication defined by $(f_1, a_1)(f_2, a_2) = (f_1 f_2^{a_1}, a_1 + a_2)$, for all $f_1, f_2 \in (K_{n-1})^{\mathbb{Z}/2\mathbb{Z}}$ and $a_1, a_2 \in \mathbb{Z}/2\mathbb{Z}$, where $f^a(a') = f(a + a') \in K_{n-1}$ for all $f \in (K_{n-1})^{\mathbb{Z}/2\mathbb{Z}}$ and $a, a' \in \mathbb{Z}/2\mathbb{Z}$. The group $K_n$ is called the *Kaloujnine 2-group* of degree $n$. Set $\Sigma = \{0, 1\}$. For $n = 1$ and $x \in \Sigma$, we set $\overline{1}(x) = 1 - x$, and $\overline{0}(x) = x$. For $n \geq 2$ let $g = (f, a) \in K_n$ and $w = xu \in \Sigma^n$, where $x \in \Sigma$ and $u \in \Sigma^{n-1}$. We then set $g(w) = a(x) f(\overline{x})(u) \in \Sigma^n$ (note that $a(x) \in \Sigma$, $f(\overline{x}) \in K_{n-1}$, and $f(\overline{x})(u) \in \Sigma^{n-1}$ is defined by induction).

(a) Show that the map $K_n \times \Sigma^n \ni (g, w) \mapsto g(w) \in \Sigma^n$ defines an action of the group $K_n$ on $\Sigma^n$.

(b) Show that this action is faithful.

(c) By virtue of (b), we may regard $K_n$ as a subgroup of $\mathrm{Sym}(\Sigma^n)$. Show, by simple counting arguments, that $K_n$ is a Sylow 2-subgroup of $\mathrm{Sym}(\Sigma^n)$.

(d) Consider the elements $g_{1,n}, g_{2,n}, \ldots, g_{n,n} \in K_n$ defined by induction as follows. First define $g_{1,1} = \overline{1} \in K_1 = \mathbb{Z}/2\mathbb{Z}$ and then, for $1 \leq m \leq n$, set $g_{1,m} = (f_{1,m}, \overline{1}) \in K_m$, where $f_{1,m} \colon \mathbb{Z}/2\mathbb{Z} \to K_{m-1}$ is given by $f_{1,m}(a) = 1_{K_{m-1}}$ for all $a \in \mathbb{Z}/2\mathbb{Z}$. Finally, for $2 \leq k \leq m \leq n$, set $g_{k,m} = (f_{k,m}, \overline{0}) \in K_m$, where $f_{k,m} \colon \mathbb{Z}/2\mathbb{Z} \to K_{m-1}$ is given by $f_{k,m}(\overline{0}) = g_{k-1,m-1}$ and $f_{k,m}(\overline{1}) = (f_{1,m-1}, \overline{0})$. Show that $g_{1,n}, g_{2,n}, \ldots, g_{n,n}$ generate $K_n$ and verify that $g_{i,n}(uxv) = u(1-x)v \in \Sigma^n$ for all $u \in \Sigma^{i-1}, x \in \Sigma$ and $v \in \Sigma^{n-i}$.

(e) Define a map $\pi_n \colon K_n \to (\mathbb{Z}/2\mathbb{Z})^n$ by induction as follows. $\pi_1 \colon K_1 = \mathbb{Z}/2\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$ is the identity map, while, for $n \geq 2$ and $(f, a) \in K_n$ we set $\pi_n(f, a) = (\sum_{b \in (\mathbb{Z}/2\mathbb{Z})} \pi_{n-1}(f(b)), a) \in (\mathbb{Z}/2\mathbb{Z})^{n-1} \times (\mathbb{Z}/2\mathbb{Z}) = (\mathbb{Z}/2\mathbb{Z})^n$ for all $f \in (K_{n-1})^{\mathbb{Z}/2\mathbb{Z}}$ and $a \in \mathbb{Z}/2\mathbb{Z}$. Show that $\pi_n$ is a surjective homomorphism. Hint: To prove that $\pi_n$ is a homomorphism use induction on $n$ and the fact that $\sum_{b \in (\mathbb{Z}/2\mathbb{Z})} \pi_{n-1}(f^a(b)) = \sum_{b \in (\mathbb{Z}/2\mathbb{Z})} \pi_{n-1}(f(b))$ for all $f \in (K_{n-1})^{\mathbb{Z}/2\mathbb{Z}}$

and $a \in \mathbb{Z}/2\mathbb{Z}$. To show surjectivity, look at the $\pi_n$-images of the elements $g_{1,n}, g_{2,n}, \dots, g_{n,n} \in K_n$ defined in (d).

(f) Deduce from (e) that $K_n/[K_n, K_n]$ (the *abelianization* of $K_n$) is isomorphic to $(\mathbb{Z}/2\mathbb{Z})^n$.

(g) Consider the map $\Phi_n \colon \Sigma^n \to \{1, 2, \dots, 2^n\}$ given by expansion in base two, that is, $\Phi_n(i_1, i_2, \dots, i_n) = i_1 + 2i_2 + 4i_3 + \cdots + 2^{n-1}i_n$, for all $i_1, i_2, \dots, i_n \in \Sigma$. Verify that, modulo the map $\Phi_3$, one has $g_{1,3} = (1\ 5)(2\ 6) \times (3\ 7)(4\ 8)$, $g_{2,3} = (1\ 3)(2\ 4)$, and $g_{3,3} = (1\ 2)$.

**6.19.** Let $G$ be the Grigorchuk group.

(a) Consider the homomorphism $\Psi_3 \colon G \to \mathrm{Sym}(8)$ defined by $\Psi_3(g) = g|_{\Sigma^3}$ (observe that the map $\Psi_3$ is well defined since $\ell(g(w)) = \ell(w)$ for all $w \in \Sigma^*$ and $g \in G$). With the notation from Exercise 6.18(f), verify that $\Psi_3(a) = (1\ 5)(2\ 6)(3\ 7)(4\ 8)$, $\Psi_3(b) = (1\ 3)(2\ 4)(5\ 6)$, $\Psi_3(c) = (1\ 3)(2\ 4)$ and $\Psi_3(d) = (5\ 6)$.

(b) Deduce from (a) and Exercise 6.18(e) that $\Psi_3(G) = K_3$, where $K_3 \subset \mathrm{Sym}(8)$ is the Kaloujnine group (cf. Exercise 6.18).

(c) By applying Proposition 6.9.3, deduce from (b) and Exercise 6.18(e) that $G/[G, G]$ (the *abelianization* of $G$) is isomorphic to $(\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})$.

**6.20.** *The word problem for the Grigorchuk group.* Let $G$ be the Grigorchuk group. Describe an algorithm which, given any word $w \in \{a, b, c, d\}^*$, determines whether $w$ represents the identity element $1_G$ or not. Hint: Given a word $w$, first count the number $\ell_a(w)$ of occurrences of the letter $a$ in $w$. Prove that if $\ell_a(w)$ is odd then $w$ does not represent $1_G$. If $\ell_a(w)$ is even, use the maps $\phi_0, \phi_1 \colon H_1 \to G$ and apply induction on the length of the word $w$.

**6.21.** Let us say that a net $(x_i)_{i \in I}$ in a set $X$ *converges to infinity* if, for every finite subset $F \subset X$, there exists an element $i_0 \in I$ such that $x_i \notin F$ for all $i \geq i_0$. Let $\varphi \colon G \to H$ be a map from a group $G$ into a group $H$. Show that $\varphi$ is a quasi-isometric embedding if and only if it satisfies the following condition: for any two nets $(u_i)_{i \in I}$ and $(v_i)_{i \in I}$ in $G$ having the same index set, the net $(u_i^{-1} v_i)_{i \in I}$ converges to infinity in $G$ if and only if the net $(\varphi(u_i)^{-1}\varphi(v_i))_{i \in I}$ converges to infinity in $H$.

**6.22.** Let $G$ be a group and let $E(G)$ denote the set consisting of all quasi-isometries $\varphi \colon G \to G$. Define a binary relation $\sim$ in $E(G)$ by declaring that $\varphi_1$ and $\varphi_2 \in E(G)$ satisfy $\varphi_1 \sim \varphi_2$ if and only if there exists a finite subset $F \subset G$ such that $\varphi_1(g)^{-1}\varphi_2(g) \in F$ for all $g \in G$.

(a) Show that $\sim$ is an equivalence relation in $E(G)$.

(b) Show that if $\varphi \in E(G)$ then there exists $\psi \in E(G)$ such that $\varphi \circ \psi \sim \mathrm{Id}_G$ and $\psi \circ \varphi \sim \mathrm{Id}_G$.

(b) Show that the composition of maps in $E(G)$ is compatible with $\sim$ and induces a group structure on the quotient set $\mathrm{QI}(G) = E(G)/\sim$.

(c) Show that if $G$ is a finite group then the group $\mathrm{QI}(G)$ is trivial.

(d) Show that the group $\mathrm{QI}(\mathbb{Z})$ contains a subgroup isomorphic to $\mathbb{R}\times\mathbb{Z}/2\mathbb{Z}$ and is therefore uncountable. Hint: Prove that the map $\varphi_\lambda\colon \mathbb{Z}\to\mathbb{Z}$ defined by $\varphi_\lambda(x) = [\lambda x]$, where $\lambda \in \mathbb{R}\setminus\{0\}$ and $[\alpha]$ denotes the *integral part* of $\alpha$, that is, the largest integer $n$ such that $n \le \alpha$, is a quasi-isometry and that one has $f_{\lambda_1} \sim f_{\lambda_2}$ if and only if $\lambda_1 = \lambda_2$.

(e) Show that if $G$ and $H$ are quasi-isometric groups then the groups $\mathrm{QI}(G)$ and $\mathrm{QI}(H)$ are isomorphic.

**6.23.** Let $A$ be a set and let $\mathcal{G} = (Q, E)$ be a finite $A$-labeled graph. Denote by $\lambda\colon E\to A$ the labeling map defined by $\lambda(e) = a$ for every edge $e = (q, a, q') \in E$. A *bi-infinite path* in $\mathcal{G}$ is a sequence $\pi = (e_n)_{n\in\mathbb{Z}}$ of edges $e_n = (q_n, a_n, q'_n) \in E$ such that $q'_n = q_{n+1}$ for all $n \in \mathbb{Z}$. We define the label of a bi-infinite path $\pi = (e_n)_{n\in\mathbb{Z}}$ as being the element $\lambda(\pi) \in A^{\mathbb{Z}}$ given by $\lambda(\pi)(n) = \lambda(e_n)$ for all $n \in \mathbb{Z}$.

(a) Denote by $X^{\mathcal{G}}$ the set of the labels $\lambda(\pi)$ of all bi-infinite paths $\pi$ in $\mathcal{G}$. Show that $X^{\mathcal{G}}$ is a subshift of $A^{\mathbb{Z}}$. It is called the subshift defined by the $A$-labeled graph $\mathcal{G}$.

(b) Show that if $\mathcal{G}$ is connected then the subshift $X^{\mathcal{G}}$ is irreducible.

**6.24.** Let $A = \{0, 1\}$ and consider the $A$-labeled graphs $\mathcal{G}_1$ and in $\mathcal{G}_2$ in Fig. 6.18. Check that the associated subshifts $X^{\mathcal{G}_1}$ and $X^{\mathcal{G}_2}$ are the even subshift (cf. Exercise 1.38) and the golden mean subshift (cf. Exercise 1.39) respectively.

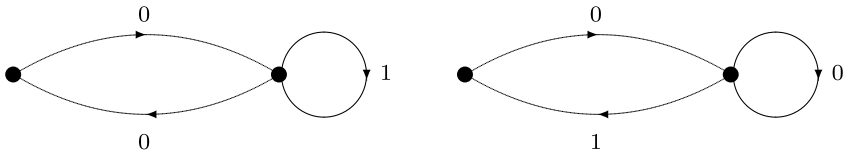



**Fig. 6.18** The $A$-labeled graphs $\mathcal{G}_1$ and $\mathcal{G}_2$

**6.25.** Let $A$ be a set and let $X \subset A^{\mathbb{Z}}$ be a subshift of finite type. Let $M$ be a positive integer such that $\{1, 2, \ldots, M\} \subset \mathbb{Z}$ is a memory set for $X$. Consider the $A$-labeled graph $\mathcal{G} = \mathcal{G}(X, M) = (Q, E)$ defined as follows: $Q = L_{M-1}(X)$ is the set of all $X$-admissible words of length $M-1$, and the edge set $E$ consists of all triples $e = (aw, a, wa') \in Q\times A\times Q$, where $a, a' \in A$ and $w \in A^{M-2}$ are such that $awa' \in L_M(X)$.

(a) Check that $X^{\mathcal{G}} = X$.

(b) Show that $X$ is irreducible if and only if $\mathcal{G}$ is connected (cf. Exercise 6.23).

**6.26.** Let $A$ be a set and let $X \subset A^{\mathbb{Z}}$ be a subshift of finite type. Let also $\tau\colon A^{\mathbb{Z}}\to A^{\mathbb{Z}}$ be a cellular automaton. Let $M$ be a positive integer such that $\{1, 2, \ldots, M\} \subset \mathbb{Z}$ is a memory set for both $X$ and $\tau$, and let $\mu\colon A^M\to A$ denote the corresponding local defining map for $\tau$. Consider the $A$-labeled

graph $\mathcal{G} = \mathcal{G}(X, \tau, M) = (Q, E)$, where $Q = L_{M-1}(X)$ is the set of all $X$-admissible words of length $M-1$ and the edge set $E$ consists of all triples $e = (aw, \mu(awa'), wa') \in Q \times A \times Q$, where $a, a' \in A$ and $w \in A^{M-2}$ are such that $awa' \in L_M(X)$. Check that $X^{\mathcal{G}} = \tau(X)$.

**6.27.** *Life on* $\mathbb{Z}$. Let $A = \{0, 1\}$ and consider the cellular automaton $\tau\colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ defined in Exercise 5.3. Let $\mathcal{G}$ be the $A$-labeled graph in Fig. 6.19. Check that $X^{\mathcal{G}} = \tau(A^G)$.

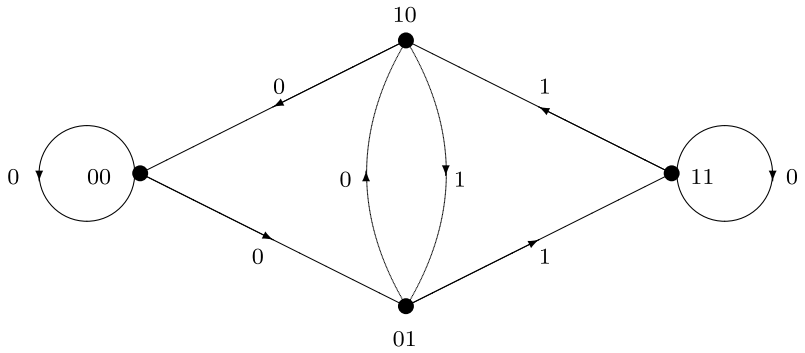



**Fig. 6.19** The $A$-labeled graph $\mathcal{G} = \mathcal{G}(A, \tau, 3)$

**6.28.** Let $A = \{0, 1\}$ and let $\tau\colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ be the majority action cellular automaton (cf. Example 1.4.3(c)) associated with the set $S = \{-1, 0, 1\}$. Let $\mathcal{G}'$ be the $A$-labeled graph in Fig. 6.20. Check that $X^{\mathcal{G}'} = \tau(A^G)$.

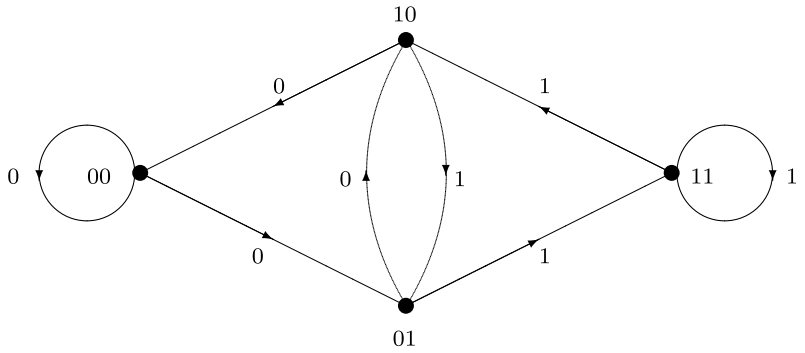



**Fig. 6.20** The $A$-labeled graph $\mathcal{G}' = \mathcal{G}(A, \tau, 3)$

**6.29.** Let $A$ be a set and let $\mathcal{G} = (Q, E)$ be a finite $A$-labeled graph.

(a) Suppose that for each pair $(a, a') \in A^2$ there exists at most one vertex $q \in Q$ which is both the terminal vertex of an edge with label $a$ and the

initial vertex of an edge with label $a'$. Show that the subshift $X^{\mathcal{G}} \subset A^{\mathbb{Z}}$ is of finite type. Hint: Show that in fact $\{0,1\} \subset \mathbb{Z}$ is a memory set for $X^{\mathcal{G}}$.

(b) Show that the hypothesis in (a) is satisfied if the labeling map $\lambda\colon E \to A$ is injective.

**6.30.** Let $A$ be a set and let $\mathcal{G} = (Q, E)$ be a finite $A$-labeled graph. We recall that given an edge $e = (q, a, q') \in E$ we denote by $\alpha(e) = q \in Q$ (resp. $\omega(e) = q' \in Q$) the initial (resp. terminal) vertex of $e$. Consider the $E$-labeled graph $\mathcal{G}' = (Q', E')$, where $Q' = Q$ and $E' = \{(\alpha(e), e, \omega(e)) \in Q \times E \times Q : e \in E\}$. Note that the labeling map $\lambda'\colon E' \to E$ on $\mathcal{G}'$ is bijective. Identify the set $\{\lambda(e) : e \in E\} \subset A$ formed by the labels of the edges of $\mathcal{G}$ with a subset of $E$ in an arbitrary way and consider the cellular automaton $\tau\colon E^{\mathbb{Z}} \to E^{\mathbb{Z}}$ with memory set $S = \{0\}$ and local defining map $\mu\colon E^S \to E$ defined by $\mu(e) = \lambda(e)$ for all $e \in E^S = E$. Observe that $\tau(x) \in A^{\mathbb{Z}}$ for all $x \in E^{\mathbb{Z}}$ and show that $\tau(X^{\mathcal{G}'}) = X^{\mathcal{G}}$.

**6.31.** Let $A$ be a set and $X \subset A^{\mathbb{Z}}$ a subshift. Show that the following conditions are equivalent:

(i) there exists a finite $A$-labeled graph $\mathcal{G}$ such that $X = X^{\mathcal{G}}$.

(ii) there exists a set $B$ containing $A$, a subshift $Y \subset B^{\mathbb{Z}}$ of finite type and a cellular automaton $\tau\colon B^{\mathbb{Z}} \to B^{\mathbb{Z}}$ such that $X = \tau(Y)$. A subshift $X \subset A^{\mathbb{Z}}$ is said to be *sofic* if it satisfies one of the two above equivalent conditions. Hint: For the implication (i) $\Rightarrow$ (ii) use Exercise 6.30. For the converse implication, use Exercise 6.25.

**6.32.** Let $A$ be a set.

(a) Show that every subshift $X \subset A^{\mathbb{Z}}$ of finite type is sofic.

(b) Suppose that $A$ has at least two distinct elements. Show that there exists a subshift $X \subset A^{\mathbb{Z}}$ which is sofic but not of finite type. Hint: The even subshift is sofic (cf. Exercise 6.24) but not of finite type (cf. Exercise 1.38(c)).

**6.33.** Let $A$ be a set. Let $X \subset A^G$ be a sofic subshift and let $\tau\colon A^G \to A^G$ be a cellular automaton. Show that $\tau(X) \subset A^G$ is a sofic subshift.

**6.34.** *A subshift which is not sofic* (cf. [LiM, Example 3.1.7]). Let $A = \{0,1,2\}$ and consider the subset $X \subset A^{\mathbb{Z}}$ consisting of all configurations $x \in \mathbb{Z}$ such that if $x(n) = 0$, $x(n+1) = x(n+2) = \cdots = x(n+h) = 1$, $x(n+h+1) = x(n+h+2) = \cdots = x(n+h+k) = 2$ and $x(n+h+k+1) = 0$ for some $n \in \mathbb{Z}$ and $h, k \in \mathbb{N}$, then necessarily $h = k$.

(a) Show that $X$ is a subshift of $A^{\mathbb{Z}}$. It is called the *context-free subshift*.

(b) Show that $X$ is not sofic. Hint: Suppose by contradiction that $X = X^{\mathcal{G}}$ for some finite $A$-labeled graph $\mathcal{G} = (Q, E)$. Let $r = |Q|$. Observe that $w = 01^{r+1}2^{r+1}0 \in L(X)$ so that there exists a path $\pi$ in $\mathcal{G}$ such that $\lambda(\pi) = w$. Let $\pi'$ denote the subpath of $\pi$ such that $\lambda(\pi') = 1^{r+1}$. Since $\ell(\pi') = r+1 > r = |Q|$, we can write $\pi' = \pi_1\pi_2\pi_3$ where $\pi_2$ is a closed path of length $s = \ell(\pi_2) > 0$ (and $\pi_1$ (resp. $\pi_3$) is a possibly empty path). It follows that $\pi'' = \pi_1\pi_2\pi_2\pi_3$ is a path in $\mathcal{G}$ and its label is $\lambda(\pi'') = 01^{r+1+s}2^{r+1}0 \notin L(X)$.

**6.35.** Let $A$ be a finite set and let $X \subset A^{\mathbb{Z}}$ be a sofic subshift. A finite $A$-labeled graph $\mathcal{G} = (Q, E)$ is said to be a *minimal presentation* of $X$ if (i) $X = X^{\mathcal{G}}$, and (ii) $|Q| \leq |Q'|$ for all $A$-labeled graphs $\mathcal{G}' = (Q', E')$ such that $X = X^{\mathcal{G}'}$. Suppose that $X$ is irreducible and that $\mathcal{G} = (Q, E)$ is a minimal presentation of $X$.

(a) Show that for every vertex $q \in Q$ there exists a word $w_q \in L(X)$ such that if a path $\pi$ in $\mathcal{G}$ satisfies $\lambda(\pi) = w_q$, then it passes through $q$. Hint: By contradiction, if this is not the case for some $q \in Q$ then the $A$-labeled graph $\mathcal{G}' = (Q', E')$ where $Q' = Q \setminus \{q\}$ and $E' = E \setminus \{e \in E : \alpha(e) = q \text{ or } \omega(e) = q\}$ satisfies $X^{\mathcal{G}'} = X$ and $|Q'| < |Q|$, contradicting the minimality of $\mathcal{G}$.

(b) Deduce from (a) that $\mathcal{G}$ is connected. Hint: Let $q, q' \in Q$. Consider the words $w_q, w_{q'} \in L(X)$ described in (a). Then by irreducibility of $X$ there exists $u \in L(X)$ such that $w_q u w_{q'} \in L(X)$. Let $\pi$ be a path in $\mathcal{G}$ such that $\lambda(\pi) = w_q u w_{q'}$; then $\pi = \pi_1 \pi_2 \pi_3$ where $\lambda(\pi_1) = w_q$, $\lambda(\pi_2) = u$ and $\lambda(\pi_3) = w_{q'}$. By definition of $w_q$ (resp. $w_{q'}$) the path $\pi_1$ (resp. $\pi_3$) passes through $q$ (resp. $q'$), say $\pi_1 = \pi_1' \pi_1''$ with $(\pi_1')^+ = q = (\pi_1'')^-$ (resp. $\pi_3 = \pi_3' \pi_3''$ with $(\pi_3')^+ = q' = (\pi_3'')^-$). Then the path $\pi_1'' \pi_2 \pi_3'$ connects $q$ to $q'$.

(c) Deduce from (b) that for every irreducible sofic subshift $Y \subset A^{\mathbb{Z}}$ there exists a connected finite $A$-labeled graph $\mathcal{G}$ such that $Y = X^{\mathcal{G}}$.

**6.36.** Let $A$ be a finite set and $X \subset A^{\mathbb{Z}}$ an irreducible sofic subshift.

(a) Show that the subset $X_f$ consisting of all configurations in $X$ whose $\mathbb{Z}$-orbit is finite (cf. Example 1.3.1(c)) is dense in $X$. Hint: Let $x \in X$ and let $n \in \mathbb{N}$. By virtue of Exercise 6.35, we can find a connected $A$-labeled graph $\mathcal{G} = (Q, E)$ such that $X = X^{\mathcal{G}}$. Let $\pi_1$ be a finite path in $\mathcal{G}$ such that $\lambda(\pi_1) = x(0)x(1)\cdots x(n-1)$ and set $q = \pi_1^-$ and $q' = \pi_1^+$. Since $\mathcal{G}$ is connected, we can find a finite path $\pi_2$ in $\mathcal{G}$ connecting $q'$ to $q$. Let $m = \ell(\pi_2)$. It follows that the path $\pi = \pi_1 \pi_2$ is closed and $t = \ell(\pi) = n + m$. If $w = \lambda(\pi)$ we deduce that $w^k \in L(X)$ for all $k \in \mathbb{N}$. Since $X$ is closed, there exists a configuration $y \in X$ such that $y(ht)y(ht+1)\cdots y((h+1)t-1) = w$ for all $h \in \mathbb{Z}$. In particular, $y(0)y(1)\cdots y(n-1) = x(0)x(1)\cdots x(n-1)$ and $y \in X_f$.

(b) Deduce from (a) that $X$ is surjunctive. Hint: Cf. Exercise 3.29.

**6.37.** Show that the Morse subshift is not sofic. Hint: An infinite minimal subshift is irreducible (cf. Exercise 3.35(b)) and contains no configuration whose $\mathbb{Z}$-orbit is finite (cf. Exercise 3.35(d)). On the other hand, by Exercise 6.36(a), every irreducible sofic subshift contains an abundance of configurations with finite $\mathbb{Z}$-orbit.

**6.38.** Let $A$ be a finite set and let $X \subset A^{\mathbb{Z}}$ be an infinite Toeplitz subshift. Show that $X$ is not sofic. Hint: The same arguments as for Exercise 6.37 apply.

**6.39.** Let $A$ be a finite set. Show that there are at most countably many distinct sofic subshifts $X \subset A^{\mathbb{Z}}$. Hint: There are at most countably many finite $A$-labeled graphs up to isomorphism.

**6.40.** Let $I$ be a finite set. Let $B = (b_{ij})_{i,j\in I}$ be a matrix with real entries $b_{ij} \geq 0$ for all $i, j \in I$. For an integer $n \geq 1$, we denote by $B^n = (b_{ij}^{(n)})_{i,j\in I}$ the $n$-th power of $B$. One says that the matrix $B$ is *irreducible* if for every $i, j \in I$ there exists an integer $n = n(i,j) \geq 1$ such that $b_{ij}^{(n)} > 0$. Suppose that $B$ is irreducible. The *period* $\mathrm{per}(i)$ of $i \in I$ is the greatest common divisor of the integers $n \geq 1$ such that $b_{ii}^{(n)} > 0$.

(a) Show that $\mathrm{per}(i) = \mathrm{per}(j)$ for all $i, j \in I$. The *period* $\mathrm{per}(B)$ of the matrix $B$ is defined as the common value of the numbers $\mathrm{per}(i)$, $i \in I$. Hint: Let $i, j \in I$. We can find $r, s \geq 1$ such that $b_{ij}^{(r)} > 0$ and $b_{ji}^{(s)} > 0$. It follows that $b_{ii}^{(r+s)} \geq b_{ij}^{(r)} b_{ji}^{(s)} > 0$ and $b_{ii}^{(r+n+s)} \geq b_{ij}^{(r)} b_{jj}^{(n)} b_{ji}^{(s)} > 0$ for all $n \geq 1$ such that $b_{jj}^{(n)} > 0$. It follows from the definition that $\mathrm{per}(i)$ divides both $r+s$ and $r+n+s$ and therefore also divides their difference $n$. This shows that $\mathrm{per}(i)$ divides $\mathrm{per}(j)$.

(b) Let $n \geq 1$ be an integer. Show that $B^n$ is irreducible if and only if $n$ and $\mathrm{per}(B)$ are relatively prime.

**6.41.** Let $A$ be a set and let $\mathcal{G} = (Q, E)$ be a finite $A$-labeled graph. The *adjacency matrix* of $\mathcal{G}$ is the $(Q \times Q)$-matrix $B_{\mathcal{G}} = (b_{qq'})_{q,q'\in Q}$ where $b_{qq'}$ is the number of edges in $E$ with initial vertex $q$ and terminal vertex $q'$.

(a) Show that $\mathcal{G}$ is connected if and only if the matrix $B_{\mathcal{G}}$ is irreducible.

(b) Suppose that $\mathcal{G}$ is connected. The *period* of $\mathcal{G}$ is the positive integer defined by $\mathrm{per}(\mathcal{G}) = \mathrm{per}(B_{\mathcal{G}})$. Show that $\mathrm{per}(\mathcal{G})$ is the greatest common divisor of the lengths of all closed paths in $\mathcal{G}$.

**6.42.** Let $A$ be a set. Given a subshift $X \subset A^{\mathbb{Z}}$ denote by $\mathrm{per}_n(X) = |\mathrm{Fix}(n\mathbb{Z}) \cap X|$ (cf. Example 1.3.1(c)) the number of $n\mathbb{Z}$-periodic configurations in $X$. Let $\mathcal{G}$ be a finite connected $A$-labeled graph.

(a) Show that for every integer $N$ there exists an integer $n \geq N$ such that $\mathrm{per}_n(X^{\mathcal{G}}) > 0$.

(b) The *period* $\mathrm{per}(X^{\mathcal{G}})$ of $X^{\mathcal{G}}$ is defined as the greatest common divisor of all integers $n \geq 1$ for which $\mathrm{per}_n(X) > 0$. Show that $\mathrm{per}(X^{\mathcal{G}}) = \mathrm{per}(\mathcal{G})$.

**6.43.** *The $N$th higher block subshift* (cf. Exercise 1.15 and Exercise 1.34). Let $A$ be a set and $N$ a positive integer. Consider the map $\Phi_N \colon A^{\mathbb{Z}} \to (A^N)^{\mathbb{Z}}$ defined by $\Phi_N(x)(n) = (x(n), x(n+1), x(n+2), \ldots, x(n+N-1))$ for all $x \in A^{\mathbb{Z}}$ and $n \in \mathbb{Z}$. Similarly, consider the map $\varphi_N \colon A^* \to (A^N)^*$ defined as follows: $\varphi_N(w) = \varepsilon$ (the empty word) if $\ell(w) < N$ and

$$\varphi(w) = (a_1, a_2, \ldots, a_N)(a_2, a_3, \ldots, a_{N+1}) \cdots (a_{m+1}, a_{m+2}, \ldots, a_{m+N})$$

if $w = a_1 a_2 \cdots a_{m+N}$, with $m \geq 0$. Let $X \subset A^{\mathbb{Z}}$ be a subshift and set $X^{[N]} = \Phi_N(X) \subset (A^N)^{\mathbb{Z}}$. Let $L \subset A^*$ be a subset and set $L^{[N]} = \varphi_N(L) \subset (A^N)^*$.

(a) Show that $X^{[N]}$ is a subshift of $(A^N)^{\mathbb{Z}}$ (it is called the $N$th *higher block subshift* of $X$) and that $L(X^{[N]}) = (L(X))^{[N]}$.

(b) Show that $X$ is irreducible (resp. topologically mixing, resp. strongly irreducible) if and only if $X^{[N]}$ is irreducible (resp. topologically mixing, resp. strongly irreducible).

(c) Show that if $X$ is of finite type then $X^{[N]}$ is of finite type.

(d) Let $F_n = \{0, 1, 2, \ldots, n-1\} \subset \mathbb{Z}$ and let $\mathcal{F} = (F_n)_{n\in\mathbb{N}}$. Show that $\mathrm{ent}_{\mathcal{F}}(X) = \mathrm{ent}_{\mathcal{F}}(X^{[N]})$.

**6.44.** Let $A$ be a set and let $\mathcal{G} = (Q, E)$ be an $A$-labeled graph. For $N$ a positive integer, the *$N$-higher edge graph* $\mathcal{G}^{[N]}$ associated with $\mathcal{G}$ is the $A^N$-labeled graph $(Q^{[N]}, E^{[N]})$ defined as follows. For $N = 1$ one has $\mathcal{G}^{[1]} = \mathcal{G}$ and, for $N \geq 2$, the vertex set $Q^{[N]}$ is the set of all paths of length $N-1$ in $\mathcal{G}$ and

$$E^{[N]} = \{(\pi, \lambda(e_1)\lambda(e_2)\cdots\lambda(e_N), \pi') \in Q^{[N]} \times A^N \times Q^{[N]} : \\ \pi = (e_1, e_2, \ldots, e_{N-1}), \pi' = (e_2, e_3, \ldots, e_{N-1}, e_N)\}.$$

(a) Show that $X^{\mathcal{G}^{[N]}} = (X^{\mathcal{G}})^{[N]}$.

(b) Deduce from (b) that if $X \subset A^{\mathbb{Z}}$ is a sofic subshift, then the subshift $X^{[N]} \subset (A^N)^{\mathbb{Z}}$ is also sofic.

**6.45.** *The $N$th higher power subshift* (cf. Exercise 1.16 and Exercise 1.35). Let $A$ be a set and $N$ a positive integer. Consider the map $\Psi_N \colon A^{\mathbb{Z}} \to (A^N)^{\mathbb{Z}}$ defined by $\Psi_N(x)(n) = (x(nN), x(nN+1), \ldots, x((n+1)N-1))$ for all $x \in A^{\mathbb{Z}}$ and $n \in \mathbb{Z}$. Similarly, consider the map $\psi_N \colon A^* \to (A^N)^*$ defined as follows: $\psi_N(w) = \varepsilon$ if $w \in A^* \setminus (\cup_{n\geq 1} A^{nN})$ and

$$\psi_N(w) = (a_1, a_2, \ldots, a_N)(a_{N+1}, a_{N+2}, \ldots, a_{2N}) \\ \cdots (a_{(n-1)N+1}, a_{(n-1)N+2}, \ldots, a_{nN})$$

if $w = a_1 a_2 \cdots a_{nN}$ for some $n \geq 1$. Let $X \subset A^{\mathbb{Z}}$ be a subshift and set $X^{(N)} = \Psi_N(X) \subset (A^N)^{\mathbb{Z}}$. Let $L \subset A^*$ be a subset and set $L^{(N)} = \psi_N(L) \subset (A^N)^*$.

(a) Show that $X^{(N)}$ is a subshift of $(A^N)^{\mathbb{Z}}$ (it is called the $N$th *higher power subshift* of $X$) and that $L(X^{(N)}) = (L(X))^{(N)}$.

(b) Show that if $X$ is of finite type then $X^{(N)}$ is of finite type.

(c) Suppose that $X$ is irreducible. Show that $X^{(N)}$ is irreducible if and only if $N$ and $\mathrm{per}(X)$ are relatively prime.

(d) Let $F_n = \{0, 1, 2, \ldots, n-1\} \subset \mathbb{Z}$ and let $\mathcal{F} = (F_n)_{n\in\mathbb{N}}$. Show that $\mathrm{ent}_{\mathcal{F}}(X) = \mathrm{ent}_{\mathcal{F}}(X^{(N)})$.

**6.46.** Let $A$ be a set and let $\mathcal{G} = (Q, E)$ be an $A$-labeled graph. For $N$ a positive integer, the *$N$-higher power graph* $\mathcal{G}^{(N)}$ associated with $\mathcal{G}$ is the $A^N$-labeled graph $(Q^{(N)}, E^{(N)})$ defined as follows. The vertex set is $Q^{(N)} = Q$ and the edge set is

$$E^{(N)} = \{(\pi^-, \lambda(\pi), \pi^+) \in Q \times A^N \times Q : \pi \text{ a path of length } N \text{ in } \mathcal{G}\},$$

where $\pi^- \in Q$ (resp. $\pi^+ \in Q$) is the initial (resp. terminal) vertex of the path $\pi$. Recall that given a labeled graph $\mathcal{H}$, we denote by $B_{\mathcal{H}}$ its adjacency matrix (cf. Exercise 6.40).

(a) Show that $B_{\mathcal{G}^{(N)}} = (B_{\mathcal{G}})^N$.

(b) Show that $X^{\mathcal{G}^{(N)}} = (X^{\mathcal{G}})^{(N)}$.

(c) Deduce from (b) that if $X \subset A^{\mathbb{Z}}$ is a sofic subshift, then the subshift $X^{(N)} \subset (A^N)^{\mathbb{Z}}$ is also sofic.

**6.47.** (cf. [Sca, Lemma]) Let $\mathcal{G} = (Q, E)$ be a finite labeled graph. Suppose that $\mathcal{G}$ is connected and let $e \in E$. Show that there exists a positive integer $n_0$ such that if $\pi$ is any path in $\mathcal{G}$ of length $n_0$, then there exists a path $\pi' = (e_1, e_2, \ldots, e_{n_0})$ of the same length $n_0$, with the same initial and terminal vertices as $\pi$, and such that $e_i = e$ for some $1 \leq i \leq n_0$. Hint: Given a path $\pi = (e_1, e_2, \ldots, e_n)$ in $\mathcal{G}$, denote by $\pi_Q = (q_0, q_1, \ldots, q_n)$ the associated sequence of visited vertices (cf. Sect. 6.2). We define a decomposition of $\pi$ as follows. Let $i_1$ be the largest index such that the vertices $q_0, q_1, \ldots, q_{i_1-1}$ are all distinct. Then $q_{i_1} = q_{j_1}$ for a suitable $j_1 < i_1$ and we set $r_1 = (e_1, e_2, \ldots, e_{j_1})$ and $c_1 = (e_{j_1+1}, e_{j_1+2}, \ldots, e_{i_1})$. Continuing this way, we obtain a decomposition of the path $\pi = r_1 c_1 r_2 c_2 \cdots r_k c_k r_{k+1}$ where the $c_1, c_2, \ldots, c_k$ are closed simple paths and $r_1, r_2, \ldots, r_{k+1}$ are simple (possibly empty) paths. With this notation, for $1 \leq s \leq r$ and a positive integer $d$, we say that the path $\pi_s = r_1 c_1 r_2 c_2 r_s r_{s+1} \cdots c_k r_k$ is obtained from $\pi$ by *collapsing* the $s$th closed simple path $c_s$. Similarly, if $\pi''$ is a closed path such that $(\pi'')^- = \pi^-$ and $d$ is a positive integer, we say that the path $(\pi'')^d \pi$ is obtained from $\pi$ by *adding* $d$ copies of $\pi''$ at the beginning of $\pi$. Now, since $\mathcal{G}$ is connected, we can find a closed path $\pi''$ with initial (= terminal) vertex $\pi^-$ containing the edge $e$. If $n_0$ is large enough then in the decomposition $\pi = r_1 c_1 r_2 c_2 \cdots r_k c_k r_{k+1}$ there exists a cycle $c$ that is repeated many times. If the length of $c$ is $\ell$, the length of $\pi''$ is $m$ and the cycle $c$ is repeated at least $m$ times, then we may collapse the first $m$ copies of $c$ and then add $\ell$ copies of $\pi''$ at the beginning of $\pi$ to obtain the desired path $\pi'$.

**6.48.** Let $\mathcal{G} = (Q, E)$ be a finite labeled graph. We denote by $P_n(\mathcal{G})$ the set of all paths of length $n$ in $\mathcal{G}$ and we define the *entropy* of $\mathcal{G}$ as

$$\operatorname{ent}(\mathcal{G}) = \lim_{n \to \infty} \frac{\log |P_n(\mathcal{G})|}{n}.$$

Show that the above limit exists and is finite. Hint: Use Lemma 6.5.1.

**6.49.** (cf. [Sca, Theorem]) Let $\mathcal{G} = (Q, E)$ be a finite connected labeled graph. Let $e \in E$ and denote by $\mathcal{H} = (Q', E')$ the labeled subgraph of $\mathcal{G}$ were $Q' = Q$ and $E' = E \setminus \{e\}$. Let $n_0$ be the positive integer given by Exercise 6.47 and set $\alpha = |P_{n_0}(\mathcal{G})|^{-1}$.

(a) Show that $|P_{(k-1)n_0}(\mathcal{G})| \geq \alpha |P_{kn_0}(\mathcal{G})|$ for all $k = 1, 2, \ldots$.

(b) We express a path $\pi \in P_{kn_0}(\mathcal{G})$ as the composition $\pi = \pi_1 \pi_2 \cdots \pi_k$, where $\pi_i \in P_{n_0}(\mathcal{G})$, $1 \leq i \leq k$. Also, for $i = 1, 2, \ldots, k$, we denote by $\varphi_i$ the set

of all paths $\pi \in P_{kn_0}(\mathcal{G})$ such that $\pi_i$ contains the edge $e$. Show that $|\varphi_1| \geq |P_{(k-1)n_0}(\mathcal{G})|$ and deduce from (a) that $|P_{kn_0}(\mathcal{G}) \setminus \varphi_1| \leq (1-\alpha)|P_{n_0}(\mathcal{G})|$. Hint: By Exercise 6.47 for any $\tilde{\pi} \in P_{(k-1)n_0}(\mathcal{G})$ there exits a path $\pi' \tilde{\pi} \in P_{kn_0}(\mathcal{G})$ such that $\pi'$ contains $e$.

(c) For $1 \leq i \leq k$ we set $P^i_{kn_0}(\mathcal{G}) = P_{kn_0}(\mathcal{G}) \setminus \bigcup_{t=1}^{i-1} \varphi_t$, $C^h = \{\pi \in P^i_{kn_0}(\mathcal{G}) : \pi_i \text{ contains } e\}$ and denote by $D^h$ the set of pairs $(\sigma_1, \sigma_2) \in P_{(i-1)n_0}(\mathcal{G}) \times P_{(k-i)n_0}(\mathcal{G})$ such that there exists $\pi = \sigma_1 \sigma \sigma_2 \in P^i_{kn_0}(\mathcal{G})$. Show that $|C^h| \geq |D^h| \geq \alpha |P^i_{kn_0}(\mathcal{G})|$. Hint: Use Exercise 6.47 to show that for every $(\sigma_1, \sigma_2) \in D^i$ there exists $\pi = \sigma_1 \pi' \sigma_2 \in P_{kn_0}(\mathcal{G})$ such that $\pi'$ contains $e$.

(d) Deduce from (c) that $|P_{kn_0}(\mathcal{G}) \setminus \bigcup_{i=1}^{k} \varphi_i| \leq (1-\alpha)^k |P_{kn_0}(\mathcal{G})|$.

(e) Observe that $P_{kn_0}(\mathcal{H}) \subset P_{kn_0}(\mathcal{G}) \setminus \bigcup_{i=1}^{k} \varphi_i$ and deduce from (d) that $\mathrm{ent}(\mathcal{H}) \leq \frac{\log(1-\alpha)}{n_0} + \mathrm{ent}(\mathcal{G})$.

(f) Deduce from (e) that $\mathrm{ent}(\mathcal{H}) < \mathrm{ent}(\mathcal{G})$.

**6.50.** Let $A$ be a finite set. Let $F_n = \{0, 1, \ldots, n\} \subset \mathbb{Z}$ and $\mathcal{F} = (F_n)_{n \in \mathbb{N}}$. Let $X \subset A^{\mathbb{Z}}$ be a subshift and let $w \in L(X)$. Set

$$X_w = \{x \in X : x(n+1)x(n+2) \cdots x(n+N) \neq w \text{ for all } n \in \mathbb{Z}\},$$

where $N = \ell(w)$.

(a) Show that $X_w \subset X$ is a subshift of $A^{\mathbb{Z}}$.

(b) Show that $(X_w)^{[N]} = (X^{[N]})_{\Phi_N(w)}$ (cf. Exercise 6.43).

(c) Suppose from now on that $X$ is irreducible of finite type. Let $M$ be a positive integer such that $\{1, 2, \ldots, M\}$ is a memory set for $X$, and let $\mathcal{G}$ be the $A$-labeled graph such that $X = X(\mathcal{G}, M)$ (cf. Exercise 6.25). Let $\mathcal{H}$ denote the labeled subgraph of $\mathcal{G}^{[N]}$ obtained by removing all edges labeled by $\Phi_N(w)$. Show that $(X^{[N]})_{\Phi_N(w)} = X_{\mathcal{H}}$.

(d) Show that $\mathrm{ent}_{\mathcal{F}}(X_w) < \mathrm{ent}_{\mathcal{F}}(X)$. Hint: Use Exercises 6.49 and 6.43(d).

**6.51.** Let $A$ be a finite set. Let $F_n = \{0, 1, \ldots, n\} \subset \mathbb{Z}$ and $\mathcal{F} = (F_n)_{n \in \mathbb{N}}$. Let $X \subset A^{\mathbb{Z}}$ be an irreducible subshift of finite type. Show that if $Y \subset A^{\mathbb{Z}}$ is such that $Y \subsetneq X$, then $\mathrm{ent}_{\mathcal{F}}(Y) < \mathrm{ent}_{\mathcal{F}}(X)$. Hint: Show that there exists $w \in L(X) \setminus L(Y)$ such that $Y \subset X_w \subset X$. Then apply Exercise 6.50(d) and Proposition 5.7.2(ii).

**6.52.** (cf. [Hed-3], [CovP], and [LiM, Theorem 8.1.16]). Let $A$ be a finite set. Let $\mathcal{F} = (F_n)_{n \in \mathbb{N}}$, where $F_n = \{0, 1, \ldots, n\} \subset \mathbb{Z}$. Also let $X \subset A^{\mathbb{Z}}$ an irreducible subshift of finite type and $\tau \colon X \to X$ a pre-injective cellular automaton.

(a) Let $M$ be a positive integer such that $\{1, 2, \ldots, M\}$ is a memory set for $X$. Set $Y = \tau(X)$ and consider the labeled graph $\mathcal{G} = \mathcal{G}(X, \tau, M)$ (cf. Exercise 6.26). Note that for every $w \in L(Y)$ there exists a path $\pi \in \mathcal{G}$ such that $w = \lambda(\pi)$. Show that, by pre-injectivity of $\tau$, for all $q, q' \in Q$ and $w \in L(Y)$ there exists at most one path $\pi$ in $\mathcal{G}$ with initial (resp. terminal) vertex $\pi^- = q$ (resp. $\pi^+ = q'$) such that $w = \lambda(\pi)$.

(b) Deduce from (a) that $|L_n(Y)| \leq |L_n(X)| \leq |L_n(Y)| \cdot |Q|^2$.

(c) Deduce from (b) that $\mathrm{ent}_{\mathcal{F}}(\tau(X)) = \mathrm{ent}_{\mathcal{F}}(X)$.

**6.53.** Let $A$ be a finite set. Let $F_n = \{0, 1, \dots, n\} \subset \mathbb{Z}$ and $\mathcal{F} = (F_n)_{n\in\mathbb{N}}$. Let $X, Y \subset A^{\mathbb{Z}}$ be two irreducible subshifts of finite type such that $\mathrm{ent}_{\mathcal{F}}(X) = \mathrm{ent}_{\mathcal{F}}(Y)$. Show that every pre-injective cellular automaton $\tau\colon X \to Y$ is surjective. Hint: Use the results of Exercise 6.51 and Exercise 6.52.

**6.54.** Let $A$ be a finite set. Let $X \subset A^G$ be an irreducible subshift of finite type. Show that every pre-injective cellular automaton $\tau\colon X \to X$ is surjective. Hint: Use Exercise 6.53 with $Y = X$.

**6.55.** (cf. [Hed-3], [CovP], and [LiM, Theorem 8.1.16]). Let $A$ be a finite set and $X \subset A^{\mathbb{Z}}$ an irreducible subshift of finite type. Let $F_n = \{0, 1, \dots, n\} \subset \mathbb{Z}$ and $\mathcal{F} = (F_n)_{n\in\mathbb{N}}$. Let $\tau\colon X \to X$ be a cellular automaton. Suppose that $\tau$ is not pre-injective.

(a) Show that there exist a positive integer $N$ and two distinct configurations $x_1, x_2 \in X$ such that $x_1|_{\mathbb{Z}\setminus\{1,2,\dots,N\}} = x_2|_{\mathbb{Z}\setminus\{1,2,\dots,N\}}$ and $\tau(x_1) = \tau(x_2)$.

(b) Up to enlarging $N$ if necessary, we may suppose that $\{1, 2, \dots, N\}$ is a memory set for both $X$ and $\tau$. Let $\mathcal{G} = \mathcal{G}(X, \tau, N) = (Q, E)$ (cf. Exercise 6.26). Show that there exist two vertices $q, q' \in Q$ and two paths $\pi_1, \pi_2$ in $\mathcal{G}$ with initial (resp. terminal) vertices $\pi_1^- = \pi_2^- = q$ (resp. $\pi_1^+ = \pi_2^+ = q'$) such that $\lambda(\pi_1) = x_1|_{\{1,2,\dots,N\}}$ and $\lambda(\pi_2) = x_2|_{\{1,2,\dots,N\}}$. One says that the two paths $\pi_1, \pi_2$ constitute a *diamond* in $\mathcal{G}$ (see Fig. 6.21).

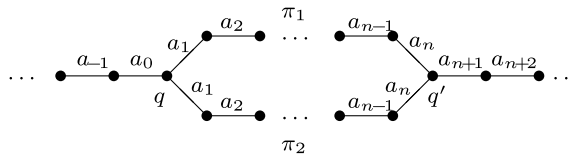



**Fig. 6.21** A diamond in an $A$-labeled graph $\mathcal{G}$

(c) Up to further enlarging $N$ if necessary, we may suppose that $N$ is relatively prime with $\mathrm{per}(X)$ (cf. Exercise 6.40 and Exercise 6.42). By Exercise 6.45(c) the $N$th power graph $\mathcal{G}^{(N)}$ is also irreducible and $\pi_1$ and $\pi_2$ are edges in $\mathcal{G}^{(N)}$ with the same initial vertex and same terminal vertex. Let $\mathcal{H}$ be the $A$-labeled subgraph of $\mathcal{G}^{(N)}$ obtained by removing the edge $\pi_1$. Deduce from Exercise 6.49 that $\mathrm{ent}_{\mathcal{F}}(X_{\mathcal{H}}) < \mathrm{ent}_{\mathcal{F}}(X_{\mathcal{G}^{(N)}})$.

(d) Show that $X_{\mathcal{H}} = Y^{(N)}$, where $Y = \tau(X)$.

(e) Deduce from (c) and (d) and from Exercise 6.45(d) that $\mathrm{ent}_{\mathcal{F}}(\tau(X)) < \mathrm{ent}_{\mathcal{F}}(X)$.

**6.56.** Let $A$ be a finite set. Let $F_n = \{0, 1, \dots, n\} \subset \mathbb{Z}$ and $\mathcal{F} = (F_n)_{n\in\mathbb{N}}$. Let $X, Y \subset A^{\mathbb{Z}}$ be two subshifts such that $X$ is irreducible of finite type and $\mathrm{ent}_{\mathcal{F}}(X) = \mathrm{ent}_{\mathcal{F}}(Y)$. Show that every surjective cellular automaton $\tau\colon X \to Y$ is pre-injective. Hint: Use the result of Exercise 6.55.

**6.57.** *The Garden of Eden theorem for irreducible subshifts of finite type over $\mathbb{Z}$ [Fio1, Corollary 2.19].* Let $A$ be a finite set and $X \subset A^{\mathbb{Z}}$ an irreducible

subshift of finite type. Let $\tau\colon X \to X$ be a cellular automaton. Show that $\tau$ is surjective if and only if it is pre-injective. Hint: Combine together the results from Exercise 6.56 with $Y = X$ and Exercise 6.54.

**6.58.** Let $A = \{0, 1\}$. Consider the subset $X \subset A^{\mathbb{Z}}$ consisting of all configurations $x \in A^{\mathbb{Z}}$ such that $\{n \in \mathbb{Z} : x(n) = 1\}$ is an interval of $\mathbb{Z}$. Let $\sigma\colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ be the cellular automaton with memory set $S = \{0, 1\}$ and local defining map $\mu\colon A^S \to A$ given by

$$\mu(y) = \begin{cases} 1 & \text{if } (y(0), y(1)) = (0, 1), \\ y(0) & \text{otherwise.} \end{cases}$$

(a) Show that $X$ is a sofic subshift.
(b) Show that $X$ is neither of finite type nor irreducible.
(c) Check that $\tau(X) \subset X$
(d) Show that the cellular automaton $\tau = \sigma|_X\colon X \to X$ is injective (and therefore pre-injective) but not surjective.
(e) Deduce from (d) that $X$ is not surjunctive.

# Chapter 7
# Local Embeddability and Sofic Groups

In this chapter we study the notions of local embeddability and soficity for groups. Roughly speaking, a group is locally embeddable into a given class of groups provided that the multiplicative table of any finite subset of the group is the same as the multiplicative table of a subset of some group in the class (cf. Definition 7.1.3). In Sect. 7.1 we discuss several stability properties of local embeddability. Subgroups of locally embeddable groups are locally embeddable (Proposition 7.1.7). Moreover, as the name suggests, local embeddability is a local property, that is, a group is locally embeddable into a class of groups if and only if all its finitely generated subgroups are locally embeddable into the class (Proposition 7.1.8). When the class is closed under finite direct products, the class of groups which are locally embeddable in the class is closed under (possibly infinite) direct products (Proposition 7.1.10). We also show that a marked group which is a limit of groups belonging to a given class is locally embeddable into this class (Theorem 7.1.16). Conversely, we prove that if the given class is closed under taking subgroups and the marking group is free, then any marked group which is locally embeddable is a limit of marked groups which are in the class (Theorem 7.1.19). If $\mathcal{C}$ is a class of groups which is closed under finite direct products, then every group which is residually $\mathcal{C}$ is locally embeddable into $\mathcal{C}$ (Corollary 7.1.14). Conversely, under the hypothesis that $\mathcal{C}$ is closed under taking subgroups, every finitely presented group which is locally embeddable into $\mathcal{C}$ is residually $\mathcal{C}$ (Corollary 7.1.21).

In Sect. 7.2 we present a characterization of local embeddability in terms of ultraproducts: a group is locally embeddable into $\mathcal{C}$ if and only if it can be embedded into an ultraproduct of a family of groups in $\mathcal{C}$.

Section 7.3 is devoted to LEF and LEA-groups. A group is called LEF (resp. LEA) if it is locally embeddable into the class of finite (resp. amenable) groups. As the class of finite (resp. amenable) groups is closed under taking subgroups and taking finite direct products, all the results obtained in the previous section can be applied. This implies in particular that every locally residually finite (resp. locally residually amenable) group is LEF (resp. LEA).

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_7, 

We give an example of a finitely generated amenable group which is LEF but not residually finite (Proposition 7.3.9). This group is not finitely presentable since every finitely presented LEF-group is residually finite.

The Hamming metric, which is a bi-invariant metric on the symmetric group of a finite set, is introduced in Sect. 7.4.

In Sect. 7.5 we define the class of sofic groups. These groups are the groups admitting local approximations by finite symmetric groups equipped with their Hamming metric. Subgroups and direct products of sofic groups are sofic. Every LEA-group is sofic (Corollary 7.5.11). In particular, residually amenable groups, and therefore amenable groups and residually finite groups, are sofic. A group is sofic if and only if it can be embedded into an ultraproduct of a family of finite symmetric groups equipped with their Hamming metrics (Theorem 7.6.6). In Sect. 7.7 we give a characterization of finitely generated sofic groups in terms of their Cayley graphs. More precisely, we show that a finitely generated group $G$ with a finite symmetric generating subset $S \subset G$ is sofic if and only if, for every integer $r \geq 0$ and every $\varepsilon > 0$, there exists a finite $S$-labeled graph $Q$ such that there is a proportion of at least $1 - \varepsilon$ of vertices $q \in Q$ such that the ball of radius $r$ centered at $q$ in $Q$ is isomorphic, as a labeled graph, to a ball of radius $r$ in the Cayley graph of $G$ associated with $S$ (Theorem 7.7.1). The last section of this chapter is devoted to the proof of the surjunctivity of sofic groups (Theorem 7.8.1).

## 7.1 Local Embeddability

**Definition 7.1.1.** Let $G$ and $C$ be two groups. Given a finite subset $K \subset G$, a map $\varphi\colon G \to C$ is called a *$K$-almost-homomorphism* of $G$ into $C$ if it satisfies the following conditions:

($K$-AH-1) $\varphi(k_1k_2) = \varphi(k_1)\varphi(k_2)$ for all $k_1, k_2 \in K$;
($K$-AH-2) the restriction of $\varphi$ to $K$ is injective.

Note that in the preceding definition, the map $\varphi$ is not required to be a homomorphism nor to be globally injective.

*Remark 7.1.2.* If $\varphi\colon G \to C$ is a $K$-almost-homomorphism and $\varphi'\colon G \to C$ is a map which coincides with $\varphi$ on $K \cup K^2$, then $\varphi'$ is also a $K$-almost-homomorphism.

Let now $\mathcal{C}$ be a *class of groups*, that is, a collection of groups satisfying the following condition: if $C \in \mathcal{C}$ and $C'$ is a group which is isomorphic to $C$, then $C' \in \mathcal{C}$. For example, $\mathcal{C}$ might be the class of finite groups, the class of nilpotent groups, the class of solvable groups, the class of amenable groups, or the class of free groups.

**Definition 7.1.3.** Let $\mathcal{C}$ be a class of groups. One says that a group $G$ is *locally embeddable* into the class $\mathcal{C}$ if, for every finite subset $K \subset G$, there exist a group $C \in \mathcal{C}$ and a $K$-almost-homomorphism $\varphi\colon G \to C$.

*Examples 7.1.4.* (a) Let $\mathcal{C}$ be a class of groups and let $G$ be a group in $\mathcal{C}$. Then $G$ is locally embeddable into $\mathcal{C}$. Indeed, the identity map $\mathrm{Id}_G\colon G \to G$ is a $K$-almost-homomorphism of $G$ into itself for every finite subset $K \subset G$.

(b) The group $\mathbb{Z}$ is locally embeddable into the class of finite groups. Indeed, let $K$ be a finite subset of $\mathbb{Z}$. Choose an integer $n \geq 0$ such that $K \subset [-n, n]$. Then, the quotient homomorphism $\varphi\colon \mathbb{Z} \to \mathbb{Z}/(2n+1)\mathbb{Z}$ is a $K$-almost-homomorphism.

*Remark 7.1.5.* Let $G$ be a group and let $\mathcal{C}$ be a class of groups which is closed under taking subgroups. Suppose that $G$ is finite and it is locally embeddable into $\mathcal{C}$. Then $G \in \mathcal{C}$. Indeed, any $G$-almost homomorphism $\varphi\colon G \to C$ from $G$ into a group $C \in \mathcal{C}$ is an injective homomorphism.

The following characterization of local embeddability will not be used in the sequel but it presents some interest on its own.

**Proposition 7.1.6.** *Let $G$ be a group and let $\mathcal{C}$ be a class of groups which is closed under taking subgroups. Then the following conditions are equivalent:*

(a) *$G$ is locally embeddable into $\mathcal{C}$;*
(b) *for every finite subset $K \subset G$, there exist a set $L$ with $K \subset L \subset G$ and a binary operation $\odot\colon L \times L \to L$ such that $(L, \odot)$ is a group in $\mathcal{C}$ and $k_1 k_2 = k_1 \odot k_2$ for all $k_1, k_2 \in K$.*

*Proof.* Suppose (a). Fix a finite subset $K \subset G$. If $G$ is finite, then $G \in \mathcal{C}$ by Remark 7.1.5. Therefore, condition (b) is satisfied in this case since we can take as $(L, \odot)$ the group $G$ itself. So let us assume that $G$ is infinite. Consider the finite subset $K' \subset G$ defined by $K' = K \cup K^2$. Since $G$ is locally embeddable into $\mathcal{C}$, we can find a group $C' \in \mathcal{C}$ and a $K'$-almost-homomorphism $\varphi'\colon G \to C'$. Note that, as $K'$ is finite, the subgroup $C \subset C'$ generated by $\varphi'(K')$ is countable. Since $G$ is infinite, it follows that there exists a surjective map $\sigma\colon G \to C$ which coincides with $\varphi'$ on $K'$. Let us define a map $\psi\colon C \to G$ as follows. If $c \in \sigma(K')$, we set $\psi(c) = k'$, where $k'$ is the unique element in $K'$ such that $\sigma(k') = c$. For $c \in C \setminus \sigma(K')$, we take as $\psi(c)$ an arbitrary element in $\sigma^{-1}(c)$. Denote by $L = \psi(C) \subset G$ the image of $\psi$. Observe that $K \subset K' \subset L$ and that $\psi$ induces a bijection from $C$ onto $L$. Let us use this bijection to transport the group structure from $C$ to $L$. If we denote by $\odot$ the corresponding group operation on $L$, this means that $\psi$ induces a group isomorphism from $C$ onto $(L, \odot)$. It follows that we have

$$\ell_1 \odot \ell_2 = \psi(\sigma(\ell_1)\sigma(\ell_2))$$

for all $\ell_1, \ell_2 \in L$. As $\mathcal{C}$ is closed under taking subgroups, $C$ and hence $(L, \odot)$ belong to $\mathcal{C}$. Moreover, for all $k_1, k_2 \in K$, we have

$$k_1 \odot k_2 = \psi(\sigma(k_1)\sigma(k_2)) = \psi(\varphi'(k_1)\varphi'(k_2)) = \psi(\varphi'(k_1k_2)) = \psi(\sigma(k_1k_2)).$$

Since $K^2 \subset K'$, we deduce that $k_1 \odot k_2 = k_1k_2$ for all $k_1, k_2 \in K$. This shows that condition (b) is satisfied.

Conversely, suppose (b). Fix a finite subset $K \subset G$. Set $C = (L, \odot)$, where $L$ is as in (b). Let $\varphi\colon G \to C$ be any map extending the identity map $\iota\colon L \to L$. Then $\varphi|_K = \iota|_K$ is injective. On the other hand, for all $k_1, k_2 \in K$, we have $k_1k_2 = k_1 \odot k_2 \in L$ and therefore

$$\begin{aligned}\varphi(k_1k_2) &= \iota(k_1k_2)\\ &= k_1k_2\\ &= k_1 \odot k_2\\ &= \varphi(k_1) \odot \varphi(k_2)\end{aligned}$$

Thus, $\varphi$ is a $K$-almost-homomorphism of $G$ into the group $C$ in $\mathcal{C}$. This shows that $G$ is locally embeddable into $\mathcal{C}$. □

**Proposition 7.1.7.** *Let $\mathcal{C}$ be a class of groups. Every subgroup of a group which is locally embeddable into $\mathcal{C}$ is locally embeddable into $\mathcal{C}$.*

*Proof.* Let $G$ be a group which is locally embeddable into $\mathcal{C}$ and let $H$ be a subgroup of $G$. Given a finite subset $K \subset H$, let $\varphi\colon G \to C$ be a $K$-almost-homomorphism of $G$ into a group $C \in \mathcal{C}$. Then the restriction map $\varphi|_H\colon H \to C$ is a $K$-almost-homomorphism of $H$ into $C$. This shows that $H$ is locally embeddable into $\mathcal{C}$. □

As the name suggests, local embeddability is a local property for groups:

**Proposition 7.1.8.** *Let $G$ be a group and let $\mathcal{C}$ be a class of groups. Then $G$ is locally embeddable into $\mathcal{C}$ if and only if every finitely generated subgroup of $G$ is locally embeddable into $\mathcal{C}$.*

*Proof.* The necessity of the condition follows from Proposition 7.1.7. Conversely, suppose that every finitely generated subgroup of $G$ is locally embeddable into $\mathcal{C}$. Let $K \subset G$ be a finite subset and denote by $H \subset G$ the subgroup generated by $K$. Then, as $H$ is locally embeddable into $\mathcal{C}$, there exists a $K$-almost-homomorphism $\varphi\colon H \to C$ of $H$ into a group $C \in \mathcal{C}$. Extend arbitrarily $\varphi$ to $G$, for example by setting $\varphi(g) = 1_C$ for all $g \in G \setminus H$. Then $\varphi\colon G \to C$ is a $K$-almost-homomorphism of $G$ into $C$. This shows that $G$ is locally embeddable into $\mathcal{C}$. □

Let $\mathcal{C}$ be a class of groups. Denote by $\overline{\mathcal{C}}$ the class consisting of all groups which are locally embeddable into $\mathcal{C}$.

**Proposition 7.1.9.** *Let $\mathcal{C}$ be a class of groups. Then $\overline{\mathcal{C}} = \overline{\overline{\mathcal{C}}}$.*

*Proof.* Let $G$ be a group in $\overline{\overline{\mathcal{C}}}$. Then, given any finite subset $K \subset G$ there exist a group $C' \in \overline{\mathcal{C}}$ and a $K$-almost-homomorphism $\varphi\colon G \to C'$. Consider the finite subset $K' = \varphi(K) \subset C'$ and let $\varphi'\colon C' \to C$ be a $K'$-almost-homomorphism of $C'$ into a group $C \in \mathcal{C}$. Then the composition map $\Phi = \varphi' \circ \varphi\colon G \to C$ is a $K$-almost-homomorphism. Indeed, for $k_1, k_2 \in K$ one has

$$\begin{aligned}\Phi(k_1k_2) &= \varphi'(\varphi(k_1k_2))\\ &= \varphi'(\varphi(k_1)\varphi(k_2))\\ &= \varphi'(\varphi(k_1))\varphi'(\varphi(k_2))\\ &= \Phi(k_1)\Phi(k_2).\end{aligned}$$

Moreover, as $\varphi|_K$ and $\varphi'|_{K'}$ are injective, and $K' = \varphi(K)$, one has that $\Phi|_K$ is also injective. It follows that $G$ is locally embeddable into $\mathcal{C}$. This shows that $\overline{\overline{\mathcal{C}}} \subset \overline{\mathcal{C}}$. The inclusion $\overline{\mathcal{C}} \subset \overline{\overline{\mathcal{C}}}$ follows from the observation in Example 7.1.4(a). This shows that $\overline{\mathcal{C}} = \overline{\overline{\mathcal{C}}}$. □

A class of groups $\mathcal{C}$ is said to be *closed under finite direct products* if one has $G_1 \times G_2 \in \mathcal{C}$ whenever $G_1 \in \mathcal{C}$ and $G_2 \in \mathcal{C}$.

**Proposition 7.1.10.** *Let $\mathcal{C}$ be a class of groups which is closed under finite direct products. Let $(G_i)_{i\in I}$ be a family of groups which are locally embeddable into $\mathcal{C}$. Then, their direct product $G = \prod_{i\in I} G_i$ is locally embeddable into $\mathcal{C}$.*

*Proof.* For each $i \in I$, let $\pi_i\colon G \to G_i$ denote the projection homomorphism. Fix a finite subset $K \subset G$. Then there exists a finite subset $J \subset I$ such that the projection homomorphism $\pi_J = \prod_{j\in J} \pi_j\colon G \to G_J = \prod_{j\in J} G_j$ is injective on $K$. Since the group $G_j$ is locally embeddable into $\mathcal{C}$, we can find, for each $j \in J$, a $\pi_j(K)$-almost homomorphism $\varphi_j\colon G_j \to C_j$ of $G_j$ into a group $C_j \in \mathcal{C}$. As the class $\mathcal{C}$ is closed under finite direct products, the group $C = \prod_{j\in J} C_j$ is also in $\mathcal{C}$. Consider the map $\varphi\colon G \to C$ defined by $\varphi = \varphi_J \circ \pi_J$, where $\varphi_J = \prod_{j\in J} \varphi_j\colon G_J \to C$. For all $k = (k_i)_{i\in I}, k' = (k'_i)_{i\in I} \in K$, we have

$$\begin{aligned}\varphi(kk') &= \varphi_J(\pi_J(kk'))\\ &= \varphi_J((k_jk'_j)_{j\in J})\\ &= (\varphi_j(k_jk'_j))_{j\in J}\\ &= (\varphi_j(k_j)\varphi_j(k'_j))_{j\in J}\\ &= (\varphi_j(k_j))_{j\in J}(\varphi_j(k'_j))_{j\in J}\\ &= \varphi(k)\varphi(k').\end{aligned}$$

Moreover, $\varphi|_K$ is injective. Indeed, given $k = (k_i)_{i\in I}, k' = (k'_i)_{i\in I} \in K$, if $\varphi(k) = \varphi(k')$ then $\varphi_j(k_j) = \varphi_j(k'_j)$ for all $j \in J$. By injectivity of $\varphi_j|_{\pi_j(K)}$, we deduce that $k_j = k'_j$ for all $j \in J$. This implies that $k = k'$ since $\pi_J|_K$ is injective. This shows that $\varphi$ is a $K$-almost-homomorphism of $G$ into $C$. It follows that $G$ is locally embeddable into $\mathcal{C}$. □

**Corollary 7.1.11.** *Let $\mathcal{C}$ be a class of groups which is closed under finite direct products. Let $(G_i)_{i\in I}$ be a family of groups which are locally embeddable into $\mathcal{C}$. Then their direct sum $G = \bigoplus_{i\in I} G_i$ is locally embeddable into $\mathcal{C}$.*

*Proof.* This follows immediately from Proposition 7.1.7 and Proposition 7.1.10, since $G$ is the subgroup of the direct product $P = \prod_{i\in I} G_i$ consisting of all $g = (g_i)_{i\in I} \in P$ for which $g_i = 1_{G_i}$ for all but finitely many $i \in I$. □

**Corollary 7.1.12.** *Let $\mathcal{C}$ be a class of groups which is closed under finite direct products. If a group $G$ is the limit of a projective system of groups which are locally embeddable into $\mathcal{C}$, then $G$ is locally embeddable into $\mathcal{C}$.*

*Proof.* Let $(G_i)_{i\in I}$ be a projective system of groups which are locally embeddable into $\mathcal{C}$ such that $G = \varprojlim G_i$. By construction of a projective limit (see Appendix E), $G$ is a subgroup of the group $\prod_{i\in I} G_i$. We deduce that $G$ is locally embeddable into $\mathcal{C}$ by using Proposition 7.1.10 and Proposition 7.1.7. □

Recall that if $\mathcal{C}$ is a class of groups, then a group $G$ is called *residually $\mathcal{C}$* if for each element $g \in G$ with $g \neq 1_G$, there exist a group $C \in \mathcal{C}$ and a surjective homomorphism $\phi\colon G \to C$ such that $\phi(g) \neq 1_C$.

**Proposition 7.1.13.** *Let $\mathcal{C}$ be a class of groups which is closed under finite direct products. Let $G$ be a group which is residually $\mathcal{C}$. Then, for every finite subset $K \subset G$, there exist a group $C \in \mathcal{C}$ and a homomorphism $\varphi\colon G \to C$ whose restriction to $K$ is injective.*

*Proof.* Fix a finite subset $K \subset G$ and consider the set

$$L = \{hk^{-1} : h, k \in K \text{ and } h \neq k\}.$$

Since $G$ is residually $\mathcal{C}$, we can find, for each $g \in L$, a group $C_g \in \mathcal{C}$ and a homomorphism $\phi_g\colon G \to C_g$ such that $\phi_g(g) \neq 1_{C_g}$. The group $C = \prod_{g\in L} C_g$ is in the class $\mathcal{C}$ since $L$ is finite and $\mathcal{C}$ is closed under finite direct products. Consider the group homomorphism $\varphi = \prod_{g\in L} \phi_g\colon G \to C$. If $h$ and $k$ are distinct elements in $K$, then $g = hk^{-1} \in L$ and $\phi_g(hk^{-1}) = \phi_g(g) \neq 1_{C_g}$. It follows that $\varphi(hk^{-1}) \neq 1_C$ and hence $\varphi(h) \neq \varphi(k)$. Thus, the restriction of $\varphi$ to $K$ is injective. □

**Corollary 7.1.14.** *Let $\mathcal{C}$ be a class of groups which is closed under finite direct products. Then every group which is residually $\mathcal{C}$ is locally embeddable into $\mathcal{C}$.*

*Proof.* Let $G$ be a group which is residually $\mathcal{C}$ and let $K \subset G$ be a finite subset. By Proposition 7.1.13, there exist a group $C \in \mathcal{C}$ and a homomorphism $\varphi\colon G \to C$ whose restriction to $K$ is injective. Such a $\varphi$ is a $K$-almost-homomorphism. Consequently, $G$ is locally embeddable into $\mathcal{C}$. □

**Corollary 7.1.15.** *Let $\mathcal{C}$ be a class of groups which is closed under finite direct products. Then every group which is locally residually $\mathcal{C}$ is locally embeddable into $\mathcal{C}$.*

*Proof.* This immediately follows from Corollary 7.1.14 and Proposition 7.1.8. □

Given a group $\Gamma$, let $\mathcal{N}(\Gamma)$ denote the space of all $\Gamma$-marked groups (cf. Sect. 3.4). Recall that $\mathcal{N}(\Gamma)$ may be identified with the set consisting of all normal subgroups of $\Gamma$ and that $\mathcal{N}(\Gamma)$ is a compact Hausdorff space for the topology induced by the prodiscrete topology on $\mathcal{P}(\Gamma) = \{0,1\}^{\Gamma}$.

**Theorem 7.1.16.** *Let $\Gamma$ be a group and let $\mathcal{C}$ be a class of groups. Let $N \in \mathcal{N}(\Gamma)$. Suppose that there exists a net $(N_i)_{i\in I}$ which converges to $N$ in $\mathcal{N}(\Gamma)$ such that $\Gamma/N_i \in \mathcal{C}$ for all $i \in I$. Then the group $\Gamma/N$ is locally embeddable into $\mathcal{C}$.*

*Proof.* Denote by $\rho\colon \Gamma \to \Gamma/N$ and $\rho_i\colon \Gamma \to \Gamma/N_i$ the quotient homomorphisms. Fix a finite subset $K \subset \Gamma/N$ and let $F \subset \Gamma$ be a finite symmetric subset such that $1_\Gamma \in F$ and $\rho(F) = K \cup K^{-1} \cup \{1_{\Gamma/N}\}$. Since the net $(N_i)_{i\in I}$ converges to $N$, we can find $i_0 \in I$ such that

$$N \cap F^4 = N_{i_0} \cap F^4. \tag{7.1}$$

Set $C = \Gamma/N_{i_0}$ and define a map $\varphi\colon \Gamma/N \to C$ by setting

$$\varphi(g) = \begin{cases} \rho_{i_0}(f') & \text{if } g = \rho(f') \text{ for some } f' \in F^2 \\ 1_C & \text{otherwise.} \end{cases} \tag{7.2}$$

Note that $\varphi$ is well defined. Indeed, suppose that $g = \rho(f_1') = \rho(f_2')$, for some $f_1', f_2' \in F^2$. Then, from $1_{\Gamma/N} = \rho(f_1')^{-1}\rho(f_2') = \rho((f_1')^{-1}f_2')$, we deduce that

$$(f_1')^{-1}f_2' \in \ker(\rho) = N. \tag{7.3}$$

As $(f_1')^{-1}f_2' \in F^4$, we deduce from (7.3) and (7.1) that $(f_1')^{-1}f_2' \in N_{i_0} = \ker(\rho_{i_0})$. Therefore, we have $\rho_{i_0}(f_1') = \rho_{i_0}(f_2')$.

Let us check now that $\varphi$ is a $K$-almost-homomorphism. Let $k_1, k_2 \in K$. Then, there exist $f_1, f_2 \in F$ such that $\rho(f_1) = k_1$ and $\rho(f_2) = k_2$. Observe that $f_1, f_2 \in F^2$ since $F \subset F^2$. Thus, we get $\varphi(k_1) = \rho_{i_0}(f_1)$ and $\varphi(k_2) = \rho_{i_0}(f_2)$ by applying (7.2). On the other hand, we also have $f_1 f_2 \in F^2$ and $k_1 k_2 = \rho(f_1)\rho(f_2) = \rho(f_1 f_2)$. Therefore, by applying again (7.2), we obtain

$$\begin{aligned} \varphi(k_1 k_2) &= \rho_{i_0}(f_1 f_2) \\ &= \rho_{i_0}(f_1)\rho_{i_0}(f_2) \\ &= \varphi(k_1)\varphi(k_2). \end{aligned}$$

Moreover, if $\varphi(k_1) = \varphi(k_2)$, then $\rho_{i_0}(f_1) = \rho_{i_0}(f_2)$ and hence $f_1^{-1}f_2 \in \ker(\rho_{i_0}) = N_{i_0}$. As $f_1^{-1}f_2 \in F^2 \subset F^4$, we deduce from (7.1) that $f_1^{-1}f_2 \in$

$N = \ker(\rho)$. Therefore $k_1 = \rho(f_1) = \rho(f_2) = k_2$. This shows that $\varphi$ is injective on $K$. We have shown that $\varphi$ is a $K$-almost-homomorphism of $\Gamma/N$ into the group $C \in \mathcal{C}$. Therefore, $\Gamma/N$ is locally embeddable into $\mathcal{C}$. □

**Corollary 7.1.17.** *Let $\Gamma$ be a group and let $\mathcal{C}$ be a class of groups. Then the set of all $N \in \mathcal{N}(\Gamma)$ such that $\Gamma/N$ is locally embeddable into $\mathcal{C}$ is closed (and hence compact) in $\mathcal{N}(\Gamma)$.*

*Proof.* Let $(N_i)_{i\in I}$ be a convergent net in $\mathcal{N}(\Gamma)$ and suppose that the groups $\Gamma/N_i$ are locally embeddable into $\mathcal{C}$ for all $i \in I$. Let $N = \lim_{i\in I} N_i$. By applying Theorem 7.1.16, we deduce that $\Gamma/N$ is locally embeddable into the class of groups which are locally embeddable into $\mathcal{C}$. It follows from Proposition 7.1.9 that $\Gamma/N$ is locally embeddable into $\mathcal{C}$. □

For the next theorem we need some (slightly technical) preliminaries.

Let $n \geq 2$ be an integer and $G, C$ two groups. Given a finite subset $K \subset G$, we say that a map $\varphi\colon G \to C$ is an *$n$-$K$-almost-homomorphism* if it satisfies the following conditions:

($n$-$K$-AH-1) $\varphi(k_1^{\varepsilon_1} k_2^{\varepsilon_2} \cdots k_t^{\varepsilon_t}) = \varphi(k_1)^{\varepsilon_1}\varphi(k_2)^{\varepsilon_2}\cdots\varphi(k_t)^{\varepsilon_t}$ for all $k_i \in K \cup \{1_G\}$, $\varepsilon_i \in \{-1, 1\}$, $1 \leq i \leq t \leq n$;
($n$-$K$-AH-2) $\varphi(1_G) = 1_C$;
($n$-$K$-AH-3) the restriction of $\varphi$ to $(K \cup K^{-1} \cup \{1_G\})^n$ is injective.

**Lemma 7.1.18.** *Let $G$ be a group and let $\mathcal{C}$ be a class of groups. Let $n \geq 2$ be an integer. Then, the following conditions are equivalent:*

(a) *$G$ is locally embeddable into $\mathcal{C}$;*
(b) *for every finite subset $K \subset G$, there exist a group $C \in \mathcal{C}$ and an $n$-$K$-almost-homomorphism $\varphi\colon G \to C$.*

*Proof.* It is clear that any $n$-$K$-almost-homomorphism is also a $K$-almost-homomorphism (take $\varepsilon_1 = \varepsilon_2 = 1$ and $t = 2$ in ($n$-$K$-AH-1), and observe that $K \subset (K \cup K^{-1} \cup \{1_G\})^n$). Thus, (b) implies (a).

Conversely, suppose (a). Given a finite subset $K \subset G$, set $K' = (K \cup K^{-1} \cup \{1_G\})^n$. As $G$ is locally embeddable into $\mathcal{C}$, there exists a $K'$-almost-homomorphism $\varphi\colon G \to C$ of $G$ into a group $C \in \mathcal{C}$. Let us show that $\varphi$ is an $n$-$K$-almost-homomorphism. As the restriction of $\varphi$ to $K'$ is injective, Property ($n$-$K$-AH-3) is trivially satisfied. On the other hand, we have $\varphi(k_1k_2) = \varphi(k_1)\varphi(k_2)$ for all $k_1, k_2 \in K'$. For $k_1 = k_2 = 1_G$, this gives us $\varphi(1_G) = \varphi(1_G)^2$, so that Property ($n$-$K$-AH-2) also holds. By taking $k_1 \in K'$ and $k_2 = k_1^{-1}$, we get $1_C = \varphi(1_G) = \varphi(k_1k_1^{-1}) = \varphi(k_1)\varphi(k_1^{-1})$. This implies that $\varphi(k_1^{-1}) = \varphi(k_1)^{-1}$ for all $k_1 \in K'$. Then, Property ($n$-$K$-AH-1) immediately follows by induction on $t$. This shows that $\varphi$ is an $n$-$K$-almost-homomorphism of $G$ into $C$. □

**Theorem 7.1.19.** *Let $F$ be a free group and let $\mathcal{C}$ be a class of groups which is closed under taking subgroups. Let $N \in \mathcal{N}(F)$ and suppose that the group $F/N$ is locally embeddable into $\mathcal{C}$. Then there exists a net $(N_i)_{i\in I}$ which converges to $N$ in $\mathcal{N}(F)$ such that $F/N_i \in \mathcal{C}$ for all $i \in I$.*

*Proof.* Let $X \subset F$ be a free base of $F$ and denote by $\rho\colon F \to F/N$ the quotient homomorphism. Let $I$ denote the set of all finite subsets $E$ of $F$ partially ordered by inclusion. Given $E \in I$, we denote by $U_E \subset X$ the subset of $X$ consisting of all elements which appear in the reduced form of some element in $E$. Since $E$ is finite, $U_E$ is also finite and there exists $n_E \in \mathbb{N}$ such that each $w \in E$ may be written in the form

$$w = x_1^{\varepsilon_1} x_2^{\varepsilon_2} \cdots x_t^{\varepsilon_t} \tag{7.4}$$

for suitable $x_i \in U_E$, $\varepsilon_i \in \{-1, 1\}$, $i = 1, 2, \ldots, t$, and $1 \leq t \leq n_E$. Set $K_E = \rho(B_{n_E}) \subset F/N$, where $B_{n_E}$ denotes the ball of radius $n_E$ centered at the identity element $1_F$ in the Cayley graph of the subgroup of $F$ generated by $U_E$. Since $F/N$ is locally embeddable into $\mathcal{C}$ it follows from Lemma 7.1.18 that there exists a group $C_E \in \mathcal{C}$ and an $n_E$-$K_E$-almost-homomorphism $\varphi_E\colon F/N \to C_E$. As $F$ is a free group with base $X$, there exists a unique homomorphism $\rho_E\colon F \to C_E$ such that $\rho_E(x) = \varphi_E(\rho(x))$ for all $x \in X$. Setting $N_E = \ker(\rho_E)$ we have $N_E \in \mathcal{N}(F)$. Moreover, since $\mathcal{C}$ is closed under taking subgroups, we have that the quotient group $F/N_E$, being isomorphic to $\rho_E(F)$, which is a subgroup of $C_E$, also belongs to $\mathcal{C}$.

Let us show that the net $(N_E)_{E \in I}$ converges to $N$. Fix a finite subset $E_0 \subset F$. Let $E \in I$ be such that $E_0 \subset E$. Let us show that

$$N \cap E_0 = N_E \cap E_0. \tag{7.5}$$

Given $w \in E_0$, we can write $w$ as in (7.4). Using the fact that $\varphi_E$ is an $n_E$-$K_E$-almost-homomorphism, we get

$$\begin{aligned}
\rho_E(w) &= \rho_E(x_1^{\varepsilon_1} x_2^{\varepsilon_2} \cdots x_t^{\varepsilon_t}) \\
&= \rho_E(x_1)^{\varepsilon_1} \rho_E(x_2)^{\varepsilon_2} \cdots \rho_E(x_t)^{\varepsilon_t} \\
&= \varphi_E(\rho(x_1))^{\varepsilon_1} \varphi_E(\rho(x_2))^{\varepsilon_2} \cdots \varphi_E(\rho(x_t))^{\varepsilon_t} \\
&= \varphi_E(\rho(x_1)^{\varepsilon_1} \rho(x_2)^{\varepsilon_2} \cdots \rho(x_t)^{\varepsilon_t}) \\
&= \varphi_E(\rho(x_1^{\varepsilon_1} x_2^{\varepsilon_2} \cdots x_t^{\varepsilon_t})) \\
&= \varphi_E(\rho(w)).
\end{aligned}$$

As $\varphi_E$ is injective on $K_E$, and $\rho(w) \in \rho(B_E) = K_E$, we deduce that $\rho_E(w) = 1_{F/N_E}$ if and only if $\rho(w) = 1_{F/N}$, equivalently, $w \in N_E$ if and only if $w \in N$. Thus (7.5) follows. This shows that $N = \lim_E N_E$. □

From Theorem 7.1.16 (with $\Gamma$ a free group $F$) and Theorem 7.1.19 we immediately deduce the following.

**Corollary 7.1.20.** *Let $F$ be a free group and let $\mathcal{C}$ be a class of groups which is closed under taking subgroups. Let $N \in \mathcal{N}(F)$. Then the following conditions are equivalent.*

(a) *the group $F/N$ is locally embeddable into $\mathcal{C}$;*

(b) *there exists a net* $(N_i)_{i\in I}$ *which converges to* $N$ *in* $\mathcal{N}(F)$ *such that* $F/N_i \in \mathcal{C}$ *for all* $i \in I$. □

**Corollary 7.1.21.** *Let* $\mathcal{C}$ *be a class of groups which is closed under taking subgroups. Then every finitely presented group which is locally embeddable into* $\mathcal{C}$ *is residually* $\mathcal{C}$.

*Proof.* Let $G$ be a finitely presented group which is locally embeddable into $\mathcal{C}$. Fix $g \in G \setminus \{1_G\}$. Since $G$ is finitely presented, there exists a free group $F$ of finite rank and a finite subset $R \subset F$ such that $G = F/N$, where $N \in \mathcal{N}(F)$ denotes the normal closure of $R$ in $F$. By Theorem 7.1.19, there exists a net $(N_i)_{i\in I}$ in $\mathcal{N}(F)$ converging to $N$ such that $F/N_i \in \mathcal{C}$ for all $i \in I$. Let $\rho\colon F \to G$ denote the quotient homomorphism and choose an element $f \in F$ such that $\rho(f) = g$. Since the net $(N_i)_{i\in I}$ converges to $N$, we can find $i_0 \in I$ such that

$$N \cap (R \cup \{f\}) = N_{i_0} \cap (R \cup \{f\}). \tag{7.6}$$

As $R \subset N$, this implies $R \subset N_{i_0}$ and hence $N \subset N_{i_0}$. Thus, there is a canonical epimorphism $\phi\colon G = F/N \to C = F/N_{i_0}$. Observe now that $f \notin N$ and therefore $f \notin N_{i_0}$ by (7.6). It follows that $\phi(g) \neq 1_C$. As $C \in \mathcal{C}$, this shows that $G$ is residually $\mathcal{C}$. □

**Corollary 7.1.22.** *Let* $\mathcal{C}$ *be a class of groups which is closed under taking subgroups and under finite direct products. Then a finitely presented group is locally embeddable into* $\mathcal{C}$ *if and only if it is residually* $\mathcal{C}$.

*Proof.* The fact that every finitely presented group which is locally embeddable into $\mathcal{C}$ is residually $\mathcal{C}$ follows from the previous corollary. The converse implication follows from Corollary 7.1.14. □

We end this section by producing an example showing that Theorem 7.1.19 becomes false if we omit the hypothesis that $F$ is free. Let us first establish the following:

**Proposition 7.1.23.** *Let* $\mathcal{C}$ *be a class of groups. Let* $G$ *be a group. Suppose that there exists a net* $(N_i)_{i\in I}$ *which converges to* $\{1_G\}$ *in* $\mathcal{N}(G)$ *such that* $G/N_i \in \mathcal{C}$ *for all* $i \in I$. *Then* $G$ *is residually* $\mathcal{C}$.

*Proof.* Let $g \in G \setminus \{1_G\}$. Let us set $F = \{1_G, g\}$. Since the net $(N_i)_{i\in I}$ converges to $\{1_G\}$ in $\mathcal{N}(G)$, there exists $i_0 \in I$ such that $N_{i_0} \cap F = \{1_G\} \cap F$. As $F \cap \{1_G\} = \{1_G\}$, this implies $g \notin N_{i_0}$. Thus, the quotient homomorphism $\phi\colon G \to G/N_{i_0}$ satisfies $\phi(g) \neq 1_{G/N_{i_0}}$. As $G/N_{i_0} \in \mathcal{C}$, this shows that $G$ is residually $\mathcal{C}$. □

Now, to exhibit the promised example showing the necessity of the freeness hypothesis on $F$ in Theorem 7.1.19, we consider the additive group $G = \mathbb{Q}$ of rational numbers and take as $\mathcal{C}$ the class of finite groups. It follows from Example 2.1.9 that $\mathbb{Q}$ is not residually finite. Thus, by the preceding propo-

sition, there is no net $(N_i)_{i\in I}$ converging to $\{1_G\}$ in $\mathcal{N}(G)$ such that $G/N_i$ is finite for all $i \in I$. On the other hand, $G$ is locally residually finite since every finitely generated abelian group is residually finite by Corollary 2.2.4. Therefore, $G$ is locally embeddable into $\mathcal{C}$ by Corollary 7.1.15.

## 7.2 Local Embeddability and Ultraproducts

Suppose that we are given a family of groups $(G_i)_{i\in I}$ and a filter $\omega$ (cf. Sect. J.1) on the index set $I$. Consider the group

$$P = \prod_{i\in I} G_i.$$

Let $\alpha = (\alpha_i)_{i\in I}$ and $\beta = (\beta_i)_{i\in I}$ be elements of $P$. We write $\alpha \sim_\omega \beta$ if $\{i \in I : \alpha_i = \beta_i\} \in \omega$.

**Proposition 7.2.1.** *One has:*

(i) $\alpha \sim_\omega \alpha$;
(ii) $\alpha \sim_\omega \beta$ *if and only if* $\beta \sim_\omega \alpha$;
(iii) *if* $\alpha \sim_\omega \beta$ *and* $\beta \sim_\omega \gamma$, *then* $\alpha \sim_\omega \gamma$;
(iv) *if* $\alpha \sim_\omega \beta$ *and* $\gamma \sim_\omega \delta$, *then* $\alpha\gamma \sim_\omega \beta\delta$;
(v) $\alpha \sim_\omega \beta$ *if and only if* $\alpha^{-1} \sim_\omega \beta^{-1}$;

*for all* $\alpha, \beta, \gamma, \delta \in P$.

*Proof.* Let $\alpha = (\alpha_i)_{i\in I}$, $\beta = (\beta_i)_{i\in I}$, $\gamma = (\gamma_i)_{i\in I}$ and $\delta = (\delta_i)_{i\in I} \in P$. We have $\alpha \sim_\omega \alpha$ since $\{i \in I : \alpha_i = \alpha_i\} = I$ belongs to $\omega$ (cf. (F-6) in Sect. J.1). This shows (i). Also, since $\{i \in I : \alpha_i = \beta_i\} = \{i \in I : \beta_i = \alpha_i\}$ we deduce (ii). Suppose now that $\alpha \sim_\omega \beta$ and $\beta \sim_\omega \gamma$. We have $\{i \in I : \alpha_i = \gamma_i\} \supset \{i \in I : \alpha_i = \beta_i\} \cap \{i \in I : \beta_i = \gamma_i\}$. Thus (iii) follows from the fact that $\omega$, being a filter, is closed under finite intersections and taking supersets (cf. (F-2) and (F-3) in Sect. J.1). Suppose now that $\alpha \sim_\omega \beta$ and $\gamma \sim_\omega \delta$, that is, $\{i \in I : \alpha_i = \beta_i\}$ and $\{i \in I : \gamma_i = \delta_i\}$ both belong to $\omega$. We have

$$\{i \in I : \alpha_i\gamma_i = \beta_i\delta_i\} \supset \{i \in I : \alpha_i = \beta_i\} \cap \{i \in I : \gamma_i = \delta_i\}.$$

As $\omega$ is closed under finite intersections and taking supersets, we deduce that $\{i \in I : \alpha_i\gamma_i = \beta_i\delta_i\} \in \omega$, that is, $\alpha\gamma \sim_\omega \beta\delta$. This shows (iv). Finally, from $\{i \in I : \alpha_i = \beta_i\} = \{i \in I : \alpha_i^{-1} = \beta_i^{-1}\}$ we deduce (v). □

Note that it follows from (i), (ii) and (iii) in Proposition 7.2.1 that $\sim_\omega$ is an equivalence relation in $P$. Consider now the subset $N_\omega \subset P$ defined by

$$N_\omega = \{\alpha \in P : \alpha \sim_\omega 1_P = (1_{G_i})_{i\in I}\}.$$

**Proposition 7.2.2.** *The set $N_\omega$ is a normal subgroup of $P$.*

*Proof.* This is an easy consequence of the properties of $\sim_\omega$ stated in Proposition 7.2.1. Indeed, we immediately deduce from Proposition 7.2.1(i) that $1_P \in N_\omega$. Suppose now that $\alpha, \beta \in N_\omega$, that is, $\alpha \sim_\omega 1_P$ and $\beta \sim_\omega 1_P$. It follows from Proposition 7.2.1(iv) that $\alpha\beta \sim_\omega 1_P 1_P = 1_P$, that is, $\alpha\beta \in N_\omega$. Similarly, from Proposition 7.2.1(v) we deduce that $\alpha^{-1} \sim_\omega 1_P^{-1} = 1_P$, that is, $\alpha^{-1} \in N_\omega$. Thus, $N_\omega$ is a subgroup of $P$. Finally, let $\gamma \in P$. Proposition 7.2.1(iv) implies that $\gamma\alpha\gamma^{-1} \sim_\omega \gamma 1_P \gamma^{-1} = 1_P$. It follows that $\gamma\alpha\gamma^{-1} \in N_\omega$ for all $\alpha \in N_\omega$ and $\gamma \in P$. This shows that $N_\omega$ is a normal subgroup of $P$. □

Observe that, given $\alpha$ and $\beta$ in $P$, one has

$$\alpha N_\omega = \beta N_\omega \iff \alpha \sim_\omega \beta. \tag{7.7}$$

Indeed, one has $\alpha N_\omega = \beta N_\omega$ if and only if $\alpha\beta^{-1} \in N_\omega$, that is, if and only if $\alpha\beta^{-1} \sim_\omega 1_P$. This is equivalent to $\alpha \sim_\omega \beta$ by Proposition 7.2.1(iv).

The quotient group $P_\omega = P/N_\omega$ is called the *reduced product* of the family of groups $(G_i)_{i\in I}$ with respect to the filter $\omega$. In the particular case when $\omega$ is an ultrafilter, one also says that $P_\omega$ is the *ultraproduct* of the family of groups $(G_i)_{i\in I}$ with respect to the ultrafilter $\omega$.

**Theorem 7.2.3.** *Let $\mathcal{C}$ be a class of groups, $(G_i)_{i\in I}$ a family of groups such that $G_i \in \mathcal{C}$ for all $i \in I$, and $\omega$ an ultrafilter on the index set $I$. Then the ultraproduct $P_\omega$ of the family of groups $(G_i)_{i\in I}$ with respect to the ultrafilter $\omega$ is locally embeddable into $\mathcal{C}$.*

*Proof.* Fix a finite subset $K \subset P_\omega$. We want to show that there exist a group $C \in \mathcal{C}$ and a $K$-almost-homomorphism $\varphi\colon P_\omega \to C$.

Choose a representative of each element $g \in P_\omega$, that is, an element $\widetilde{g} = (\widetilde{g}_i)_{i\in I} \in P$ such that $g = \widetilde{g}N_\omega$. If $h$ and $k$ are arbitrary elements of $K$, we have $\widetilde{hk}N_\omega = \widetilde{h}\widetilde{k}N_\omega$ and therefore $\widetilde{hk} \sim_\omega \widetilde{h}\widetilde{k}$ by (7.7). It follows that the set $I_{h,k} = \{i \in I : \widetilde{hk}_i = \widetilde{h}_i\widetilde{k}_i\}$ belongs to $\omega$. Thus, since $\omega$ is closed under finite intersections (cf.(F-5) in Sect. J.1), we have that

$$I_K = \bigcap_{h,k\in K} I_{h,k}$$

also belongs to $\omega$. On the other hand, if $h$ and $k$ are distinct elements of $K$, the subset $\{i \in I : \widetilde{h}_i = \widetilde{k}_i\}$ does not belongs to $\omega$. As $\omega$ is an ultrafilter, this implies that $I'_{h,k} = I \setminus \{i \in I : \widetilde{h}_i = \widetilde{k}_i\} = \{i \in I : \widetilde{h}_i \neq \widetilde{k}_i\}$ belongs to $\omega$ (cf. (UF) in Sect. J.1). Using again the fact that $\omega$ is closed under finite intersections, we deduce that

$$I'_K = \bigcap_{\substack{h,k\in K \\ h\neq k}} I'_{h,k}$$

also belongs to $\omega$. Since $\omega$ is closed under finite intersections and $\varnothing \notin \omega$ (cf. (F-1) in Sect. J.1), we can find an index $i_0 \in I$ such that

$$i_0 \in I_K \cap I'_K.$$

Consider the map $\varphi\colon P_\omega \to G_{i_0}$ defined by $\varphi(g) = \widetilde{g}_{i_0}$ for all $g \in P_\omega$. The map $\varphi$ satisfies the following properties:

(1) $\varphi(hk) = \widetilde{hk}_{i_0} = \widetilde{h}_{i_0}\widetilde{k}_{i_0} = \varphi(h)\varphi(k)$ for all $h, k \in K$ (because $i_0 \in I_K$);
(2) $\varphi(h) = \widetilde{h}_{i_0} \neq \widetilde{k}_{i_0} = \varphi(k)$ for all $h, k \in K$ such that $h \neq k$ (because $i_0 \in I'_K$).

This shows that $\varphi$ is a $K$-almost homomorphism. It follows that $P_\omega$ is locally embeddable into the class $\mathcal{C}$. □

*Remark 7.2.4.* When the ultrafilter $\omega$ is principal, then the group $P_\omega$ itself belongs to $\mathcal{C}$. Indeed, in this case, there is an element $i_0 \in I$ such that $\omega$ consists of all subsets of $I$ containing $i_0$. This implies that $\alpha \sim_\omega \beta$ if and only if $\alpha_{i_0} = \beta_{i_0}$ for all $\alpha = (\alpha_i)_{i\in I}, \beta = (\beta_i)_{i\in I} \in P$. Therefore, $N_\omega$ consists of all $\alpha \in P$ such that $\alpha_{i_0} = 1_{G_{i_0}}$. It follows that the group $P_\omega$ is isomorphic to the group $G_{i_0}$ and therefore that $P_\omega \in \mathcal{C}$.

**Theorem 7.2.5.** *Let $\mathcal{C}$ be a class of groups and let $G$ be a group. The following conditions are equivalent:*

(a) *$G$ is locally embeddable into $\mathcal{C}$;*
(b) *there exists a family of groups $(G_i)_{i\in I}$ such that $G_i \in \mathcal{C}$ for all $i \in I$ and an ultrafilter $\omega$ on $I$ such that $G$ is isomorphic to a subgroup of the ultraproduct $P_\omega$ of the family $(G_i)_{i\in I}$ with respect to the ultrafilter $\omega$.*

*Proof.* If $G$ satisfies (b), then $G$ is locally embeddable into $\mathcal{C}$ since $P_\omega$ is locally embeddable into $\mathcal{C}$ by Theorem 7.2.3 and every subgroup of a group which is locally embeddable into $\mathcal{C}$ is itself locally embeddable into $\mathcal{C}$ by Proposition 7.1.7.

Conversely, suppose that $G$ is locally embeddable into $\mathcal{C}$. Consider the set $I$ consisting of all finite subsets of $G$.

For each $K \in I$ we define the set

$$I_K = \{K' \in I : K \subset K'\}.$$

Observe that $I_K \neq \varnothing$ as $K \in I_K$. Moreover, the family of nonempty subsets $(I_K)_{K\in I}$ is closed under finite intersections, since

$$I_{K_1} \cap I_{K_2} = I_{K_1 \cup K_2}$$

for all $K_1, K_2 \in I$. It follows from Proposition J.1.3 and Theorem J.1.6 that there exists an ultrafilter $\omega$ on $I$ such that $I_K \in \omega$ for all $K \in I$.

As $G$ is locally embeddable into $\mathcal{C}$, we can find, for each $K \in I$, a group $G_K$ in $\mathcal{C}$ and a $K$-almost-homomorphism $\varphi_K\colon G \to G_K$.

Consider the ultraproduct $P_\omega$ of the family $(G_K)_{K\in I}$ with respect to $\omega$. By Theorem 7.2.3, $P_\omega$ is locally embeddable into $\mathcal{C}$. Let $\widetilde{\varphi}\colon G \to P = \prod_{K\in I} G_K$ denote the product map $\widetilde{\varphi} = \prod_{K\in I} \varphi_K$. Thus, we have

$$\widetilde{\varphi}(g) = (\varphi_K(g))_{K\in I}$$

for all $g \in G$. Let $\rho\colon P \to P_\omega = P/N_\omega$ denote the canonical epimorphism. Let us show that the composite map $\Phi = \rho \circ \widetilde{\varphi}\colon G \to P_\omega$ is an injective homomorphism. This will prove that $G$ is isomorphic to a subgroup of $P_\omega$.

Let $g, h \in G$. Consider the element $K_0 = \{g, h\} \in I$. If $K \in I$ satisfies $K_0 \subset K$, then

$$\varphi_K(gh) = \varphi_K(g)\varphi_K(h)$$

since $\varphi_K$ is a $K$-almost-homomorphism. Thus, the set

$$\{K \in I : \varphi_K(gh) = \varphi_K(g)\varphi_K(h)\}$$

contains $I_{K_0}$ and therefore belongs to $\omega$. This implies that

$$\widetilde{\varphi}(gh) \sim_\omega \widetilde{\varphi}(g)\widetilde{\varphi}(h).$$

Therefore, we have $\Phi(gh) = \Phi(g)\Phi(h)$ by (7.7). This shows that $\Phi$ is a homomorphism.

On the other hand, if $g \neq h$, we have

$$\varphi_K(g) \neq \varphi_K(h)$$

for all $K \in I$ such that $K_0 \subset K$. This implies that $\{K \in I : \varphi_K(g) \neq \varphi_K(h)\} \in \omega$. As $\omega$ is a filter, it follows that $\{K \in I : \varphi_K(g) = \varphi_K(h)\} \notin \omega$, that is,

$$\widetilde{\varphi}(g) \not\sim_\omega \widetilde{\varphi}(h).$$

Therefore we have $\Phi(g) \neq \Phi(h)$, again by (7.7).

Consequently, $\Phi$ is injective. This shows that (a) implies (b). □

*Remark 7.2.6.* Suppose that $G$ is a group and that $\mathcal{C}$ is a class of groups which is closed under taking subgroups. If $G$ is locally embeddable into $\mathcal{C}$ and $G \notin \mathcal{C}$ then for any family of groups $(G_i)_{i\in I}$ and any ultrafilter $\omega$ on the index set $I$ satisfying condition (b) in Theorem 7.2.5, we deduce from Remark 7.2.4 that the ultrafilter $\omega$ is necessarily non-principal.

## 7.3 LEF-Groups and LEA-Groups

Let us rewrite the results obtained in Sect. 7.1 in the particular case when $\mathcal{C}$ is either the class of finite groups or the class of amenable groups. Note that these classes are closed under taking subgroups and taking finite direct

products (this is trivial for finite groups and follows from Proposition 4.5.1 and Corollary 4.5.6 for amenable groups), so that all the results established in Sect. 7.1 apply to these two classes.

We shall use the following terminology, which is very popular in the field. A group which is locally embeddable into the class of finite groups is briefly called an *LEF*-group. Similarly, a group which is locally embeddable into the class of amenable groups is called an *LEA*-group. Note that every LEF-group is LEA since the class of finite groups is contained in the class of amenable groups by Proposition 4.4.6.

We deduce from Proposition 7.1.6 the following characterization of LEF-groups.

**Proposition 7.3.1.** *Let $G$ be a group. Then the following conditions are equivalent:*

(a) *$G$ is an LEF-group;*
(b) *for every finite subset $K \subset G$, there exist a finite set $L$ with $K \subset L \subset G$ and a binary operation $\odot\colon L \times L \to L$ such that $(L, \odot)$ is a group and $k_1 k_2 = k_1 \odot k_2$ for all $k_1, k_2 \in K$.*

□

Analogously, we deduce from Proposition 7.1.6 the following characterization of LEA-groups.

**Proposition 7.3.2.** *Let $G$ be a group. Then the following conditions are equivalent:*

(a) *$G$ is an LEA-group;*
(b) *for every finite subset $K \subset G$, there exist a set $L$ with $K \subset L \subset G$ and a binary operation $\odot\colon L \times L \to L$ such that $(L, \odot)$ is an amenable group and $k_1 k_2 = k_1 \odot k_2$ for all $k_1, k_2 \in K$.*

□

From Theorem 7.2.5 we deduce the following characterization of LEF (resp. LEA) groups in terms of ultraproducts.

**Corollary 7.3.3.** *Let $G$ be a group. The following conditions are equivalent:*

(a) *$G$ is an LEF (resp. LEA) group;*
(b) *there exists a family of finite (resp. amenable) groups $(G_i)_{i\in I}$ and an ultrafilter $\omega$ on $I$ such that $G$ is isomorphic to a subgroup of the ultraproduct $P_\omega$ of the family $(G_i)_{i\in I}$ with respect to the ultrafilter $\omega$.*

□

From Corollary 7.1.15, we get:

**Proposition 7.3.4.** *Every locally residually finite group is LEF and therefore LEA.* □

**Corollary 7.3.5.** *All finite groups, all residually finite groups, all profinite groups, all free groups, all abelian groups, and all locally finite groups are LEF and therefore LEA.*

*Proof.* In order to complete the proof, it suffices to recall that free groups (resp. profinite groups) are residually finite by Theorem 2.3.1 (resp. Corollary 2.2.8) and that abelian groups are locally residually finite by Corollary 2.2.4. □

Similarly, by applying Corollary 7.1.15 to the class of amenable groups, we get:

**Proposition 7.3.6.** *Every locally residually amenable group is LEA.* □

From Proposition 7.1.7, Proposition 7.1.8, Proposition 7.1.10, Corollary 7.1.11, Corollary 7.1.12, Corollary 7.1.17, and Theorem 7.1.19, we get:

**Proposition 7.3.7.** *The following assertions hold:*

(i) *every subgroup of an LEF-group (resp. LEA-group) is an LEF-group (resp. LEA-group);*
(ii) *a group is LEF (resp. LEA) if and only if all its finitely generated subgroups are LEF (resp. LEA);*
(iii) *let $(G_i)_{i\in I}$ be a family of LEF-groups (resp. LEA-groups). Then their direct product $\prod_{i\in I} G_i$ is an LEF-group (resp. LEA-group);*
(iv) *let $(G_i)_{i\in I}$ be a family of LEF-groups (resp. LEA-groups). Then their direct sum $\bigoplus_{i\in I} G_i$ is an LEF-group (resp. LEA-group);*
(v) *let $(G_i)_{i\in I}$ be a projective system of LEF-groups (resp. LEA-groups). Then their projective limit $G = \varprojlim G_i$ is an LEF-group (resp. LEA-group);*
(vi) *Let $\Gamma$ be a group. Then the set of all $N \in \mathcal{N}(\Gamma)$ such that $\Gamma/N$ is LEF (resp. LEA) is closed (and hence compact) in $\mathcal{N}(\Gamma)$.*
(vii) *let $F$ be a free group and let $N \in \mathcal{N}(F)$. Then $F/N$ is an LEF-group (resp. LEA-group) if and only if there exists a net $(N_i)_{i\in I}$ in $\mathcal{N}(F)$ with $F/N_i$ finite (resp. amenable) for all $i \in I$ such that $N = \lim_i N_i$.*

□

Finally, we deduce from Corollary 7.1.21 the following:

**Proposition 7.3.8.** *A finitely presented group is LEF (resp. LEA) if and only if it is residually finite (resp. residually amenable).* □

As we have observed at the end of the previous section, the additive group $\mathbb{Q}$ is LEF but not residually finite. On the other hand, it follows from Proposition 7.3.8 that every finitely presented LEF-group is residually finite. The group $\mathbb{Q}$ is not finitely generated. However, there exist finitely generated LEF-groups which are not residually finite. An example of such a group is provided by the group $G_1$ introduced in Sect. 2.6:

**Proposition 7.3.9.** *The group $G_1$ of Sect. 2.6 is a finitely generated amenable LEF-group which is not residually finite.*

*Proof.* By construction, the group $G_1$ is finitely generated since it is defined as being a subgroup of $\mathrm{Sym}(\mathbb{Z})$ generated by two elements, namely the transposition $(0\ 1)$ and the translation $n \mapsto n+1$.

The group $G_1$ is not residually finite by Proposition 2.6.1.

Recall that, denoting by $\mathrm{Sym}_0(\mathbb{Z})$ the normal subgroup of $\mathrm{Sym}(\mathbb{Z})$ consisting of all permutations of $\mathbb{Z}$ with finite support, the group $G_1$ is the semidirect product of $\mathrm{Sym}_0(\mathbb{Z})$ with the infinite cyclic group generated by the translation $T\colon n \mapsto n+1$ (see Lemma 2.6.4). Therefore, each element $g \in G_1$ can be uniquely written in the form $g = T^{i(g)}\sigma(g)$, where $i(g) \in \mathbb{Z}$ and $\sigma(g) \in \mathrm{Sym}_0(\mathbb{Z})$. Note that we have $i(gh) = i(g) + i(h)$ and $\sigma(gh) = T^{-i(h)}\sigma(g)T^{i(h)}\sigma(h)$ for all $g, h \in G_1$.

To prove that $G_1$ is an LEF-group, consider a finite subset $K \subset G_1$. Let us show that there exist a finite group $F$ and a $K$-almost-homomorphism of $G_1$ into $F$.

Let $\ell = \max_{k\in K} |i(k)|$ and choose an integer $r \geq \ell$ such that the supports of the elements $T^{-j}\sigma(k)T^j \in \mathrm{Sym}_0(\mathbb{Z})$ are contained in the interval $[-r, r]$ for all $-\ell \leq j \leq \ell$ and $k \in K$. Let us set $R = 4r$,

$$X = \{-R, -R+1, \ldots, -1, 0, 1, \ldots, R-1, R\},$$

and $F = \mathrm{Sym}(X)$. Consider the $(2R+1)$-cycle $\gamma \in F$ given by

$$\gamma = (-R \ \ -R+1 \ \ -R+2 \ \ \cdots \ \ R-1 \ \ R).$$

For all $\sigma \in \mathrm{Sym}_0(\mathbb{Z})$ whose support is contained in $X$, let $\overline{\sigma} \in F$ denote the element defined by $\overline{\sigma}(x) = \sigma(x)$ for all $x \in X$. Observe that if $\sigma, \sigma' \in \mathrm{Sym}_0(\mathbb{Z})$ have both their support contained in $X$, then so does $\sigma\sigma'$ and that we have

$$\overline{\sigma\sigma'} = \overline{\sigma}\,\overline{\sigma'}. \tag{7.8}$$

Moreover, for all $k \in K$ and $j \in \mathbb{Z}$ such that $-\ell \leq j \leq \ell$, the supports of $\sigma(k)$ and $T^{-j}\sigma(k)T^j$ are contained in $X$ and we have

$$\overline{T^{-j}\sigma(k)T^j} = \gamma^{-j}\overline{\sigma(k)}\gamma^j. \tag{7.9}$$

Indeed, suppose first that $x \in [-2r, 2r]$. Then $x + j \in [-3r, 3r] \subset X$ and we have

$$\begin{aligned}[T^{-j}\sigma(k)T^j](x) &= [T^{-j}\sigma(k)](x+j)\\ &= T^{-j}\left(\sigma(k)(x+j)\right)\\ &= \sigma(k)(x+j) - j,\end{aligned}$$

and

$$\begin{aligned}[\gamma^{-j}\overline{\sigma(k)}\gamma^j](x) &= [\gamma^{-j}\overline{\sigma(k)}](x+j)\\ &= \gamma^{-j}\left(\overline{\sigma(k)}(x+j)\right)\\ &= \gamma^{-j}\left(\sigma(k)(x+j)\right)\\ &= \sigma(k)(x+j)-j.\end{aligned}$$

Suppose now that $x \in X \setminus [-2r, 2r]$. Then $[T^{-j}\sigma(k)T^j](x) = x$ since the support of $T^{-j}\sigma(k)T^j$ is contained in $[-\ell+r, r+\ell] \subset [-2r, 2r]$. On the other hand, if we denote, for $y \in \mathbb{Z}$, by $\widehat{y}$ the unique element in $\{y + (2R+1)n : n \in \mathbb{Z}\} \cap [-R, R]$, we have

$$\widehat{x+j} \in X \setminus [-r, r] \tag{7.10}$$

and

$$\begin{aligned}[\gamma^{-j}\overline{\sigma(k)}\gamma^j](x) &= [\gamma^{-j}\overline{\sigma(k)}](\widehat{x+j})\\ &= \gamma^{-j}\left(\sigma(k)(\widehat{x+j})\right)\\ \text{by (7.10)} \quad &= \gamma^{-j}(\widehat{x+j})\\ &= x.\end{aligned}$$

This shows (7.9).

Define a map $\varphi\colon G_1 \to F$ by setting $\varphi(g) = \gamma^{i(g)}\overline{\sigma(g)}$ if the support of $\sigma(g)$ is contained in $X$, and $\varphi(g) = 1_F$ otherwise. Let us show that $\varphi$ is a $K$-almost-homomorphism.

Let $k_1, k_2 \in K$. We have

$$\begin{aligned}\varphi(k_1k_2) &= \varphi[T^{i(k_1)+i(k_2)}(T^{-i(k_2)}\sigma(k_1)T^{i(k_2)}\sigma(k_2))]\\ &= \gamma^{i(k_1)+i(k_2)}\overline{T^{-i(k_2)}\sigma(k_1)T^{i(k_2)}\sigma(k_2)}\\ \text{(by (7.8))} \quad &= \gamma^{i(k_1)+i(k_2)}\cdot\overline{T^{-i(k_2)}\sigma(k_1)T^{i(k_2)}}\,\overline{\sigma(k_2)}\\ \text{(by (7.9))} \quad &= \gamma^{i(k_1)+i(k_2)}\cdot\gamma^{-i(k_2)}\overline{\sigma(k_1)}\gamma^{i(k_2)}\cdot\overline{\sigma(k_2)}\\ &= \gamma^{i(k_1)}\overline{\sigma(k_1)}\gamma^{i(k_2)}\overline{\sigma(k_2)}\\ &= \varphi[T^{i(k_1)}\sigma(k_1)]\varphi[T^{i(k_2)}\sigma(k_2)]\\ &= \varphi(k_1)\varphi(k_2).\end{aligned}$$

Let us show that $\varphi|_K$ is injective. Let $k_1, k_2 \in K$ and suppose that $\varphi(k_1) = \varphi(k_2)$. This implies that $\gamma^{i(k_1)}\overline{\sigma(k_1)} = \gamma^{i(k_2)}\overline{\sigma(k_2)}$, that is, $\gamma^{i(k_1)-i(k_2)} = \overline{\sigma(k_2)}\cdot\overline{\sigma(k_1)}^{-1}$. But the support of $\overline{\sigma(k_2)}\cdot\overline{\sigma(k_1)}^{-1}$ is contained in $[-r, r]$ while the support of $\gamma^i$, $i \in \mathbb{Z}$, is the whole set $X$ if $i$ is not a multiple of $2R+1$. As $|i(k_1) - i(k_2)| \le 2\ell < 2R+1$, we deduce that $i(k_1) - i(k_2) = 0$ and $\overline{\sigma(k_2)}\cdot\overline{\sigma(k_1)}^{-1} = 1_F$. This implies $i(k_1) = i(k_2)$ and $\sigma(k_1) = \sigma(k_2)$, and therefore $k_1 = k_2$. This shows that the restriction of $\varphi$ to $K$ is injective. It

follows that $\varphi$ is a $K$-almost homomorphism and therefore that $G_1$ is an LEF group.

Finally, observe that the group $\mathrm{Sym}_0(\mathbb{Z})$ is locally finite and therefore amenable by Corollary 4.5.12. Now, $G_1$ is the semidirect product of the amenable group $\mathrm{Sym}_0(\mathbb{Z})$ with an infinite cyclic (and therefore amenable) group so that, by Proposition 4.5.5, it is an amenable group as well. □

*Remarks 7.3.10.* (a) From Proposition 7.3.4 and Proposition 7.3.9, we deduce that the class of locally residually finite groups is strictly contained in the class of LEF-groups.

(b) Proposition 7.3.8 and Proposition 7.3.9 imply that the group $G_1$ is not finitely presentable.

## 7.4 The Hamming Metric

Let $G$ be a group.

A metric $d$ on $G$ is called *left-invariant* (resp. *right-invariant*) if $d(hg_1, hg_2) = d(g_1, g_2)$ (resp. $d(g_1h, g_2h) = d(g_1, g_2)$) for all $g_1, g_2, h \in G$. A metric on $G$ which is both left and right-invariant is called *bi-invariant*.

Note that a left-invariant (resp. right-invariant) metric $d$ on $G$ is entirely determined by the map $g \mapsto d(1_G, g)$, $g \in G$, since $d(h, k) = d(1_G, h^{-1}k)$ (resp. $d(h, k) = d(1_G, kh^{-1})$) for all $h, k \in G$.

Let now $F$ be a nonempty finite set and consider the symmetric group $\mathrm{Sym}(F)$. For $\alpha \in \mathrm{Sym}(F)$, we denote by $\mathrm{Fix}(\alpha)$ the set $\{x \in F : \alpha(x) = x\}$ of fixed points of $\alpha$. The support of $\alpha$ is the set $\{x \in F : \alpha(x) \neq x\} = F \setminus \mathrm{Fix}(\alpha)$, so that we have

$$|\{x \in F : \alpha(x) \neq x\}| = |F| - |\operatorname{Fix}(\alpha)|. \tag{7.11}$$

Consider the map $d_F \colon \mathrm{Sym}(F) \times \mathrm{Sym}(F) \to \mathbb{R}$ defined by

$$d_F(\alpha_1, \alpha_2) = \frac{|\{x \in F : \alpha_1(x) \neq \alpha_2(x)\}|}{|F|} \tag{7.12}$$

for all $\alpha_1, \alpha_2 \in \mathrm{Sym}(F)$. Observe that the set $\{x \in F : \alpha_1(x) \neq \alpha_2(x)\} = \{x \in F : x \neq \alpha_1^{-1}\alpha_2(x)\}$ is the support of $\alpha_1^{-1}\alpha_2$, so that (7.11) gives us

$$d_F(\alpha_1, \alpha_2) = 1 - \frac{|\operatorname{Fix}(\alpha_1^{-1}\alpha_2)|}{|F|}. \tag{7.13}$$

**Proposition 7.4.1.** *Let $F$ be a nonempty finite set. Then $d_F$ is a bi-invariant metric on* $\mathrm{Sym}(F)$.

*Proof.* It is immediate from the definition that $d_F(\alpha_1, \alpha_2) \geq 0$ and $d_F(\alpha_1, \alpha_2) = d_F(\alpha_2, \alpha_1)$ for all $\alpha_1, \alpha_2 \in \mathrm{Sym}(F)$. Moreover, the equality $d_F(\alpha_1, \alpha_2) = 0$ holds if and only if $\alpha_1(x) = \alpha_2(x)$ for all $x \in F$, that is, if and only if $\alpha_1 = \alpha_2$.

Let now $\alpha_1, \alpha_2, \alpha_3 \in \mathrm{Sym}(F)$. If $x \in F$ satisfies $\alpha_1(x) \neq \alpha_2(x)$, then $\alpha_1(x) \neq \alpha_3(x)$ or $\alpha_2(x) \neq \alpha_3(x)$. Thus, we have the inclusion

$$\{x \in F : \alpha_1(x) \neq \alpha_2(x)\} \subset \{x \in F : \alpha_1(x) \neq \alpha_3(x)\} \cup \{x \in F : \alpha_2(x) \neq \alpha_3(x)\}.$$

This implies

$$\begin{aligned} d_F(\alpha_1, \alpha_2) &= \frac{1}{|F|} |\{x \in F : \alpha_1(x) \neq \alpha_2(x)\}| \\ &\leq \frac{1}{|F|} |\{x \in F : \alpha_1(x) \neq \alpha_3(x)\} \cup \{x \in F : \alpha_2(x) \neq \alpha_3(x)\}| \\ &\leq \frac{1}{|F|} \left(|\{x \in F : \alpha_1(x) \neq \alpha_3(x)\}| + |\{x \in F : \alpha_2(x) \neq \alpha_3(x)\}|\right) \\ &= d_F(\alpha_1, \alpha_3) + d_F(\alpha_2, \alpha_3). \end{aligned}$$

This shows that $d_F$ also satisfies the triangle inequality. Therefore, $d_F$ is a metric on $\mathrm{Sym}(F)$.

It remains to show that $d_F$ is bi-invariant. Let $\alpha_1, \alpha_2, \beta \in \mathrm{Sym}(F)$. Since $\beta$ is bijective, we have

$$\{x \in F : \beta\alpha_1(x) \neq \beta\alpha_2(x)\} = \{x \in F : \alpha_1(x) \neq \alpha_2(x)\}.$$

This implies that

$$\begin{aligned} d_F(\beta\alpha_1, \beta\alpha_2) &= \frac{1}{|F|} |\{x \in F : \beta\alpha_1(x) \neq \beta\alpha_2(x)\}| \\ &= \frac{1}{|F|} |\{x \in F : \alpha_1(x) \neq \alpha_2(x)\}| \\ &= d_F(\alpha_1, \alpha_2). \end{aligned}$$

Thus, $d_F$ is left-invariant. On the other hand, we have

$$\{x \in F : \alpha_1\beta(x) \neq \alpha_2\beta(x)\} = \beta^{-1}(\{x \in F : \alpha_1(x) \neq \alpha_2(x)\}),$$

which implies

$$\begin{aligned} d_F(\alpha_1\beta, \alpha_2\beta) &= \frac{1}{|F|} |\{x \in F : \alpha_1\beta(x) \neq \alpha_2\beta(x)\}| \\ &= \frac{1}{|F|} |\beta^{-1}(\{x \in F : \alpha_1(x) \neq \alpha_2(x)\})| \\ &= \frac{1}{|F|} |\{x \in F : \alpha_1(x) \neq \alpha_2(x)\}| \\ &= d_F(\alpha_1, \alpha_2). \end{aligned}$$

Consequently, $d_F$ is also right-invariant. □

**Definition 7.4.2.** Let $F$ be a nonempty finite set. The bi-invariant metric $d_F$ is called the *(normalized) Hamming metric* on $\mathrm{Sym}(F)$.

Suppose now that $m$ is a positive integer and that $F_1, F_2, \ldots, F_m$ are nonempty finite sets. Consider the Cartesian product $F = F_1 \times F_2 \times \cdots \times F_m$

and the natural group homomorphism $\Phi\colon \prod_{i=1}^m \mathrm{Sym}(F_i) \to \mathrm{Sym}(F)$ defined by

$$\Phi(\alpha)(x) = (\alpha_1(x_1), \alpha_2(x_2), \ldots, \alpha_m(x_m))$$

for all $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m) \in \prod_{i=1}^m \mathrm{Sym}(F_i)$ and $x = (x_1, x_2, \ldots, x_m) \in F$.

**Proposition 7.4.3.** *With the above notation, one has*

$$d_F(\Phi(\alpha), \Phi(\beta)) = 1 - \prod_{i=1}^m (1 - d_{F_i}(\alpha_i, \beta_i)) \tag{7.14}$$

*for all* $\alpha = (\alpha_i)_{1\le i\le m}$ *and* $\beta = (\beta_i)_{1\le i\le m}$ *in* $\prod_{i=1}^m \mathrm{Sym}(F_i)$.

*Proof.* First observe that if $\alpha = (\alpha_i)_{1\le i\le m} \in \prod_{i=1}^m \mathrm{Sym}(F_i)$, then we have $\mathrm{Fix}(\Phi(\alpha)) = \prod_{i=1}^m \mathrm{Fix}(\alpha_i)$, and therefore

$$\begin{aligned} d_F(\mathrm{Id}_F, \Phi(\alpha)) &= 1 - \frac{|\mathrm{Fix}(\alpha)|}{|F|} \\ &= 1 - \frac{\prod_{i=1}^m |\mathrm{Fix}(\alpha_i)|}{\prod_{i=1}^m |F_i|} \\ &= 1 - \prod_{i=1}^m \frac{|\mathrm{Fix}(\alpha_i)|}{|F_i|} \\ &= 1 - \prod_{i=1}^m (1 - d_{F_i}(\mathrm{Id}_{F_i}, \alpha_i)). \end{aligned}$$

We deduce that, for all $\alpha = (\alpha_i)_{1\le i\le m}, \beta = (\beta_i)_{1\le i\le m} \in \prod_{i=1}^m \mathrm{Sym}(F_i)$, we have

$$\begin{aligned} d_F(\Phi(\alpha), \Phi(\beta)) &= d(\mathrm{Id}_F, \Phi(\alpha)^{-1}\Phi(\beta)) \quad \text{(by left-invariance of } d_F) \\ &= d_F(\mathrm{Id}_F, \Phi(\alpha^{-1}\beta)) \\ &= 1 - \prod_{i=1}^m (1 - d_{F_i}(\mathrm{Id}_{F_i}, \alpha_i^{-1}\beta_i) \\ &= 1 - \prod_{i=1}^m (1 - d_{F_i}(\alpha_i, \beta_i)) \end{aligned}$$

where the last equality follows from the left-invariance of $d_{F_i}$. □

**Corollary 7.4.4.** *Let $m$ be a positive integer and let $F$ be a nonempty finite set. Consider the homomorphism*

$$\Psi\colon \mathrm{Sym}(F) \to \mathrm{Sym}(F^m) \tag{7.15}$$

*defined by*

$$\Psi(\alpha)(x_1, x_2, \ldots, x_m) = (\alpha(x_1), \alpha(x_2), \ldots, \alpha(x_m))$$

*for all* $\alpha \in \mathrm{Sym}(F)$ *and* $x_1, x_2, \ldots, x_m \in F$. *Then one has*

$$d_{F^m}(\Psi(\alpha), \Psi(\beta)) = 1 - (1 - d_F(\alpha, \beta))^m \tag{7.16}$$

*for all* $\alpha, \beta \in \mathrm{Sym}(F)$. □

## 7.5 Sofic Groups

**Definition 7.5.1.** Let $G$ be a group, $K \subset G$ a finite subset, and $\varepsilon > 0$. Let $F$ be a nonempty finite set. A map $\varphi\colon G \to \mathrm{Sym}(F)$ is called a $(K, \varepsilon)$-*almost-homomorphism* if it satisfies the following conditions:

($(K, \varepsilon)$-AH-1) for all $k_1, k_2 \in K$, one has $d_F(\varphi(k_1 k_2), \varphi(k_1)\varphi(k_2)) \leq \varepsilon$;
($(K, \varepsilon)$-AH-2) for all $k_1, k_2 \in K$, $k_1 \neq k_2$, one has $d_F(\varphi(k_1), \varphi(k_2)) \geq 1 - \varepsilon$,

where $d_F$ denotes the normalized Hamming metric on $\mathrm{Sym}(F)$.

**Definition 7.5.2.** A group $G$ is called *sofic* if it satisfies the following condition: for every finite subset $K \subset G$ and every $\varepsilon > 0$, there exist a nonempty finite set $F$ and a $(K, \varepsilon)$-almost-homomorphism $\varphi\colon G \to \mathrm{Sym}(F)$.

**Proposition 7.5.3.** *Every finite group is sofic.*

*Proof.* Let $G$ be a finite group. Consider the map $L\colon G \to \mathrm{Sym}(G)$ defined by $L(g)(h) = gh$ for all $g, h \in G$. As $L$ is a homomorphism (cf. the proof of Cayley's theorem (Theorem C.1.2)), we have

$$d_G(L(g_1 g_2), L(g_1)L(g_2)) = 0 \tag{7.17}$$

for all $g_1, g_2 \in G$. Moreover, for all distinct $g_1, g_2 \in G$ we have $L(g_1)(h) = g_1 h \neq g_2 h = L(g_2)(h)$ for all $h \in G$ so that

$$d_G(L(g_1), L(g_2)) = 1. \tag{7.18}$$

This shows that $L$ is a $(K, \varepsilon)$-almost-homomorphism for all $K \subset G$ and $\varepsilon > 0$. It follows that $G$ is sofic. □

**Proposition 7.5.4.** *Every subgroup of a sofic group is sofic.*

*Proof.* Let $G$ be a sofic group and let $H$ be a subgroup of $G$. Fix a finite subset $K \subset H$ and $\varepsilon > 0$. As $G$ is sofic, there exists a nonempty finite set $F$ and a $(K, \varepsilon)$-almost-homomorphism $\varphi\colon G \to \mathrm{Sym}(F)$. Then the restriction map $\varphi|_H\colon H \to \mathrm{Sym}(F)$ is a $(K, \varepsilon)$-almost-homomorphism. This shows that $H$ is sofic. □

**Proposition 7.5.5.** *Every locally sofic group is sofic.*

*Proof.* Let $G$ be a locally sofic group. Let $K \subset G$ be a finite subset and $\varepsilon > 0$. Denote by $H$ the subgroup of $G$ generated by $K$. Then, as $H$ is sofic, there exist a nonempty finite set $F$ and a $(K, \varepsilon)$-almost-homomorphism $\psi\colon H \to \mathrm{Sym}(F)$. Extend arbitrarily $\psi$ to a map $\varphi\colon G \to \mathrm{Sym}(F)$, for example by setting $\varphi(g) = \mathrm{Id}_F$ for all $g \in G \setminus H$. It is clear that $\varphi$ is a $(K, \varepsilon)$-almost-homomorphism. This shows that $G$ is sofic. □

From Proposition 7.5.3 and Proposition 7.5.5, we immediately deduce that every locally finite group is sofic. As any locally finite group is amenable by Corollary 4.5.12, this is actually covered by the following:

**Proposition 7.5.6.** *Every amenable group is sofic.*

*Proof.* Suppose that $G$ is an amenable group. Let $K \subset G$ be a finite subset and $\varepsilon > 0$. Set $S = (\{1_G\} \cup K \cup K^{-1})^2$. Since $G$ is amenable, it follows from Theorem 4.9.1 and Proposition 4.7.1(a) that there exists a nonempty finite subset $F \subset G$ such that

$$|F \setminus sF| \leq \frac{\varepsilon}{|S|}|F| \tag{7.19}$$

for all $s \in S$. Consider the set $E = \bigcap_{s\in S} sF$. Observe that $E \subset F$ since $1_G \in S$. In fact, as $S = S^{-1}$, we get

$$sE \subset F \tag{7.20}$$

for all $s \in S$. Moreover, we have

$$\begin{aligned} |F \setminus E| &= |F \setminus \bigcap_{s\in S} sF| \\ &= |\bigcup_{s\in S} (F \setminus sF)| \\ &\leq \sum_{s\in S} |F \setminus sF| \\ &\leq \varepsilon|F| \quad \text{by (7.19)}. \end{aligned} \tag{7.21}$$

This implies

$$|E| \geq (1 - \varepsilon)|F|. \tag{7.22}$$

For each $g \in G$, we have $|F| = |gF|$ and hence $|F \setminus gF| = |gF \setminus F|$. Therefore, we can find a bijective map $\alpha_g\colon gF \setminus F \to F \setminus gF$. Consider the map $\varphi\colon G \to \mathrm{Sym}(F)$ defined by setting

$$\varphi(g)(f) = \begin{cases} gf & \text{if } gf \in F \\ \alpha_g(gf) & \text{otherwise} \end{cases}$$

for all $g \in G$ and $f \in F$ (see Fig. 7.1).

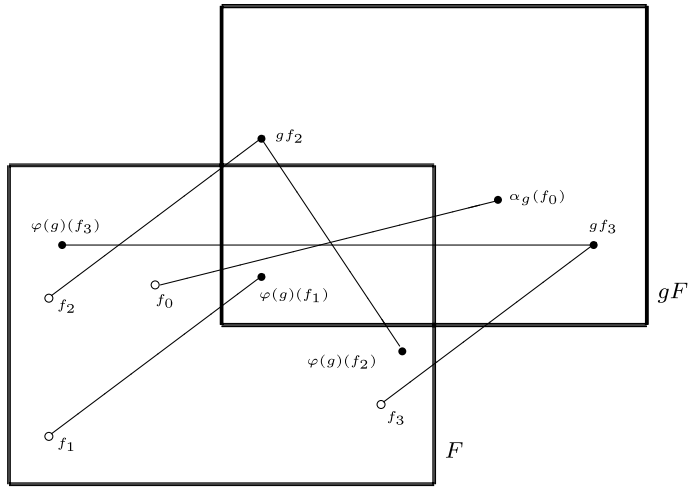



**Fig. 7.1** The maps $\alpha_g \colon gF \setminus F \to F \setminus gF$ and $\varphi(g) \in \mathrm{Sym}(F)$. We have $\varphi(g)(f_1) = gf_1$, $\varphi(g)(f_2) = \alpha_g(gf_2)$ and $\varphi(g)(f_3) = \alpha_g(gf_3)$

Now, suppose that $k_1, k_2 \in K$ and $f \in E$. Then we have $k_2, k_1k_2 \in S$, so that $k_2f, k_1k_2f \in F$ by (7.20). This implies $\varphi(k_1k_2)(f) = k_1k_2f$ and $(\varphi(k_1)\varphi(k_2))(f) = \varphi(k_1)(\varphi(k_2)(f)) = \varphi(k_1)(k_2f) = k_1k_2f$. Therefore, the permutations $\varphi(k_1k_2)$ and $\varphi(k_1)\varphi(k_2)$ coincide on $E$. From (7.21), we deduce that

$$d_F(\varphi(k_1k_2), \varphi(k_1)\varphi(k_2)) \leq \frac{|F \setminus E|}{|F|} \leq \varepsilon$$

for all $k_1, k_2 \in K$.

On the other hand, if $k_1, k_2 \in K$, $k_1 \neq k_2$, and $f \in E$, then we have $k_1f, k_2f \in F$ so that $\varphi(k_1)(f) = k_1f \neq k_2f = \varphi(k_2)(f)$. By using (7.22), we deduce that

$$d_F(\varphi(k_1), \varphi(k_2)) \geq \frac{|E|}{|F|} \geq 1 - \varepsilon$$

for all $k_1, k_2 \in K$ with $k_1 \neq k_2$. Thus, the map $\varphi \colon G \to \mathrm{Sym}(F)$ is a $(K, \varepsilon)$-almost-homomorphism. This shows that $G$ is a sofic group. □

Observe that, when $G$ is finite, the proof of Proposition 7.5.6 reduces to that of Proposition 7.5.3 by taking $F = G$.

**Proposition 7.5.7.** *Let $(G_i)_{i \in I}$ be a family of sofic groups. Then, their direct product $G = \prod_{i \in I} G_i$ is sofic.*

*Proof.* For each $i \in I$, let $\pi_i \colon G \to G_i$ denote the projection homomorphism. Fix a finite subset $K \subset G$ and $\varepsilon > 0$. Then there exists a finite subset $J \subset I$

such that the projection $\pi_J = \prod_{j\in J} \pi_j \colon G \to G_J = \prod_{j\in J} G_j$ is injective on $K$. Choose a constant $0 < \eta < 1$ small enough so that

$$1 - (1-\eta)^{|J|} \leq \varepsilon \tag{7.23}$$

and

$$\eta \leq \varepsilon. \tag{7.24}$$

Since the group $G_j$ is sofic for each $j \in J$, we can find a nonempty finite set $F_j$ and a $(\pi_j(K), \eta)$-almost homomorphism $\varphi_j \colon G_j \to \mathrm{Sym}(F_j)$. Consider the nonempty finite set $F = \prod_{j\in J} F_j$ and the map $\varphi \colon G \to \mathrm{Sym}(F)$ defined by

$$\varphi(g)(f) = (\varphi_j(g_j)(f_j))_{j\in J}$$

for all $g = (g_i)_{i\in I} \in G$, and $f = (f_j)_{j\in J} \in F$. Then, for all $k, k' \in K$, we have, by applying (7.14),

$$\begin{aligned} d_F(\varphi(kk'), \varphi(k)\varphi(k')) &= 1 - \prod_{j\in J} \left(1 - d_{F_j}(\varphi_j(k_j k'_j), \varphi_j(k_j)\varphi_j(k'_j)\right) \\ &\leq 1 - (1-\eta)^{|J|} \\ &\leq \varepsilon \quad \text{(by (7.23))}. \end{aligned}$$

On the other hand, if $k$ and $k'$ are distinct elements in $K$, then there exists $j_0 \in J$ such that $k_{j_0} \neq k'_{j_0}$. This implies, again by using (7.14),

$$\begin{aligned} d_F(\varphi(k), \varphi(k')) &= 1 - \prod_{j\in J} (1 - d_{F_j}(\varphi_j(k_j), \varphi_j(k'_j)) \\ &\geq 1 - (1 - d_{F_{j_0}}(\varphi_{j_0}(k_{j_0}), \varphi_{j_0}(k'_{j_0})) \\ &\geq 1 - \eta \\ &\geq 1 - \varepsilon \quad \text{(by (7.24))}. \end{aligned}$$

This shows that $\varphi$ is a $(K, \varepsilon)$-almost-homomorphism of $G$. It follows that $G$ is sofic. □

**Corollary 7.5.8.** *Let $(G_i)_{i\in I}$ be a family of sofic groups. Then their direct sum $G = \bigoplus_{i\in I} G_i$ is sofic.*

*Proof.* This follows immediately from Proposition 7.5.4 and Proposition 7.5.7, since $G$ is the subgroup of the direct product $P = \prod_{i\in I} G_i$ consisting of all $g = (g_i) \in P$ for which $g_i = 1_{G_i}$ for all but finitely many $i \in I$. □

**Corollary 7.5.9.** *The limit of a projective system of sofic groups is sofic.*

*Proof.* Let $(G_i)_{i\in I}$ be a projective system of sofic groups such that $G = \varprojlim G_i$. By construction of a projective limit (see Appendix E), $G$ is a sub-

group of the group $\prod_{i\in I} G_i$. We deduce that $G$ is sofic by using Proposition 7.5.7 and Proposition 7.5.4. □

**Proposition 7.5.10.** *Every group which is locally embeddable into the class of sofic groups is sofic.*

*Proof.* Let $G$ be a group which is locally embeddable into the class of sofic groups. Let $K \subset G$ be a finite subset and $\varepsilon > 0$. By definition of local embeddability, there exists a sofic group $G'$ and a $K$-almost-homomorphism $\varphi\colon G \to G'$. Set $K' = \varphi(K)$. By soficity of $G'$, there exists a nonempty finite set $F$ and a $(K', \varepsilon)$-almost-homomorphism $\varphi'\colon G' \to \mathrm{Sym}(F)$. Let us prove that the composite map $\Phi = \varphi' \circ \varphi\colon G \to \mathrm{Sym}(F)$ is a $(K, \varepsilon)$-almost-homomorphism. Let $k_1, k_2 \in K$. Then we have

$$\begin{aligned} d_F(\Phi(k_1k_2), \Phi(k_1)\Phi(k_2)) &= d_F(\varphi'(\varphi(k_1k_2)), \varphi'(\varphi(k_1))\varphi'(\varphi(k_2))) \\ &= d_F(\varphi'(\varphi(k_1)\varphi(k_2)), \varphi'(\varphi(k_1))\varphi'(\varphi(k_2))) \\ &\leq \varepsilon \quad (\text{as } \varphi(k_1), \varphi(k_2) \in K'). \end{aligned}$$

Finally, let $k_1, k_2 \in K$ be such that $k_1 \neq k_2$. Since $\varphi|_K$ is injective, we have $\varphi(k_1) \neq \varphi(k_2)$ and therefore

$$d_F(\Phi(k_1), \Phi(k_2)) = d_F(\varphi'(\varphi(k_1)), \varphi'(\varphi(k_2))) \geq 1 - \varepsilon.$$

Thus, $\Phi$ is a $(K, \varepsilon)$-almost-homomorphism. It follows that $G$ is sofic. □

Since every amenable group is sofic by Proposition 7.5.6, an immediate consequence of Proposition 7.5.10 is the following:

**Corollary 7.5.11.** *Every LEA-group is sofic. In particular, every LEF-group, every locally residually amenable group, every locally residually finite group, every residually amenable group, and every residually finite group is sofic.* □

As the class of sofic groups is closed under direct products by Proposition 7.5.7, it follows from Corollary 7.1.15 that every locally residually sofic group is locally embeddable into the class of sofic groups. By applying Proposition 7.5.10, we get:

**Corollary 7.5.12.** *Every locally residually sofic group is sofic.* □

It follows from Proposition 7.5.10 that the class of groups which are locally embeddable into the class of sofic groups coincide with the class of sofic groups. By applying Corollary 7.1.17, we then deduce the following:

**Corollary 7.5.13.** *Let $\Gamma$ be a group. Then the set of $\Gamma$-marked groups $N \in \mathcal{N}(\Gamma)$ such that $\Gamma/N$ is sofic is closed (and hence compact) for the prodiscrete topology on $\mathcal{N}(\Gamma) \subset \mathcal{P}(\Gamma) = \{0,1\}^\Gamma$.* □

We end this section by showing that extensions of sofic groups by amenable groups are sofic. This is a generalization of Proposition 7.5.6 since every amenable group is an extension of the trivial group by itself.

**Proposition 7.5.14.** *Let $G$ be a group. Suppose that $G$ contains a normal subgroup $N$ such that $N$ is sofic and $G/N$ is amenable. Then $G$ is sofic.*

*Proof.* Let $K \subset G$ be a finite subset and $0 < \varepsilon < 1$. Denote by $\overline{g}$ the image of an element $g \in G$ under the canonical epimorphism of $G$ onto $G/N$. Fix a set $T \subset G$ of representatives for the cosets of $N$ in $G$ and denote by $\sigma\colon G/N \to T$ the map which associates with each element in $G/N$ its representative in $T$. Note that $\sigma(\overline{g})^{-1}g \in N$ for all $g \in G$. Also set $\varepsilon' = 1 - \sqrt{1-\varepsilon}$, so that $0 < \varepsilon' < \varepsilon$ and $(1-\varepsilon')^2 = 1 - \varepsilon$.

Since $G/N$ is amenable and hence sofic by Proposition 7.5.6, there exist a nonempty finite set $F_1$ and a $(\overline{K}, \varepsilon')$-almost-homomorphism $\varphi_1\colon G/N \to \mathrm{Sym}(F_1)$. In fact, in the proof of Proposition 7.5.6 it is shown that we can take $F_1 \subset G/N$ such that there exists a subset $E_1 \subset F_1$ with $|E_1| \geq (1-\varepsilon')|F_1|$ satisfying $\varphi_1(\overline{k})(f_1) = \overline{k}f_1 \in F_1$ and $\varphi_1(\overline{hk})(f_1) = \overline{hk}f_1 \in F_1$ for all $\overline{h}, \overline{k} \in \overline{K}$ and $f_1 \in E_1$.

Set $M = N \cap (\sigma(F_1)^{-1} \cdot K \cdot \sigma(F_1)) \subset N$. As $N$ is sofic, we can find a finite set $F_2$ and an $(M, \varepsilon')$-almost-homomorphism $\varphi_2\colon N \to \mathrm{Sym}(F_2)$. Thus, for all $m, m' \in M$ we can find a set $E_2 \subset F_2$ such that $|E_2| \geq (1-\varepsilon)|F_2|$ and

$$\varphi_2(mm')(f_2) = \varphi_2(m)(\varphi_2(m')(f_2)) \quad \text{for all } f_2 \in E_2. \tag{7.25}$$

Set $F = F_1 \times F_2$ and $E = E_1 \times E_2$ and observe that

$$|E| = |E_1| \cdot |E_2| \geq (1-\varepsilon')^2 |F_1| \cdot |F_2| = (1-\varepsilon)|F|. \tag{7.26}$$

Consider the map $\Phi\colon G \to \mathrm{Sym}(F)$ defined by setting

$$\Phi(g)(f_1, f_2) = (\varphi_1(\overline{g})(f_1), \varphi_2(\sigma(\overline{g}f_1)^{-1} g \sigma(f_1))(f_2))$$

for all $g \in G$ and $(f_1, f_2) \in F$. Let us show that $\Phi$ is a $(K, \varepsilon)$-almost-homomorphism.

Let $h, k \in K$ and $(f_1, f_2) \in E$. Recall that the elements $\overline{k}f_1 = \varphi_1(\overline{k})(f_1)$ and $\overline{hk}f_1 = \varphi_1(\overline{hk})(f_1) = \varphi_1(\overline{h})(\varphi_1(\overline{k})(f_1))$ both belong to $F_1$ for all $f_1 \in E_1$. It follows that

$$\begin{aligned}
\Phi(h)\Phi(k)(f_1, f_2) &= \Phi(h)(\varphi_1(\overline{k})(f_1), \varphi_2(\sigma(\overline{k}f_1)^{-1} k \sigma(f_1))(f_2)) \\
&= (\varphi_1(\overline{h})(\varphi_1(\overline{k})(f_1)), \varphi_2(\sigma(\overline{h}\varphi_1(\overline{k})(f_1))^{-1} h \sigma(\varphi_1(\overline{k})(f_1))) \\
&\qquad (\varphi_2(\sigma(\overline{k}f_1)^{-1} k \sigma(f_1))(f_2))) \\
&= (\varphi_1(\overline{hk})(f_1), \varphi_2(\sigma(\overline{hk}f_1)^{-1} h \sigma(\overline{k}f_1)) \\
&\qquad (\varphi_2(\sigma(\overline{k}f_1)^{-1} k \sigma(f_1))(f_2))) \\
&=_* (\varphi_1(\overline{hk})f_1, \varphi_2(\sigma(\overline{hk}f_1)^{-1} hk \sigma(f_1))(f_2)) \\
&= \Phi(hk)(f_1, f_2)
\end{aligned}$$

where $=_*$ follows from (7.25) since $m = \sigma(\overline{hk}f_1)^{-1}h\sigma(\overline{k}f_1)$ and $m' = \sigma(\overline{k}f_1)^{-1}k\sigma(f_1)$ both belong to $(\sigma(F_1)^{-1}K\sigma(F_1)) \cap N = M$. It follows from (7.26) that $d_F(\Phi(hk), \Phi(h)\Phi(k)) \leq \varepsilon$.

Let now $h, k \in K$ be such that $h \neq k$. We distinguish two cases.

If $\overline{h} \neq \overline{k}$ then, as $\varphi_1$ is a $(\overline{K}, \varepsilon)$-almost-homomorphism, we have $d_{F_1}(\varphi_1(\overline{h}), \varphi_1(\overline{k})) \geq 1 - \varepsilon$. It follows that there exists a subset $B \subset F_1$ such that $|B| \geq (1-\varepsilon)|F_1|$ such that $\varphi_1(\overline{h})(b) \neq \varphi_1(\overline{k})(b)$ for all $b \in B$. Setting $B' = B \times F_2$ we have that

$$|B'| = |B| \cdot |F_2| \geq (1-\varepsilon)|F_1| \cdot |F_2| = (1-\varepsilon)|F| \tag{7.27}$$

and

$$\Phi(h)(b') \neq \Phi(k)(b') \text{ for all } b' \in B'. \tag{7.28}$$

Suppose now that $\overline{h} = \overline{k}$. As $\varphi_2$ is an $(M, \varepsilon)$-almost-homomorphism one has $d_{F_2}(\varphi_2(m), \varphi_2(m')) > 1 - \varepsilon$ for all $m, m' \in M$ such that $m \neq m'$. It follows that for all distinct $m, m' \in M$ there exists a subset $D \subset F_2$ such that $|D| \geq (1-\varepsilon)|F_2|$ and $\varphi_2(m)(d) \neq \varphi_2(m')(d)$ for all $d \in D$.

From $\overline{h} = \overline{k}$ we deduce that for all $f_1 \in F_1$ one has that if $m = \sigma(\overline{h}f_1)^{-1} \times h\sigma(f_1)$ and $m' = \sigma(\overline{k}f_1)^{-1}k\sigma(f_1)$, then $m, m' \in M$ and $m = \sigma(\overline{h}f_1)^{-1}h\sigma(f_1) = \sigma(\overline{k}f_1)^{-1}h\sigma(f_1) \neq \sigma(\overline{k}f_1)^{-1}k\sigma(f_1) = m'$. Thus $\varphi_2(\sigma(\overline{h}f_1)^{-1}h\sigma(f_1))(d) \neq \varphi_2(\sigma(\overline{k}f_1)^{-1}k\sigma(f_1))(d)$ for all $d \in D$. Set $D' = D \times F_2$ so that

$$|D'| = |D| \cdot |F_2| \geq (1-\varepsilon)|F_1| \cdot |F_2| = (1-\varepsilon)|F|. \tag{7.29}$$

It follows that for all $d' \in D'$ one has

$$\Phi(h)(d') \neq \Phi(k)(d') \text{ for all } d' \in D'. \tag{7.30}$$

From (7.28) and (7.30), and taking into account (7.27) and (7.29) respectively, we deduce that in either cases $d_F(\Phi(h), \Phi(k)) \geq 1 - \varepsilon$.

This shows that $\Phi\colon G \to \mathrm{Sym}(F)$ is a $(K, \varepsilon)$-almost-homomorphism. Thus $G$ is sofic. □

## 7.6 Sofic Groups and Metric Ultraproducts of Finite Symmetric Groups

Suppose that we are given a triple $T = (I, \omega, \mathcal{F})$ consisting of the following data: a set $I$, an ultrafilter $\omega$ on $I$, and a family $\mathcal{F} = (F_i)_{i \in I}$ of nonempty finite sets indexed by $I$. Our first goal in this section is to associate with such a triple $T$ a sofic group $G_T$. We start by forming the direct product group

$$P_T = \prod_{i \in I} \mathrm{Sym}(F_i).$$

Let $\alpha = (\alpha_i)_{i\in I}$ and $\beta = (\beta_i)_{i\in I}$ be elements of $P_T$. Since $0 \leq d_{F_i}(\alpha_i, \beta_i) \leq 1$ for all $i \in I$, it follows from Corollary J.2.6 that the Hamming distances $d_{F_i}(\alpha_i, \beta_i)$ have a limit

$$\delta_\omega(\alpha, \beta) = \lim_{i\to\omega} d_{F_i}(\alpha_i, \beta_i) \in [0, 1]$$

along the ultrafilter $\omega$.

**Proposition 7.6.1.** *One has:*

(i) $\delta_\omega(\alpha, \alpha) = 0$;
(ii) $\delta_\omega(\beta, \alpha) = \delta_\omega(\alpha, \beta)$;
(iii) $\delta_\omega(\alpha, \beta) \leq \delta_\omega(\alpha, \gamma) + \delta_\omega(\gamma, \beta)$;
(iv) $\delta_\omega(\gamma\alpha, \gamma\beta) = \delta_\omega(\alpha, \beta)$;
(v) $\delta_\omega(\alpha\gamma, \beta\gamma) = \delta_\omega(\alpha, \beta)$.

*for all* $\alpha, \beta, \gamma \in P_T$.

*Proof.* Let $\alpha = (\alpha_i)_{i\in I}$, $\beta = (\beta_i)_{i\in I}$, $\gamma = (\gamma_i)_{i\in I} \in P_T$. For each $i \in I$, we have $d_{F_i}(\alpha_i, \alpha_i) = 0$, $d_{F_i}(\beta_i, \alpha_i) = d_{F_i}(\alpha_i, \beta_i)$, $d_{F_i}(\alpha_i, \beta_i) \leq d_{F_i}(\alpha_i, \gamma_i) + d_{F_i}(\gamma_i, \beta_i)$, $d_{F_i}(\gamma_i\alpha_i, \gamma_i\beta_i) = d_{F_i}(\alpha_i, \beta_i)$, and $d_{F_i}(\alpha_i\gamma_i, \beta_i\gamma_i) = d_{F_i}(\alpha_i, \beta_i)$ since $d_{F_i}$ is a bi-invariant metric on $\mathrm{Sym}(F_i)$. This gives us properties (i), (ii), (iii), (iv), and (v) for $\delta_\omega$ by taking limits along $\omega$ (cf. Corollary J.2.10). □

Consider now the subset $N_T \subset P_T$ defined by

$$N_T = \{\alpha \in P_T : \delta_\omega(1_{P_T}, \alpha) = 0\}.$$

**Proposition 7.6.2.** *The set* $N_T$ *is a normal subgroup of* $P_T$.

*Proof.* This is an easy consequence of the properties of $\delta_\omega$ stated in Proposition 7.6.1. Indeed, we deduce from Property (i) that $\delta_\omega(1_{P_T}, 1_{P_T}) = 0$, that is, $1_{P_T} \in N_T$. On the other hand, by using successively (iv), (iii), and (ii), we get

$$\delta_\omega(1_{P_T}, \alpha^{-1}\beta) = \delta_\omega(\alpha, \beta) \leq \delta_\omega(\alpha, 1_{P_T}) + \delta_\omega(1_{P_T}, \beta) = \delta_\omega(1_{P_T}, \alpha) + \delta_\omega(1_{P_T}, \beta)$$

for all $\alpha, \beta \in P_T$. This implies that $\alpha^{-1}\beta \in N_T$ if $\alpha, \beta \in N_T$. Thus, $N_T$ is a subgroup of $P_T$. Finally, Properties (iv) and (v) imply that

$$\delta_\omega(1_{P_T}, \gamma\alpha\gamma^{-1}) = \delta_\omega(\gamma^{-1}, \alpha\gamma^{-1}) = \delta_\omega(\gamma^{-1}\gamma, \alpha) = \delta_\omega(1_{P_T}, \alpha)$$

for all $\alpha, \gamma \in P_T$. It follows that $\gamma\alpha\gamma^{-1} \in N_T$ for all $\alpha \in N_T$ and $\gamma \in P_T$. This shows that $N_T$ is a normal subgroup of $P_T$. □

Observe that, given $\alpha = (\alpha_i)_{i\in I}$ and $\beta = (\beta_i)_{i\in I}$ in $P_T$, one has

$$\alpha N_T = \beta N_T \iff \delta_\omega(\alpha, \beta) = 0. \tag{7.31}$$

Indeed, one has $\alpha N_T = \beta N_T$ if and only if $\alpha^{-1}\beta \in N_T$, that is, if and only if $\delta_\omega(1_{P_T}, \alpha^{-1}\beta) = 0$. This is equivalent to $\delta_\omega(\alpha, \beta) = 0$ by the left-invariance of $\delta_\omega$.

**Theorem 7.6.3.** *The group $G_T = P_T/N_T$ is sofic.*

*Proof.* Fix a finite subset $K \subset G_T$ and $\varepsilon > 0$. We want to show that there exist a nonempty finite set $F$ and a $(K, \varepsilon)$-almost-homomorphism $\varphi\colon G_T \to \mathrm{Sym}(F)$.

Choose a representative of each element $g \in G_T$, that is, an element $\widetilde{g} = (\widetilde{g}_i)_{i\in I} \in P_T$ such that $g = \widetilde{g}N_T$.

We first introduce some constants that will be used in the proof. If $h$ and $k$ are distinct elements in $K$, then $\delta_\omega(\widetilde{h}, \widetilde{k}) > 0$ by (7.31). Let us set

$$\eta = \min_{\substack{h,k\in K\\ h\neq k}} \frac{\delta_\omega(\widetilde{h}, \widetilde{k})}{2}. \tag{7.32}$$

Note that $0 < \eta \leq 1/2$.

Now choose an integer $m \geq \log\varepsilon/\log(1-\eta)$, so that

$$1 - (1-\eta)^m \geq 1 - \varepsilon. \tag{7.33}$$

Finally, choose a real number $\xi$ with $0 < \xi < 1$ sufficiently small to make

$$1 - (1-\xi)^m \leq \varepsilon. \tag{7.34}$$

If $h$ and $k$ are arbitrary elements of $K$, we have $\widetilde{hk}N_T = \widetilde{h}\widetilde{k}N_T$ and therefore $\delta_\omega(\widetilde{hk}, \widetilde{h}\widetilde{k}) = 0$ by (7.31). It follows that the set

$$A(h,k) = \{i \in I : d_{F_i}(\widetilde{hk}_i, \widetilde{h}_i\widetilde{k}_i) \leq \xi\} \tag{7.35}$$

belongs to $\omega$. On the other hand, if $h$ and $k$ are distinct elements of $K$, we have $\delta_\omega(\widetilde{h}, \widetilde{k}) \geq 2\eta$ by (7.32). As $\eta > 0$, this implies that the set

$$C(h,k) = \{i \in I : d_{F_i}(\widetilde{h}_i, \widetilde{k}_i) \geq \eta\} \tag{7.36}$$

belongs to $\omega$. As any finite intersection of elements of $\omega$ is in $\omega$ and therefore nonempty, we deduce that there exists an index $j \in I$ such that

$$j \in \left(\bigcap_{h,k\in K} A(h,k)\right) \bigcap \left(\bigcap_{\substack{h,k\in K\\ h\neq k}} C(h,k)\right).$$

Consider the map $\psi\colon G_T \to \mathrm{Sym}(F_j)$ defined by $\psi(g) = \widetilde{g}_j$ for all $g \in G_T$. It immediately follows from (7.35) and (7.36) that the map $\psi$ satisfies the following properties:

(1) $d_{F_j}(\psi(hk), \psi(h)\psi(k)) \leq \xi$ for all $h, k \in K$;
(2) $d_{F_j}(\psi(h), \psi(k)) \geq \eta$ for all $h, k \in K$ such that $h \neq k$.

Consider now the Cartesian product

$$F = \underbrace{F_j \times F_j \times \cdots \times F_j}_{m \text{ times}}$$

and the homomorphism

$$\Psi \colon \mathrm{Sym}(F_j) \to \mathrm{Sym}(F)$$

defined by

$$\Psi(\sigma)(x_1, x_2, \ldots, x_m) = (\sigma(x_1), \sigma(x_2), \ldots, \sigma(x_m))$$

for all $\sigma \in \mathrm{Sym}(F_j)$ and $(x_1, x_2, \ldots, x_m) \in F$. By Corollary 7.4.4 we have

$$d_F(\Psi(\sigma), \Psi(\sigma')) = 1 - (1 - d_{F_j}(\sigma, \sigma'))^m$$

for all $\sigma, \sigma' \in \mathrm{Sym}(F_j)$. It follows that the composite map

$$\varphi = \Psi \circ \psi \colon G \to \mathrm{Sym}(F)$$

satisfies the following properties:

(1') $d_F(\varphi(hk), \varphi(h)\varphi(k)) \leq 1 - (1 - \xi)^m$ for all $h, k \in K$;
(2') $d_F(\varphi(h), \varphi(k)) \geq 1 - (1 - \eta)^m$ for all $h, k \in K$ such that $h \neq k$.

As $1-(1-\xi)^m \leq \varepsilon$ by (7.34) and $1-(1-\eta)^m \geq 1-\varepsilon$ by (7.33), we deduce that $\varphi$ is a $(K, \varepsilon)$-almost-homomorphism. This shows that $G_T$ is a sofic group. □

*Remark 7.6.4.* When the ultrafilter $\omega$ is principal, then the group $G_T$ is finite. Indeed, in this case, there is an element $i_0 \in I$ such that $\omega$ consists of all subsets of $I$ containing $i_0$. This implies that $\delta_\omega(\alpha, \beta) = d_{F_{i_0}}(\alpha_{i_0}, \beta_{i_0})$ for all $\alpha = (\alpha_i)_{i \in I}, \beta = (\beta_i)_{i \in I} \in P_T$. Therefore, $N_T$ consists of all $\alpha \in P_T$ such that $\alpha_{i_0} = 1_{\mathrm{Sym}(F_{i_0})}$. It follows that the group $G_T$ is isomorphic to the group $\mathrm{Sym}(F_{i_0})$.

*Remark 7.6.5.* Let $\alpha, \alpha', \beta, \beta' \in P_T$ such that $\alpha N_T = \alpha' N_T$ and $\beta N_T = \beta' N_T$. By applying Proposition 7.6.1(iii) and (7.31), we get

$$d_\omega(\alpha, \beta) \leq d_\omega(\alpha, \alpha') + d_\omega(\alpha', \beta') + d_\omega(\beta', \beta) = d_\omega(\alpha', \beta').$$

By exchanging the roles of $\alpha$ and $\alpha'$ and of $\beta$ and $\beta'$, we obtain $d_\omega(\alpha', \beta') \leq d_\omega(\alpha, \beta)$. It follows that $d_\omega(\alpha, \beta) = d_\omega(\alpha', \beta')$. Therefore, if $g, h \in G_T$ and $\alpha, \beta \in P_T$ are such that $g = \alpha N_T$ and $h = \beta N_T$, the quantity

$$\Delta_\omega(g, h) = \delta_\omega(\alpha, \beta) \tag{7.37}$$

is well defined. Moreover, the map $\Delta_\omega \colon G_T \times G_T \to [0,1]$ given by (7.37) is a bi-invariant metric on $G_T$. This follows immediately from Proposition 7.6.1 taking into account that $N_T$ is a normal subgroup.

**Theorem 7.6.6.** *Let $G$ be a group. The following conditions are equivalent:*

(a) *$G$ is sofic;*
(b) *there exists a triple $T = (I, \omega, \mathcal{F})$, where $I$ is a set, $\omega$ is an ultrafilter on $I$, and $\mathcal{F} = (F_i)_{i \in I}$ is a family of nonempty finite sets indexed by $I$, such that $G$ is isomorphic to a subgroup of the group $G_T = P_T/N_T$.*

*Proof.* If $G$ satisfies (b), then $G$ is sofic since $G_T$ is sofic by Theorem 7.6.3 and every subgroup of a sofic group is itself sofic by Proposition 7.5.4.

Conversely, suppose that $G$ is sofic. Consider the set $I$ consisting of all pairs $(K, \varepsilon)$, where $K$ is a finite subset of $G$ and $\varepsilon > 0$. We partially order the set $I$ by setting $(K, \varepsilon) \preceq (K', \varepsilon')$ if $K \subset K'$ and $\varepsilon' \leq \varepsilon$.

For each $i = (K, \varepsilon) \in I$ we define the set

$$I_i = \{j \in I : i \preceq j\} \subset I.$$

Observe that $I_i \neq \varnothing$ as $i \in I_i$. Moreover, the family of nonempty subsets $\{I_i\}_{i \in I}$ is closed under finite intersections, since

$$I_{(K_1, \varepsilon_1)} \cap I_{(K_2, \varepsilon_2)} = I_{(K_1 \cup K_2, \min\{\varepsilon_1, \varepsilon_2\})}$$

for all $(K_1, \varepsilon_1), (K_2, \varepsilon_2) \in I$. It follows from Proposition J.1.3 and Theorem J.1.6 that there exists an ultrafilter $\omega$ on $I$ such that $I_{(K,\varepsilon)} \in \omega$ for all $(K, \varepsilon) \in I$.

As $G$ is sofic, we can find, for each $i = (K, \varepsilon) \in I$, a nonempty finite set $F_i$ and a $(K, \varepsilon)$-almost-homomorphism $\varphi_i \colon G \to \mathrm{Sym}(F_i)$. Consider the triple $T = (I, \omega, \mathcal{F})$, where $\mathcal{F} = (F_i)_{i \in I}$, and the associated sofic group $G_T = P_T/N_T$. Let $\widetilde{\varphi} \colon G \to P_T$ denote the product map $\widetilde{\varphi} = \prod_{i \in I} \varphi_i$. Thus, we have

$$\widetilde{\varphi}(g) = (\varphi_i(g))_{i \in I}$$

for all $g \in G$. Let $\rho \colon P_T \to G_T = P_T/N_T$ denote the canonical epimorphism. Let us show that the composite map $\Phi = \rho \circ \widetilde{\varphi} \colon G \to G_T$ is an injective homomorphism. This will prove that $G$ is isomorphic to a subgroup of $G_T$.

Let $g, h \in G$ and let $\eta > 0$. Consider the element $i_0 = (\{g, h\}, \eta) \in I$. If $i = (K, \varepsilon) \in I$ satisfies $i_0 \preceq i$, then

$$d_{F_i}(\varphi_i(gh), \varphi_i(g)\varphi_i(h)) \leq \varepsilon \leq \eta$$

since $\varphi_i$ is a $(K, \varepsilon)$-almost-homomorphism. Thus, the set

$$\{i \in I : d_{F_i}(\varphi_i(gh), \varphi_i(g)\varphi_i(h)) \leq \eta\}$$

contains $I_{i_0}$ and therefore belongs to $\omega$. This implies that

$$\delta_\omega(\widetilde{\varphi}(gh), \widetilde{\varphi}(g)\widetilde{\varphi}(h)) = \lim_{i\to\omega} d_{F_i}(\varphi_i(gh), \varphi_i(g)\varphi_i(h)) = 0.$$

Therefore, we have $\Phi(gh) = \Phi(g)\Phi(h)$ by (7.31). This shows that $\Phi$ is a homomorphism.

On the other hand, if $g \neq h$, we have

$$d_{F_i}(\varphi_i(g), \varphi_i(h)) \geq 1 - \varepsilon \geq 1 - \eta$$

for all $i \in I$ such that $i_0 \preceq i$. This implies that

$$\delta_\omega(\widetilde{\varphi}(g), \widetilde{\varphi}(h)) = \lim_{i\to\omega} d_{F_i}(\varphi_i(g), \varphi_i(h)) = 1. \tag{7.38}$$

Therefore, we have $\Phi(g) \neq \Phi(h)$ by (7.31). Consequently, $\Phi$ is injective. This shows that (a) implies (b). □

*Remarks 7.6.7.* (a) If $G$ is an infinite sofic group and $T = (I, \omega, \mathcal{F})$ is a triple satisfying condition (b) in Theorem 7.6.6, then the ultrafilter $\omega$ is necessarily non-principal by Remark 7.6.4.

(b) Let $G$ be a sofic group and let $\Phi\colon G \to G_T$ be as in the proof of Theorem 7.6.6. Using the notation from Remark 7.6.5, we deduce from (7.38) that $\Delta_\omega(\Phi(g), \Phi(h)) = 1$ for all $g, h \in G$ such that $g \neq h$. It follows that the restriction of the bi-invariant metric $\Delta_\omega$ to the subgroup $\Phi(G) \subset G_T$ is the discrete metric.

## 7.7 A Characterization of Finitely Generated Sofic Groups

In this section we give a geometric characterization of finitely generated sofic groups in terms of a finiteness condition on their Cayley graphs.

Let $G$ be a finitely generated group and let $S$ be a finite symmetric generating subset of $G$. Given $r \in \mathbb{N}$, we denote by $B_S(r)$ the ball of radius $r$ centered at the vertex corresponding to the identity element $1_G$ of $G$ in the Cayley graph $\mathcal{C}_S(G)$ of $G$ with respect to $S$, with the induced $S$-labeled graph structure (cf. Sects. 6.1, 6.2 and 6.3).

Let also $\mathcal{Q} = (Q, E)$ be an $S$-labeled graph. Given $q \in Q$ and $r \in \mathbb{N}$, we denote by $B(q, r)$ the ball of radius $r$ centered at $q$ with the induced $S$-labeled graph structure.

Given $r \in \mathbb{N}$ we denote by $Q(r)$ the set of all $q \in Q$ such that there exists an $S$-labeled graph isomorphism

$$\psi_{q,r}\colon B_S(r) \to B(q, r) \tag{7.39}$$

satisfying

$$\psi_{q,r}(1_G) = q. \tag{7.40}$$

Observe that if such a map $\psi_{q,r}$ exists it is unique. We have the inclusions

$$Q = Q(0) \supset Q(1) \supset Q(2) \supset \cdots \supset Q(r) \supset Q(r+1) \supset \cdots \tag{7.41}$$

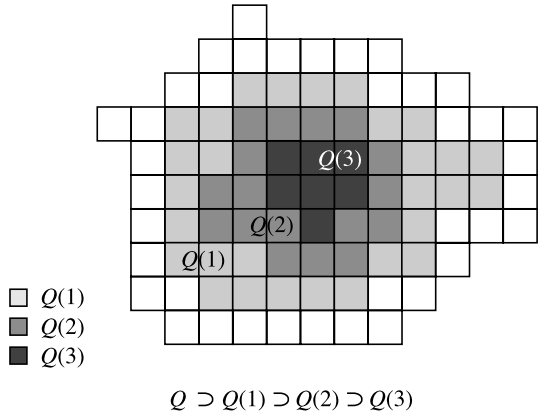



**Fig. 7.2** The inclusions $Q \supset Q(1) \supset Q(2) \supset Q(3)$

Note also that since $\mathcal{C}_S(G)$ (and therefore the induced $S$-labeled subgraph $B_S(r)$) is edge-symmetric (with respect to the involution $s \mapsto s^{-1}$ on $S$ (cf. Sect. 6.3)) and $\psi_{q,r}$ is an $S$-labeled graph isomorphism, then $B(q,r) = \psi_{q,r}(B_S(r))$ is edge-symmetric as well.

**Theorem 7.7.1.** *Let $G$ be a finitely generated group and let $S$ be a finite symmetric generating subset of $G$. The following conditions are equivalent:*

(a) *the group $G$ is sofic;*
(b) *for all $\varepsilon > 0$ and $r \in \mathbb{N}$, there exists a finite $S$-labeled graph $\mathcal{Q} = (Q, E)$ such that*

$$|Q(r)| \geq (1-\varepsilon)|Q|, \tag{7.42}$$

*where $Q(r) \subset Q$ denotes the set consisting of all vertices $q \in Q$ for which there exists an $S$-labeled graph isomorphism $\psi_{q,r} \colon B_S(r) \to B(q,r)$ from the ball $B_S(r)$ in the Cayley graph $\mathcal{C}_S(G)$ of $G$ with respect to $S$ onto the ball $B(q,r)$ in $\mathcal{Q}$ satisfying $\psi_{q,r}(1_G) = q$.*

Before starting the proof of the theorem we present some preliminary results.

**Lemma 7.7.2.** *Let $\mathcal{Q} = (Q, E)$ be an $S$-labeled graph and $r_0, i \in \mathbb{N}$. Suppose that $q_0 \in Q((i+1)r_0)$. Then $B(q_0, r_0) \subset Q(ir_0)$.*

*Proof.* Let $q' \in B(q, r_0)$ and let us show that $q' \in Q(ir_0)$. It follows from the triangle inequality that the ball $B(q', ir_0)$ is entirely contained in the ball $B(q_0, (i+1)r_0)$. Moreover, since $\psi_{q_0,(i+1)r_0}$ is isometric, setting $g =$

$\psi^{-1}_{q_0,(i+1)r_0}(q')$, we have $g \in B_S(r_0)$ so that $gh \in B((i+1)r_0)$ for all $h \in B(ir_0)$. It follows that the map

$$\psi_{q',ir_0} \colon B_S(ir_0) \to B(q', ir_0) \tag{7.43}$$

defined by $\psi_{q',ir_0}(h) = \psi_{q_0,(i+1)r_0}(gh)$ for all $h \in B_S(ir_0)$ yields an $S$-labeled graph isomorphism satisfying $\psi_{q',ir_0}(1_G) = \psi_{q_0,(i+1)r_0}(g) = q'$. This shows that $q' \in Q(ir_0)$. We deduce that $B(q_0, r_0) \subset Q(ir_0)$. □

**Lemma 7.7.3.** *Let $\mathcal{Q} = (Q, E)$ be an $S$-labeled graph and $r_0 \in \mathbb{N}$. Let $q_1, q_2 \in Q(2r_0)$ such that $q_1 \neq q_2$ and $g \in B_S(r_0)$. Then we have*

$$\psi_{q_1,2r_0}(g) \neq \psi_{q_2,2r_0}(g). \tag{7.44}$$

*Proof.* If $g = 1_G$ we have $\psi_{q_1,2r_0}(g) = \psi_{q_1,2r_0}(1_G) = q_1 \neq q_2 = \psi_{q_2,2r_0}(1_G) = \psi_{q_2,2r_0}(g)$. Suppose now that $g \neq 1_G$. Suppose by contradiction that $\psi_{q_1,2r_0}(g) = \psi_{q_2,2r_0}(g) = q_0$. Since $\psi_{q_1,2r_0}$ is isometric we have $q_0 \in B(q_1, r_0)$. It follows from Lemma 7.7.2 that $q_0 \in Q(r_0)$.

As $g \in B_S(r_0)$, we can find $1 \leq r' \leq r_0$ and $s_1, s_2, \ldots, s_{r'} \in S$ such that $g = s_1 s_2 \cdots s_{r'}$. Consider the path

$$\pi = ((1_G, s_1, s_1), (s_1, s_2, s_1 s_2), \ldots, (s_1 s_2 \cdots s_{r'-1}, s_{r'}, g))$$

and observe that it is contained in $B_S(r_0)$. Now, $\pi$ is mapped by $\psi_{q_1,2r_0}$ and $\psi_{q_2,2r_0}$ into two paths $\pi_1$ and $\pi_2$ in $\mathcal{Q}$ with initial vertices $\pi_1^- = q_1$, $\pi_2^- = q_2$ and same terminal vertex $\pi_1^+ = \psi_{q_1,2r_0}(g) = q_0 = \psi_{q_2,2r_0}(g) = \pi_2^+$. Note that since $\psi_{q_1,2r_0}$ and $\psi_{q_2,2r_0}$ are isometric $\pi_1$ and $\pi_2$ are both contained in $B(q_0, r_0)$. Moreover, since $\psi_{q_1,2r_0}$ and $\psi_{q_2,2r_0}$ are label-preserving, $\pi_1$ and $\pi_2$ have the same label $s_1 s_2 \cdots s_{r'}$. The inverse images of $\pi_1$ and $\pi_2$ under the $S$-label preserving graph isomorphism $\psi_{q_0,r_0} \colon B_S(r_0) \to B(q_0, r_0)$ are both equal to $\pi$. Indeed in a Cayley graph there exists a unique path which ends at a given vertex and with a given label. It follows that $\pi_1 = \pi_2$ and therefore $q_1 = \pi_1^- = \pi_2^- = q_2$. This contradicts our assumptions. We deduce that $\psi_{q_1,2r_0}(g) \neq \psi_{q_2,2r_0}(g)$. □

**Lemma 7.7.4.** *Let $\mathcal{Q} = (Q, E)$ be an $S$-labeled graph and $r_0 \in \mathbb{N}$. Let $h, k \in B_S(r_0)$ and $q_0 \in Q(2r_0)$. We have*

$$\psi_{q_0,2r_0}(h) \in Q(r_0) \tag{7.45}$$

*and*

$$\psi_{q_0,2r_0}(hk) = \psi_{\psi_{q_0,2r_0}(h),r_0}(k), \tag{7.46}$$

*where $\psi_{q,r}$, $q \in Q(r)$, $r \in \mathbb{N}$, is as in (7.39) and (7.40).*

*Proof.* Since $\psi_{q_0,2r_0}$ is isometric we have $\psi_{q_0,2r_0}(h) \in B(q_0, r_0)$. Then (7.45) follows from Lemma 7.7.2. Let us show (7.46). First note that (7.46) makes sense by virtue of (7.45). If $k = 1_G$ (7.46) follows from $\psi_{q_0,2r_0}(hk) =$

$\psi_{q_0,2r_0}(h) = \psi_{\psi_{q_0,2r_0}(h),r_0}(1_G) = \psi_{\psi_{q_0,2r_0}(h),r_0}(k)$. Now suppose that $k \neq 1_G$. Then we can find $1 \leq r' \leq r_0$ and $s_1, s_2, \ldots, s_{r'} \in S$ such that $k = s_1 s_2 \cdots s_{r'}$. Consider the path

$$\pi_1 = ((h, s_1, hs_1), (hs_1, s_2, hs_1 s_2), \ldots, (hs_1 s_2 \cdots s_{r'-1}, s_{r'}, hk))$$

and observe that it is contained in $B_S(2r_0)$, since $h, k \in B_S(r_0)$. The path $\pi_1$ is mapped by $\psi_{q_0,2r_0}$ into the path

$$\begin{aligned}\overline{\pi}_1 = ((\psi_{q_0,2r_0}(h), s_1, \psi_{q_0,2r_0}(hs_1)), (\psi_{q_0,2r_0}(hs_1), s_2, \psi_{q_0,2r_0}(hs_1 s_2)), \ldots \\ \ldots, (\psi_{q_0,2r_0}(hs_1 s_2 \cdots s_{r'-1}), s_{r'}, \psi_{q_0,2r_0}(hk))).\end{aligned}$$

As $\psi_{q_0,2r_0}$ is isometric, we have that $q' = \psi_{q_0,2r_0}(h)$ belongs to $B(q_0, r_0)$ and since $q_0 \in Q(2r_0)$ we deduce from Lemma 7.7.2 that $q' \in Q(r_0)$. Consider the inverse image of the path $\overline{\pi}_1$ under the $S$-labeled graph isomorphism $\psi_{q',r_0}$. Since $\psi^{-1}_{q',r_0}((\overline{\pi}_1)^-) = \psi^{-1}_{q',r_0}(\psi_{q_0,2r_0}(h)) = \psi^{-1}_{q',r_0}(q') = 1_G$ and $\psi_{q',r_0}$ preserves the label, this inverse image is necessarily equal to the path

$$\pi_2 = ((1_G, s_1, s_1), (s_1, s_2, s_1 s_2), \ldots, (s_1 s_2 \cdots s_{r'-1}, s_{r'}, k)),$$

since in a Cayley graph there exists a unique path which starts at a given vertex and with a given label. It follows that $\psi^{-1}_{q',r_0}(\psi_{q_0,2r_0}(hk)) = \psi^{-1}_{q',r_0}(\overline{\pi}_1^+) = \pi_2^+ = k$, that is, $\psi_{q_0,2r_0}(hk) = \psi_{q',r_0}(k)$. Since $q' = \psi_{q_0,2r_0}(h)$, we deduce (7.46). □

We are now in position to prove Theorem 7.7.1.

*Proof of Theorem 7.7.1.* Suppose that $G$ is sofic. Fix $\varepsilon > 0$ and $r \in \mathbb{N}$. Set $K = B_S(2r+1)$ and

$$\varepsilon' = \varepsilon(1 + |B_S(r)| \cdot |S| + |B_S(r)|^2)^{-1}. \tag{7.47}$$

Since $G$ is sofic, there exists a nonempty finite set $F$ and a $(K, \varepsilon')$-almost-homomorphism $\varphi\colon G \to \mathrm{Sym}(F)$. We construct an $S$-labeled graph $\mathcal{Q} = (Q, E)$ as follows. We take as vertex set $Q = F$. Then, as set of edges we take the set $E \subset Q \times S \times Q$ consisting of all the triples $(q, s, \varphi(s^{-1})(q))$, were $q \in Q$ and $s \in S$. Note that $\mathcal{Q}$ may have loops and multiple edges and that $\mathcal{Q}$ is not necessarily edge-symmetric with respect to the involution $s \mapsto s^{-1}$ on $S$. Observe however that, if $q \in Q$ and $s \in S$ are fixed, then there exists a unique edge in $\mathcal{Q}$ with initial vertex $q$ and label $s$.

For each $q \in Q$ denote by $\psi_q\colon G \to Q$ the map defined by setting $\psi_q(g) = \varphi(g^{-1})(q)$ for all $g \in G$. Denote by $Q_0$ the subset of $Q$ consisting of all $q \in Q$ satisfying the following conditions:

$(*)$ $\psi_q(1_G) = q$,
$(**)$ $\psi_q(gs) = \psi_{\psi_q(g)}(s)$ for all $g \in B_S(r)$ and $s \in S$,
$(***)$ $\psi_q(g) \neq \psi_q(h)$ for all $g, h \in B_S(r)$ with $g \neq h$.

Suppose that $q \in Q_0$. Let $g \in B_S(r)$. If $g = 1_G$ we have $\psi_q(g) = \psi_q(1_G) = q \in B(q,r)$, by $(*)$. If $g \neq 1_G$, then there exist $1 \leq r' \leq r$ and $s_1, s_2, \ldots, s_{r'} \in S$ such that $g = s_1 s_2 \cdots s_{r'}$. Consider the sequence of edges

$$\begin{aligned}
e_1 &= (q, s_1, \varphi(s_1^{-1})(q)) = (q, s_1, \psi_q(s_1)),\\
e_2 &= (\psi_q(s_1), s_2, \varphi(s_2^{-1})\psi_q(s_1))\\
&= (\psi_q(s_1), s_2, \psi_{\psi_q(s_1)}(s_2))\\
&= (\psi_q(s_1), s_2, \psi_q(s_1 s_2)) \quad \text{(by } (**)\text{)},\\
&\cdots\cdots\\
e_{r'} &= (\psi_q(s_1 s_2 \cdots s_{r'-1}), s_{r'}, \varphi(s_{r'}^{-1})(\psi_q(s_1 s_2 \cdots s_{r'-1})))\\
&= (\psi_q(s_1 s_2 \cdots s_{r'-1}), s_{r'}, \psi_{\psi_q(s_1 s_2 \cdots s_{r'-1})}(s_{r'}))\\
&= (\psi_q(s_1 s_2 \cdots s_{r'-1}), s_{r'}, \psi_q(s_1 s_2 \cdots s_{r'-1} s_{r'})) \quad \text{(by } (**)\text{)}\\
&= (\psi_q(s_1 s_2 \cdots s_{r'-1}), s_{r'}, \psi_q(g)).
\end{aligned}$$

The path $\pi = (e_1, e_2, \ldots, e_{r'})$ connects $q$ to $\psi_q(g)$ and has length $\ell(\pi) \leq r$. This shows that the graph distance from $q$ to $\psi_q(g)$ in $\mathcal{Q}$ does not exceed $r$ so that $\psi_q(B_S(r)) \subset B(q,r)$. Conversely, let $q' \in B(q,r)$. If $q' = q$ then by $(*)$ we have $q' = q = \psi_q(1_G) \in \psi_q(B_S(r))$. If $q \neq q'$ then there exist $1 \leq r' \leq r$ and a sequence of edges $(q, s_1, q_1), (q_1, s_2, q_2), \ldots, (q_{r'-1}, s_{r'}, q') \in E$. Using $(**)$ as above, we get $q' = \psi_q(g)$, where $g = s_1 s_2 \cdots s_{r'} \in B_S(r)$. This shows that $B(q,r) \subset \psi_q(B_S(r))$. Thus

$$B(q,r) = \psi_q(B_S(r)).$$

Moreover, by condition $(***)$, the map $\psi_q|_{B_S(r)}$ is injective. Finally, from $(**)$ we deduce that if $g \in B_S(r)$ and $s \in S$ then we have

$$(\psi_q(g), s, \psi_q(gs)) = (\psi_q(g), s, \psi_{\psi_q(g)}(s)) = (\psi_q(g), s, \varphi(s^{-1})\psi_q(g)) \in E.$$

Thus, the map $\psi_{q,r} = \psi_q|_{B_S(r)} \colon B_S(r) \to B(q,r)$ is an $S$-labeled graph isomorphism such that $\psi_{q,r}(1_G) = \psi_q(1_G) = q$, where the last equality follows from $(*)$. We deduce that $Q_0 \subset Q(r)$.

Let's now estimate from below the cardinality of $Q_0$. We denote by $d_Q$ the normalized Hamming metric on $\mathrm{Sym}(Q)$.

In order to estimate the cardinality of the set of $q \in Q$ for which condition $(*)$ is satisfied, let us first observe that

$$d_Q(\mathrm{Id}_Q, \varphi(1_G)) \leq \varepsilon'. \tag{7.48}$$

Indeed, since $\varphi$ is a $(K, \varepsilon')$-almost-homomorphism, taking $k_1 = k_2 = 1_G$ in property (i) of Definition 7.5.1, we deduce that $d_Q(\varphi(1_G), \varphi(1_G)\varphi(1_G)) \leq \varepsilon'$ which, by left-invariance of $d_Q$, implies (7.48). From (7.48) we deduce the

existence of a subset $Q' \subset Q$ of cardinality $|Q'| \leq \varepsilon' |Q|$ such that

$$\varphi(1_G)(q) = q \tag{7.49}$$

for all $q \in Q \setminus Q'$. It follows that $\psi_q(1_G) = \varphi(1_G^{-1})(q) = \varphi(1_G)(q) = q$, for all $q \in Q \setminus Q'$ so that condition $(*)$ is satisfied in $Q \setminus Q'$.

Let now estimate the cardinality of the set of $q \in Q$ for which condition $(**)$ is satisfied. Let $g \in B_S(r)$ and $s \in S$. As $\varphi$ is a $(K, \varepsilon')$-almost-homomorphism we have $d_Q(\varphi(s^{-1}g^{-1}), \varphi(s^{-1})\varphi(g^{-1})) \leq \varepsilon'$ so that there exists a subset $Q''(g,s) \subset Q$ with $|Q''(g,s)| \leq \varepsilon' |Q|$ such that

$$\begin{aligned} \psi_q(gs) &= \varphi((gs)^{-1})(q) \\ &= \varphi(s^{-1}g^{-1})(q) \\ &= \varphi(s^{-1})\varphi(g^{-1})(q) \\ &= \varphi(s^{-1})(\psi_q(g)) \\ &= \psi_{\psi_q(g)}(s) \end{aligned}$$

for all $q \in Q \setminus Q''(g,s)$. Setting

$$Q'' = \bigcup_{\substack{g \in B_S(r) \\ s \in S}} Q''(g,s)$$

we have $|Q''| \leq |B_S(r)| \cdot |S| \varepsilon' |Q|$ and condition $(**)$ holds for all $q \in Q \setminus Q''$.

Fix now two distinct elements $g, h \in B_S(r)$. Since $\varphi$ is a $(K, \varepsilon')$-almost-homomorphism and $g^{-1}, h^{-1} \in K$, by property (ii) of Definition 7.5.1 we deduce that $d_Q(\varphi(g^{-1}), \varphi(h^{-1})) \geq 1 - \varepsilon'$. Thus we can find a subset $Q'''(g,h) \subset Q$ of cardinality $|Q'''(g,h)| \leq \varepsilon' |Q|$ such that

$$\psi_q(g) = \varphi(g^{-1})(q) \neq \varphi(h^{-1})(q) = \psi_q(h) \tag{7.50}$$

for all $q \in Q \setminus Q'''(g,h)$.

Let now $g \neq h$ vary in $B_S(r)$ and set

$$Q''' = \bigcup_{\substack{g,h \in B_S(r) \\ g \neq h}} Q'''(g,h).$$

Observe that $|Q'''| \leq |B_S(r)|^2 \varepsilon' |Q|$. It follows from (7.50) that condition $(***)$ holds for all $q \in Q \setminus Q'''$.

In conclusion, conditions $(*)$, $(**)$ and $(***)$ are satisfied for all $q \in Q$ outside of $Q' \cup Q'' \cup Q'''$. We have

$$|Q' \cup Q'' \cup Q'''| \leq (1 + |B_S(r)| \cdot |S| + |B_S(r)|^2) \varepsilon' |Q| = \varepsilon |Q|,$$

where the equality follows from (7.47). We deduce that

$$|Q(r)| \geq |Q_0| \geq (1-\varepsilon)|Q|.$$

Thus $G$ satisfies condition (b). This shows (a) $\Rightarrow$ (b).

Conversely, suppose (b). Fix a finite set $K \subset G$ and $\varepsilon > 0$. Let $r_0 \in \mathbb{N}$ be such that $K \cup K^2 \subset B_S(r_0)$. Let $\mathcal{Q} = (Q, E)$ be the finite $S$-labeled graph given by condition (b) corresponding to $r = 2r_0$ and $\varepsilon$.

Let $g \in B_S(r_0)$. Since the map from $Q(2r_0)$ into $Q$ defined by $q \mapsto \psi_{q,2r_0}(g)$ is injective (by Lemma 7.44), we have that

$$|\{\psi_{q,2r_0}(g) : q \in Q(2r_0)\}| = |Q(2r_0)| \tag{7.51}$$

and therefore

$$|Q \setminus \{\psi_{q,2r_0}(g) : q \in Q(2r_0)\}| = |Q \setminus Q(2r_0)|. \tag{7.52}$$

As a consequence, there exists a bijection $\alpha_g \colon Q \setminus Q(2r_0) \to Q \setminus \{\psi_{q,2r_0}(g) : q \in Q(2r_0)\}$. Since $\psi_{q,2r_0}(1_G) = q$ for all $q \in Q(2r_0)$, we have $\{\psi_{q,2r_0}(1_G) : q \in Q(2r_0)\} = Q(2r_0)$ and therefore we can take

$$\alpha_{1_G} = \mathrm{Id}_{Q \setminus Q(2r_0)}\,. \tag{7.53}$$

Consider now the map $\varphi \colon G \to \mathrm{Sym}(Q)$ defined by

$$\varphi(g)(q) = \begin{cases} \psi_{q,2r_0}(g^{-1}) & \text{if } g \in B_S(r_0) \text{ and } q \in Q(2r_0); \\ \alpha_{g^{-1}}(q) & \text{if } g \in B_S(r_0) \text{ and } q \in Q \setminus Q(2r_0); \\ q & \text{otherwise.} \end{cases} \tag{7.54}$$

Note that $\varphi(g) \in \mathrm{Sym}(Q)$ for all $g \in G$, by construction.

Let us show that the map $\varphi \colon G \to \mathrm{Sym}(Q)$ is a $(K, \varepsilon)$-almost-homomorphism. Let $k_1, k_2 \in K \subset B_S(r_0)$ and $q \in Q(2r_0)$. We have

$$\begin{aligned} \varphi(k_1 k_2)(q) &= \psi_{q,2r_0}(k_2^{-1} k_1^{-1}) \\ &= \psi_{\psi_{q,2r_0}(k_2^{-1}),r_0}(k_1^{-1}) \quad \text{(by (7.46))} \\ &= \varphi(k_1)(\psi_{q,2r_0}(k_2^{-1})) \\ &= [\varphi(k_1)\varphi(k_2)](q). \end{aligned}$$

This shows that on $Q(2r_0)$ we have $\varphi(k_1k_2) = \varphi(k_1)\varphi(k_2)$. As

$$|Q(2r_0)| = |Q(r)| \geq (1-\varepsilon)|Q| \tag{7.55}$$

we deduce that $d_Q(\varphi(k_1k_2), \varphi(k_1)\varphi(k_2)) \leq \varepsilon$.

Finally, suppose that $k_1 \neq k_2$. We have $\varphi(k_1)(q) = \psi_{q,2r_0}(k_1^{-1}) \neq \psi_{q,2r_0}(k_2^{-1}) = \varphi(k_2)(q)$, since $k_1^{-1}, k_2^{-1} \in B_S(2r_0)$ and $\psi_{q,2r_0}$ is injective.

From (7.55) we deduce that $d_Q(\varphi(k_1), \varphi(k_2)) \geq 1 - \varepsilon$. It follows that $\varphi$ is a $(K, \varepsilon)$-almost-homomorphism. Therefore $G$ is sofic. This shows that (b) $\Rightarrow$ (a). □

## 7.8 Surjunctivity of Sofic Groups

In this section we prove that sofic groups are surjunctive. Note that this result covers the fact that locally residually amenable groups are surjunctive, which had been previously established in Corollary 5.9.3. Indeed, all locally residually amenable groups are sofic by Corollary 7.5.11.

**Theorem 7.8.1 (Gromov-Weiss).** *Every sofic group is surjunctive.*

Let us first establish the following:

**Lemma 7.8.2.** *Let $G$ be a group, $A$ a finite set, and equip $A^G$ with the prodiscrete topology. Let $X \subset A^G$ be a closed $G$-invariant subset and let $f\colon X \to A^G$ be a continuous $G$-equivariant map. Then there exists a cellular automaton $\tau\colon A^G \to A^G$ such that $f = \tau|_X$.*

*Proof.* From the continuity of $f$ we deduce the existence of a finite set $S \subset G$ such that if two configurations $y, z \in X$ coincide on $S$ then $f(y)(1_G) = f(z)(1_G)$. Let $a_0 \in A$ and consider the map $\mu\colon A^S \to A$ defined by setting

$$\mu(u) = \begin{cases} f(x)(1_G) & \text{if there exists } x \in X \text{ such that } x|_S = u \\ a_0 & \text{otherwise} \end{cases}$$

for all $u \in A^S$. Then $\mu$ is well defined and if we denote by $\tau\colon A^G \to A^G$ the cellular automaton with memory set $S$ and local defining map $\mu$, by $G$-equivariance of $f$ we clearly have $\tau|_X = f$. □

*Proof of Theorem 7.8.1.* Let $G$ be a sofic group. Let $A$ be a finite set of cardinality $|A| \geq 2$ and let $\tau\colon A^G \to A^G$ be an injective cellular automaton. We want to show that $\tau$ is surjective. Every subgroup of a sofic group is sofic by Proposition 7.5.4. On the other hand it follows from Proposition 3.2.2 that a group is surjunctive if all its finitely generated subgroups are surjunctive. Thus we can assume that $G$ is finitely generated.

Let then $S \subset G$ be a finite symmetric generating subset of $G$. As usual, for $r \in \mathbb{N}$, we denote by $B_S(r) \subset G$ the ball of radius $r$ centered at $1_G$ in the Cayley graph of $G$ with respect to $S$. We set $Y = \tau(A^G)$. Observe that $Y$ is $G$-invariant and, by Lemma 3.3.2, it is closed in $A^G$.

The inverse map $\tau^{-1}\colon Y \to A^G$ is $G$-equivariant and, by compactness of $A^G$, it is also continuous. By Lemma 7.8.2, there exists a cellular automaton $\sigma\colon A^G \to A^G$ such that $\sigma|_Y = \tau^{-1}\colon Y \to A^G$. Choose $r_0$ large enough so

that the ball $B_S(r_0)$ is a memory set for both $\tau$ and $\sigma$. Let $\mu\colon A^{B_S(r_0)} \to A$ and $\nu\colon A^{B_S(r_0)} \to A$ denote the corresponding local defining maps for $\tau$ and $\sigma$ respectively.

We proceed by contradiction. Suppose that $\tau$ is not surjective, that is, $Y \subsetneq A^G$. Then, since $Y$ is closed in $A^G$, there exists a finite subset $\Omega \subset G$ such that $\pi_\Omega(Y) \subsetneq A^\Omega$, where, for a subset $E \subset G$, we denote by $\pi_E\colon A^G \to A^E$ the projection map. It is not restrictive, up to taking a larger $r_0$, again if necessary, to suppose that $\Omega \subset B_S(r_0)$. Thus, $\pi_{B_S(r_0)}(Y) \subsetneq A^{B_S(r_0)}$.

Fix $\varepsilon > 0$ such that

$$\varepsilon < 1 - \frac{|B(2r_0)| \cdot \log|A|}{|B(2r_0)| \cdot \log|A| - \log(1 - |A|^{-|B_S(r_0)|})}. \tag{7.56}$$

Note that $\log(1 - |A|^{-|B_S(r_0)|}) < 0$, so that the right hand side of (7.56) is well defined and positive.

Since $G$ is sofic, it follows from Theorem 7.7.1 that we can find a finite $S$-labeled graph $Q$ such that

$$|Q(3r_0)| \geq (1 - \varepsilon)|Q|, \tag{7.57}$$

where we recall that $Q(r)$, $r \in \mathbb{N}$, denotes the set of all $q \in Q$ such that there exists an $S$-labeled graph isomorphism $\psi_{q,r}\colon B_S(r) \to B(q,r)$ satisfying $\psi_{q,r}(1_G) = q$ (cf. Theorem 7.7.1).

We have the inclusions

$$Q(r_0) \supset Q(2r_0) \supset \cdots \supset Q(ir_0) \supset Q((i+1)r_0) \supset \cdots.$$

(cf. (7.41); see also Fig. 7.2). Also recall from Lemma 7.7.2 that $B(q, r_0) \subset Q(ir_0)$ for all $Q((i+1)r_0)$ and $i \in \mathbb{N}$.

For each integer $i \geq 1$, we define the map $\mu_i\colon A^{Q(ir_0)} \to A^{Q((i+1)r_0)}$ by setting, for all $u \in A^{Q(ir_0)}$ and $q \in Q((i+1)r_0)$,

$$\mu_i(u)(q) = \mu\left(u|_{B(q,r_0)} \circ \psi_{q,r_0}\right)(1_G),$$

where $\psi_{q,r_0}$ is the unique isomorphism of $S$-labeled graphs from $B_S(r_0) \subset G$ to $B(q, r_0) \subset Q$ sending $1_G$ to $q$.

Similarly, we define the map $\nu_i\colon A^{Q(ir_0)} \to A^{Q((i+1)r_0)}$ by setting, for all $u \in A^{Q(ir_0)}$ and $q \in Q((i+1)r_0)$,

$$\nu_i(u)(q) = \nu\left(u|_{B(q,r_0)} \circ \psi_{q,r_0}\right)(1_G).$$

From the fact that $\sigma \circ \tau = \sigma|_Y \circ \tau = \tau^{-1} \circ \tau$ is the identity map on $A^G$, we deduce that the composite $\nu_{i+1} \circ \mu_i\colon A^{Q(ir_0)} \to A^{Q((i+2)r_0)}$ satisfies $(\nu_{i+1} \circ \mu_i)(u)(q) = u(q)$ for all $u \in A^{Q(ir_0)}$ and $q \in Q((i+2)r_0)$. In other words, denoting by $\rho_i\colon A^{Q(ir_0)} \to A^{Q((i+2)r_0)}$ the restriction map, we have that $\nu_{i+1} \circ \mu_i = \rho_i$ for all $i \geq 1$. In particular, we have $\nu_2 \circ \mu_1 = \rho_1$. Thus, setting

$Z = \mu_1(A^{Q(r_0)}) \subset A^{Q(2r_0)}$, we deduce that $\nu_2(Z) = \rho_1(A^{Q(r_0)}) = A^{Q(3r_0)}$. It follows that $|Z| \geq |A|^{|Q(3r_0)|}$, so that, by taking logarithms, we get

$$\log |Z| \geq |Q(3r_0)| \cdot \log |A|. \tag{7.58}$$

In order to estimate the cardinality of $Z$ from above, we first show the existence of a subset $Q' \subset Q(3r_0)$ satisfying

$$|Q'| \geq \frac{|Q(3r_0)|}{|B(2r_0)|} \tag{7.59}$$

and such that the balls $B(q', r_0)$ with $q' \in Q'$ are all disjoint.

Indeed, let $Q'$ be a maximal subset of $Q(3r_0)$ such that the balls $B(q', r_0)$ with $q' \in Q'$ are all disjoint. If $q \in Q(3r_0) \setminus Q'$ is at distance greater than $2r_0$ from $Q'$, then $B(q, r_0) \cap B(q', r_0) = \varnothing$ for all $q' \in Q'$, contradicting the maximality of $Q'$. Therefore $Q(3r_0)$ is contained in the union of the balls $B(q', 2r_0)$, with $q' \in Q'$. This implies

$$|Q(3r_0)| \leq |Q'| \cdot |B(2r_0)|,$$

which gives (7.59).

Let then $Q' \subset Q(3r_0)$ be as above and set $\overline{Q'} = \coprod_{q' \in Q'} B(q', r_0)$. Note that $\overline{Q'} \subset Q(2r_0)$ and that

$$|\overline{Q'}| = |Q'| \cdot |B_S(r_0)|. \tag{7.60}$$

Given a subset $E \subset Q$ we denote by $\pi_E \colon A^Q \to A^E$ the projection map. Now observe that, for all $q \in Q(2r_0)$, we have a natural bijection $\pi_{B(q,r_0)}(Z) \to \pi_{B_S(r_0)}(Y)$ given by $u \mapsto u \circ \psi_{q,r_0}$, where $\psi_{q,r_0}$ denotes, as above, the unique isomorphism of $S$-labeled graphs from $B_S(r_0)$ to $B(q, r_0)$ sending $1_G$ to $q$. Since $\pi_{B_S(r_0)}(Y) \subsetneq A^{B_S(r_0)}$, this implies that

$$\left|\pi_{B(q',r_0)}(Z)\right| = \left|\pi_{B_S(r_0)}(Y)\right| \leq |A|^{|B_S(r_0)|} - 1, \tag{7.61}$$

for all $q' \in Q'$.

We have

$$\begin{aligned} Z &\subset \pi_{\overline{Q'}}(Z) \times \pi_{Q(2r_0)\setminus\overline{Q'}}(Z) \\ &\subset \left( \prod_{q' \in Q'} \pi_{B(q',r_0)}(Z) \right) \times \pi_{Q(2r_0)\setminus\overline{Q'}}(Z) \\ &\subset \left( \prod_{q' \in Q'} \pi_{B(q',r_0)}(Z) \right) \times A^{Q(2r_0)\setminus\overline{Q'}} \end{aligned}$$

and we deduce

$$\begin{aligned}
|Z| &\leq \left(|A|^{|B_S(r_0)|}-1\right)^{|Q'|}\cdot|A|^{|Q(2r_0)|-|\overline{Q'}|} \quad \text{(by (7.61))}\\
&\leq \left(|A|^{|B_S(r_0)|}-1\right)^{|Q'|}\cdot|A|^{|Q|-|\overline{Q'}|} \quad \text{(as } Q(2r_0)\subset Q\text{)}\\
&= \left(1-|A|^{-|B_S(r_0)|}\right)^{|Q'|}\cdot|A|^{|Q'|\cdot|B_S(r_0)|}\cdot|A|^{|Q|-|\overline{Q'}|}\\
&= \left(1-|A|^{-|B_S(r_0)|}\right)^{|Q'|}\cdot|A|^{|Q|} \quad \text{(by (7.60))}.
\end{aligned}$$

By taking logarithms we get

$$\log|Z| \leq |Q'|\cdot\log\left(1-|A|^{-|B_S(r_0)|}\right)+|Q|\cdot\log|A|.$$

As we remarked earlier, we have $\log(1-|A|^{-|B_S(r_0)|})<0$ so that, by (7.59), we obtain

$$\log|Z| \leq \frac{|Q(3r_0)|}{|B(2r_0)|}\cdot\log\left(1-|A|^{-|B_S(r_0)|}\right)+|Q|\cdot\log|A|$$

and, by (7.58),

$$|Q(3r_0)|\cdot\log|A| \leq \frac{|Q(3r_0)|}{|B(2r_0)|}\cdot\log(1-|A|^{-|B_S(r_0)|})+|Q|\cdot\log|A|.$$

This gives us

$$|Q(3r_0)| \leq \frac{|B(2r_0)|\cdot\log|A|}{|B(2r_0)|\cdot\log|A|-\log(1-|A|^{-|B_S(r_0)|})}|Q|$$

and, by (7.56),

$$|Q(3r_0)| < (1-\varepsilon)|Q|$$

which contradicts (7.57). □

## Notes

The problem of approximation of infinite groups by finite ones goes back, in the framework of purely algebraic constructions, to A.I. Mal'cev [Mal1, Mal3]. These ideas were developed further by A.M. Vershik [Ver] and A.M. Stëpin [Stë1, Stë2] also in the wider context of operator algebras and ergodic theory. The notion of local embeddability was introduced by Vershik and E.I. Gordon in [VeG]. They studied the groups which are locally embeddable into the class of finite groups (LEF groups). Among other things, they established a relationship between this notion of approximation by finite groups and

the convergence in the topological space of marked groups. Moreover, some stability properties for the class of LEF groups were presented. They also showed that finitely presented LEF groups are residually finite. Nilpotent groups and metabelian groups are LEF groups since they are locally residually finite [Hall2]. Vershik and Gordon [VeG] gave an example of a solvable (and therefore amenable) group which is not an LEF group.

The basic results on general local embeddability presented in Sect. 7.1 are natural generalizations of those for LEF groups established in [VeG]. Vershik and Gordon also considered some notions of local approximation for group actions, namely equivariant approximation of actions on groups, free approximation of actions on arbitrary sets, and uniform free approximation of actions on spaces with quasi-invariant measures. Equivariant approximation was used to show that semi-direct products of LEF groups with equivariantly approximable actions are LEF groups. It was then deduced that the non-residually finite group $G_1$ of Sect. 2.6 considered by Mal'cev in [Mal1] is an LEF group. Also, Vershik and Gordon proved the following characterizations of local embeddability into the class of finite groups. A group admits a (faithful) freely approximable action if and only if it is an LEF group. A countable group admits a uniformly freely approximable action on some measurable space if and only if it is an LEF group (note that the "only if" part was already established with slightly different terminology by A. Stëpin in [Stë1], see also [Stë2]).

Groups which are locally embeddable into the class of amenable groups where considered by M. Gromov in [Gro5] under the name of *initially subamenable* groups. More precisely, in Sect. 4.G of [Gro5], a finitely generated group $G$ is said to be initially subamenable if for any finite generating subset $S \subset G$ there exists a sequence of $F$-marked amenable groups converging to $G$, where $F$ denotes the free group based on $S$. The fact that a finitely generated group is initially subamenable if and only if it is LEA follows from Proposition 7.3.7(4). In Sect. 6.E" of [Gro5], Gromov introduced the notion of an *initially subamenable* graph (which, for Cayley graphs, reduces to condition (b) in Theorem 7.7.1) and, in the Example in Sect. 6.E"', he claims that the Cayley graph of a finitely generated group is initially subamenable if and only if the group is initially subamenable. Later, groups whose Cayley graphs are initially subamenable (according with Gromov's definition in Sect. 6.E" of [Gro5]) were called *sofic* groups by B. Weiss [Weiss], this terminology coming from the Hebrew word סוֹפִי which means *finite* and derives from סוֹף which means *end*.

Gromov [Gro5] and Weiss [Weiss] proved that sofic groups are surjunctive (Theorem 7.8.1). G. Elek and E. Szabó [ES2] proved that a group is sofic if and only if it is a subgroup of an ultraproduct of finite symmetric groups equipped with their Hamming distances (Theorem 7.6.6). They used this result to prove that any countable sofic group can be embedded into a countable simple sofic group. In [ES3] Elek and Szabó proved that the class of sofic groups is closed under taking direct products, subgroups, projective limits and direct limits, free products, and extensions by amenable groups. Moreover, by modifying

the example of the group $G_1$ in Sect. 2.6 (which is LEF but not residually finite), Elek and Szabó constructed an example of a finitely generated LEF-group which is not residually amenable.

A detailed exposition of the theory of hyperlinear and sofic groups is presented in V.G. Pestov's survey [Pes] and in the notes by Pestov and A. Kwiatkowska [PeK]. To define *hyperlinear groups*, it suffices to replace the groups $\mathrm{Sym}(F)$, where $F$ is a nonempty finite set, occurring in the definition of sofic groups by the unitary groups $\mathrm{U}(H)$, where $H$ is a finite-dimensional Hilbert space, equipped with their normalized Hilbert-Schmidt metric. This class of groups has its origin in the theory of operator algebras and is related to the Connes Embedding Conjecture. Indeed, by a profound result of E. Kirchberg, F. Radulescu and N. Ozawa, a group is hyperlinear if and only if it satisfies the *Connes embedding conjecture for groups*, that is, if and only if its von Neumann algebra embeds into an ultrapower of the hyperfinite $\mathrm{II}_1$ factor $R$ (see [Pes] and the references therein). Elek and Szabó [ES2] proved that every sofic group is hyperlinear. Thus, we have the implications

$$\begin{array}{ccccccccc}
\text{finite} & \Longrightarrow & \text{resid. finite} & \Longrightarrow & \text{LEF} & & & & \\
\Downarrow & & \Downarrow & & \Downarrow & & & & \\
\text{amenable} & \Longrightarrow & \text{resid. amenable} & \Longrightarrow & \text{LEA} & \Longrightarrow & \text{sofic} & \Longrightarrow & \text{hyperlinear}
\end{array}$$

There exist finitely presented groups which are LEA but not LEF. For example, the Baumslag-Solitar group $BS(2,3)$ is known to be residually solvable. Thus, it is residually amenable and therefore LEA. As the group $BS(2,3)$ is finitely presented and not residually finite, it is not LEF by Proposition 7.3.8. In [Abe], H. Abels gave an example of a finitely presented solvable group which is not residually finite. Abels' group is amenable (and therefore LEA) but not LEF.

There exist finitely presented groups which are not LEA. For example, if $G$ is a finitely presented non-amenable simple group, such as one of Thompson's groups $V$ and $T$ (see [CFP]) or one of the Burger-Mozes groups [BuM], then $G$ is not residually amenable and therefore not LEA by Proposition 7.3.8.

Y. de Cornulier [Cor] provided an example of a finitely presented sofic group which is not LEA. Cornulier's example is sofic because it is an extension of a locally residually finite group by an abelian group and it is not LEA because it is non-amenable and isolated as a free-marked group. A. Thom [Tho] provided an example of a group which is hyperlinear but not LEA, but it is unknown whether Thom's example is sofic or not. The existence of non-sofic groups, of non-hyperlinear groups, and of hyperlinear groups which are not sofic, remain open problems.

L. Bowen [Bow] obtained the following extension of the Kolmogorov-Ornstein-Weiss theorem for Bernoulli shifts. Recall that given a finite set $A$, a strict probability distribution $p = (p_a)_{a\in A}$ on $A$ (that is, $p_a > 0$ for all $a \in A$ and $\sum_{a\in A} p_a = 1$) and a countable group $G$, then, denoting by

$\mu_p = \prod_{g\in G} p$ the corresponding probability measure on the product space $A^G$, the triple $(G, A^G, \mu_p)$ is called a *Bernoulli shift*. The associated *entropy* is the nonnegative number $H(p) = -\sum_{a\in A} p_a \log p_a$. Two Bernoulli shifts $(G, A^G, \mu_p)$ and $(G, B^G, \mu_q)$ are said to be isomorphic (or measurably conjugate) if there exist two $G$-invariant subsets $X \subset A^G$ and $Y \subset B^G$ such that $\mu_p(A^G \setminus X) = \mu_q(B^G \setminus Y) = 0$ and a $G$-equivariant bijective measurable map $\theta\colon X \to Y$ with measurable inverse $\theta^{-1}\colon Y \to X$ such that $\theta_*\mu_p = \mu_q$. Bowen [Bow, Theorem 1.1] proved that if $G$ is a countable sofic group, then two isomorphic Bernoulli shifts $(G, A^G, \mu_p)$ and $(G, B^G, \mu_q)$ have the same entropy, i.e. $H(p) = H(q)$. This result had been established when $G = \mathbb{Z}$ by N. Kolmogorov [Ko1, Ko2] in 1958–59 and then extended to countable amenable groups by D. Ornstein and B. Weiss [OrW] in 1987.

L. Glebsky and L.M. Rivera [GlR] introduced the concept of a *weakly-sofic* group. This is a natural extension of the definition of soficity where the Hamming metric on symmetric groups is replaced by general bi-invariant metrics on finite groups. Glebsky and Rivera showed that the existence of a non-weakly-sofic group is equivalent to a conjecture on the closure in the profinite topology of products of conjugacy classes in free groups of finite rank.

## Exercises

**7.1.** Let $G$ and $C$ be two groups. Suppose that $K$ is a finite symmetric subset of $G$ such that $1_G \in K$. Show that a map $\varphi\colon G \to C$ is a $K$-almost homomorphism if and only if it satisfies $\varphi(k_1k_2) = \varphi(k_1)\varphi(k_2)$ for all $k_1, k_2 \in K$ and $\varphi(k) \neq 1_C$ for all $k \in K \setminus \{1_G\}$.

**7.2.** Show that every group which is locally embeddable into the class of abelian groups is itself abelian.

**7.3.** Show that every group which is locally embeddable into the class of metabelian groups is itself metabelian.

**7.4.** Show that every abelian group is locally embeddable into the class of finite cyclic groups.

**7.5.** Let $\mathcal{C}$ be a class of groups. Suppose that a group $G$ contains a family $(H_i)_{i\in I}$ of subgroups satisfying the following properties: (1) $G = \bigcup_{i\in I} H_i$; (2) For all $i, j \in I$ there exists $k \in I$ such that $H_i \cup H_j \subset H_k$; (3) $H_i$ is locally embeddable into $\mathcal{C}$ for all $i \in I$. Show that $G$ is locally embeddable into $\mathcal{C}$.

**7.6.** Let $\mathcal{C}$ be a class of groups which is closed under finite direct products. Show that every group which is residually locally embeddable into $\mathcal{C}$ is locally embeddable into $\mathcal{C}$.

**7.7.** Let $\mathcal{C}$ be a class of groups which is closed under taking subgroups and taking finite direct products. Let $G$ be a group which is residually $\mathcal{C}$. Show that there exists a net $(N_i)_{i\in I}$ which converges to $\{1_G\}$ in $\mathcal{N}(G)$ such that $G/N_i \in \mathcal{C}$ fro all $i \in I$. Hint: Take as $I$ the set of finite subsets of $G$ partially ordered by inclusion and use Proposition 7.1.13.

**7.8.** Show that if $G$ is a finitely presented infinite simple group then $G$ is not LEF.

**7.9.** Show that if $G$ is a finitely presented non-amenable simple group then $G$ is not LEA.

**7.10.** (cf. [VeG]) Let $G_1$ and $G_2$ be two groups. Recall that, by Proposition 7.3.1, $G_i$ $(i = 1, 2)$ is an LEF-group if and only if the following holds:

$(*)$ for every finite subset $K_i \subset G_i$ there exist a finite set $L_i$ such that $K_i \subset L_i \subset G_i$ and a binary operation $\odot_i \colon L_i \times L_i \to L_i$ such that $(L_i, \odot_i)$ is a group and $k_i k_i' = k_i \odot k_i'$ for all $k_i, k_i' \in K_i$.

Suppose that $G_1$ and $G_2$ are LEF-groups. An action of $G_2$ on $G_1$ by group automorphisms, i.e., a group homomorphism $\varphi \colon G_2 \to \mathrm{Aut}(G_1)$, is said to be *equivariantly approximable*, if for all finite subsets $K_1 \subset G_1$ and $K_2 \subset G_2$ there exist finite groups $(L_1, \odot_1)$ and $(L_2, \odot_2)$ as in $(*)$ and an action of $L_2$ on $L_1$ by group automorphisms $\psi \colon L_2 \to \mathrm{Aut}(L_1)$ such that if $k_1 \in K_1$, $k_2 \in K_2$ and $\varphi(k_2)(k_1) \in K_1$ then $\varphi(k_2)(k_1) = \psi(k_2)(k_1)$.

Show that if $G_1$ and $G_2$ are LEF-groups and $\varphi \colon G_2 \to \mathrm{Aut}(G_1)$ is an equivariantly approximable action, then the semidirect product $G_1 \rtimes_\varphi G_2$ is an LEF-group.

**7.11.** Use the previous exercise to show that the group $G_1$ of Sect. 2.6 is an LEF-group (cf. Proposition 7.3.9).

**7.12.** Show that every residually free group is torsion-free.

**7.13.** A group $G$ is called *fully residually free* if for any finite subset $K \subset G$, there exist a free group $F$ and a group homomorphism $\phi \colon G \to F$ whose restriction to $K$ is injective.

(a) Show that every fully residually free group is residually free.

(b) Show that every fully residually free group is locally embeddable into the class of free groups.

(c) A group $G$ is called *commutative-transitive* if whenever $a, b, c \in G \backslash \{1_G\}$ satisfy $ab = ba$ and $bc = cb$, then $ac = ca$. Prove that every group which is locally embeddable into the class of free groups is commutative-transitive. Hint: First prove that if $F$ is a free group and $a, b \in F$ satisfy $ab = ba$, then there exist an element $x \in F$ and integers $m, n \in \mathbb{Z}$ such that $a = x^m$ and $b = x^n$.

(d) Show that if $F$ is a nonabelian free group, then the group $F \times \mathbb{Z}$ is residually free but not locally embeddable into the class of free groups.

**7.14.** Give an example of a finitely generated LEF-group which is neither residually finite nor amenable. Hint: Take for instance the group $G = G_1 \times F_2$, where $G_1$ is the group introduced in Sect. 2.6 and $F_2$ denotes a free group of rank two.

**7.15.** Let $F$ be a nonempty finite set. Denote by $\mathrm{GL}(\mathbb{R}^F)$ the automorphism group of the real vector space $\mathbb{R}^F = \{x\colon F \to \mathbb{R}\}$. For $\alpha \in \mathrm{Sym}(F)$, define $\lambda(\alpha)\colon \mathbb{R}^F \to \mathbb{R}^F$ by $\lambda(\alpha)(x) = x \circ \alpha^{-1}$ for all $x \in \mathbb{R}^F$.

(a) Show that $\lambda(\alpha) \in \mathrm{GL}(\mathbb{R}^F)$ for all $\alpha \in \mathrm{Sym}(F)$.

(b) Show that the map $\lambda\colon \mathrm{Sym}(F) \to \mathrm{GL}(\mathbb{R}^F)$ is an injective group homomorphism.

(c) Show that the Hamming metric $d_F$ on $\mathrm{Sym}(F)$ satisfies

$$d_F(\alpha, \beta) = \frac{1}{|F|} \mathrm{Tr}(\lambda(\alpha^{-1}\beta))$$

for all $\alpha, \beta \in \mathrm{Sym}(F)$, where $\mathrm{Tr}(\cdot)$ denotes the trace.

**7.16.** Let $H$ be a real or complex Hilbert space of finite dimension $n \geq 1$. Let $L(H)$ denote the vector space consisting of all linear maps $u\colon H \to H$. If $u \in L(H)$, we denote by $\mathrm{Tr}(u)$ the trace of $u$ and by $u^*$ its adjoint. For $u, v \in L(H)$, we set

$$\langle u, v \rangle_{HS} = \frac{1}{n} \mathrm{Tr}(u \circ v^*).$$

(a) Show that $\langle \cdot, \cdot \rangle_{HS}$ is a scalar product on $L(H)$. Let $\| \cdot \|_{HS}$ denote the associated norm.

(b) Let $\mathrm{U}(H) = \{u \in L(H) : u \circ u^* = \mathrm{Id}_H\}$. Show that $\mathrm{U}(H)$ is a group for the composition of maps.

(c) Show that the map $d_{HS}\colon \mathrm{U}(H) \times \mathrm{U}(H) \to \mathbb{R}$ defined by $d_{HS}(u, v) = \|u - v\|_{HS}$ for all $u, v \in \mathrm{U}(H)$ is a bi-invariant metric on $\mathrm{U}(H)$. (The group $\mathrm{U}(H)$ is called the *unitary group* of $H$ and $d_{HS}$ is called the *normalized Hilbert-Schmidt metric* on $\mathrm{U}(H)$.)

**7.17.** Let $G$ be a group, $K \subset G$ a finite subset, $C$ a finite group, and $\varphi\colon G \to C$ a $K$-almost-homomorphism. Denote by $L\colon C \to \mathrm{Sym}(C)$ the Cayley homomorphism, that is, the map defined by $L(g)(h) = gh$ for all $g, h \in C$ and set $\Phi = L \circ \varphi\colon G \to \mathrm{Sym}(C)$. Show that $\Phi$ is a $(K, \varepsilon)$-almost-homomorphism for all $\varepsilon > 0$.

**7.18.** Let $G$ be a group and let $K$ be a finite subset of $G$. Let $F$ be a nonempty finite set. Show that if $0 < \varepsilon < 2/|F|$ then every $(K, \varepsilon)$-almost-homomorphism $\varphi\colon G \to \mathrm{Sym}(F)$ is a $K$-almost-homomorphism of $G$ into the group $\mathrm{Sym}(F)$.

**7.19.** Suppose that a group $G$ contains a family $(H_i)_{i \in I}$ of subgroups satisfying the following properties: (1) $G = \bigcup_{i \in I} H_i$; (2) For all $i, j \in I$ there exists $k \in I$ such that $H_i \cup H_j \subset H_k$; (3) $H_i$ is sofic for all $i \in I$. Show that $G$ is sofic.

**7.20.** Show that every virtually sofic group is sofic. Hint: Use Proposition 7.5.14.

**7.21.** By using Exercise 6.5 and Exercise 6.6, give a direct proof of the fact that the direct product of finitely many groups which satisfy condition (b) in Theorem 7.7.1 also satisfies it.

**7.22.** Let $G$ be a finitely generated group and suppose that $S$ and $S'$ are two finite symmetric generating subsets of $G$. Show that the pair $(G, S)$ satisfies condition (b) in Theorem 7.7.1 if and only if $(G, S')$ does.

**7.23.** Give a direct proof of the fact that every finitely generated residually finite group satisfies the condition (b) in Theorem 7.7.1.

**7.24.** Give a direct proof of the fact that every finitely generated amenable group satisfies the condition (b) in Theorem 7.7.1.

# Chapter 8
# Linear Cellular Automata

In this chapter we study linear cellular automata, namely cellular automata whose alphabet is a vector space and which are linear with respect to the induced vector space structure on the set of configurations. If the alphabet vector space and the underlying group are fixed, the set of linear cellular automata is a subalgebra of the endomorphism algebra of the configuration space (Proposition 8.1.4). An important property of linear cellular automata is that the image of a finitely supported configuration by a linear cellular automaton also has finite support (Proposition 8.2.3). Moreover, a linear cellular automaton is entirely determined by its restriction to the space of finitely-supported configurations (Proposition 8.2.4) and it is pre-injective if and only if this restriction is injective (Proposition 8.2.5). The algebra of linear cellular automata is naturally isomorphic to the group algebra of the underlying group with coefficients in the endomorphism algebra of the alphabet vector space (Theorem 8.5.2). Linear cellular automata may be also regarded as endomorphisms of the space of finitely-supported configurations, viewed as a module over the group algebra of the underlying group with coefficients in the ground field (Proposition 8.7.5). This representation of linear cellular automata is always one-to-one and, when the alphabet vector space is finite-dimensional, it is also onto (Theorem 8.7.6). The image of a linear cellular automaton is closed in the space of configurations for the prodiscrete topology, provided that the alphabet is finite dimensional (Theorem 8.8.1). We exhibit an example showing that if one drops the finite dimensionality of the alphabet, then the image of a linear cellular automaton may fail to be closed. In Sect. 8.9 we prove a linear version of the Garden of Eden theorem. For the proof, we introduce the mean dimension of a vector subspace of the configuration space. We show that for a linear cellular automaton with finite-dimensional alphabet, both pre-injectivity and surjectivity are equivalent to the maximality of the mean dimension of the image of the cellular automaton (Theorem 8.9.6). We exhibit two examples of linear cellular automata with finite-dimensional alphabet over the free group of rank two, one which is pre-injective but not surjective, and one which is surjective but

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_8, 

not pre-injective. This shows that the linear version of the Garden of Eden theorem fails to hold for groups containing nonabelian free subgroups (see Sects. 8.10 and 8.11). Provided the alphabet is finite dimensional, the inverse of every bijective linear cellular automaton is also a linear cellular automaton (Corollary 8.12.2). In Sect. 8.13 we study the pre-injectivity and surjectivity of the discrete Laplacian over the real numbers and prove a Garden of Eden type theorem (Theorem 8.13.2) for such linear cellular automata with no amenability assumptions on the underlying group. As an application, we deduce a characterization of locally finite groups in terms of real linear cellular automata (Corollary 8.13.4). In Sect. 8.14 we define linear surjunctivity and prove that all sofic groups are linearly surjunctive (Theorem 8.14.4). The notion of stable finiteness for rings is introduced in Sect. 8.15. A stably finite ring is a ring for which one-sided invertible square matrices are also two-sided invertible. It is shown that linear surjunctivity is equivalent to stable finiteness of the associated group algebra (Corollary 8.15.6). As a consequence, we deduce that group algebras of sofic groups are stably finite for any ground field (Corollary 8.15.8). In the last section, we prove that the absence of zero-divisors in the group algebra of an arbitrary group is equivalent to the fact that every non-identically-zero linear cellular automaton with one-dimensional alphabet is pre-injective (Corollary 8.16.12).

We recall that in this book all rings are assumed to be associative (but not necessarily commutative) with a unity element, and that a field is a nonzero commutative ring in which each nonzero element is invertible.

## 8.1 The Algebra of Linear Cellular Automata

Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$.

The set $V^G$ consisting of all configurations $x\colon G \to V$ over the group $G$ and the alphabet $V$ has a natural structure of vector space over $\mathbb{K}$ in which addition and scalar multiplication are given by

$$(x + x')(g) = x(g) + x'(g) \quad \text{and} \quad (kx)(g) = kx(g)$$

for all $x, x' \in V^G$, $k \in \mathbb{K}$, and $g \in G$. With the prodiscrete topology, $V^G$ becomes a topological vector space (cf. Sect. F.1). The $G$-shift (see Sect. 1.1) is then $\mathbb{K}$-linear and continuous, that is, for each $g \in G$, the map $x \mapsto gx$ is a continuous endomorphism of $V^G$.

A *linear cellular automaton* over the group $G$ and the alphabet $V$ is a cellular automaton $\tau\colon V^G \to V^G$ which is $\mathbb{K}$-linear, i.e., which satisfies

$$\tau(x + x') = \tau(x) + \tau(x') \quad \text{and} \quad \tau(kx) = k\tau(x)$$

for all $x, x' \in V^G$ and $k \in \mathbb{K}$.

**Proposition 8.1.1.** *Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$. Let $\tau\colon V^G \to V^G$ be a cellular automaton with memory set $S \subset G$ and local defining map $\mu\colon V^S \to V$. Then $\tau$ is linear if and only if $\mu$ is $\mathbb{K}$-linear.*

*Proof.* Suppose first that $\tau$ is linear. Let $y, y' \in V^S$ and $k \in \mathbb{K}$. Denote by $x$ and $x'$ two configurations in $V^G$ extending $y$ and $y'$ respectively, i.e., such that $x|_S = y$ and $x'|_S = y'$. We then have

$$\begin{aligned}\mu(y + y') &= \tau(x + x')(1_G) = (\tau(x) + \tau(x'))(1_G) = \tau(x)(1_G) + \tau(x')(1_G) \\ &= \mu(y) + \mu(y').\end{aligned}$$

Similarly, as $(kx)|_S = ky$, we have

$$\mu(ky) = \tau(kx)(1_G) = k\tau(x)(1_G) = k\mu(y).$$

This shows that $\mu$ is $\mathbb{K}$-linear.

Conversely, suppose that $\mu$ is $\mathbb{K}$-linear. Then, for all $x, x' \in V^G$, $k \in \mathbb{K}$ and $g \in G$ we have

$$\begin{aligned}\tau(x + x')(g) &= \mu((g^{-1}(x + x'))|_S) = \mu((g^{-1}x)|_S + (g^{-1}x')|_S) \\ &= \mu((g^{-1}x)|_S) + \mu((g^{-1}x')|_S) \\ &= \tau(x)(g) + \tau(x')(g) \\ &= (\tau(x) + \tau(x'))(g)\end{aligned}$$

and

$$\tau(kx)(g) = \mu((g^{-1}(kx))|_S) = \mu(k(g^{-1}x)|_S) = k\mu((g^{-1}x)|_S) = k\tau(x)(g).$$

This shows that $\tau(x + x') = \tau(x) + \tau(x')$ and $\tau(kx) = k\tau(x)$. It follows that $\tau$ is linear. □

The following result is a linear analogue of the Curtis-Hedlund theorem (Theorem 1.8.1).

**Theorem 8.1.2.** *Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$. Let $\tau\colon V^G \to V^G$ be a $G$-equivariant and $\mathbb{K}$-linear map. Then the following conditions are equivalent:*

(a) *the map $\tau$ is a linear cellular automaton;*
(b) *the map $\tau$ is uniformly continuous (with respect to the prodiscrete uniform structure on $V^G$);*
(c) *the map $\tau$ is continuous (with respect to the prodiscrete topology on $V^G$);*
(d) *the map $\tau$ is continuous (with respect to the prodiscrete topology on $V^G$) at the constant configuration $x = 0$.*

*Proof.* The implication (a) $\Rightarrow$ (b) immediately follows from Theorem 1.9.1. Since the topology associated with the prodiscrete uniform structure is the prodiscrete topology (cf. Example B.1.4(a)) and every uniformly continuous map is continuous (cf. Proposition B.2.2), we also have (b) $\Rightarrow$ (c). The implication (c) $\Rightarrow$ (d) is trivial. Therefore, we are only left to show that (d) $\Rightarrow$ (a). Suppose that $\tau$ is continuous at 0. Then, the map $V^G \to V$ defined by $x \mapsto \tau(x)(1_G)$ is continuous at 0 since the projection maps $V^G \to V$ defined by $x \mapsto x(g)$ are continuous (for the prodiscrete topology) for all $g \in G$ and the composition of continuous maps is continuous. We deduce that there exists a finite subset $M \subset G$ such that if $x \in V^G$ satisfies $x(m) = 0$ for all $m \in M$, then $\tau(x)(1_G) = 0$. By linearity, we have that if two configurations $x$ and $y$ coincide on $M$ then $\tau(x)(1_G) = \tau(y)(1_G)$. Thus there exists a linear map $\mu\colon V^M \to V$ such that $\tau(x)(1_G) = \mu(x|_M)$ for all $x \in V^G$. As $\tau$ is $G$-equivariant, we deduce that $\tau(x)(g) = \tau(g^{-1}x)(1_G) = \mu((g^{-1}x)|_M)$ for all $x \in V^G$ and $g \in G$. This shows that $\tau$ is the (linear) cellular automaton with memory set $M$ and local defining map $\mu$. Thus (d) implies (a). □

*Examples 8.1.3.* (a) Let $G$ be a group, let $S$ be a nonempty finite subset of $G$, and let $\mathbb{K}$ be a field. The discrete Laplacian $\Delta_S\colon \mathbb{K}^G \to \mathbb{K}^G$ (cf. Example 1.4.3(b)) is a linear cellular automaton.

(b) Let $G$ be a group, $V$ a vector space over a field $\mathbb{K}$, and $f \in \mathrm{End}_{\mathbb{K}}(V)$. Then the map $\tau\colon V^G \to V^G$ defined by $\tau(x) = f \circ x$ for all $x \in V^G$ is a linear cellular automaton (cf. Example 1.4.3(d)).

(c) Let $G$ be a group, $V$ a vector space over a field $\mathbb{K}$, and $s_0$ an element of $G$. Let $R_{s_0}\colon G \to G$ be the right multiplication by $s_0$ in $G$, that is, the map defined by $R_{s_0}(g) = gs_0$ for all $g \in G$. Then the map $\tau\colon V^G \to V^G$ defined by $\tau(x) = x \circ R_{s_0}$ is a linear cellular automaton (cf. Example 1.4.3(e)).

(d) Let $G = \mathbb{Z}$ and $\mathbb{K}$ be a field. Consider the vector space $V = \mathbb{K}[t]$ of all polynomials in the indeterminate $t$ with coefficients in $\mathbb{K}$. A configuration $x \in V^G$ may be viewed as a sequence $x = (x_n)_{n\in\mathbb{Z}}$, where $x_n = x_n(t)$ is a polynomial for all $n \in \mathbb{Z}$. Let $S = \{0, 1\}$ and consider the $\mathbb{K}$-linear map $\mu\colon V^S \to V$ defined by $\mu(p, q) = p - tq'$ for all $(p, q) \in V^S = V \times V$, where $q' \in V$ denotes the derivative of the polynomial $q$. The linear cellular automaton $\tau\colon V^{\mathbb{Z}} \to V^{\mathbb{Z}}$ with memory set $S$ and local defining map $\mu$ is then given by $\tau(x) = y$, where $y_n = x_n - tx'_{n+1} \in V$, $n \in \mathbb{Z}$, for all $x = (x_n)_{n\in\mathbb{Z}} \in V^{\mathbb{Z}}$.

We recall that an *algebra* over a field $\mathbb{K}$ (or a $\mathbb{K}$-*algebra*) is a vector space $\mathcal{A}$ over $\mathbb{K}$ endowed with a product $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$ such that $\mathcal{A}$ is a ring with respect to the sum and the product and such that the following associative law holds for the product and the multiplication by scalars:

$$(ha)(kb) = (hk)(ab)$$

for all $h, k \in \mathbb{K}$ and $a, b \in \mathcal{A}$.

A subset $\mathcal{B}$ of a $\mathbb{K}$-algebra $\mathcal{A}$ is called a *subalgebra* of $\mathcal{A}$ if $\mathcal{B}$ is both a vector subspace and a subring of $\mathcal{A}$.

If $V$ is a vector space over a field $\mathbb{K}$, the set $\mathrm{End}_{\mathbb{K}}(V)$ consisting of all endomorphisms of the vector space $V$ has a natural structure of a $\mathbb{K}$-algebra for which

$$(f+f')(v) = f(v) + f'(v),$$
$$(ff')(v) = (f \circ f')(v) = f(f'(v))$$

and

$$(kf)(v) = kf(v)$$

for all $f, f' \in \mathrm{End}_{\mathbb{K}}(V)$, $k \in \mathbb{K}$, and $v \in V$. The identity map $\mathrm{Id}_V$ is the unity element of $\mathrm{End}_{\mathbb{K}}(V)$.

Given a group $G$ and a vector space $V$ over a field $\mathbb{K}$, we denote by $\mathrm{LCA}(G;V)$ the set of all linear cellular automata over the group $G$ and the alphabet $V$. It immediately follows from the definition of a linear cellular automaton that $\mathrm{LCA}(G;V) \subset \mathrm{End}_{\mathbb{K}}(V^G)$.

**Proposition 8.1.4.** *Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$. Then, $\mathrm{LCA}(G;V)$ is a subalgebra of $\mathrm{End}_{\mathbb{K}}(V^G)$.*

*Proof.* Let $\tau_1, \tau_2 \in \mathrm{LCA}(G;V)$. Let $S_1$ and $S_2$ be memory sets for $\tau_1$ and $\tau_2$. Then, the set $S = S_1 \cup S_2$ is also a memory set for $\tau_1$ and $\tau_2$ (cf. Sect. 1.5). Let $\mu_1 \colon V^S \to V$ and $\mu_2 \colon V^S \to V$ be the corresponding local defining maps and set $\mu = \mu_1 + \mu_2$. For all $x \in V^G$ and $g \in G$ we have

$$\begin{aligned}(\tau_1+\tau_2)(x)(g) &= \tau_1(x)(g) + \tau_2(x)(g)\\ &= \mu_1(g^{-1}x|_S) + \mu_2(g^{-1}x|_S)\\ &= \mu(g^{-1}x|_S).\end{aligned}$$

This shows that $\tau_1 + \tau_2$ is a cellular automaton with memory set $S$ and local defining map $\mu$. Since the map $\tau_1 + \tau_2$ is $\mathbb{K}$-linear, we deduce that $\tau_1 + \tau_2 \in \mathrm{LCA}(G;V)$.

On the other hand, let $k \in \mathbb{K}$ and let $\tau \in \mathrm{LCA}(G;V)$ with memory set $S$ and local defining map $\mu \colon V^S \to V$. Then, for all $x \in V^G$ and $g \in G$, we have

$$(k\tau)(x)(g) = k\tau(x)(g) = k\mu(g^{-1}x|_S) = (k\mu)(g^{-1}x|_S).$$

Therefore the $\mathbb{K}$-linear map $k\tau$ is a cellular automaton with memory set $S$ and local defining map $k\mu$. It follows that $k\tau \in \mathrm{LCA}(G;V)$.

We clearly have $\mathrm{Id}_{V^G} \in \mathrm{LCA}(G;V)$ (cf. Example 1.4.3(d)). Finally, it follows from Proposition 1.4.9 that if $\tau_1, \tau_2 \in \mathrm{LCA}(G;V)$ then the $\mathbb{K}$-linear map $\tau_1\tau_2 = \tau_1 \circ \tau_2$ is also a cellular automaton and hence $\tau_1\tau_2 \in \mathrm{LCA}(G;V)$. This shows that $\mathrm{LCA}(G;V)$ is a subalgebra of $\mathrm{End}_{\mathbb{K}}(V^G)$. □

## 8.2 Configurations with Finite Support

Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$.

The *support* of a configuration $x \in V^G$ is the set $\{g \in G : x(g) \neq 0_V\}$. We denote by $V[G]$ the subset of $V^G$ consisting of all configurations with finite support.

**Proposition 8.2.1.** *Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$. Then the set $V[G]$ is a vector subspace of $V^G$. Moreover, $V[G]$ is dense in $V^G$ for the prodiscrete topology.*

*Proof.* If $k_1, k_2 \in \mathbb{K}$ and $x_1, x_2 \in V^G$, then the support of $k_1x_1 + k_2x_2$ is contained in the union of the support of $x_1$ and the support of $x_2$. Therefore, if $x_1$ and $x_2$ have finite support, so does $k_1x_1 + k_2x_2$. Consequently, $V[G]$ is a vector subspace of $V^G$.

Let $x \in V^G$ and let $W \subset V^G$ be a neighborhood of $x$ for the prodiscrete topology. By definition of the prodiscrete topology, there exists a finite subset $\Omega \subset G$ such that $W$ contains all configurations which coincide with $x$ on $\Omega$. It follows that the configuration $y \in V[G]$ which coincides with $x$ on $\Omega$ and is identically zero outside of $\Omega$ is in $W$. This shows that $V[G]$ is dense in $V^G$. □

**Proposition 8.2.2.** *Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$. Let $x, x' \in V^G$. Then the configurations $x$ and $x'$ are almost equal if and only if $x - x' \in V[G]$.*

*Proof.* By definition, $x$ and $x'$ are almost equal if and only if the set $\{g \in G : x(g) \neq x'(g)\}$ is finite. This is equivalent to $x - x' \in V[G]$ since $\{g \in G : x(g) \neq x'(g)\} = \{g \in G : (x - x')(g) \neq 0_V\}$. □

**Proposition 8.2.3.** *Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$. Let $\tau \in \mathrm{LCA}(G; V)$. Then one has $\tau(V[G]) \subset V[G]$.*

*Proof.* Denote by $S \subset G$ a memory set for $\tau$ and let $\mu \colon V^S \to V$ be the corresponding local defining map. Let $x \in V[G]$ and let $T \subset G$ denote the support of $x$. For all $g \in G$, we have $\tau(x)(g) = \mu((g^{-1}x)|_S)$. As $\mu$ is linear (Proposition 8.1.1) and the support of $g^{-1}x$ is $g^{-1}T$, we deduce that $\tau(x)(g) = 0$ if $g^{-1}T \cap S = \varnothing$. It follows that the support of $\tau(x)$ is contained in the finite set $TS^{-1} \subset G$. This shows that $\tau(x) \in V[G]$. □

Observe that if $\tau \in \mathrm{LCA}(G; V)$, then the restriction map

$$\tau|_{V[G]} \colon V[G] \to V[G]$$

is $\mathbb{K}$-linear, that is, $\tau|_{V[G]} \in \mathrm{End}_{\mathbb{K}}(V[G])$.

Let $\mathcal{A}$ and $\mathcal{B}$ be two algebras over a field $\mathbb{K}$. A map $F\colon \mathcal{A} \to \mathcal{B}$ is called a $\mathbb{K}$-*algebra homomorphism* if $F$ is both a vector space homomorphism (i.e., a $\mathbb{K}$-linear map) and a ring homomorphism. This is equivalent to the fact that $F$ satisfies $F(a + a') = F(a) + F(a')$, $F(aa') = F(a)F(a')$, $F(ka) = kF(a)$ for all $a, a' \in \mathcal{A}$ and $k \in \mathbb{K}$, and $F(1_{\mathcal{A}}) = 1_{\mathcal{B}}$.

**Proposition 8.2.4.** *Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$. Then the map $\Lambda\colon \mathrm{LCA}(G;V) \to \mathrm{End}_{\mathbb{K}}(V[G])$ defined by $\Lambda(\tau) = \tau|_{V[G]}$, where $\tau|_{V[G]}\colon V[G] \to V[G]$ is the restriction of $\tau$ to $V[G]$, is an injective $\mathbb{K}$-algebra homomorphism.*

*Proof.* The fact that $\Lambda$ is an algebra homomorphism immediately follows from the definition of the algebra operations on $\mathrm{LCA}(G;V)$ and $\mathrm{End}_{\mathbb{K}}(V[G])$. Suppose that $\tau \in \mathrm{LCA}(G;V)$ satisfies $\Lambda(\tau) = 0$. This means that $\tau(y) = 0$ for all $y \in V[G]$. As $\tau\colon V^G \to V^G$ is continuous by Proposition 1.4.8 and $V[G]$ is dense in $V^G$ by Proposition 8.2.1 for the prodiscrete topology, we deduce that $\tau(x) = 0$ for all $x \in V^G$, that is, $\tau = 0$. This shows that $\Lambda$ is injective. □

**Proposition 8.2.5.** *Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$. Let $\tau \in \mathrm{LCA}(G;V)$. Then the following conditions are equivalent:*

(a) *$\tau$ is pre-injective;*
(b) *$\tau|_{V[G]}\colon V[G] \to V[G]$ is injective.*

*Proof.* Suppose that $\tau$ is pre-injective. Let $x \in \ker(\tau|_{V[G]})$. Then $x \in V[G]$ and $\tau(x) = \tau|_{V[G]}(x) = 0$. As the trivial configuration 0 and the configuration $x$ are almost equal and $\tau(0) = 0$ by linearity of $\tau$, the pre-injectivity of $\tau$ implies that $x = 0$. This shows that $\tau|_{V[G]}$ is injective.

Conversely, suppose that $\tau|_{V[G]}$ is injective. Let $x, x' \in V^G$ be two configurations which are almost equal and such that $\tau(x) = \tau(x')$. Then $x - x' \in V[G]$ by Proposition 8.2.2 and we have $\tau|_{V[G]}(x - x') = \tau(x - x') = \tau(x) - \tau(x') = 0$. As $\tau|_{V[G]}$ is injective, this implies $x - x' = 0$, that is, $x = x'$. Therefore, $\tau$ is pre-injective. □

## 8.3 Restriction and Induction of Linear Cellular Automata

In this section, we show that the operations of restriction and induction for cellular automata that were introduced in Sect. 1.7 preserve linearity.

Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$. Let $H$ be a subgroup of $G$. We denote by $\mathrm{LCA}(G, H; V) = \mathrm{LCA}(G;V) \cap \mathrm{CA}(G, H; V)$ the set of all linear cellular automata $\tau\colon V^G \to V^G$ admitting a memory set $S$ such that $S \subset H$. Recall from Sect. 1.7, that, given a cellular automaton $\tau\colon V^G \to V^G$ with memory set $S \subset H$, we denote by $\tau_H\colon V^H \to V^H$

the restriction cellular automaton. Similarly, given a cellular automaton $\sigma\colon V^H \to V^H$, we denote by $\sigma^G\colon V^G \to V^G$ the induced cellular automaton.

**Proposition 8.3.1.** *Let $\tau \in \mathrm{CA}(G, H; V)$. Then, $\tau \in \mathrm{LCA}(G, H; V)$ if and only if $\tau_H \in \mathrm{LCA}(H; V)$.*

*Proof.* It follows immediately from the definitions of restriction and induction that a cellular automaton $\tau \in \mathrm{CA}(G, H; V)$ is linear if and only if $\tau_H \in \mathrm{CA}(H; V)$ is linear. □

Let $\mathcal{A}$ and $\mathcal{B}$ be two algebras over a field $\mathbb{K}$. A map $F\colon \mathcal{A} \to \mathcal{B}$ is called a $\mathbb{K}$-*algebra isomorphism* if $F$ is a bijective $\mathbb{K}$-algebra homomorphism. It is clear that if $F\colon \mathcal{A} \to \mathcal{B}$ is a $\mathbb{K}$-algebra isomorphism then its inverse map $F^{-1}\colon \mathcal{B} \to \mathcal{A}$ is also a $\mathbb{K}$-algebra isomorphism.

**Proposition 8.3.2.** *The set $\mathrm{LCA}(G, H; V)$ is a subalgebra of $\mathrm{LCA}(G; V)$. Moreover, the map $\tau \mapsto \tau_H$ is a $\mathbb{K}$-algebra isomorphism from $\mathrm{LCA}(G, H; V)$ onto $\mathrm{LCA}(H; V)$ whose inverse is the map $\sigma \mapsto \sigma^G$.*

*Proof.* Let $\tau_1, \tau_2 \in \mathrm{LCA}(G, H; V)$ with memory sets $S_1, S_2 \subset H$ respectively. Then the linear cellular automaton $\tau_1 + \tau_2$ admits $S_1 \cup S_2$ as a memory set. As $S_1 \cup S_2 \subset H$, we have $\tau_1 + \tau_2 \in \mathrm{LCA}(G, H; V)$. If $k \in \mathbb{K}$ and $\tau \in \mathrm{LCA}(G, H; V)$, with memory set $S \subset H$, then $S$ is also a memory set for $k\tau$ and therefore $k\tau \in \mathrm{LCA}(G, H; V)$. This shows that $\mathrm{LCA}(G, H; V)$ is a vector subspace of $\mathrm{LCA}(G; V)$. Since $\mathrm{LCA}(G, H; V)$ is a submonoid of $\mathrm{LCA}(G; V)$ by Proposition 1.7.1, we deduce that $\mathrm{LCA}(G, H; V)$ is a subalgebra of $\mathrm{LCA}(G; V)$.

To simplify notation, denote by $\Phi\colon \mathrm{LCA}(G, H; V) \to \mathrm{LCA}(H; V)$ and $\Psi\colon \mathrm{LCA}(H; V) \to \mathrm{LCA}(G, H; V)$ the maps defined by $\Phi(\tau) = \tau_H$ and $\Psi(\sigma) = \sigma^G$ respectively. It is clear from the definitions that $\Psi \circ \Phi\colon \mathrm{LCA}(G, H; V) \to \mathrm{LCA}(G, H; V)$ and $\Phi \circ \Psi\colon \mathrm{LCA}(H; V) \to \mathrm{LCA}(H; V)$ are the identity maps. Therefore, $\Phi$ is bijective with inverse $\Psi$. It remains to show that $\Phi$ is a $\mathbb{K}$-algebra homomorphism.

Let $\tau_1, \tau_2 \in \mathrm{LCA}(G, H; V)$ and $k_1, k_2 \in \mathbb{K}$. Let $x \in V^H$ and let $\widetilde{x} \in V^G$ extending $x$. By applying (1.13), we have

$$\Phi(k_1\tau_1 + k_2\tau_2)(x)(h) = (k_1\tau_1 + k_2\tau_2)(\widetilde{x})(h) = k_1\tau_1(\widetilde{x})(h) + k_2\tau_2(\widetilde{x})(h)$$

for all $h \in H$. We deduce that $\Phi(k_1\tau_1 + k_2\tau_2)(x) = (k_1\Phi(\tau_1) + k_2\Phi(\tau_2))(x)$ for all $x \in V^H$, that is, $\Phi(k_1\tau_1 + k_2\tau_2) = k_1\Phi(\tau_1) + k_2\Phi(\tau_2)$. This shows that $\Phi$ is $\mathbb{K}$-linear.

Finally, it follows from Proposition 1.7.2 that $\Phi(\mathrm{Id}_{V^G}) = \mathrm{Id}_{V^H}$ and $\Phi(\tau_1\tau_2) = \Phi(\tau_1)\Phi(\tau_2)$ for all $\tau_1, \tau_2 \in \mathrm{LCA}(G, H; V)$. We have shown that $\Phi$ is a $\mathbb{K}$-algebra isomorphism. □

## 8.4 Group Rings and Group Algebras

Let $G$ be a group and let $R$ be a ring. We denote by $0_R$ (resp. $1_R$) the zero (resp. unity) element of $R$. We regard $R$ as a left $R$-module over itself. Then $R^G = \{\alpha\colon G \to R\}$ has a natural structure of left $R$-module with addition and scalar multiplication given by

$$(\alpha+\beta)(g) = \alpha(g)+\beta(g) \quad \text{and } (r\alpha)(g) = r\alpha(g)$$

for all $\alpha, \beta \in R^G$, $r \in R$, and $g \in G$. Define the *support* of an element $\alpha \in R^G$ as being the set $\{g \in G : \alpha(g) \neq 0_R\}$. Let $R[G]$ denote the set consisting of all elements $\alpha \in R^G$ which have finite support. Then $R[G]$ is a free submodule of $R^G$. We have

$$R[G] = \bigoplus_{g\in G} R \subset \prod_{g\in G} R = R^G,$$

so that the elements $\delta_g\colon G \to R$, $g \in G$, defined by

$$\delta_g(h) = \begin{cases} 1_R & \text{if } h = g \\ 0_R & \text{if } h \neq g \end{cases} \tag{8.1}$$

freely generate $R[G]$ as a left $R$-module. Note that the decomposition of an element $\alpha \in R[G]$ in this basis is simply given by the formula

$$\alpha = \sum_{g\in G} \alpha(g)\delta_g.$$

Let now $\alpha$ and $\beta$ be two elements of $R[G]$ and denote their supports by $S$ and $T$. The *convolution product* of $\alpha$ and $\beta$ is the element $\alpha\beta \in R^G$ defined by

$$(\alpha\beta)(g) = \sum_{h\in G} \alpha(h)\beta(h^{-1}g) = \sum_{h\in S} \alpha(h)\beta(h^{-1}g) \tag{8.2}$$

for all $g \in G$. Note that $\alpha\beta \in R[G]$ as the support of $\alpha\beta$ is contained in $ST = \{st : s \in S, t \in T\}$. By using the change of variables $h_1 = h$ and $h_2 = h^{-1}g$, the convolution product (8.2) may be also expressed as follows:

$$(\alpha\beta)(g) = \sum_{\substack{h_1,h_2\in G \\ h_1h_2=g}} \alpha(h_1)\beta(h_2). \tag{8.3}$$

**Proposition 8.4.1.** *Let $G$ be a group and let $R$ be a ring. Then the addition and the convolution product gives a ring structure to $R[G]$.*

*Proof.* We know that $(R[G], +)$ is an abelian group. Let $\alpha, \beta, \gamma \in R[G]$ and $g \in G$. We have

$$\begin{aligned}
[(\alpha\beta)\gamma](g) &= \sum_{h\in G}(\alpha\beta)(h)\gamma(h^{-1}g)\\
&= \sum_{h\in G}\sum_{k\in G}\alpha(k)\beta(k^{-1}h)\gamma(h^{-1}g)\\
\text{(by setting } s=k^{-1}h) \quad &= \sum_{s\in G}\sum_{k\in G}\alpha(k)\beta(s)\gamma(s^{-1}k^{-1}g)\\
&= \sum_{k\in G}\alpha(k)\Big(\sum_{s\in G}\beta(s)\gamma(s^{-1}k^{-1}g)\Big)\\
&= \sum_{k\in G}\alpha(k)(\beta\gamma)(k^{-1}g)\\
&= [\alpha(\beta\gamma)](g).
\end{aligned}$$

This shows that $(\alpha\beta)\gamma = \alpha(\beta\gamma)$. Thus, the convolution product is associative. On the other hand, given $\alpha \in R[G]$ we have, for all $g \in G$,

$$[\delta_{1_G}\alpha](g) = \sum_{h\in G}\delta_{1_G}(h)\alpha(h^{-1}g) = \alpha(g) = \sum_{k\in G}\alpha(k)\delta_{1_G}(k^{-1}g) = [\alpha\delta_{1_G}](g).$$

Thus, $\delta_{1_G}\alpha = \alpha\delta_{1_G} = \alpha$ for all $\alpha \in R[G]$. This shows that $\delta_{1_G}$ is an identity element for the convolution product. Finally,

$$\begin{aligned}
[(\alpha+\beta)\gamma](g) &= \sum_{h\in G}(\alpha+\beta)(h)\gamma(h^{-1}g)\\
&= \sum_{h\in G}[\alpha(h)+\beta(h)]\gamma(h^{-1}g)\\
&= \sum_{h\in G}\alpha(h)\gamma(h^{-1}g) + \sum_{k\in G}\beta(k)\gamma(k^{-1}g)\\
&= (\alpha\gamma)(g) + (\beta\gamma)(g).
\end{aligned}$$

Thus, we have $(\alpha+\beta)\gamma = \alpha\gamma + \beta\gamma$. Similarly, one shows that $\alpha(\beta+\gamma) = \alpha\beta+\alpha\gamma$. It follows that the distributive laws also hold in $R[G]$. Consequently, $R[G]$ is a ring with unity element $1_{R[G]} = \delta_{1_G}$. □

The ring $R[G]$ is called the *group ring* of $G$ with coefficients in $R$.

Given a ring $R$, denote by $U(R)$ the multiplicative group consisting of all invertible elements in $R$.

**Proposition 8.4.2.** *Let $G$ be a group and let $R$ be a ring. Then one has $\delta_g \in \mathrm{U}(R[G])$ for all $g \in G$. Moreover the map $\phi\colon G \to \mathrm{U}(R[G])$ given by $\phi(g) = \delta_g$ is a group homomorphism.*

*Proof.* Let $g_1, g_2, g \in G$. Then $(\delta_{g_1}\delta_{g_2})(g) = \sum_{h\in G}\delta_{g_1}(h)\delta_{g_2}(h^{-1}g)$ equals 1 if $g_1^{-1}g = g_2$, that is, if $g = g_1g_2$, and equals 0 otherwise. This shows that

$$\delta_{g_1}\delta_{g_2} = \delta_{g_1g_2}. \tag{8.4}$$

We deduce that $\delta_g \delta_{g^{-1}} = \delta_{gg^{-1}} = \delta_{1_G} = 1_{R[G]}$ and, similarly, $\delta_{g^{-1}} \delta_g = 1_{R[G]}$. This shows that $\delta_g$ belongs to $\mathrm{U}(R[G])$. The fact that $\phi$ is a group homomorphism follows from (8.4). □

Let $G$ be a group and let $R$ be a ring. For $\alpha \in R[G]$ define $\alpha^* \colon G \to R$ by setting $\alpha^*(g) = \alpha(g^{-1})$ for all $g \in G$. If $S$ is the support of $\alpha$, then the support of $\alpha^*$ is $S^{-1}$. Thus one has $\alpha^* \in R[G]$.

**Proposition 8.4.3.** *Let $G$ be a group and let $R$ be a ring. Let $\alpha, \beta \in R[G]$. Then one has*

(i) $(1_{R[G]})^* = 1_{R[G]}$,
(ii) $(\alpha^*)^* = \alpha$,
(iii) $(\alpha + \beta)^* = \alpha^* + \beta^*$.

*Moreover, if the ring $R$ is commutative, then one has*

(iv) $(\alpha\beta)^* = \beta^* \alpha^*$.

*Proof.* (i) We have $(1_{R[G]})^* = 1_{R[G]}$ since the support of $1_{R[G]} = \delta_{1_G}$ is $\{1_G\}$.
(ii) For all $g \in G$, we have

$$(\alpha^*)^*(g) = \alpha((g^{-1})^{-1}) = \alpha(g).$$

Therefore $(\alpha^*)^* = \alpha$.
(iii) For all $g \in G$, we have

$$(\alpha+\beta)^*(g) = (\alpha+\beta)(g^{-1}) = \alpha(g^{-1})+\beta(g^{-1}) = \alpha^*(g)+\beta^*(g) = (\alpha^*+\beta^*)(g).$$

Therefore $(\alpha + \beta)^* = \alpha^* + \beta^*$.
(iv) Suppose that the ring $R$ is commutative. For all $g \in G$, we have

$$\begin{aligned}
(\alpha\beta)^*(g) &= (\alpha\beta)(g^{-1}) \\
&= \sum_{h \in G} \alpha(h)\beta(h^{-1}g^{-1}) \\
&= \sum_{h \in G} \beta(h^{-1}g^{-1})\alpha(h) \quad \text{(since } R \text{ is commutative)} \\
&= \sum_{h \in G} \beta^*(gh)\alpha^*(h^{-1}) \\
&= \sum_{k \in G} \beta^*(k)\alpha^*(k^{-1}g) \\
&= (\beta^*\alpha^*)(g).
\end{aligned}$$

Therefore $(\alpha\beta)^* = \beta^*\alpha^*$. □

If $R = (R, +, \cdot)$ is a ring, its *opposite ring* is the ring $R^{op} = (R, +, *)$ having the same underlying set and addition as $R$ and with multiplication $*$ defined by $r * s = sr$ for all $r, s \in R$.

From Proposition 8.4.3, we deduce that, if $G$ is a group and $R$ is a commutative ring, then the map $\alpha \mapsto \alpha^*$ is a ring isomorphism between the ring $R[G]$ and its opposite ring $(R[G])^{op}$. Thus we have

**Corollary 8.4.4.** *Let $G$ be a group and let $R$ be a commutative ring. Then the ring $R[G]$ is isomorphic to its opposite ring $(R[G])^{op}$.* □

*Remarks 8.4.5.* Let $G$ be a group and let $R$ be a ring.

(a) It is immediate to verify that the map $R \to R[G]$ defined by $r \mapsto r\delta_{1_G}$ is an injective ring homomorphism. This is often used to regard $R$ as a subring of $R[G]$.

(b) Observe that $(r\alpha)\beta = r(\alpha\beta)$ for all $r \in R$ and $\alpha, \beta \in R[G]$. Therefore, we can use the notation $r\alpha\beta = (r\alpha)\beta = r(\alpha\beta)$. If the ring $R$ is commutative, then one has the additional property $(r_1\alpha)(r_2\beta) = r_1r_2\alpha\beta$ for all $r_1, r_2 \in R$ and $\alpha, \beta \in R[G]$.

(c) Suppose that the group $G$ is abelian and that the ring $R$ is commutative. Then the ring $R[G]$ is commutative. Indeed, it immediately follows from (8.3) that we have $\alpha\beta = \beta\alpha$ for all $\alpha, \beta \in R[G]$ under these hypotheses.

Suppose now that we are given an algebra $\mathcal{A}$ over a field $\mathbb{K}$. Then $\mathcal{A}[G]$ is both a $\mathbb{K}$-vector space and a ring; Moreover, we have $(k_1\alpha)(k_2\beta) = k_1k_2\alpha\beta$ for all $k_1, k_2 \in \mathbb{K}$ and $\alpha, \beta \in \mathcal{A}[G]$. Therefore, $\mathcal{A}[G]$ is an algebra over $\mathbb{K}$. The $\mathbb{K}$-algebra $\mathcal{A}[G]$ is called the *group algebra* of $G$ with coefficients in the $\mathbb{K}$-algebra $\mathcal{A}$. In the particular case $\mathcal{A} = \mathbb{K}$, this gives the $\mathbb{K}$-algebra $\mathbb{K}[G]$.

## 8.5 Group Ring Representation of Linear Cellular Automata

Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$.

Consider the $\mathbb{K}$-algebra $\mathrm{End}_{\mathbb{K}}(V)[G]$, that is, the group algebra of $G$ with coefficients in the $\mathbb{K}$-algebra $\mathrm{End}_{\mathbb{K}}(V)$ (see Sect. 8.4).

For each $\alpha \in \mathrm{End}_{\mathbb{K}}(V)[G]$, we define a map $\tau_\alpha \colon V^G \to V^G$ by setting

$$\tau_\alpha(x)(g) = \sum_{h \in G} \alpha(h)(x(gh)) \tag{8.5}$$

for all $x \in V^G$ and $g \in G$. Observe that $\alpha(h) \in \mathrm{End}_{\mathbb{K}}(V)$ and $x(gh) \in V$ for all $x \in V^G$ and $g, h \in G$. Note also that there is only a finite number of nonzero terms in the sum appearing in the right hand side of (8.5) since $\alpha$ has finite support. It follows that $\tau_\alpha$ is well defined.

For each $v \in V$, define the configuration $c_v \in V[G]$ by

$$c_v(g) = \begin{cases} v & \text{if } g = 1_G, \\ 0 & \text{otherwise.} \end{cases} \tag{8.6}$$

Note that the map from $V$ to $V[G]$ given by $v \mapsto c_v$ is injective and $\mathbb{K}$-linear.

**Proposition 8.5.1.** *Let $\alpha \in \mathrm{End}_{\mathbb{K}}(V)[G]$. Then one has:*

(i) $\tau_\alpha \in \mathrm{LCA}(G; V)$;
(ii) $\alpha(g)(v) = \tau_\alpha(c_v)(g)$ *for all* $v \in V$ *and* $g \in G$;
(iii) *the support of* $\alpha$ *is the minimal memory set of* $\tau_\alpha$.

*Proof.* Let $S \subset G$ denote the support of $\alpha$. Consider the map $\mu\colon V^S \to V$ given by

$$\mu(y) = \sum_{s \in S} \alpha(s)(y(s))$$

for all $y \in V^S$. Then, for all $x \in V^G$ and $g \in G$, we have

$$\begin{aligned} \tau_\alpha(x)(g) &= \sum_{h \in G} \alpha(h)(x(gh)) \\ &= \sum_{h \in G} \alpha(h)(g^{-1}x(h)) \\ &= \sum_{s \in S} \alpha(s)(g^{-1}x(s)) \\ &= \mu((g^{-1}x)|_S). \end{aligned} \tag{8.7}$$

Thus $\tau_\alpha$ is the cellular automaton with memory set $S$ and local defining map $\mu$. It is clear from (8.5) that $\tau_\alpha$ is a $\mathbb{K}$-linear map. Thus, we have $\tau_\alpha \in \mathrm{LCA}(G; V)$. This shows (i).

Let $v \in V$ and let $c_v \in V[G]$ as in (8.6). By applying (8.5), we get, for all $g \in G$,

$$\tau_\alpha(c_v)(g^{-1}) = \sum_{h \in G} \alpha(h)(c_v(g^{-1}h)) = \alpha(g)(v).$$

This gives us (ii).

Finally, let $S_0 \subset G$ denote the minimal memory set of $\tau_\alpha$. We have seen in the proof of (i) that $S$ is a memory set for $\tau_\alpha$. Therefore, we have $S_0 \subset S$. On the other hand, we deduce from (ii) that, for all $v \in V$ and $g \in G$, we have

$$\alpha(g)(v) = \tau_\alpha(c_v)(g) = \mu_0((g^{-1}c_v)|_{S_0}),$$

where $\mu_0\colon V^{S_0} \to V$ denote the local defining map for $\tau_\alpha$ associated with $S_0$. Since the configuration $g^{-1}c_v$ is identically zero on $G \setminus \{g\}$, it follows that $\alpha(g) = 0$ if $g \notin S_0$. This implies $S \subset S_0$. Thus, we have $S = S_0$. This shows (iii). □

**Theorem 8.5.2.** *Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$. Then the map $\Psi\colon \mathrm{End}_{\mathbb{K}}(V)[G] \to \mathrm{LCA}(G; V)$ defined by $\alpha \mapsto \tau_\alpha$ is a $\mathbb{K}$-algebra isomorphism.*

*Proof.* Let $\alpha, \beta \in \mathrm{End}_{\mathbb{K}}(V)[G]$ and $k \in \mathbb{K}$. By applying (8.5), we get

$$\begin{aligned}\tau_{\alpha+\beta}(x)(g) &= \sum_{h\in G}[(\alpha+\beta)(h)](x(gh))\\ &= \sum_{h\in G}(\alpha(h)+\beta(h))(x(gh))\\ &= \sum_{h\in G}[\alpha(h)(x(gh))+\beta(h)(x(gh))]\\ &= \sum_{h\in G}\alpha(h)(x(gh))+\sum_{h\in G}\beta(h)(x(gh))\\ &= \tau_\alpha(x)(g)+\tau_\beta(x)(g)\end{aligned}$$

and, similarly,

$$\begin{aligned}\tau_{k\alpha}(x)(g) &= \sum_{h\in G}[(k\alpha)(h)](x(gh))\\ &= \sum_{h\in G}k[\alpha(h)](x(gh))\\ &= k\sum_{h\in G}[\alpha(h)](x(gh))\\ &= k\tau_\alpha(x)(g)\end{aligned}$$

for all $x \in V^G$ and $g \in G$. Thus $\tau_{\alpha+\beta} = \tau_\alpha + \tau_\beta$ and $\tau_{k\alpha} = k\tau_\alpha$. This shows that $\Psi$ is a $\mathbb{K}$-linear map.

Let us show that $\Psi$ is a ring homomorphism. Let $\alpha, \beta \in \mathrm{End}_{\mathbb{K}}(V)[G]$. Let $x \in V^G$ and set $y = \tau_\beta(x)$. For all $g, h \in G$, we have

$$y(gh) = \sum_{t\in G}\beta(t)(x(ght)). \tag{8.8}$$

It follows that

$$\begin{aligned}\tau_\alpha(\tau_\beta(x))(g) &= \tau_\alpha(y)(g)\\ &= \sum_{h\in G}\alpha(h)(y(gh))\\ \text{(by (8.8))} \quad &= \sum_{h\in G}\alpha(h)\Big(\sum_{t\in G}\beta(t)(x(ght))\Big)\\ \text{(by setting } z=ht) \quad &= \sum_{h\in G}\alpha(h)\Big(\sum_{z\in G}\beta(h^{-1}z)(x(gz))\Big)\\ &= \sum_{z\in G}\Big(\sum_{h\in G}\alpha(h)\beta(h^{-1}z)\Big)(x(gz))\\ &= \sum_{z\in G}[\alpha\beta](z)(x(gz))\\ &= \tau_{\alpha\beta}(x)(g).\end{aligned}$$

Thus we have $\tau_{\alpha\beta} = \tau_\alpha \circ \tau_\beta$, that is, $\Psi(\alpha\beta) = \Psi(\alpha)\Psi(\beta)$. Observe that $1_{\mathrm{End}_{\mathbb{K}}(V)[G]} = \delta_{1_G}$, where $\delta_{1_G} \colon G \to \mathrm{End}_{\mathbb{K}}(V)$ is given by $\delta_{1_G}(h) = \mathrm{Id}_V$ if $h = 1_G$ and $\delta_{1_G}(h) = 0$ otherwise. Thus, if $x \in V^G$ we have

$$\begin{aligned}\Psi(1_{\mathrm{End}_{\mathbb{K}}(V)[G]})(x)(g) &= \Psi(\delta_{1_G})(x)(g)\\ &= \sum_{h\in G} \delta_{1_G}(h)(x(gh))\\ &= x(g)\end{aligned}$$

for all $g \in G$. It follows that $\Psi(1_{\mathrm{End}_{\mathbb{K}}(V)[G]}) = \mathrm{Id}_{V^G}$. This proves that $\Psi$ is a ring homomorphism.

Let us show now that $\Psi$ is injective. Let $\alpha \in \ker(\Psi)$. Then,

$$\tau_\alpha(x) = 0 \tag{8.9}$$

for all $x \in V^G$. Taking $x = c_v$ in (8.9), where, for $v \in V$, the element $c_v \in V[G]$ is as in (8.6), we deduce from Proposition 8.5.1(ii) that $\alpha(g)(v) = 0$ for all $v \in V$ and $g \in G$. Thus, we have $\alpha(g) = 0$ for all $g \in G$ and therefore $\alpha = 0$. It follows that $\ker(\Psi) = \{0\}$, that is, $\Psi$ is injective.

Let us show that $\Psi$ is surjective. Suppose that $\tau \in \mathrm{LCA}(G; V)$ has memory set $S$ and local defining map $\mu \colon V^S \to V$. As $\mu$ is $\mathbb{K}$-linear (cf. Proposition 8.1.1), there exist $\mathbb{K}$-linear maps $\alpha_s \colon V \to V$, $s \in S$, such that $\mu(y) = \sum_{s\in S} \alpha_s(y(s))$ for all $y \in V^S$. For all $x \in V^G$ and $g \in G$, we have

$$\tau(x)(g) = \mu((g^{-1}x)|_S) = \sum_{s\in S} \alpha_s\left(g^{-1}x(s)\right) = \sum_{s\in S} \alpha_s\left(x(gs)\right).$$

This shows that $\tau = \tau_\alpha$, where $\alpha \in \mathrm{End}_{\mathbb{K}}(V)[G]$ is defined by

$$\alpha(h) = \begin{cases} \alpha_s & \text{if } h = s \in S,\\ 0 & \text{otherwise.}\end{cases}$$

Thus $\Psi$ is surjective. □

If the vector space $V$ is one-dimensional over the field $\mathbb{K}$, then each endomorphism of $V$ is of the form $v \mapsto kv$, for some $k \in \mathbb{K}$. It follows that the $\mathbb{K}$-algebra $\mathrm{End}_{\mathbb{K}}(V)$ is canonically isomorphic to $\mathbb{K}$ in this case. Thus, we get:

**Corollary 8.5.3.** *Let $G$ be a group and let $V$ be a one-dimensional vector space over a field $\mathbb{K}$. Then the map $\Psi \colon \mathbb{K}[G] \to \mathrm{LCA}(G; V)$ defined by*

$$\Psi(\alpha)(x)(g) = \sum_{h\in G} \alpha(h)x(gh)$$

*for all $\alpha \in \mathbb{K}[G]$, $x \in V^G$, and $g \in G$, is a $\mathbb{K}$-algebra isomorphism.* □

When $G$ is an abelian group, then the $\mathbb{K}$-algebra $\mathbb{K}[G]$ is commutative for any field $\mathbb{K}$ by Remark 8.4.5(c). Thus, as an immediate consequence of the preceding corollary, we get:

**Corollary 8.5.4.** *Let $G$ be an abelian group and let $V$ be a one-dimensional vector space over a field $\mathbb{K}$. Then the $\mathbb{K}$-algebra $\mathrm{LCA}(G;V)$ is commutative.* □

*Examples 8.5.5.* (a) Let $G$ be a group, $S$ a nonempty finite subset of $G$, and let $\mathbb{K}$ be a field. Consider the element

$$\alpha = |S|\delta_{1_G} - \sum_{s\in S} \delta_s \in \mathbb{K}[G],$$

Then, taking $V = \mathbb{K}$, we have, for all $x \in \mathbb{K}^G$ and $g \in G$,

$$\Psi(\alpha)(x)(g) = \sum_{h\in G} \alpha(h)x(gh) = |S|x(g) - \sum_{s\in S} x(gs).$$

It follows that $\Psi(\alpha)$ is the discrete Laplacian $\Delta_S \colon \mathbb{K}^G \to \mathbb{K}^G$ associated with $S$ (cf. Example 8.1.3(a)).

(b) Let $G$ be a group, $V$ a vector space over a field $\mathbb{K}$ and $f \in \mathrm{End}_{\mathbb{K}}(V)$. Consider the element $\alpha \in \mathrm{End}_{\mathbb{K}}(V)[G]$ defined by $\alpha = f\delta_{1_G}$. Let $x \in V^G$ and $g \in G$. We have

$$\Psi(\alpha)(x)(g) = \sum_{h\in G} \alpha(h)(x(gh)) = f(x(g)).$$

It follows that $\Psi(\alpha)$ is the linear cellular automaton $\tau \in \mathrm{LCA}(G;V)$ defined by $\tau(x) = f \circ x$ for all $x \in V^G$ (cf. Example 8.1.3(b)).

(c) Let $G$ be a group, $V$ a vector space over a field $\mathbb{K}$, and $s_0$ an element of $G$. Consider the group algebra element $\alpha \in \mathrm{End}_{\mathbb{K}}(V)[G]$ defined by $\alpha = \delta_{s_0}$. For all $x \in V^G$ and $g \in G$, we have

$$\Psi(\alpha)(x)(g) = \sum_{h\in G} \alpha(h)(x(gh)) = x(gs_0).$$

It follows that $\Psi(\alpha)$ is the linear cellular automaton $\tau \in \mathrm{LCA}(G;V)$ defined by $\tau(x) = x \circ R_{s_0}$, where $R_{s_0} \colon G \to G$ is the right multiplication by $s_0$ (cf. Example 8.1.3(c)).

(d) Let $G = \mathbb{Z}$, let $\mathbb{K}$ be a field, and let $V = \mathbb{K}[t]$ be the vector space consisting of all polynomials in the indeterminate $t$ with coefficients in $\mathbb{K}$. Denote by $D$ and $U$ the elements in $\mathrm{End}_{\mathbb{K}}(V)$ defined respectively by $D(p) = p'$ and $U(p) = tp$, for all $p \in V$. Consider the element

$$\alpha = \delta_0 - (U \circ D)\delta_1 \in \mathrm{End}_{\mathbb{K}}(V)[\mathbb{Z}].$$

For all $x = (x_n)_{n\in\mathbb{Z}} \in V^{\mathbb{Z}}$, we have $\Psi(\alpha)(x) = (y_n)_{n\in\mathbb{Z}} \in V^{\mathbb{Z}}$, where

$$y_n = \sum_{m\in\mathbb{Z}} \alpha(m)(x_{n+m}) = x_n - tx'_{n+1}$$

for all $n \in \mathbb{Z}$. It follows that $\Psi(\alpha)$ is the linear cellular automaton $\tau \in \mathrm{LCA}(\mathbb{Z}; \mathbb{K}[t])$ described in Example 8.1.3(d).

## 8.6 Modules over a Group Ring

Let $R$ be a ring and let $M$ be a left $R$-module. We denote by $\mathrm{End}_R(M)$ the endomorphism ring of $M$. Recall that $\mathrm{End}_R(M)$ is the set consisting of all maps $f\colon M \to M$ satisfying

$$f(x + x') = f(x) + f(x') \quad \text{and} \quad f(rx) = rf(x)$$

for all $r \in R$ and $x, x' \in M$. The ring operations in $\mathrm{End}_R(M)$ are given by

$$(f + f')(x) = f(x) + f'(x) \quad \text{and} \quad (ff')(x) = (f \circ f')(x) = f(f'(x))$$

for all $f, f' \in \mathrm{End}_R(M)$ and $x \in M$, and the unity element of $\mathrm{End}_R(M)$ is the identity map $\mathrm{Id}_M$.

Suppose that there is a group $G$ which acts on $M$ by endomorphisms. We then define a structure of left $R[G]$-module on $M$ as follows.

For $\alpha \in R[G]$ and $x \in M$, define the element $\alpha x \in M$ by

$$\alpha x = \sum_{g\in G} \alpha(g)gx. \tag{8.10}$$

Note that this definition makes sense. Indeed, one has $\alpha(g) \in R$ and $gx \in M$ for each $g \in G$. On the other hand, there is only finitely many nonzero terms in the right hand side of (8.10) since the support of $\alpha$ is finite.

**Proposition 8.6.1.** *One has:*

(i) $1_{R[G]}x = x$,
(ii) $\alpha(x + x') = \alpha x + \alpha x'$,
(iii) $(\alpha + \beta)x = \alpha x + \beta x$,
(iv) $\alpha(\beta x) = (\alpha\beta)x$

*for all $\alpha, \beta \in R[G]$ and $x, x' \in M$.*

*Proof.* (i) We have

$$1_{R[G]}x = \delta_{1_G}x = x.$$

(ii) We have

$$\begin{aligned}\alpha(x+x') &= \sum_{g\in G}\alpha(g)g(x+x') = \sum_{g\in G}(\alpha(g)gx+\alpha(g)gx')\\ &= \sum_{g\in G}\alpha(g)gx + \sum_{g\in G}\alpha(g)gx' = \alpha x+\alpha x'.\end{aligned}$$

(iii) We have

$$\begin{aligned}(\alpha+\beta)x &= \sum_{g\in G}(\alpha+\beta)(g)gx = \sum_{g\in G}(\alpha(g)+\beta(g))gx\\ &= \sum_{g\in G}\alpha(g)gx + \sum_{g\in G}\beta(g)gx = \alpha x+\beta x.\end{aligned}$$

(iv) We have, by using (8.3),

$$\begin{aligned}\alpha(\beta x) &= \sum_{h_1\in G}\alpha(h_1)h_1(\beta x)\\ &= \sum_{h_1\in G}\alpha(h_1)h_1\Bigg(\sum_{h_2\in G}\beta(h_2)h_2x\Bigg)\\ &= \sum_{h_1\in G}\sum_{h_2\in G}\alpha(h_1)\beta(h_2)h_1h_2x\\ &= \sum_{g\in G}\Bigg(\sum_{\substack{h_1,h_2\in G\\ h_1h_2=g}}\alpha(h_1)\beta(h_2)\Bigg)gx\\ &= \sum_{g\in G}(\alpha\beta)(g)gx\\ &= (\alpha\beta)x.\end{aligned}$$

□

It follows from Proposition 8.6.1 that the addition on $M$ and the multiplication $(\alpha, x)\mapsto \alpha x$ defined by (8.10) gives us a left $R[G]$-module structure on $M$. Observe that this $R[G]$-module structure extends the $R$-module structure on $M$ if we regard $R$ as a subring of $R[G]$ via the map $r\mapsto r\delta_{1_G}$ (cf. Remark 8.4.5(a)).

**Proposition 8.6.2.** *Let $f\colon M\to M$ be a map. Then the following conditions are equivalent:*

(a) $f\in \mathrm{End}_{R[G]}(M)$;
(b) $f\in \mathrm{End}_R(M)$ *and* $f$ *is $G$-equivariant.*

*Proof.* Suppose that $f$ satisfies (b). Then $f(x+x') = f(x)+f(x')$ for all $x, x'\in M$ since $f\in \mathrm{End}_R(M)$. On the other hand, if $\alpha\in R[G]$ and $x\in M$,

we have, by using the $G$-equivariance and the $R$-linearity of $f$,

$$f(\alpha x) = f\left(\sum_{g\in G} \alpha(g)gx\right) = \sum_{g\in G} \alpha(g)f(gx) = \sum_{g\in G} \alpha(g)gf(x) = \alpha f(x).$$

It follows that $f \in \mathrm{End}_{R[G]}(M)$. This shows that (b) implies (a).

Conversely, suppose that $f$ satisfies (a). Let $x, x' \in M$ and $r \in R$. Then we have $f(x + x') = f(x) + f(x')$, $f(rx) = f(r\delta_{1_G}x) = r\delta_{1_G}f(x) = rf(x)$, and $f(gx) = f(\delta_g x) = \delta_g f(x) = gf(x)$ since $f \in \mathrm{End}_{R[G]}(M)$. This shows that $f \in \mathrm{End}_R(M)$ and that $f$ is $G$-equivariant. Thus, (a) implies (b). □

*Remark 8.6.3.* Conversely, suppose that we are given a left $R[G]$-module $M$. Then, by applying the above construction, we recover the initial $R[G]$-module structure on $M$ if we start from the $R$-module structure on $M$ obtained by restricting the scalars and the $R$-linear action of $G$ on $M$ defined by $gx = \delta_g x$ for all $g \in G$ and $x \in M$.

## 8.7 Matrix Representation of Linear Cellular Automata

Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$.

Since the $G$-shift action on $V^G$ is $\mathbb{K}$-linear, it induces a structure of left $\mathbb{K}[G]$-module on $V^G$ which extends the $\mathbb{K}$-vector space structure on $V^G$ (see Sect. 8.6). Note that by applying (8.10) we get, for all $\alpha \in \mathbb{K}[G]$ and $x \in V^G$,

$$\alpha x = \sum_{h\in G} \alpha(h)hx, \tag{8.11}$$

that is,

$$(\alpha x)(g) = \sum_{h\in G} \alpha(h)x(h^{-1}g)$$

for all $g \in G$. Thus, $\alpha x$ may be regarded as the "convolution product" of $\alpha$ and $x$ (compare with (8.2)).

Let $\tau\colon V^G \to V^G$ be a linear cellular automaton. Then $\tau$ is $\mathbb{K}$-linear by definition. On the other hand, $\tau$ is $G$-equivariant by Proposition 1.4.4. Thus, it follows from Proposition 8.6.2 that $\tau$ is an endomorphism of the $\mathbb{K}[G]$-module $V^G$. Consequently, we have $\mathrm{LCA}(G; V) \subset \mathrm{End}_{\mathbb{K}[G]}(V^G)$. It is clear that $\mathrm{End}_{\mathbb{K}[G]}(V^G)$ is a subalgebra of the $\mathbb{K}$-algebra $\mathrm{End}_{\mathbb{K}}(V^G)$. Since $\mathrm{LCA}(G; V)$ is a subalgebra of $\mathrm{End}_{\mathbb{K}}(V^G)$ by Proposition 8.1.4, we get:

**Proposition 8.7.1.** *The set* $\mathrm{LCA}(G; V)$ *is a subalgebra of the* $\mathbb{K}$*-algebra* $\mathrm{End}_{\mathbb{K}[G]}(V^G)$. □

*Remark 8.7.2.* When $G$ is a finite group, we have $\mathrm{LCA}(G; V) = \mathrm{End}_{\mathbb{K}[G]}(V^G)$. Indeed, in this case, every $u \in \mathrm{End}_{\mathbb{K}[G]}(V^G)$ is a linear cellular automaton

with memory set $G$ and local defining map $\mu\colon V^G \to V$ given by $\mu(y) = u(y)(1_G)$, since $u(x)(g) = g^{-1}u(x)(1_G) = u(g^{-1}x)(1_G)$ for all $x \in V^G$ and $g \in G$.

Consider now the vector subspace $V[G] \subset V^G$ consisting of all configurations with finite support. Observe that if $x \in V^G$ has support $\Omega \subset G$ and $g \in G$, then the support of the configuration $gx$ is $g\Omega$. Therefore, $V[G]$ is a submodule of the $\mathbb{K}[G]$-module $V^G$.

Recall that, for $v \in V$, the configuration $c_v \in V[G]$ is defined by

$$c_v(g) = \begin{cases} v & \text{if } g = 1_G, \\ 0 & \text{otherwise.} \end{cases}$$

Note that if $g \in G$ and $v \in V$, then $gc_v = \delta_g c_v$ is the configuration which takes the value $v$ at $g$ and is identically 0 on $G \setminus \{g\}$. It follows that every $x \in V[G]$ can be written as

$$x = \sum_{g \in G} g c_{x(g)}. \tag{8.12}$$

**Proposition 8.7.3.** *Suppose that $(e_i)_{i\in I}$ is a basis of the $\mathbb{K}$-vector space $V$. Then the family of configurations $(c_{e_i})_{i\in I}$ is a free basis for the left $\mathbb{K}[G]$-module $V[G]$.*

*Proof.* Let $x \in V[G]$. As $(e_i)_{i\in I}$ is a $\mathbb{K}$-basis for $V$, we can find elements $\alpha_i \in \mathbb{K}[G]$, $i \in I$, such that

$$x(g) = \sum_{i \in I} \alpha_i(g) e_i$$

for all $g \in G$. This gives us

$$\begin{aligned} x &= \sum_{g \in G} g c_{x(g)} \\ &= \sum_{g \in G} g \left( \sum_{i \in I} \alpha_i(g) c_{e_i} \right) \\ &= \sum_{g \in G} \left( \sum_{i \in I} \alpha_i(g) g c_{e_i} \right) \\ &= \sum_{i \in I} \left( \sum_{g \in G} \alpha_i(g) g c_{e_i} \right) \\ &= \sum_{i \in I} \alpha_i c_{e_i}, \end{aligned}$$

where the last equality follows from (8.11). If $x = 0 \in V[G]$, then $x(g) = 0$ for all $g \in G$ and therefore $\alpha_i = 0$ for all $i \in I$. This shows that $(c_{e_i})_{i\in I}$ is a basis for the left $\mathbb{K}[G]$-module $V[G]$. □

As every vector space admits a basis, we deduce the following

**Corollary 8.7.4.** *The left $\mathbb{K}[G]$-module $V[G]$ is free.* □

If $\tau\colon V^G \to V^G$ is a linear cellular automaton, we have $\tau(V[G]) \subset V[G]$ by Proposition 8.2.3. Moreover, it follows from Proposition 8.2.4 that the map $\Lambda\colon \mathrm{LCA}(G;V) \to \mathrm{End}_{\mathbb{K}}(V[G])$ which associates with each $\tau \in \mathrm{LCA}(G;V)$ its restriction $\tau|_{V[G]}\colon V[G] \to V[G]$ is an injective homomorphism of $\mathbb{K}$-algebras. As we have $\tau|_{V[G]} \in \mathrm{End}_{\mathbb{K}[G]}(V[G]) \subset \mathrm{End}_{\mathbb{K}}(V[G])$ for all $\tau \in \mathrm{LCA}(G;V)$, we get:

**Proposition 8.7.5.** *Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$. Then the map $\Phi\colon \mathrm{LCA}(G;V) \to \mathrm{End}_{\mathbb{K}[G]}(V[G])$ defined by $\Phi(\tau) = \tau|_{V[G]}$, where $\tau|_{V[G]}\colon V[G] \to V[G]$ is the restriction of $\tau$ to $V[G]$, is an injective $\mathbb{K}$-algebra homomorphism.* □

When the alphabet is finite-dimensional, we have the following:

**Theorem 8.7.6.** *Let $G$ be a group and let $V$ be a finite-dimensional vector space over a field $\mathbb{K}$. Then the map $\Phi\colon \mathrm{LCA}(G;V) \to \mathrm{End}_{\mathbb{K}[G]}(V[G])$ defined by $\Phi(\tau) = \tau|_{V[G]}$, where $\tau|_{V[G]}\colon V[G] \to V[G]$ is the restriction of $\tau$ to $V[G]$, is a $\mathbb{K}$-algebra isomorphism.*

*Proof.* By the preceding proposition, it suffices to show that $\Phi$ is surjective. Let $u \in \mathrm{End}_{\mathbb{K}[G]}(V[G])$. Consider an element $x \in V[G]$. We have

$$x = \sum_{g\in G} g c_{x(g)}$$

by (8.12). This implies

$$u(x) = u\left(\sum_{g\in G} g c_{x(g)}\right) = \sum_{g\in G} g u(c_{x(g)})$$

Since $u$ is $\mathbb{K}[G]$-linear. We deduce that

$$u(x)(1_G) = \sum_{g\in G} u(c_{x(g)})(g^{-1}). \tag{8.13}$$

Suppose now that $\dim_{\mathbb{K}}(V) = d$ and let $e_1, e_2, \ldots, e_d$ be a $\mathbb{K}$-basis for $V$. Denote by $T_i \subset G$ the support of $u(c_{e_i})$, $1 \le i \le d$, and consider the finite subset $S \subset G$ defined by $S = \bigcup_{1\le i\le d} T_i^{-1}$.

If $v \in V$, we can write $v = \sum_{i=1}^{d} k_i e_i$ with $k_i \in \mathbb{K}$, $1 \le i \le d$. We then have $c_v = \sum_{i=1}^{d} k_i c_{e_i}$ and hence

$$u(c_v) = u\left(\sum_{i=1}^{d} k_i c_{e_i}\right) = \sum_{i=1}^{d} k_i u(c_{e_i}).$$

We deduce that $u(c_v)(g^{-1}) = 0$ if $g \notin S$. Thus, using (8.13), we get

$$u(x)(1_G) = \sum_{g \in S} u(c_{x(g)})(g^{-1}).$$

This shows that $u(x)(1_G)$ only depends on the restriction of $x$ to $S$. More precisely, there is a $\mathbb{K}$-linear map $\mu\colon V^S \to V$ such that

$$u(x)(1_G) = \mu(x|_S) \quad \text{for all } x \in V[G].$$

Using the $\mathbb{K}[G]$-linearity of $u$, we get

$$\begin{aligned} u(x)(g) &= (g^{-1}u(x))(1_G) \\ &= u(g^{-1}x)(1_G) \\ &= \mu((g^{-1}x)|_S) \end{aligned}$$

for all $x \in V[G]$ and $g \in G$. Therefore $u$ is the restriction to $V[G]$ of the linear cellular automaton $\tau\colon V^G \to V^G$ with memory set $S$ and local defining map $\mu$. This shows that $\Phi$ is surjective. □

*Remarks 8.7.7.* (a) When $G$ is a finite group and $V$ is a (not necessarily finite-dimensional) vector space over a field $\mathbb{K}$, one has $V[G] = V^G$ and $\Phi\colon \mathrm{LCA}(G;V) \to \mathrm{End}_{\mathbb{K}[G]}(V[G]) = \mathrm{End}_{\mathbb{K}[G]}(V^G)$ is the identity map by Remark 8.7.2.

(b) Suppose now that $G$ is an infinite group and that $V$ is an infinite-dimensional vector space over a field $\mathbb{K}$. Then the map $\Phi\colon \mathrm{LCA}(G;V) \to \mathrm{End}_{\mathbb{K}[G]}(V[G])$ is not surjective. To see this, take a $\mathbb{K}$-basis $(e_i)_{i\in I}$ for $V$. As $G$ and $I$ are infinite, we can find a family $(T_i)_{i\in I}$ of finite subsets of $G$ such that $\bigcup_{i\in I} T_i$ is infinite. Choose, for each $i \in I$, a configuration $z_i \in V[G]$ whose support is $T_i$. It follows from Proposition 8.7.3 that $V[G]$ is a free left $\mathbb{K}[G]$-module admitting the family $(c_{e_i})_{i\in I}$ as a basis. Therefore, there is an element $u \in \mathrm{End}_{\mathbb{K}[G]}(V[G])$ such that $u(c_{e_i}) = z_i$ for all $i \in I$. Let us show that $u$ is not in the image of $\Phi$. We proceed by contradiction. Suppose that $u$ is in the image of $\Phi$. This means that there is a linear cellular automaton $\tau\colon V^G \to V^G$ such that $u = \tau|_{V[G]}$. If $S$ is a memory set for $\tau$, this implies that, for $x \in V[G]$, the value of $u(x)(1_G)$ depends only on the restriction of $x$ to $S$. As $S$ is finite and $\bigcup_{i\in I} T_i^{-1}$ is infinite, we can find elements $g_0 \in G$ and $i_0 \in I$ such that $g_0 \notin S$ and $g_0 \in T_{i_0}^{-1}$. Consider the configuration $x_0 \in V[G]$ which takes the value $e_{i_0}$ at $g_0$ and is identically 0 on $G \setminus \{g_0\}$. Observe that $x_0 = g_0 c_{e_{i_0}}$, so that

$$u(x_0) = u(g_0 c_{e_{i_0}}) = g_0 u(c_{e_{i_0}}) = g_0 z_{i_0}.$$

Thus, we have

$$u(x_0)(1_G) = (g_0 z_{i_0})(1_G) = z_{i_0}(g_0^{-1}).$$

We deduce that $u(x_0)(1_G) \neq 0$ as $g_0^{-1}$ is in the support $T_{i_0}$ of $z_{i_0}$. This gives us a contradiction since $x_0$ coincides with the 0 configuration on $S$. This shows that $\Phi$ is not surjective.

(c) By combining the two preceding remarks with Theorem 8.7.6, we deduce that the map $\Phi\colon \mathrm{LCA}(G;V) \to \mathrm{End}_{\mathbb{K}[G]}(V[G])$ is surjective if and only if $G$ is finite or $V$ is finite-dimensional.

Let us recall the following basic facts from linear algebra. Let $R$ be a ring and let $d \geq 1$ be an integer. We denote by $\mathrm{Mat}_d(R)$ the ring consisting of all $d \times d$ matrices $A = (a_{ij})_{1\leq i,j\leq d}$ with entries $a_{ij} \in R$. Recall that the addition and the multiplication in $\mathrm{Mat}_d(R)$ are given by

$$A + B = (a_{ij} + b_{ij})_{1\leq i,j\leq d} \quad \text{and} \quad AB = \left(\sum_{k=1}^{d} a_{ik}b_{kj}\right)_{1\leq i,j\leq d}$$

for all $A = (a_{ij})_{1\leq i,j\leq d}, B = (b_{ij})_{1\leq i,j\leq d} \in \mathrm{Mat}_d(R)$. Suppose now that $M$ is a free left $R$-module with basis $e_1, e_2, \ldots, e_d$. If $u \in \mathrm{End}_R(M)$, then we have $u(e_i) = \sum_{j=1}^{d} a_{ij}e_j$, where $a_{ij} \in R$ for $1 \leq i, j \leq d$. Then the map $u \mapsto (a_{ij})_{1\leq i,j\leq d}$ is a ring isomorphism from $\mathrm{End}_R(M)$ onto $\mathrm{Mat}_d(R^{op})$, where $R^{op}$ denotes the opposite ring of $R$. Moreover, when $R$ is an algebra over a field $\mathbb{K}$, it is an isomorphism of $\mathbb{K}$-algebras.

Let $G$ be a group and let $V$ be a vector space of finite dimension $d \geq 1$ over a field $\mathbb{K}$. By Proposition 8.7.3, the left $\mathbb{K}[G]$-module $V[G]$ admits a free basis of cardinality $d$. As the $\mathbb{K}$-algebras $\mathbb{K}[G]$ and $(\mathbb{K}[G])^{op}$ are isomorphic by Corollary 8.4.4, we deduce from Theorem 8.7.6 the following:

**Corollary 8.7.8.** *Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$ of finite dimension $\dim_{\mathbb{K}}(V) = d \geq 1$. Then the $\mathbb{K}$-algebras $\mathrm{LCA}(G;V)$ and $\mathrm{Mat}_d(\mathbb{K}[G])$ are isomorphic.* □

*Remark 8.7.9.* For $d = 1$, Corollary 8.7.8 tells us that if $G$ is a group and $V$ is a one-dimensional vector space over a field $\mathbb{K}$, then the $\mathbb{K}$-algebras $\mathrm{LCA}(G;V)$ and $\mathbb{K}[G]$ are isomorphic. Note that this last result also follows from Corollary 8.5.3.

## 8.8 The Closed Image Property

As we have seen in Lemma 3.3.2, the image of a cellular automaton $\tau\colon A^G \to A^G$ is always closed in $A^G$ for the prodiscrete topology if the alphabet $A$ is finite. When $A$ is infinite, the image of $\tau$ may fail to be closed in $A^G$ (cf. Example 8.8.3). However, it turns out that if $A = V$ is a finite-dimensional

vector space over a field $\mathbb{K}$ and $\tau\colon V^G \to V^G$ is a linear cellular automaton, then the image of $\tau$ is closed in $V^G$ even when the field $\mathbb{K}$ is infinite.

**Theorem 8.8.1.** *Let $G$ be a group and let $V$ be a finite-dimensional vector space over a field $\mathbb{K}$. Let $\tau\colon V^G \to V^G$ be a linear cellular automaton. Then $\tau(V^G)$ is closed in $V^G$ for the prodiscrete topology.*

*Proof.* We split the proof into two steps. Suppose first that $G$ is countable. Then we can find a sequence $(A_n)_{n\in\mathbb{N}}$ of finite subsets of $G$ such that $G = \bigcup_{n\in\mathbb{N}} A_n$ and $A_n \subset A_{n+1}$ for all $n \in \mathbb{N}$. Let $S$ be a memory set for $\tau$ and let $B_n = A_n^{-S}$ denote the $S$-interior of $A_n$ (cf. Sect. 5.4). Note that $G = \bigcup_{n\in\mathbb{N}} B_n$ and $B_n \subset B_{n+1}$ for all $n \in \mathbb{N}$.

It follows from Proposition 5.4.3 that if $x$ and $x'$ are elements in $V^G$ such that $x$ and $x'$ coincide on $A_n$ then the configurations $\tau(x)$ and $\tau(x')$ coincide on $B_n$. Therefore, given $x_n \in V^{A_n}$ and denoting by $\widetilde{x}_n \in V^G$ a configuration extending $x_n$, the pattern

$$y_n = \tau(\widetilde{x}_n)|_{B_n} \in V^{B_n}$$

does not depend on the particular choice of the extension $\widetilde{x}_n$. Thus we can define a map $\tau_n\colon V^{A_n} \to V^{B_n}$ by setting $\tau_n(x_n) = y_n$ for all $x_n \in V^{A_n}$. It is clear that $\tau_n$ is $\mathbb{K}$-linear.

Let $y \in V^G$ and suppose that $y$ is in the closure of $\tau(V^G)$. Then, for all $n \in \mathbb{N}$ there exists $z_n \in V^G$ such that

$$y|_{B_n} = \tau(z_n)|_{B_n}. \tag{8.14}$$

Consider, for each $n \in \mathbb{N}$, the affine subspace $L_n \subset V^{A_n}$ defined by $L_n = \tau_n^{-1}(y|_{B_n})$. We have $L_n \neq \varnothing$ for all $n$ by (8.14). For $n \leq m$, the restriction map $V^{A_m} \to V^{A_n}$ induces an affine map $\pi_{n,m}\colon L_m \to L_n$. Consider, for all $n \leq m$, the affine subspace $K_{n,m} \subset L_n$ defined by $K_{n,m} = \pi_{n,m}(L_m)$. We have $K_{n,m'} \subset K_{n,m}$ for all $n \leq m \leq m'$ since $\pi_{n,m'} = \pi_{n,m} \circ \pi_{m,m'}$. As the sequence $K_{n,m}$ $(m = n, n+1, \dots)$ is a decreasing sequence of finite-dimensional affine subspaces, it stabilizes, i.e., for each $n \in \mathbb{N}$ there exist a non-empty affine subspace $J_n \subset L_n$ and an integer $m_0 = m_0(n) \geq n$ such that $K_{n,m} = J_n$ if $m_0 \leq m$. For all $n \leq n' \leq m$, we have $\pi_{n,n'}(K_{n',m}) \subset K_{n,m}$ since $\pi_{n,n'} \circ \pi_{n',m} = \pi_{n,m}$. Therefore, $\pi_{n,n'}$ induces by restriction an affine map $\rho_{n,n'}\colon J_{n'} \to J_n$ for all $n \leq n'$. We claim that $\rho_{n,n'}$ is surjective. To see this, let $u \in J_n$. Let us choose $m$ large enough so that $J_n = K_{n,m}$ and $J_{n'} = K_{n',m}$. Then we can find $v \in L_m$ such that $u = \pi_{n,m}(v)$. We have $u = \rho_{n,n'}(w)$, where $w = \pi_{n',m}(v) \in K_{n',m} = J_{n'}$. This proves the claim. Now, using the surjectivity of $\rho_{n,n+1}$ for all $n$, we construct by induction a sequence of elements $x_n \in J_n$, $n \in \mathbb{N}$, as follows. We start by choosing an arbitrary element $x_0 \in J_0$. Then, assuming $x_n$ has been constructed, we take as $x_{n+1}$ an arbitrary element in $\rho_{n,n+1}^{-1}(x_n)$. Since $x_{n+1}$ coincides with $x_n$ on $A_n$, there exists $x \in V^G$ such that $x|_{A_n} = x_n$ for all $n$. We have

$\tau(x)|_{B_n} = \tau_n(x_n) = y_n = y|_{B_n}$ for all $n$. Since $G = \cup_{n\in\mathbb{N}} B_n$, we deduce that $\tau(x) = y$. This ends the proof in the case when $G$ is a countable group.

We now drop the countability assumption on $G$ and prove the theorem in the general case. Let $S \subset G$ be a memory set for $\tau$ and denote by $H \subset G$ the subgroup of $G$ generated by $S$. Then $H$ is countable since it is finitely generated. Consider the restriction cellular automaton $\tau_H \colon V^H \to V^H$ and observe that, by the first step of the proof,

$$\tau_H(V^H) \text{ is closed in } V^H \tag{8.15}$$

for the prodiscrete topology.

With the notation from Sect. 1.7 we have, by virtue of (1.16), that

$$\tau(V^G) = \prod_{c\in G/H} \tau_c(V^c). \tag{8.16}$$

Also, it follows from (1.18) that, for all $c \in G/H$ and $g \in c$,

$$\tau_c(V^c) = (\phi_g^*)^{-1}\left(\tau_H(V^H)\right). \tag{8.17}$$

As the map $\phi_g^* \colon V^c \to V^H$ is a uniform isomorphism and therefore a homeomorphism, it follows from (8.15) and (8.17) that $\tau_c(V^c)$ is closed in $V^c$ for all $c \in G/H$. As the product of closed subspaces is closed in the product topology (cf. Proposition A.4.3), we deduce from (8.16) that $\tau(V^G)$ is closed in $V^G$. □

**Corollary 8.8.2.** *Let $G$ be a group and let $V$ be a finite-dimensional vector space over a field $\mathbb{K}$. Let $\tau \colon V^G \to V^G$ be a linear cellular automaton. Suppose that every configuration with finite support $y \in V[G]$ lies in the image of $\tau$. Then $\tau$ is surjective.*

*Proof.* By our hypothesis, we have $V[G] \subset \tau(V^G) \subset V^G$. As $V[G]$ is dense in $V^G$ by Proposition 8.2.1 and $\tau(V^G)$ is closed in $V^G$ by Theorem 8.8.1, we deduce that $\tau(V^G) = V^G$. Thus, $\tau$ is surjective. □

In the example below we show that if we drop the finite dimensionality of the alphabet vector space $V$, then the image of a linear cellular automaton may fail to be closed in $V^G$.

*Example 8.8.3.* Let $G = \mathbb{Z}$ and let $\mathbb{K}$ be a field. Consider the $\mathbb{K}$-algebra $V = \mathbb{K}[t]$ consisting of all polynomials in the indeterminate $t$ with coefficients in $\mathbb{K}$. A configuration in $V^{\mathbb{Z}}$ is therefore a sequence $x = (x_n)_{n\in\mathbb{Z}}$, where $x_n = x_n(t)$ is a polynomial for all $n \in \mathbb{Z}$. Consider the $\mathbb{K}$-linear map $\tau \colon V^{\mathbb{Z}} \to V^{\mathbb{Z}}$ defined by setting $\tau(x) = y$ where $y = (y_n)_{n\in\mathbb{Z}}$ is given by

$$y_n = x_{n+1} - tx_n$$

for all $n \in \mathbb{Z}$. Then $\tau$ is a linear cellular automaton with memory set $\{0, 1\}$ and local defining map $\mu \colon V^{\{0,1\}} \to V$ given by $\mu(x_0, x_1) = x_1 - tx_0$.

Consider the configuration $z = (z_n)_{n\in\mathbb{Z}}$, where $z_n = 1$ for all $n \in \mathbb{Z}$.

Let $\Omega$ be a finite subset of $\mathbb{Z}$ and choose an integer $M \in \mathbb{Z}$ such that $\Omega \subset [M, \infty)$. Consider the configuration $x = (x_n)_{n\in\mathbb{Z}}$ defined by

$$x_n = \begin{cases} 1 + t + \cdots + t^{n-M} & \text{if } n \geq M, \\ 0 & \text{if } n < M. \end{cases}$$

Observe that $x_{n+1} = tx_n + 1$ for all $n \geq M$, so that the configuration $y = \tau(x)$ coincides with $z$ on $[M, \infty)$ and hence on $\Omega$. Thus $z$ is in the closure of $\tau(V^{\mathbb{Z}})$ in $V^{\mathbb{Z}}$. On the other hand, the configuration $z$ is not in the image of $\tau$. Indeed, $z = \tau(x)$ for some $x \in V^{\mathbb{Z}}$ would imply $x_{n+1} = tx_n + 1$ and hence $\deg(x_{n+1}) > \deg(x_n)$ for all $n \in \mathbb{Z}$, which is clearly impossible. This shows that $\tau(V^{\mathbb{Z}})$ is not closed in $V^{\mathbb{Z}}$ for the prodiscrete topology.

Observe that every $y \in V[\mathbb{Z}]$ is in the image of $\tau$. Indeed, if $y = (y_n)_{n\in\mathbb{Z}} \in V[Z]$ has support contained in $[M, \infty)$ for some $M \in \mathbb{Z}$, we can construct $x = (x_n)_{n\in\mathbb{Z}} \in V^{\mathbb{Z}}$ with $\tau(x) = y$ inductively by setting $x_n = 0$ for all $n \leq M$ and $x_{n+1} = tx_n + y_n$ for all $n \geq M$. This shows that we cannot omit the hypothesis saying that $V$ is finite-dimensional in Corollary 8.8.2.

## 8.9 The Garden of Eden Theorem for Linear Cellular Automata

In this section, we present a linear version of Theorem 5.8.1. We first introduce the notion of mean dimension, which plays here the role which was played by the entropy in the finite alphabet case, and we present some of its basic properties. In the definition of mean dimension, the dimension of finite-dimensional vector spaces replaces the cardinality of finite sets. The proof of the properties of the mean dimension and the proof of the linear version of the Garden of Eden Theorem follow the same lines as their finite alphabet counterparts (see Sects. 5.7 and 5.8).

From now on, in this section, $G$ is an amenable group, $\mathcal{F} = (F_j)_{j\in J}$ a right Følner net for $G$, and $V$ a finite-dimensional vector space over a field $\mathbb{K}$.

For $E \subset G$, we denote by $\pi_E \colon V^G \to V^E$ the canonical projection (restriction map). We thus have $\pi_E(x) = x|_E$ for all $x \in V^G$. Note that $\pi_E$ is $\mathbb{K}$-linear so that, if $X$ is a vector subspace of $V^G$, then $\pi_E(X)$ is a vector subspace of $V^E$.

**Definition 8.9.1.** Let $X$ be a vector subspace of $V^G$. The *mean dimension* $\mathrm{mdim}_{\mathcal{F}}(X)$ of $X$ with respect to the right Følner net $\mathcal{F} = (F_j)_{j\in J}$ is defined by

$$\mathrm{mdim}_{\mathcal{F}}(X) = \limsup_j \frac{\dim(\pi_{F_j}(X))}{|F_j|}, \tag{8.18}$$

where $\dim(\pi_{F_j}(X))$ denotes the dimension of the $\mathbb{K}$-vector subspace $\pi_{F_j}(X) \subset V^{F_j}$.

*Remark 8.9.2.* In the particular case when $\mathbb{K}$ is a finite field, every finite-dimensional vector space $W$ over $\mathbb{K}$ is finite of cardinality $|W| = |\mathbb{K}|^{\dim(W)}$. Therefore, in this case, we have

$$\mathrm{mdim}_{\mathcal{F}}(X) = \frac{1}{\log |\mathbb{K}|} \mathrm{ent}_{\mathcal{F}}(X)$$

for every vector subspace $X \subset V^G$.

Here are some immediate properties of mean dimension.

**Proposition 8.9.3.** *One has:*

(i) $\mathrm{mdim}_{\mathcal{F}}(V^G) = \dim(V)$;
(ii) $\mathrm{mdim}_{\mathcal{F}}(X) \leq \mathrm{mdim}_{\mathcal{F}}(Y)$ *if* $X \subset Y$ *are vector subspaces of* $V^G$;
(iii) $\mathrm{mdim}_{\mathcal{F}}(X) \leq \dim V$ *for all vector subspaces* $X \subset V^G$.

*Proof.* (i) If $X = V^G$, then, for every $j$, we have $\pi_{F_j}(X) = V^{F_j}$ and therefore

$$\frac{\dim(\pi_{F_j}(X))}{|F_j|} = \frac{\dim(V^{F_j})}{|F_j|} = \frac{|F_j| \dim(V)}{|F_j|} = \dim(V).$$

This implies that $\mathrm{mdim}_{\mathcal{F}}(X) = \dim(V)$.

(ii) If $X \subset Y$ are vector subspaces of $V^G$, then $\pi_{F_j}(X) \subset \pi_{F_j}(Y)$ are vector subspaces of $V^{F_j}$ and hence $\dim(\pi_{F_j}(X)) \leq \dim(\pi_{F_j}(Y))$ for all $j$. This implies $\mathrm{mdim}_{\mathcal{F}}(X) \leq \mathrm{mdim}_{\mathcal{F}}(Y)$.

(iii) This follows immediately from (i) and (ii). □

An important property of linear cellular automata is the fact that applying a linear cellular automaton to a vector subspace of configurations cannot increase the mean dimension of the subspace. More precisely, we have the following:

**Proposition 8.9.4.** *Let* $\tau \colon V^G \to V^G$ *be a linear cellular automaton and let* $X$ *be a vector subspace of* $V^G$. *Then one has*

$$\mathrm{mdim}_{\mathcal{F}}(\tau(X)) \leq \mathrm{mdim}_{\mathcal{F}}(X).$$

*Proof.* Let us set $Y = \tau(X)$ and observe that $Y$ is a vector subspace of $V^G$, by linearity of $\tau$. Let $S \subset G$ be a memory set for $\tau$. After replacing $S$ by $S \cup \{1_G\}$, we can assume that $1_G \in S$. Let $\Omega$ be a finite subset of $G$. Observe first that $\tau$ induces a map

$$\tau_\Omega \colon \pi_\Omega(X) \to \pi_{\Omega^{-S}}(Y)$$

defined as follows. If $u \in \pi_\Omega(X)$, then

$$\tau_\Omega(u) = (\tau(x))|_{\Omega^{-S}},$$

where $x$ is an element of $X$ such that $x|_\Omega = u$. Note that the fact that $\tau_\Omega(u)$ does not depend on the choice of such an $x$ follows from Proposition 5.4.3. Clearly $\tau_\Omega$ is surjective. Indeed, if $v \in \pi_{\Omega^{-S}}(Y)$, then there exists $x \in X$ such that $(\tau(x))|_{\Omega^{-S}} = v$. Then, setting $u = \pi_\Omega(x)$ we have, by construction, $\tau_\Omega(u) = v$. Since $\tau_\Omega$ is $\mathbb{K}$-linear and surjective, we have

$$\dim(\pi_{\Omega^{-S}}(Y)) \leq \dim(\pi_\Omega(X)). \tag{8.19}$$

Observe now that $\Omega^{-S} \subset \Omega$, since $1_G \in S$ (cf. Proposition 5.4.2(iv)). Thus $\pi_\Omega(Y)$ is a vector subspace of $\pi_{\Omega^{-S}}(Y) \times V^{\Omega \setminus \Omega^{-S}}$. This implies

$$\begin{aligned}
\dim(\pi_\Omega(Y)) &\leq \dim(\pi_{\Omega^{-S}}(Y) \times V^{\Omega \setminus \Omega^{-S}}) \\
&= \dim(\pi_{\Omega^{-S}}(Y)) + \dim(V^{\Omega \setminus \Omega^{-S}}) \\
&= \dim(\pi_{\Omega^{-S}}(Y)) + |\Omega \setminus \Omega^{-S}| \dim(V) \\
&\leq \dim(\pi_\Omega(X)) + |\Omega \setminus \Omega^{-S}| \dim(V),
\end{aligned}$$

by (8.19). As $\Omega \setminus \Omega^{-S} \subset \partial_S(\Omega)$, we deduce that

$$\dim(\pi_\Omega(Y)) \leq \dim(\pi_\Omega(X)) + |\partial_S(\Omega)| \dim(V).$$

By taking $\Omega = F_j$, this gives us

$$\frac{\dim(\pi_{F_j}(Y))}{|F_j|} \leq \frac{\dim(\pi_{F_j}(X))}{|F_j|} + \frac{|\partial_S(F_j)|}{|F_j|} \dim(V).$$

Since

$$\lim_j \frac{|\partial_S(F_j)|}{|F_j|} = 0$$

by Proposition 5.4.4, we finally get

$$\mathrm{mdim}_{\mathcal{F}}(Y) = \limsup_j \frac{\dim(\pi_{F_j}(Y))}{|F_j|} \leq \limsup_j \frac{\dim(\pi_{F_j}(X))}{|F_j|} = \mathrm{mdim}_{\mathcal{F}}(X).$$

□

By Proposition 8.9.3, the maximal value for the mean dimension of a vector subspace $X \subset V^G$ is $\dim(V)$. The following result gives a sufficient condition on $X$ which guarantees that its mean dimension is strictly less than $\dim(V)$.

**Proposition 8.9.5.** *Let $X$ be a $G$-invariant vector subspace of $V^G$. Suppose that there exists a finite subset $E \subset G$ such that $\pi_E(X) \subsetneq V^E$. Then one has* $\mathrm{mdim}_{\mathcal{F}}(X) < \dim(V)$.

*Proof.* Let $E' = \{g_1 g_2^{-1} : g_1, g_2 \in E\}$. By Proposition 5.6.3, we may find an $(E, E')$-tiling $T \subset G$.

For each $j \in J$, let us define, as in Proposition 5.6.4, the subset $T_j \subset T$ by $T_j = T \cap F_j^{-E} = \{g \in T : gE \subset F_j\}$ and set

$$F_j^* = F_j \setminus \coprod_{g \in T_j} gE.$$

Since $\pi_E(X) \subsetneqq V^E$ and $X$ is $G$-invariant, we have $\pi_{gE}(X) \subsetneqq V^{gE}$ for all $g \in G$. We thus have

$$\dim(\pi_{gE}(X)) \leq \dim(V^{gE}) - 1 = |gE| \dim(V) - 1 \quad \text{for all } g \in T. \tag{8.20}$$

As

$$\pi_{F_j}(X) \subset V^{F_j^*} \times \prod_{g \in T_j} \pi_{gE}(X),$$

we get

$$\begin{aligned}
\dim(\pi_{F_j}(X)) &\leq \dim(V^{F_j^*} \times \prod_{g \in T_j} \pi_{gE}(X)) \\
&= |F_j^*| \dim(V) + \sum_{g \in T_j} \dim(\pi_{gE}(X)) \\
&\leq |F_j^*| \dim(V) + \sum_{g \in T_j} (|gE| \dim(V) - 1) \quad \text{(by (8.20))} \\
&= \left( |F_j^*| + \sum_{g \in T_j} |gE| \right) \dim(V) - |T_j| \\
&= |F_j| \dim(V) - |T_j|,
\end{aligned}$$

since

$$|F_j| = |F_j^*| + \sum_{g \in T_j} |gE| \quad \text{and} \quad |gE| = |E|.$$

Now, by Proposition 5.6.4, there exist $\alpha > 0$ and $j_0 \in J$ such that $|T_j| \geq \alpha |F_j|$ for all $j \geq j_0$. Thus

$$\frac{\dim(\pi_{F_j}(X))}{|F_j|} \leq \dim(V) - \alpha \quad \text{for all } j \geq j_0.$$

This implies that

$$\mathrm{mdim}_{\mathcal{F}}(X) = \limsup_j \frac{\dim(\pi_{F_j}(X))}{|F_j|} \leq \dim(V) - \alpha < \dim(V).$$

□

We are now in position to state the linear version of the Garden of Eden theorem.

**Theorem 8.9.6.** *Let $G$ be an amenable group and let $V$ be a finite-dimensional vector space over a field $\mathbb{K}$. Let $\mathcal{F} = (F_j)_{j\in J}$ be a right Følner net for $G$. Let $\tau\colon V^G \to V^G$ be a linear cellular automaton. Then the following conditions are equivalent:*

(a) *$\tau$ is surjective;*
(b) $\mathrm{mdim}_{\mathcal{F}}(\tau(V^G)) = \dim(V)$*;*
(c) *$\tau$ is pre-injective.*

Since every injective cellular automaton is pre-injective, we immediately deduce the following:

**Corollary 8.9.7.** *Let $G$ be an amenable group and let $V$ be a finite-dimensional vector space over a field $\mathbb{K}$. Then every injective linear cellular automaton $\tau\colon V^G \to V^G$ is surjective.* □

We divide the proof of Theorem 8.9.6 into several lemmas.

**Lemma 8.9.8.** *Let $\tau\colon V^G \to V^G$ be a linear cellular automaton. Suppose that $\tau$ is not surjective. Then* $\mathrm{mdim}_{\mathcal{F}}(\tau(V^G)) < \dim(V)$.

*Proof.* Let $X = \tau(V^G)$ and choose a configuration $y \in V^G \setminus X$. Since $V^G \setminus X$ is an open subset of $V^G$ for the prodiscrete topology by Theorem 8.8.1, we can find a finite subset $\Omega \subset G$ such that $\pi_\Omega(y) \notin \pi_\Omega(X)$. Therefore we have $\pi_\Omega(X) \subsetneqq V^\Omega$. Observe that $X$ is a $G$-invariant vector subspace of $V^G$, as $\tau$ is $G$-equivariant and linear. By applying Proposition 8.9.5, we deduce that $\mathrm{mdim}_{\mathcal{F}}(X) < \dim(V)$. □

**Lemma 8.9.9.** *Let $\tau\colon V^G \to V^G$ be a linear cellular automaton. Suppose that*

$$\mathrm{mdim}_{\mathcal{F}}(\tau(V^G)) < \dim(V). \tag{8.21}$$

*Then $\tau$ is not pre-injective.*

*Proof.* Let us set $X = \tau(V^G)$. Let $S$ be a memory set for $\tau$ such that $1_G \in S$.

As $\pi_{F_j^{+S}}(X)$ is a vector subspace of $\pi_{F_j}(X) \times V^{F_j^{+S}\setminus F_j}$, we have

$$\begin{aligned}\dim(\pi_{F_j^{+S}}(X)) &\leq \dim(\pi_{F_j}(X)) + |F_j^{+S} \setminus F_j| \dim(V)\\ &\leq \dim(\pi_{F_j}(X)) + |\partial_S(F_j)| \dim(V).\end{aligned}$$

We thus have

$$\frac{\dim(\pi_{F_j^{+S}}(X))}{|F_j|} \leq \frac{\dim(\pi_{F_j}(X))}{|F_j|} + \frac{|\partial_S(F_j)|}{|F_j|} \dim(V).$$

It follows from (8.21) and Proposition 5.4.4 that we can find $j_0 \in J$ such that $\dim(\pi_{F_{j_0}^{+S}}(X)) < |F_{j_0}| \dim(V)$. Let $Z$ denote the (finite-dimensional) vector subspace of $V^G$ consisting of all configurations whose support is contained in $F_{j_0}$. Observe that $\tau(x)$ vanishes outside of $F_{j_0}^{+S}$ for every $x \in Z$ by Proposition 5.4.3. Thus we have

$$\begin{aligned}\dim(\tau(Z)) &= \dim(\pi_{F_{j_0}^{+S}}(\tau(Z))) \\ &\leq \dim(\pi_{F_{j_0}^{+S}}(X)) < |F_{j_0}| \dim(V) = \dim(Z).\end{aligned}$$

This implies that the restriction of $\tau$ to $Z$ is not injective. As all configurations in $Z$ have finite support, we deduce that $\tau$ is not pre-injective (cf. Proposition 8.2.5). □

**Lemma 8.9.10.** *Let $\tau\colon V^G \to V^G$ be a linear cellular automaton. Suppose that $\tau$ is not pre-injective. Then* $\mathrm{mdim}_{\mathcal{F}}(\tau(V^G)) < \dim(V)$.

*Proof.* Since the linear cellular automaton $\tau$ is not pre-injective, we can find, by Proposition 8.2.5, an element $x_0 \in V^G$ with nonempty finite support $\Omega \subset G$ such that $\tau(x_0) = 0$. Let $S$ be a memory set for $\tau$ such that $1_G \in S$ and $S = S^{-1}$. Let $E = \Omega^{+S^2}$. By Proposition 5.6.3, we can find a finite subset $F \subset G$ such that $G$ contains a $(E, F)$-tiling $T$. Note that for each $g \in G$, the support of $gx_0$ is $g\Omega \subset gE$. Let us choose, for each $g \in T$, a hyperplane $H_g \subset V^{g\Omega}$ which does not contain the restriction to $g\Omega$ of $gx_0$. Consider the vector subspace $X \subset V^G$ consisting of all $x \in V^G$ such that the restriction of $x$ to $g\Omega$ belongs to $H_g$ for each $g \in T$. We claim that $\tau(V^G) = \tau(X)$. Indeed, let $z \in V^G$. Then, for each $g \in T$, there exists a scalar $k_g \in \mathbb{K}$ such that the restriction to $g\Omega$ of $z + k_g(gx_0)$ belongs to $H_g$. Let $z' \in V^G$ be such that $\pi_{g\Omega}(z') = \pi_{g\Omega}(z + k_g(gx_0))$ for each $g \in T$ and $z' = z$ outside of $\coprod_{g\in T} g\Omega$. We have $z' \in X$ by construction. On the other hand, since $z'$ and $z$ coincide outside $\coprod_{g\in T} g\Omega$, we have $\tau(z') = \tau(z)$ outside $\coprod_{g\in T} g\Omega^{+S}$. Now, if $h \in g\Omega^{+S}$ for some $g \in T$, then $hS \subset g\Omega^{+S^2} = gE$ and therefore $\tau(z')(h) = \tau(z + k_g(gx_0))(h) = \tau(z)(h)$ since $gx_0$ lies in the kernel of $\tau$. Thus $\tau(z) = \tau(z')$ and the claim follows.

We deduce that

$$\mathrm{mdim}_{\mathcal{F}}(\tau(V^G)) = \mathrm{mdim}_{\mathcal{F}}(\tau(X)) \leq \mathrm{mdim}_{\mathcal{F}}(X) < \dim(V)$$

where the first inequality follows from Proposition 8.9.4 and the second one from Proposition 8.9.5. □

*Proof of Theorem 8.9.6.* Condition (a) implies (b) since we have $\mathrm{mdim}_{\mathcal{F}}(V^G) = \dim(V)$ by Proposition 8.9.3(i). The converse implication follows from Lemma 8.9.8. On the other hand, condition (c) implies (b) by Lemma 8.9.9. Finally, (b) implies (c) by Lemma 8.9.10. □

We end this section by showing that Corollary 8.9.7 as well as both implications (a) ⇒ (c) and (c) ⇒ (a) in Theorem 8.9.6 fail to hold when the vector space $V$ is infinite-dimensional.

*Example 8.9.11.* Let $G$ be any group and let $V$ be an infinite-dimensional vector space over a field $\mathbb{K}$. Let us choose a basis subset $B$ for $V$. Every map $\varphi\colon B \to B$ uniquely extends to a $\mathbb{K}$-linear map $\widetilde{\varphi}\colon V \to V$. The product map $\tau_\varphi = \widetilde{\varphi}^G\colon V^G \to V^G$ is a linear cellular automaton with memory set $S = \{1_G\}$ and local defining map $\widetilde{\varphi}$. Since $B$ is infinite, we can find a map $\varphi_1\colon B \to B$ which is surjective but not injective and a map $\varphi_2\colon B \to B$ which is injective but not surjective.

Consider first the cellular automaton $\tau_1 = \tau_{\varphi_1}$. Let us show that $\tau_1$ is surjective but not pre-injective. Observe that $\widetilde{\varphi}_1$ is surjective. As a product of surjective maps is a surjective map, it follows that $\tau_1$ is surjective. On the other hand, as $\widetilde{\varphi}_1$ is not injective, there exists $v \in \ker(\widetilde{\varphi}_1) \setminus \{0\}$. Consider the configuration $x \in V^G$ defined by $x(1_G) = v$ and $x(g) = 0$ for all $g \in G \setminus \{1_G\}$. Then $x \in V[G] \cap \ker(\tau_1)$ and $x \neq 0$. This shows that $\tau_1$ is not pre-injective.

Consider now the cellular automaton $\tau_2 = \tau_{\varphi_2}$. Let us show that $\tau_2$ is injective (and therefore pre-injective) but not surjective. Suppose that $x \in \ker(\tau_2)$, that is, $0 = \tau_2(x)(g) = \widetilde{\varphi}_2(x(g))$ for all $g \in G$. As $\widetilde{\varphi}_2$ is injective, we deduce that $x(g) = 0$ for all $g \in G$, in other words, $x = 0$. This shows that $\tau_2$ is injective. On the other hand, as $\widetilde{\varphi}_2$ is not surjective, there exists $v \in V \setminus \widetilde{\varphi}_2(V)$. Consider the constant configuration $x \in V^G$ where $x(g) = v$ for all $g \in G$. It is clear that $x \in V^G \setminus \tau_2(V^G)$. This shows that $\tau_2$ is not surjective.

## 8.10 Pre-injective but not Surjective Linear Cellular Automata

In this section, we give examples of linear cellular automata with finite-dimensional alphabet which are pre-injective but not surjective. We recall that the underlying groups for such automata cannot be amenable by Theorem 8.9.6.

**Proposition 8.10.1.** *Let $G = F_2$ be the free group of rank two and let $V$ be a two-dimensional vector space over a field $\mathbb{K}$. Then there exists a linear cellular automaton $\tau\colon V^G \to V^G$ which is pre-injective but not surjective.*

*Proof.* We may assume $V = \mathbb{K}^2$. Let $a$ and $b$ denote the canonical generators of $G = F_2$. Let $p_1$ and $p_2$ be the elements of $\mathrm{End}_{\mathbb{K}}(V)$ defined respectively by $p_1(v) = (k_1, 0)$ and $p_2(v) = (k_2, 0)$ for all $v = (k_1, k_2) \in V$. Consider the linear cellular automaton $\tau\colon V^G \to V^G$ given by

$$\tau(x)(g) = p_1(x(ga)) + p_2(x(gb)) + p_1(x(ga^{-1})) + p_2(x(gb^{-1}))$$

for all $x \in V^G$ and $g \in G$. This cellular automaton is the one described in Sect. 5.11 for $H = \mathbb{K}$. By Proposition 5.11.1, $\tau$ is pre-injective but not surjective. □

If $H$ is a subgroup of a group $G$ and $\tau\colon V^H \to V^H$ is a linear cellular automaton over $H$ which is pre-injective and not surjective, then the induced cellular automaton $\tau^G\colon V^G \to V^G$ is also linear, pre-injective and not surjective (see Proposition 1.7.4 and Proposition 5.2.2). Therefore, an immediate consequence of Proposition 8.10.1 is the following:

**Corollary 8.10.2.** *Let $G$ be a group containing a free subgroup of rank two and let $V$ be a two-dimensional vector space over a field $\mathbb{K}$. Then there exists a linear cellular automaton $\tau\colon V^G \to V^G$ which is pre-injective but not surjective.* □

This shows that implication (c) ⇒ (a) in Theorem 8.9.6 becomes false if $G$ is a group containing a free subgroup of rank two.

## 8.11 Surjective but not Pre-injective Linear Cellular Automata

We now present examples of linear cellular automata with finite-dimensional alphabet which are surjective but not pre-injective. We recall that, by Theorem 8.9.6, the underlying groups for such examples are necessarily non-amenable.

**Proposition 8.11.1.** *Let $G = F_2$ be the free group of rank two and let $V$ be a two-dimensional vector space over a field $\mathbb{K}$. Then there exists a linear cellular automaton $\tau\colon V^G \to V^G$ which is surjective but not pre-injective.*

*Proof.* Let $a, b$ denote the canonical generators of $G$. We can assume that $V = \mathbb{K}^2$. Consider the elements $q_1, q_2 \in \mathrm{End}_{\mathbb{K}}(V)$ respectively defined by $q_1(v) = (k_1, 0)$ and $q_2(v) = (0, k_1)$ for all $v = (k_1, k_2) \in V$.

Let $\tau\colon V^G \to V^G$ be the linear cellular automaton given by

$$\tau(x)(g) = q_1(x(ga)) + q_1(x(ga^{-1})) + q_2(x(gb)) + q_2(x(gb^{-1}))$$

for all $x \in V^G$, $g \in G$. A memory set for $\tau$ is the set $S = \{a, b, a^{-1}, b^{-1}\}$.

Let $k_0$ be any nonzero element in $\mathbb{K}$. Consider the configuration $x_0 \in V^G$ defined by

$$x_0(g) = \begin{cases} (0, k_0) & \text{if } g = 1_G \\ (0, 0) & \text{otherwise.} \end{cases}$$

Then, $x_0$ is almost equal to 0 but $x_0 \neq 0$. As $\tau(x_0) = 0$, we deduce that $\tau$ is not pre-injective.

However, $\tau$ is surjective. To see this, let $z = (z_1, z_2) \in \mathbb{K}^G \times \mathbb{K}^G = V^G$. Let us show that there exists $x \in V^G$ such that $\tau(x) = z$. We define $x(g)$ by induction on the graph distance (cf. Sect. 6.2), which we denote by $|g|$, of $g \in G$ from $1_G$ in the Cayley graph of $G$. We first set $x(1_G) = (0, 0)$.

Then, for $s \in S$ we set

$$x(s) = \begin{cases} (z_1(1_G), 0) & \text{if } s = a \\ (z_2(1_G), 0) & \text{if } s = b \\ (0, 0) & \text{otherwise.} \end{cases} \tag{8.22}$$

Suppose that $x(g)$ has been defined for all $g \in G$ with $|g| \leq n$, for some $n \geq 1$. For $g \in G$ with $|g| = n$, let $g' \in G$ and $s' \in S$ be the unique elements such that $|g'| = n - 1$ and $g = g's'$. Then, for $s \in S$ with $s's \neq 1_G$, we set

$$x(gs) = \begin{cases} (z_1(g) - x_1(g'), 0) & \text{if } s' \in \{a, a^{-1}\} \text{ and } s = s' \\ (z_2(g), 0) & \text{if } s' \in \{a, a^{-1}\} \text{ and } s = b \\ (z_1(g), 0) & \text{if } s' \in \{b, b^{-1}\} \text{ and } s = a \\ (z_2(g) - x_2(g'), 0) & \text{if } s' \in \{b, b^{-1}\} \text{ and } s = s' \\ (0, 0) & \text{otherwise.} \end{cases} \tag{8.23}$$

Let us check that $\tau(x) = z$. We have

$$\begin{aligned} \tau(x)(1_G) &= q_1(x(a)) + q_1(x(a^{-1})) + q_2(x(b)) + q_2(x(b^{-1})) \\ \text{(by (8.22))} \quad &= q_1(z_1(1_G), 0) + q_1(0, 0) + q_2(z_2(1_G), 0) + q_2(0, 0) \\ &= (z_1(1_G), 0) + (0, 0) + (0, z_2(1_G)) + (0, 0) \\ &= (z_1(1_G), z_2(1_G)) \\ &= z(1_G). \end{aligned}$$

Let now $g = g's' \in G$ with $|g| = |g'| + 1 > 2$. Suppose, for instance, that $s' = a$. We then have:

$$\begin{aligned} \tau(x)(g) &= \tau(x)(g'a) \\ \text{(by (8.23))} \quad &= q_1(x(g'a^2)) + q_1(x(g')) + q_2(x(g'ab)) + q_2(x(g'ab^{-1})) \\ &= q_1(z_1(g) - x_1(g'), 0) + q_1(x_1(g'), x_2(g')) + q_2(z_2(g), 0) + q_2(0, 0) \\ &= (z_1(g) - x_1(g'), 0) + (x_1(g'), 0) + (0, z_2(g)) + (0, 0) \\ &= (z_1(g), z_2(g)) \\ &= z(g). \end{aligned}$$

The cases when $s' = a^{-1}, b, b^{-1}$ are similar. It follows that $\tau(x) = z$. This shows that $\tau$ is surjective. □

If $H$ is a subgroup of a group $G$, and $\tau\colon V^H \to V^H$ is a linear cellular automaton over $H$ which is surjective and not pre-injective, then the induced cellular automaton $\tau^G\colon V^G \to V^G$ is also linear, surjective and not

pre-injective (see Proposition 1.7.4 and Proposition 5.2.2). Therefore, an immediate consequence of Proposition 8.11.1 is the following:

**Corollary 8.11.2.** *Let $G$ be a group containing a free subgroup of rank two and let $V$ be a two-dimensional vector space over a field $\mathbb{K}$. Then there exists a linear cellular automaton $\tau\colon V^G \to V^G$ which is surjective but not pre-injective.* □

As a consequence, we deduce that implication (a) ⇒ (c) in Theorem 8.9.6 becomes false if the group $G$ contains a free subgroup of rank two.

## 8.12 Invertible Linear Cellular Automata

**Theorem 8.12.1.** *Let $G$ be a group and let $V$ be a finite-dimensional vector space over a field $\mathbb{K}$. Let $\tau\colon V^G \to V^G$ be an injective linear cellular automaton. Then there exists a linear cellular automaton $\sigma\colon V^G \to V^G$ such that $\sigma \circ \tau = \mathrm{Id}_{V^G}$.*

*Proof.* We split the proof into two steps. Suppose first that $G$ is countable. Since $\tau$ is injective, it induces a bijective map $\widetilde{\tau}\colon V^G \to Y$ where $Y = \tau(V^G)$ is the image of $\tau$. The fact that $\tau$ is $\mathbb{K}$-linear and $G$-equivariant implies that $Y$ is a $G$-invariant vector subspace of $V^G$ and that the inverse map $\widetilde{\tau}^{-1}\colon Y \to V^G$ is $\mathbb{K}$-linear and $G$-equivariant. Let us show that the following local property is satisfied by $\widetilde{\tau}^{-1}$: there exists a finite subset $T \subset G$ such that

(∗) for $y \in Y$, the element $\widetilde{\tau}^{-1}(y)(1_G)$ only depends on the restriction of $y$ to $T$.

Let us assume by contradiction that there exists no such $T$.

Let $S$ be a memory set for $\tau$ such that $1_G \in S$. Since $G$ is countable, we can find a sequence $(A_n)_{n\in\mathbb{N}}$ of finite subsets of $G$ such that $G = \bigcup_{n\in\mathbb{N}} A_n$, $S \subset A_0$ and $A_n \subset A_{n+1}$ for all $n \in \mathbb{N}$. Let $B_n = A_n^{-S}$ denote the $S$-interior of $A_n$ (cf. Sect. 5.4). Note that $G = \bigcup_{n\in\mathbb{N}} B_n$ and $B_n \subset B_{n+1}$ for all $n \in \mathbb{N}$.

Since there exists no finite subset $T \subset G$ satisfying condition (∗), we can find, for each $n \in \mathbb{N}$, two configurations $y'_n, y''_n \in Y$ such that $y'_n|_{B_n} = y''_n|_{B_n}$ and $\widetilde{\tau}^{-1}(y'_n)(1_G) \neq \widetilde{\tau}^{-1}(y''_n)(1_G)$. By linearity of $\widetilde{\tau}^{-1}$, the configuration $y_n = y'_n - y''_n \in Y$ satisfies $y_n|_{B_n} = 0$ and $\widetilde{\tau}^{-1}(y_n)(1_G) \neq 0$.

It follows from Proposition 5.4.3 that if $x$ and $x'$ are elements in $V^G$ such that $x$ and $x'$ coincide on $A_n$ then the configurations $\tau(x)$ and $\tau(x')$ coincide on $B_n$. Therefore, given $x_n \in V^{A_n}$ and denoting by $\widetilde{x}_n \in V^G$ a configuration extending $x_n$, the pattern

$$u_n = \tau(\widetilde{x}_n)|_{B_n} \in V^{B_n}$$

does not depend on the particular choice of the extension $\widetilde{x}_n$ of $x_n$. Thus we can define a map $\tau_n\colon V^{A_n} \to V^{B_n}$ by setting $\tau_n(x_n) = u_n$. It is clear that $\tau_n$ is $\mathbb{K}$-linear.

Consider, for each $n \in \mathbb{N}$, the vector subspace $L_n \subset V^{A_n}$ defined by $L_n = \mathrm{Ker}(\tau_n)$. where for $n \leq m$, the restriction map $V^{A_m} \to V^{A_n}$ induces a $\mathbb{K}$-linear map $\pi_{n,m} \colon L_m \to L_n$. Indeed, if $u \in L_m$, then we have $u|_{A_n} \in L_n$ since $\tau_m(u) = 0$ and therefore $\tau_n(u|_{A_n}) = (\tau_m(u))|_{B_n} = 0$. Consider now, for all $n \leq m$, the vector subspace $K_{n,m} \subset L_n$ defined by $K_{n,m} = \pi_{n,m}(L_m)$. We have $K_{n,m'} \subset K_{n,m}$ for all $n \leq m \leq m'$ since $\pi_{n,m'} = \pi_{n,m} \circ \pi_{m,m'}$. Therefore, if we fix $n$, the sequence $K_{n,m}$, where $m = n, n+1, \ldots$, is a decreasing sequence of vector subspaces of $L_n$. As $L_n \subset V^{A_n}$ is finite-dimensional, this sequence stabilizes, i.e., there exist a vector subspace $J_n \subset L_n$ and an integer $k_n \geq n$ such that $K_{n,m} = J_n$ for all $m \geq k_n$.

For all $n \leq n' \leq m$, we have $\pi_{n,n'}(K_{n',m}) \subset K_{n,m}$ since $\pi_{n,n'} \circ \pi_{n',m} = \pi_{n,m}$. Therefore, $\pi_{n,n'}$ induces by restriction a linear map $\rho_{n,n'} \colon J_{n'} \to J_n$ for all $n \leq n'$. We claim that $\rho_{n,n'}$ is surjective. To see this, let $u \in J_n$. Let us choose $m$ large enough so that $J_n = K_{n,m}$ and $J_{n'} = K_{n',m}$. Then we can find $v \in L_m$ such that $u = \pi_{n,m}(v)$. We have $u = \rho_{n,n'}(w)$, where $w = \pi_{n',m}(v) \in K_{n',m} = J_{n'}$. This proves the claim.

Now, using the surjectivity of $\rho_{n,n+1}$ for all $n$, we construct by induction a sequence of elements $z_n \in J_n$, $n \in \mathbb{N}$, as follows. We start by taking as $z_0$ the restriction of $x_{k_0}$ to $A_0$. Observe that $x_{k_0} \in L_{k_0}$ and hence $z_0 = \pi_{0,k_0}(x_{k_0}) \in J_0$. Then, assuming that $z_n$ has been constructed, we take as $z_{n+1}$ an arbitrary element in $\rho_{n,n+1}^{-1}(z_n)$. Since $z_{n+1}$ coincides with $z_n$ on $A_n$, there exists a unique element $z \in V^G$ such that $z|_{A_n} = z_n$ for all $n$. We have $z \in Y$, since $z_n \in \pi_{A_n}(Y)$ for all $n$ and $X$ is closed in $V^G$. As $z(1_G) = z_0(1_G) = x_{k_0}(1_G) \neq 0$, we have $z \neq 0$. On the other hand, $\tau(z) = 0$ since $\tau(z)|_{B_n} = \tau_n(z_n) = 0$ for all $n$ by construction. This contradicts the injectivity of $\tau$. Thus, there exists a finite subset $T \subset G$ satisfying $(*)$. Consider the linear map $\nu \colon V^T \to V$ defined by $\nu(z) = \widetilde{\tau}^{-1}(\widetilde{z})(1_G)$, where $\widetilde{z} \in V^G$ is any configuration extending the pattern $z \in V^T$. Then the linear cellular automaton $\sigma \colon V^G \to V^G$ with memory set $T$ and local defining map $\nu$ clearly satisfies $\sigma \circ \tau = \mathrm{Id}_{V^G}$. Indeed, if $x \in V^G$, then denoting by $y = \tau(x) \in Y$ its image by $\tau$, we have $x = \widetilde{\tau}^{-1}(y)$ and

$$\begin{aligned}(\sigma \circ \tau)(x)(1_G) &= \sigma(y)(1_G) \\ &= \nu(y|_T) \\ &= \widetilde{\tau}^{-1}(y)(1_G) \\ &= x(1_G)\end{aligned}$$

showing that $(\sigma \circ \tau)(x) = x$, by $G$-equivariance of $\sigma \circ \tau$. It follows that $\sigma \circ \tau = \mathrm{Id}_{V^G}$ and this proves the statement when $G$ is a countable group.

We now drop the countability assumption on $G$ and prove the theorem in the general case. Let $S \subset G$ be a memory set for $\tau$ and denote by $\mu \colon V^S \to V$ the corresponding local defining map. Let $H \subset G$ be the subgroup generated by $S$. Note that $H$ is countable. Consider the set $G/H$ of all left cosets of $H$ in $G$. For $c \in G/H$ denote by

$$\pi_c \colon V^G \to V^c = \prod_{g \in c} V$$

the projection map. We have $V^G = \prod_{c \in G/H} V^c$ and, for every $x \in V^G$ we write $x = (x_c)_{c \in G/H}$, where $x_c = \pi_c(x) \in V^c$.

For $c \in G/H$ and $g \in c$ denote by $\phi_g \colon H \to c$ the linear map defined by $\phi_g(h) = gh$ for all $h \in H$. Consider the map $\phi_g^* \colon V^c \to V^H$ defined by $\phi_g^*(z) = z \circ \phi_g$. Then, if $x = (x_c)_{c \in G/H} \in V^G$ and $c \in G/H$, we have

$$(\phi_g * (x_c))(h) = x_c(gh) = x(gh) = (g^{-1}x)(h) = (g^{-1}x)_H(h)$$

for all $h \in H$, that is,

$$\phi_g^*(x_c) = (g^{-1}x)_H. \tag{8.24}$$

For $c \in G/H$, define the map $\tau_c \colon V^c \to V^c$ by setting

$$\tau_c(z)(g) = \mu((\phi_g^*(z))|_S)$$

for all $z \in V^c$ and $g \in c$. Note that $\tau_H \colon V^H \to V^H$ is the restriction of $\tau$ to the subgroup $H$. We then have

$$\tau = \prod_{c \in G/H} \tau_c. \tag{8.25}$$

From (8.25) we immediately deduce that $Y = \prod_{c \in G/H} Y_c$, where $Y_c = \pi_c(Y) = \tau_c(V^c)$, and that $\tau_c$ is injective for all $c \in G/H$.

By the first part of the present proof, there exists a linear cellular automaton $\sigma_H \colon V^H \to V^H$ such that

$$\sigma_H \circ \tau_H = \mathrm{Id}_{V^H}. \tag{8.26}$$

Let $T \subset H$ be a memory set for $\sigma_H$ and let $\nu \colon V^T \to V$ be the corresponding local defining map. Consider the linear cellular automaton $\sigma \colon V^G \to V^G$ defined by setting

$$\sigma(y)(g) = \nu((g^{-1}y)|_T)$$

for all $y \in V^G$ and $g \in G$. Note that $\sigma = (\sigma_H)^G$ is the induced cellular automaton of $\sigma_H$ from $H$ to $G$, so that, if $\sigma_c \colon V^c \to V^c$ is the map defined by setting $\sigma_c(z)(g) = \nu((\phi_g^*(z))|_T)$ for all $z \in V^c$ and $g \in c$ then

$$\sigma = \prod_{c \in G/H} \sigma_c. \tag{8.27}$$

Given $c \in G/H$, $z \in V^c$ and $g \in c$, we have

$$\begin{aligned}(\sigma_c \circ \tau_c)(z)(g) &= \sigma_c(\tau_c(z))(g)\\ &= [\phi_g^* \sigma_c(\tau_c(z))](1_G)\\ \text{(by (8.24))} \quad &= \sigma_H(\phi_g^*(\tau_c(z)))(1_G)\\ \text{(again by (8.24))} \quad &= \sigma_H(\tau_H(\phi_g^*(z)))(1_G)\\ \text{(by (8.26))} \quad &= (\phi_g^*(z))(1_G)\\ &= z(g).\end{aligned}$$

This shows that $(\sigma_c \circ \tau_c)(z) = z$ for all $z \in V^c$. We deduce that $\sigma_c \circ \tau_c = \mathrm{Id}_{V^c}$ for all $c \in C$. It follows from (8.25) and (8.27) that $\sigma \circ \tau = \mathrm{Id}_{V^G}$. □

We recall that given a group $G$ and a set $A$, a cellular automaton $\tau\colon A^G \to A^G$ is said to be invertible if $\tau$ is bijective and the inverse map $\tau^{-1}\colon A^G \to A^G$ is also a cellular automaton. When $A$ is a finite set, every bijective cellular automaton $\tau\colon A^G \to A^G$ is invertible by Theorem 1.10.2. The following is a linear analogue.

**Corollary 8.12.2.** *Let $G$ be a group and let $V$ be a finite-dimensional vector space over a field $\mathbb{K}$. Then every bijective linear cellular automaton $\tau\colon V^G \to V^G$ is invertible.*

*Proof.* Let $\tau\colon V^G \to V^G$ be a bijective linear cellular automaton. It follows from Theorem 8.12.1 that there exists a linear cellular automaton $\sigma\colon V^G \to V^G$ such that $\sigma \circ \tau = \mathrm{Id}_{V^G}$. This implies that $\tau^{-1} = \sigma$ is a cellular automaton. Thus $\tau$ is an invertible cellular automaton. □

*Remark 8.12.3.* Let $G = \mathbb{Z}$ and let $\mathbb{K}$ be a field. Consider the infinite dimensional $\mathbb{K}$-vector space $V = \mathbb{K}[[t]]$ consisting of all formal power series in one indeterminate $t$ with coefficients in $\mathbb{K}$. Thus, an element of $V$ is just a sequence $v = (k_i)_{i\in\mathbb{N}}$ of elements of $\mathbb{K}$ written in the form

$$v = k_0 + k_1 t + k_2 t^2 + k_3 t^3 + \cdots = \sum_{i\in\mathbb{N}} k_i t^i.$$

Consider the map $\tau\colon V^{\mathbb{Z}} \to V^{\mathbb{Z}}$ defined by

$$\tau(x)(n) = x(n) - tx(n+1)$$

for all $x \in V^{\mathbb{Z}}$, $n \in \mathbb{Z}$. In Example 1.10.3 we have showed that $\tau$ is a bijective cellular automaton whose inverse map $\tau^{-1}\colon V^{\mathbb{Z}} \to V^{\mathbb{Z}}$ is not a cellular automaton. It is clear that $\tau$ is $\mathbb{K}$-linear. This shows that Corollary 8.12.2 becomes false if we omit the finite-dimensionality of the alphabet $V$.

Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$. We have seen in Sect. 1.10 that the set $\mathrm{ICA}(G; V)$ consisting of all invertible cellular automata $\tau\colon V^G \to V^G$ is a group for the composition of maps. The subset of $\mathrm{ICA}(G; V)$ consisting of all invertible linear cellular automata is a subgroup

of $\mathrm{ICA}(G;V)$ since it is the intersection of $\mathrm{ICA}(G;V)$ with the automorphism group of the $\mathbb{K}$-vector space $V^G$.

Given a ring $R$ and an integer $d \geq 1$, we denote by $\mathrm{GL}_d(R)$ the group of invertible elements of the matrix ring $\mathrm{Mat}_d(R)$.

An immediate consequence of Corollary 8.12.2 and Corollary 8.7.8 is the following

**Corollary 8.12.4.** *Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$ of finite dimension $\dim_{\mathbb{K}}(V) = d \geq 1$. Then the set consisting of all bijective linear cellular automata $\tau \colon V^G \to V^G$ is a subgroup of $\mathrm{ICA}(G;V)$ isomorphic to $\mathrm{GL}_d(\mathbb{K}[G])$.* □

## 8.13 Pre-injectivity and Surjectivity of the Discrete Laplacian

Let $G$ be a group and let $\mathbb{K}$ be a field. Given a nonempty finite subset $S$ of $G$, we recall that the discrete Laplacian associated with $G$ and $S$ is the linear cellular automaton $\Delta_S \colon \mathbb{K}^G \to \mathbb{K}^G$ (with memory set $S \cup \{1_G\}$) defined by

$$\Delta_S(f)(g) = |S| f(g) - \sum_{s \in S} f(gs)$$

for all $f \in \mathbb{K}^G$ and $g \in G$, where $|S|$ denotes the cardinality of $S$ (cf. Example 1.4.3(b) and Example 8.1.3(a)).

Let us observe that $\Delta_S$ is never injective since all constant maps $f \colon G \to \mathbb{K}$ are in the kernel of $\Delta_S$.

**Proposition 8.13.1.** *Let $G$ be a group and let $\mathbb{K}$ be a field. Let $S \subset G$ be a non-empty finite subset and suppose that the subgroup $H \subset G$ generated by $S$ is finite. Then $\Delta_S$ is neither pre-injective nor surjective.*

*Proof.* Let $(\Delta_S)_H \colon \mathbb{K}^H \to \mathbb{K}^H$ denote the restriction cellular automaton of $\Delta_S$ to $H$. Observe that $(\Delta_S)_H$ is the discrete Laplacian on $\mathbb{K}^H$ associated with $H$ and $S$. As we have seen above, $(\Delta_S)_H$ is not injective. As $H$ is finite, it follows that $(\Delta_S)_H$ is not pre-injective. On the other hand, the non-injectivity of $(\Delta_S)_H$ implies its non-surjectivity since $(\Delta_S)_H$ is $\mathbb{K}$-linear and $\mathbb{K}^H$ is finite-dimensional. By applying Proposition 5.2.2 (resp. Proposition 1.7.4), we deduce that $\Delta_S$ is not pre-injective (resp. not surjective). □

In the remaining of this section, we assume that the field $\mathbb{K}$ is the field $\mathbb{R}$ of real numbers. The following result is a Garden of Eden type theorem for the discrete laplacian. Note that there is no amenability hypothesis for the underlying group.

**Theorem 8.13.2.** *Let $G$ be a group and let $S$ be a nonempty finite subset of $G$. Let $\Delta_S \colon \mathbb{R}^G \to \mathbb{R}^G$ denote the associated discrete real laplacian. Then the following conditions are equivalent:*

(a) *$\Delta_S$ is surjective;*
(b) *the subgroup of $G$ generated by $S$ is infinite;*
(c) *$\Delta_S$ is pre-injective.*

Let us first establish the following simple fact. Recall that given a group $G$, the *subsemigroup* generated by a subset $S \subset G$ is the smallest subset $P$ of $G$ containing $S$ which is closed under the group operation, i.e., such that $p_1 p_2 \in P$ for all $p_1, p_2 \in P$. Clearly, $P$ consists of all elements of the form $p = s_1 s_2 \cdots s_n$ where $n \geq 1$ and $s_i \in S$ for $1 \leq i \leq n$.

**Lemma 8.13.3.** *Let $G$ be a group and let $S$ be a subset of $G$. Then the following conditions are equivalent:*

(a) *the subsemigroup of $G$ generated by $S$ is infinite;*
(b) *the subgroup of $G$ generated by $S$ is infinite.*

*Proof.* Let $P$ (resp. $H$) denote the subsemigroup (resp. subgroup) of $G$ generated by $S$. The implication (a) $\Rightarrow$ (b) is obvious since $P \subset H$.

Suppose that $S$ is nonempty and that $P$ is finite. Then, for each $s \in S$, the set $\{s^n : n \geq 1\} \subset P$ is also finite. This implies that every element of $S$ has finite order. Therefore, for each $s \in S$, we can find an integer $n \geq 2$ such that $s^n = 1_G$. It follows that $s^{-1} = s^{n-1} \in S$. Consequently, we have $P^{-1} = P$ and $1_G \in P$. This implies $H = P$, so that $H$ is finite. This shows (b) $\Rightarrow$ (a). □

*Proof of Theorem 8.13.2.* The implication (a) $\Rightarrow$ (b) immediately follows from Proposition 8.13.1.

Suppose that the subgroup generated by $S$ is infinite. Let $f$ be an element of $\mathbb{R}[G]$ such that $\Delta_S(f) = 0$. Let us show that $f = 0$. Let $g_0 \in G$ such that $|f(g_0)| = \max_{g \in G} |f(g)|$. Note that such a $g_0$ exists since $f$ takes only finitely many values. As

$$0 = \Delta_S(f)(g_0) = |S| f(g_0) - \sum_{s \in S} f(g_0 s),$$

we get

$$|f(g_0)| \leq \frac{1}{|S|} \sum_{s \in S} |f(g_0 s)| \tag{8.28}$$

by applying the triangle inequality. We deduce from (8.28) and the definition of $g_0$ that $|f(g_0 s)| = |f(g_0)|$ for all $s \in S$. By iterating the previous argument, we get $|f(g_0 p)| = |f(g_0)|$ for all $p \in P$, where $P$ denotes the subsemigroup of $G$ generated by $S$. As $P$ is infinite (cf. Lemma 8.13.3), this implies that $|f(h)| = |f(g_0)| = \max_{g \in G} |f(g)|$ for infinitely many $h \in G$. Since $f$ has finite

support, it follows that $f = 0$. Therefore, $\Delta_S$ is pre-injective. This proves the implication (b) $\Rightarrow$ (c).

To complete the proof, it suffices to prove that (c) implies (a).

Suppose first that $G$ is an amenable group. It follows from the implication (c) $\Rightarrow$ (a) in Theorem 8.9.6 that if $\Delta_S$ is pre-injective, then it is surjective. This shows the implication (c) $\Rightarrow$ (a) for $G$ amenable.

Suppose now that $G$ is non-amenable (and hence infinite) and that $S$ is a generating subset for $G$. This implies that $G$ is countable. Consider the Hilbert space $\ell^2(G) \subset \mathbb{R}^G$ consisting of all square-summable real functions on $G$ and the continuous linear map $\Delta_S^{(2)} \colon \ell^2(G) \to \ell^2(G)$ obtained by restriction of $\Delta_S$ to $\ell^2(G)$ (cf. Sect. 6.12).

By the Kesten-Day amenability criterion (Theorem 6.12.9), the non-amenability of $G$ implies that 0 does not belong to the spectrum $\sigma(\Delta_S^{(2)})$ of $\Delta_S^{(2)}$. Thus, $\Delta_S^{(2)}$ is bijective and hence

$$\mathbb{R}[G] \subset \ell^2(G) = \Delta_S^{(2)}(\ell^2(G)) = \Delta_S(\ell^2(G)) \subset \Delta_S(\mathbb{R}^G).$$

By applying Corollary 8.8.2, we deduce that $\Delta_S(\mathbb{R}^G) = \mathbb{R}^G$. This shows that $\Delta_S$ is surjective in the case when $G$ is non-amenable and $S$ generates $G$.

Finally, suppose now that $G$ is an arbitrary non-amenable group and that $\Delta_S$ is pre-injective. Denote by $H$ the subgroup of $G$ generated by $S$. Then the restriction cellular automaton $(\Delta_S)_H \colon \mathbb{R}^H \to \mathbb{R}^H$, which is the discrete laplacian associated with $H$ and $S$, is pre-injective by Proposition 5.2.2. The subgroup $H$ may be amenable or not. However, it follows from the two preceding cases that $(\Delta_S)_H$ is surjective. By applying Proposition 1.7.4, we deduce that $\Delta_S$ is surjective as well. This completes the proof that (a) implies (c). □

As a consequence of Theorem 8.13.2, we obtain the following characterization of locally finite groups in terms of real linear cellular automata:

**Corollary 8.13.4.** *Let $G$ be a group and let $V$ be a real vector space of finite dimension $d \geq 1$. Then the following conditions are equivalent:*

(a) *$G$ is locally finite;*
(b) *every surjective linear cellular automaton $\tau \colon V^G \to V^G$ is injective.*

*Proof.* Suppose (a). Let $\tau \colon V^G \to V^G$ be a surjective linear cellular automaton with memory set $S \subset G$. As $G$ is locally finite, the subgroup $H$ generated by $S$ is finite. Consider the linear cellular automaton $\tau_H \colon V^H \to V^H$ obtained from $\tau$ by restriction. Observe that $\tau_H$ is surjective by Proposition 1.7.4(ii). Since $V^H$ is finite-dimensional, it follows that $\tau_H$ is injective. By applying Proposition 1.7.4(i), we deduce that $\tau$ is also injective. This shows that (a) implies (b).

Now, suppose that $G$ is not locally finite. Let us show that there exists a linear cellular automaton $\tau \colon V^G \to V^G$ which is surjective but not injective.

We can assume that $V = \mathbb{R}^d$. Since $G$ is not locally finite, we can find a finite subset $S \subset G$ such that the subgroup $H$ generated by $S$ is infinite. By Theorem 8.13.2, the discrete Laplacian $\Delta_S \colon \mathbb{R}^G \to \mathbb{R}^G$ is surjective. As mentioned above $\Delta_S$ is not injective since all constant configurations are in its kernel. Consider now the product map $\tau = (\Delta_S)^d \colon (\mathbb{R}^d)^G \to (\mathbb{R}^d)^G$, where we use the natural identification $(\mathbb{R}^d)^G = (\mathbb{R}^G)^d$. Clearly, $\tau$ is a linear cellular automaton admitting $S \cup \{1_G\}$ as a memory set. On the other hand, $\tau$ is surjective but not injective since any product of surjective (resp. non-injective) maps is a surjective (resp. non-injective) map. This shows that (b) implies (a). □

## 8.14 Linear Surjunctivity

In analogy with the finite alphabet case, we introduce the following definition.

**Definition 8.14.1.** A group $G$ is said to be *L-surjunctive* if, for any field $\mathbb{K}$ and any finite-dimensional vector space $V$ over $\mathbb{K}$, every injective linear cellular automaton $\tau \colon V^G \to V^G$ is surjective.

**Proposition 8.14.2.** *Every subgroup of an L-surjunctive group is L-surjunctive.*

*Proof.* Suppose that $H$ is a subgroup of a L-surjunctive group $G$. Let $V$ be a finite-dimensional vector space over a field $\mathbb{K}$ and let $\tau \colon V^H \to V^H$ be an injective linear cellular automaton over $H$. Consider the cellular automaton $\tau^G \colon V^G \to V^G$ over $G$ obtained from $\tau$ by induction (see Sect. 1.7). The fact that $\tau$ is injective implies that $\tau^G$ is injective by Proposition 1.7.4(i). Also, $\tau^G$ is linear by Proposition 8.3.1. Since $G$ is L-surjunctive, it follows that $\tau^G$ is surjective. By applying Proposition 1.7.4(ii), we deduce that $\tau$ is surjective. This shows that $H$ is L-surjunctive. □

**Proposition 8.14.3.** *Let $G$ be a group. Then the following conditions are equivalent:*

(a) *$G$ is L-surjunctive;*
(b) *every finitely generated subgroup of $G$ is L-surjunctive.*

*Proof.* The fact that (a) implies (b) follows from Proposition 8.14.2. Conversely, let $G$ be a group all of whose finitely generated subgroups are L-surjunctive. Let $V$ be a finite dimensional vector space over a field $\mathbb{K}$ and let $\tau \colon V^G \to V^G$ be an injective linear cellular automaton with memory set $S$. Let $H$ denote the subgroup of $G$ generated by $S$ and consider the linear cellular automaton $\tau_H \colon V^H \to V^H$ obtained by restriction of $\tau$ (see Sect. 1.7 and Proposition 8.3.1). The fact that $\tau$ is injective implies that $\tau_H$ is injective by Proposition 1.7.4(i). As $H$ is finitely generated, it is L-surjunctive by

our hypothesis on $G$. It follows that $\tau_H$ is surjective. By applying Proposition 1.7.4(ii), we deduce that $\tau$ is also surjective. This shows that (b) implies (a). □

Note that the preceding proposition may be stated by saying that a group is L-surjunctive if and only if it is locally L-surjunctive.

Every finite group is obviously L-surjunctive. Indeed, if $G$ is a finite group and $V$ is a finite-dimensional vector space, then the vector space $V^G$ is also finite-dimensional and therefore every injective endomorphism of $V^G$ is surjective. More generally, it follows from Corollary 8.9.7 that every amenable group is L-surjunctive. In fact, we have the following result, which is a linear analogue of Theorem 7.8.1,

**Theorem 8.14.4.** *Every sofic group is L-surjunctive.*

*Proof.* Let $G$ be a sofic group. Let $V$ be a finite-dimensional vector space over a field $\mathbb{K}$ of dimension $\dim_{\mathbb{K}}(V) = d \geq 1$ and let $\tau \colon V^G \to V^G$ be an injective linear cellular automaton. We want to show that $\tau$ is surjective. Every subgroup of a sofic group is sofic by Proposition 7.5.4. On the other hand, it follows from Proposition 3.2.2 that a group is L-surjunctive if all its finitely generated subgroups are L-surjunctive. Thus we can assume that $G$ is finitely generated.

Let then $S \subset G$ be a finite symmetric generating subset of $G$. As usual, for $r \in \mathbb{N}$, we denote by $B_S(r) \subset G$ the ball of radius $r$ centered at $1_G$ in the Cayley graph associated with $(G, S)$. We set $Y = \tau(V^G)$. Observe that $Y$ is a $G$-invariant vector subspace of $V^G$. On the other hand, it follows from Theorem 8.8.1 that $Y$ is closed in $V^G$ with respect to the prodiscrete topology.

By Theorem 8.12.1, there exists a linear cellular automaton $\sigma \colon V^G \to V^G$ such that

$$\sigma \circ \tau = \mathrm{Id}_{V^G} .$$

Choose $r_0$ large enough so that the ball $B_S(r_0)$ is a memory set for both $\tau$ and $\sigma$. Let $\mu \colon V^{B_S(r_0)} \to V$ and $\nu \colon V^{B_S(r_0)} \to V$ denote the corresponding local defining maps for $\tau$ and $\sigma$ respectively.

We proceed by contradiction. Suppose that $\tau$ is not surjective, that is, $Y \subsetneq V^G$. Then, since $Y$ is closed in $V^G$, there exists a finite subset $\Omega \subset G$ such that $Y|_\Omega \subsetneq V^\Omega$. It is not restrictive, up to taking a larger $r_0$, again if necessary, to suppose that $\Omega \subset B_S(r_0)$. Thus, $Y|_{B_S(r_0)} \subsetneq V^{B_S(r_0)}$.

Let $\varepsilon > 0$ be such that

$$\varepsilon < \frac{1}{d|B(2r_0)| + 1}. \tag{8.29}$$

Note that from (8.29) we have $1 - \varepsilon > 1 - \frac{1}{d|B(2r_0)|+1}$ which yields

$$(1 - \varepsilon)^{-1} < 1 + \frac{1}{d|B(2r_0)|}. \tag{8.30}$$

Since $G$ is sofic, we can find a finite $S$-labeled graph $(Q, E, \lambda)$ such that

$$|Q(3r_0)| \geq (1-\varepsilon)|Q|, \tag{8.31}$$

where we recall that $Q(r)$, $r \in \mathbb{N}$, denotes the set of all $q \in Q$ such that there exists an $S$-labeled graph isomorphism $\psi_{q,r} \colon B_S(r) \to B(q,r)$ satisfying $\psi_{q,r}(1_G) = q$ (cf. Theorem 7.7.1).

Note the inclusions

$$Q(r_0) \supset Q(2r_0) \supset \cdots \supset Q(ir_0) \supset Q((i+1)r_0) \supset \cdots .$$

(cf. (7.41); see also Fig. 7.2). Also recall from Lemma 7.7.2 that $B(q, r_0) \subset Q(ir_0)$ for all $Q((i+1)r_0)$ and $i \geq 0$.

For each integer $i \geq 1$, we define the map $\mu_i \colon V^{Q(ir_0)} \to V^{Q((i+1)r_0)}$ by setting, for all $u \in V^{Q(ir_0)}$ and $q \in Q((i+1)r_0)$,

$$\mu_i(u)(q) = \mu\left(u|_{B(q,r_0)} \circ \psi_{q,r_0}\right)(1_G),$$

where $\psi_{q,2r_0}$ is the unique isomorphism of $S$-labeled graphs from $B_S(r_0) \subset G$ to $B(q, r_0) \subset Q$ sending $1_G$ to $q$ (cf. 7.39 and 7.40).

Similarly, we define the map $\nu_i \colon V^{Q(ir_0)} \to V^{Q((i+1)r_0)}$ by setting, for all $u \in V^{Q(ir_0)}$ and $q \in Q((i+1)r_0)$,

$$\nu_i(u)(q) = \nu\left(u|_{B(q,r_0)} \circ \psi_{q,r_0}\right)(1_G).$$

From the fact that $\tau^{-1} \circ \tau$ is the identity map on $V^G$, we deduce that the composite $\nu_{i+1} \circ \mu_i \colon V^{Q(ir_0)} \to V^{Q((i+2)r_0)}$ is the identity on $V^{Q((i+2)r_0)}$. More precisely, denoting by $\rho_i \colon V^{Q(ir_0)} \to V^{Q((i+2)r_0)}$ the restriction map, we have that $\nu_{i+1} \circ \mu_i = \rho_i$ for all $i \geq 1$. In particular, we have $\nu_2 \circ \mu_1 = \rho_1$. Thus, setting $Z = \mu_1(V^{Q(r_0)}) \subset V^{Q(2r_0)}$, we deduce that $\nu_2(Z) = \rho_1(V^{Q(r_0)}) = V^{Q(3r_0)}$. It follows that

$$\dim(Z) \geq d|Q(3r_0)|. \tag{8.32}$$

Let $Q' \subset Q(3r_0)$ be as in (7.59) and set $\overline{Q'} = \coprod_{q' \in Q'} B(q', r_0)$. Note that $\overline{Q'} \subseteq Q(2r_0)$ so that

$$|Q(2r_0)| = |Q'| \cdot |B_S(r_0)| + |Q(2r_0) \setminus \overline{Q'}|. \tag{8.33}$$

Now observe that, for all $q \in Q(2r_0)$, we have a natural isomorphism of vector spaces $Z|_{B(q,r_0)} \to Y|_{B_S(r_0)}$ given by $u \mapsto u \circ \psi_{q,r_0}$, where $\psi_{q,r_0}$ denotes as above the unique isomorphism of $S$-labeled graphs from $B_S(r_0)$ to $B(q, r_0)$ such that $\psi_{q,r_0}(1_G) = q$. Since $Y|_{B_S(r_0)} \subsetneq V^{B_S(r_0)}$, this implies that

$$\dim\left(Z|_{B(q,r_0)}\right) = \dim(Y|_{B_S(r_0)}) \leq d \cdot |B_S(r_0)| - 1, \tag{8.34}$$

for all $q \in Q'$.

Thus we have

$$\begin{aligned}\dim(Z) &\le \dim\big(Z|_{\overline{Q'}}\big) + \dim\big(Z|_{Q(2r_0)\setminus\overline{Q'}}\big)\\ &\le |Q'|\cdot(d\cdot|B_S(r_0)|-1) + d\cdot|Q(2r_0)\setminus\overline{Q'}|\\ &= d\Big(|Q(2r_0)| - \frac{|Q'|}{d}\Big)\end{aligned}$$

where the last equality follows from (8.33). Comparing this with (8.32) we obtain

$$|Q(3r_0)| \le |Q(2r_0)| - \frac{|Q'|}{d}.$$

Thus,

$$\begin{aligned}|Q| \ge |Q(2r_0)| &\ge |Q(3r_0)| + \frac{|Q'|}{d}\\ &\ge |Q(3r_0)| + \frac{|Q(3r_0)|}{d|B(2r_0)|} \quad \text{by (7.59)},\\ &= |Q(3r_0)|\Big(1 + \frac{1}{d|B(2r_0)|}\Big)\\ &> |Q(3r_0)|(1-\varepsilon)^{-1}\end{aligned}$$

where the last inequality follows from (8.30). This yields

$$|Q(3r_0)| < (1-\varepsilon)|Q|$$

which contradicts (8.31). This shows that $\tau(V^G) = Y = V^G$, that is, $\tau$ is surjective. It follows that the group $G$ is L-surjunctive. □

## 8.15 Stable Finiteness of Group Algebras

Let $R$ be a ring and denote by $1_R$ its unity element.

If $a, b \in R$ satisfy $ab = 1_R$, then one says that $b$ is a *right-inverse* of $a$ and that $a$ is a *left-inverse* of $b$. An element $a \in R$ is said to be *right-invertible* (resp. *left-invertible*) if it admits a right-inverse (resp. a left-inverse). Every invertible element in $R$ is both right-invertible and left-invertible. Conversely, suppose that an element $a \in R$ admits a right-inverse $b'$ and a left-inverse $b''$. Then, we have $ab' = 1_R$ and $b''a = 1_R$, so that $b' = (b''a)b' = b''(ab') = b''$. This shows that $a$ is invertible with inverse $a^{-1} = b' = b''$. Thus, if an element is both right-invertible and left-invertible, then it is invertible.

One says that the ring $R$ is *directly finite* if every right-invertible element (or, equivalently, every left-invertible element) is invertible. This is equivalent to saying that if any two elements $a, b \in R$ satisfy $ab = 1_R$, then they also satisfy $ba = 1_R$.

The ring $R$ is said to be *stably finite* if the matrix ring $\mathrm{Mat}_d(R)$ is directly finite for any $d \geq 1$. Observe that every stably finite ring $R$ is directly finite since $\mathrm{Mat}_1(R) = R$.

**Proposition 8.15.1.** *Every finite ring is stably finite.*

*Proof.* Let $R$ be a finite ring. Suppose that $a, b \in R$ satisfy $ab = 1_R$. Consider the map $f\colon R \to R$ defined by $f(r) = ar$. We have $f(br) = a(br) = (ab)r = 1_R r = r$ for all $r \in R$. Therefore, the map $f$ is surjective. As $R$ is finite, this implies that $f$ is also injective. Since $f(ba) = a(ba) = (ab)a = a = f(1_R)$, we deduce that $ba = 1_R$. This shows that every finite ring is directly finite.

If the ring $R$ is finite, then the ring $\mathrm{Mat}_d(R)$ is also finite for every $d \geq 1$. Consequently, every finite ring is stably finite. □

**Proposition 8.15.2.** *Every commutative ring is stably finite.*

*Proof.* Let $d \geq 1$ and suppose that $a \in \mathrm{Mat}_d(R)$ is right-invertible. Then there exists $b \in \mathrm{Mat}_d(R)$ such that $ab = I_d$. This implies that

$$\det(a)\det(b) = \det(ab) = \det(I_d) = 1_R.$$

It follows that $\det(a)$ is an invertible element in $R$. Therefore $a$ is invertible in $\mathrm{Mat}_d(R)$. This shows that $\mathrm{Mat}_d(R)$ is directly finite for any $d \geq 1$. Consequently, $R$ is stably finite. □

Let us give an example of a ring which is not directly finite.

*Example 8.15.3.* Let $R$ be a nonzero ring and consider the free left $R$-module $M = \oplus_{n\in\mathbb{N}} R$. Every element in $M$ can be represented in the form $m = (m_n)_{n\in\mathbb{N}}$ where $m_n \in R$ for all $n \in \mathbb{N}$ and $m_n = 0_R$ for all but finitely many $n \in \mathbb{N}$. Consider the maps $a\colon M \to M$ and $b\colon M \to M$ defined by setting $a(m) = m'$ (resp. $b(m) = m''$) where $m'_n = m_{n-1}$ for $n \geq 1$ and $m'_0 = 0_R$ (resp. $m''_n = m_{n+1}$ for all $n \in \mathbb{N}$). Then $a$ and $b$ are obviously $R$-linear, in other words, $a, b \in \mathrm{End}_R(M)$ and $ab = \mathrm{Id}_M = 1_{\mathrm{End}_R(M)}$. Consider the element $\overline{m} \in M$ such that $\overline{m}_0 = 1_R$ and $\overline{m}_n = 0_R$ for all $n \geq 1$. As $a(\overline{m}) = 0$ we have $ba(\overline{m}) = 0$. This shows that $ba \neq 1_{\mathrm{End}_R(M)}$. It follows that the ring $\mathrm{End}_R(M)$ is not directly finite.

Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$. Recall that the set $\mathrm{LCA}(G; V)$ consisting of all linear cellular automata $\tau\colon V^G \to V^G$ has a natural structure of $\mathbb{K}$-algebra in which the multiplication is given by the composition of maps.

**Proposition 8.15.4.** *Let $G$ be a group and let $V$ be a vector space over a field $\mathbb{K}$. Let $\tau\colon V^G \to V^G$ be a linear cellular automaton. Then the following hold:*

(i) *if $\tau$ is left-invertible in $\mathrm{LCA}(G; V)$, then $\tau$ is injective;*

(ii) *if $\tau$ is right-invertible in* $\mathrm{LCA}(G;V)$*, then $\tau$ is surjective;*
(iii) *if $\tau$ is invertible in* $\mathrm{LCA}(G;V)$*, then $\tau$ is bijective.*

*If, in addition, the vector space $V$ is finite-dimensional, then:*

(iv) *$\tau$ is left-invertible in* $\mathrm{LCA}(G;V)$ *if and only if $\tau$ is injective;*
(v) *$\tau$ is invertible in* $\mathrm{LCA}(G;V)$ *if and only if $\tau$ is bijective.*

*Proof.* (i) Suppose that $\tau \in \mathrm{LCA}(G;V)$ is left-invertible. This means that there exists $\sigma \in \mathrm{LCA}(G;V)$ such that $\sigma \circ \tau = \mathrm{Id}_{V^G}$. As $\mathrm{Id}_{V^G}$ is injective, this implies that $\tau$ is injective.

(ii) Suppose that $\tau \in \mathrm{LCA}(G;V)$ is right-invertible. This means that there exists $\sigma \in \mathrm{LCA}(G;V)$ such that $\tau \circ \sigma = \mathrm{Id}_{V^G}$. As $\mathrm{Id}_{V^G}$ is surjective, this implies that $\tau$ is surjective.

(iii) This immediately follows from (i) and (ii).

(iv) This immediately follows from (i) and Theorem 8.12.1.

(v) This immediately follows from (iii) and Corollary 8.12.2. □

*Remarks 8.15.5.* (a) If $V$ is infinite-dimensional, it may happen that a bijective linear cellular automaton $\tau\colon V^G \to V^G$ is not left-invertible in $\mathrm{LCA}(G;V)$. Therefore, the converses of assertions (i) and (iii) in Proposition 8.15.4 are false if we do not add the hypothesis that $V$ is finite-dimensional. To see this, consider the bijective linear cellular automaton $\tau\colon V^{\mathbb{Z}} \to V^{\mathbb{Z}}$ described in Remark 8.12.3. Then, there is no $\sigma \in \mathrm{LCA}(G;V)$ such that $\sigma \circ \tau = \mathrm{Id}_{V^G}$ since otherwise the inverse map of $\tau$ would coincide with $\sigma$, which is impossible as $\tau$ is not an invertible cellular automaton.

(b) The converse of assertion (ii) in Proposition 8.15.4 does not hold even under the additional hypothesis that $V$ is finite-dimensional. For instance, take an arbitrary field $\mathbb{K}$ and consider the linear cellular automaton $\tau\colon \mathbb{K}^{\mathbb{Z}} \to \mathbb{K}^{\mathbb{Z}}$ defined by $\tau(x)(n) = x(n+1) - x(n)$ for all $x \in \mathbb{K}^{\mathbb{Z}}$ and $n \in \mathbb{Z}$. Observe that $\tau$ is surjective. Indeed, given $y \in \mathbb{K}^{\mathbb{Z}}$, the configuration $x \in \mathbb{K}^{\mathbb{Z}}$ defined by

$$x(n) = \begin{cases} 0 & \text{if } n = 0, \\ y(0) + y(1) + \cdots + y(n-1) & \text{if } n > 0, \\ y(n) + y(n+1) + \cdots + y(-1) & \text{if } n < 0, \end{cases}$$

clearly satisfies $\tau(x) = y$. However, $\tau$ is not right-invertible in $\mathrm{LCA}(\mathbb{Z};\mathbb{K})$. To see this, observe that, as the $\mathbb{K}$-algebra $\mathrm{LCA}(\mathbb{Z};\mathbb{K})$ is commutative by Corollary 8.5.4, the right-invertibility of $\tau$ would imply the bijectivity of $\tau$ by Proposition 8.15.4(iii). But $\tau$ is not injective as all constant configurations are mapped to 0.

**Corollary 8.15.6.** *Let $G$ be a group and let $\mathbb{K}$ be a field. Let $V$ be a vector space over $\mathbb{K}$ of finite dimension* $\dim_{\mathbb{K}}(V) = d \geq 1$*. Then the following conditions are equivalent:*

(a) *every injective linear cellular automaton $\tau\colon V^G \to V^G$ is surjective;*
(b) *the $\mathbb{K}$-algebra* $\mathrm{LCA}(G;V)$ *is directly finite;*

(c) *the* $\mathbb{K}$*-algebra* $\mathrm{Mat}_d(\mathbb{K}[G])$ *is directly finite.*

*Proof.* The equivalence between conditions (b) and (c) follow from the fact that the $\mathbb{K}$-algebras $\mathrm{LCA}(G; V)$ and $\mathrm{Mat}_d(\mathbb{K}[G])$ are isomorphic by Corollary 8.7.8.

Suppose (a). Let $\tau \in \mathrm{LCA}(G; V)$ be a left-invertible element in $\mathrm{LCA}(G; V)$. Then $\tau$ is injective by Proposition 8.15.4(i). Condition (a) implies then that $\tau$ is surjective and hence bijective. By applying Proposition 8.15.4(v), we deduce that $\tau$ is invertible in $\mathrm{LCA}(G; V)$. This shows that (a) implies (b).

Conversely, suppose (b). Let $\tau\colon V^G \to V^G$ be an injective linear cellular automaton. Then $\tau$ is left-invertible in $\mathrm{LCA}(G; V)$ by Proposition 8.15.4(iv). It follows from condition (b) that $\tau$ is invertible in $\mathrm{LCA}(G; V)$. By using Proposition 8.15.4(ii), we deduce that $\tau$ is surjective. This shows that (b) implies (a). □

**Corollary 8.15.7.** *Let* $G$ *be a group. Then the following conditions are equivalent:*

(a) *the group* $G$ *is L-surjunctive;*
(b) *for any field* $\mathbb{K}$*, the group algebra* $\mathbb{K}[G]$ *is stably finite.* □

From Theorem 8.14.4 and Corollary 8.15.7, we deduce the following:

**Corollary 8.15.8.** *Let* $G$ *be a sofic group and let* $\mathbb{K}$ *be a field. Then the group algebra* $\mathbb{K}[G]$ *is stably finite.* □

## 8.16 Zero-Divisors in Group Algebras and Pre-injectivity of One-Dimensional Linear Cellular Automata

Let $R$ be a ring.

A nonzero element $a \in R$ is said to be a *left zero-divisor* (resp. a *right zero-divisor*) if there exists a nonzero element $b$ in $R$ such that $ab = 0$ (resp. $ba = 0$). Note that the existence of a left zero-divisor in $R$ is equivalent to the existence of a right zero-divisor. One says that the ring $R$ *has no zero-divisors* if $R$ admits no left (or, equivalently, no right) zero-divisors.

*Examples 8.16.1.* (a) Let $R = \mathbb{Z}/n\mathbb{Z}$, $n \geq 2$, be the (commutative) ring of integers modulo $n$. Then $R$ has no zero-divisors if and only if $n$ is a prime number.

(b) Let $G$ be a group and let $R$ be a nonzero ring. Suppose that $G$ contains an element $g_0$ of finite order $n \geq 2$ and let $H = \{1_G, g_0, g_0^2, \ldots, g_0^{n-1}\}$ denote the subgroup of $G$ generated by $g_0$. Consider the elements $\alpha, \beta \in R[G]$ defined respectively by

$$\alpha(g) = \begin{cases} 1 & \text{if } g = 1_G, \\ -1 & \text{if } g = g_0, \\ 0 & \text{if } g \notin \{1_G, g_0\}. \end{cases} \quad \text{and} \quad \beta(g) = \begin{cases} 1 & \text{if } g \in H, \\ 0 & \text{otherwise.} \end{cases}$$

One easily checks that $\alpha \neq 0$, $\beta \neq 0$, and $\alpha\beta = \beta\alpha = 0$. Thus, $\alpha$ and $\beta$ are both left and right zero-divisors in $R[G]$.

Let $R$ be a ring. An element $x \in R$ is called an *idempotent* if it satisfies $x^2 = x$. An idempotent $x \in R$ is said to be *proper* if $x \neq 0$ and $x \neq 1$.

**Proposition 8.16.2.** *Let $R$ be a ring. Then the following hold:*

(i) *if $R$ has no zero-divisors, then $R$ has no proper idempotents;*
(ii) *if $R$ has no proper idempotents, then $R$ is directly finite.*

*Proof.* (i) Every idempotent $x \in R$ satisfies $x(x-1) = x^2 - x = 0$. If $R$ has no zero-divisors, this implies $x = 0$ or $x = 1$.

(ii) Suppose that $R$ has no proper idempotents. Let $a, b \in R$ such that $ab = 1$. Then we have $(ba)^2 = b(ab)a = ba$ so that $ba$ is an idempotent. Therefore, $ba = 0$ or $ba = 1$. If $ba = 0$, then $a = (ab)a = a(ba) = 0$ so that $0 = 1$ and $R$ is reduced to 0. Therefore, we have $ba = 1$ in all cases. This shows that $R$ is directly finite. □

*Remarks 8.16.3.* (a) A ring without proper idempotents may admit zero-divisors. For example, the ring $\mathbb{Z}/8\mathbb{Z}$ has no proper idempotents. However, the classes of 2, 4, and 6 are zero-divisors in $\mathbb{Z}/8\mathbb{Z}$.

(b) A directly finite ring may admit proper idempotents. For example, take a nonzero commutative ring $R$. Then the ring $\mathrm{Mat}_2(R)$ is directly finite by Proposition 8.15.2. However, the matrices $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ are proper idempotents in $\mathrm{Mat}_2(R)$.

A group $G$ is called a *unique-product group* if, given any two non-empty finite subsets $A, B \subset G$, there exists an element $g \in G$ which can be uniquely expressed as a product $g = ab$ with $a \in A$ and $b \in B$.

*Remark 8.16.4.* A unique-product group is necessarily torsion-free. Indeed, suppose that $G$ is a group containing an element $g_0$ of order $n \geq 2$. Take $A = B = \{1_G, g_0, g_0^2, \ldots, g_0^{n-1}\}$. Then $AB = A$ and there is no $g \in AB$ which can be uniquely written in the form $g = ab$ with $a \in A$ and $b \in B$.

Recall that a *total ordering* on a set $X$ is a binary relation $\leq$ on $X$ which is reflexive, antisymmetric, transitive, and such that one has $x \leq y$ or $y \leq x$ for all $x, y \in X$. A group $G$ is called *orderable* if it admits a left-invariant total ordering, that is, a total ordering $\leq$ such that $g_1 \leq g_2$ implies $gg_1 \leq gg_2$ for all $g, g_1, g_2 \in G$.

*Examples 8.16.5.* (a) Every subgroup of an orderable group is orderable. Indeed, if $H$ is a subgroup of a group $G$ and $\leq$ is a left-invariant total ordering on $G$, then the restriction of $\leq$ to $H$ is a left-invariant total ordering on $H$.

(b) The additive groups $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ are orderable since the usual ordering on $\mathbb{R}$ is translation-invariant.

(c) The direct product of two orderable groups is an orderable group. Indeed, let $G_1$ and $G_2$ be two groups and suppose that $\leq_1$ and $\leq_2$ are left-invariant total orderings on $G_1$ and $G_2$ respectively. Then the *lexicographic ordering* $\leq$ on $G_1 \times G_2$ defined by

$$(g_1, g_2) \leq (h_1, h_2) \iff \begin{cases} g_1 \leq_1 g_2 \\ \text{or} \\ g_1 = g_2 \text{ and } h_1 \leq_2 h_2 \end{cases}$$

is a left-invariant total ordering on $G_1 \times G_2$.

(d) More generally, the direct product of any family (finite or infinite) of orderable groups is an orderable group. Indeed, let $(G_i)_{i \in I}$ be a family of orderable groups. Let $\leq_i$ be a left-invariant total ordering on $G_i$ for each $i \in I$. Let us fix a well-ordering on $I$, i.e., a total ordering such that every nonempty subset of $I$ admits a minimal element (the fact that any set can be well-ordered is a basic fact in set theory which may be deduced from the Axiom of Choice). Then the lexicographic ordering on $G = \prod_{i \in I} G_i$, which is defined by setting $g \leq h$ for $g = (g_i), h = (h_i) \in G$ if and only if $g = h$ or $g_{i_0} < h_{i_0}$ where $i_0 = \min\{i \in I : g_i \neq h_i\}$, is a left-invariant total ordering on $G$. As $\oplus_{i \in I} G_i$ is a subgroup of $\prod_{i \in I} G_i$, it follows that the direct sum of any family of orderable groups is an orderable group. Since a free abelian group is isomorphic to a direct sum of copies of $\mathbb{Z}$, we deduce in particular that every free abelian group is orderable.

(e) In fact, every torsion-free abelian group is orderable. Indeed, if $G$ is a torsion-free abelian group then $G$ embeds in the $\mathbb{Q}$-vector space $G \otimes_{\mathbb{Z}} \mathbb{Q}$ via the map $g \mapsto g \otimes 1$. On the other hand, the additive group underlying a $\mathbb{Q}$-vector space $V$ is isomorphic to a direct sum of copies of $\mathbb{Q}$ (since $V$ admits a $\mathbb{Q}$-basis) and hence orderable.

(f) Suppose that $G$ is a group containing a normal subgroup $N$ such that both $N$ and $G/N$ are orderable groups. Then the group $G$ is orderable. Indeed, let $\leq_1$ (resp. $\leq_2$) be a left-invariant total ordering on $N$ (resp. $G/N$). Then one easily checks that the binary relation $\leq$ on $G$ defined by setting

$$g \leq h \iff \begin{cases} \rho(g) \leq_2 \rho(h) \\ \text{or} \\ \rho(g) = \rho(h) \text{ and } 1_G \leq_1 g^{-1}h \end{cases}$$

is a left-invariant total ordering on $G$.

(g) Let $R$ be any ring. Consider the Heisenberg group

$$H_R = \left\{ M(x,y,z) = \begin{pmatrix} 1 & y & z \\ 0 & 1 & x \\ 0 & 0 & 1 \end{pmatrix} \quad : \quad x,y,z \in R \right\}.$$

(cf. Example 4.6.5). The kernel of the group homomorphism of $H_R$ onto $R^2$ given by $M(x,y,z) \mapsto (x,y)$ is the normal subgroup

$$N = \left\{ \begin{pmatrix} 1 & 0 & z \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad : \quad z \in R \right\}.$$

As the groups $N$ and $H_R/N$ are both abelian, it follows from Examples (e) and (f) above that $H_R$ is orderable.

Given a set $X$ equipped with a total ordering $\leq$, we denote by $\mathrm{Sym}(X,\leq)$ the group of order-preserving permutations of $X$, that is, the subgroup of $\mathrm{Sym}(X)$ consisting of all bijective maps $f\colon X \to X$ such that $x \leq y$ implies $f(x) \leq f(y)$ for all $x,y \in X$.

**Proposition 8.16.6.** *Let $G$ be a group. Then the following conditions are equivalent:*

(a) *$G$ is orderable;*
(b) *there exists a set $X$ equipped with a total ordering $\leq$ such that $G$ is isomorphic to a subgroup of $\mathrm{Sym}(X,\leq)$.*

*Proof.* Suppose that $\leq$ is a left invariant total ordering on $G$. Then the action of $G$ on itself given by left multiplication is order-preserving. Thus $G$ is isomorphic to a subgroup of $\mathrm{Sym}(G,\leq)$. This shows that (a) implies (b).

Conversely, suppose that $X$ is a set equipped with a total ordering $\leq$. Choose a well ordering $\leq_W$ on $X$. Then we can define a lexicographic ordering $\leq_L$ on $\mathrm{Sym}(X,\leq)$ by setting $f \leq_L g$ for $f,g \in \mathrm{Sym}(X,\leq)$ if and only if $f = g$ or $f(x_0) \leq g(x_0)$ where $x_0 = \min\{x \in X : f(x) \neq g(x)\}$. Clearly $\leq_L$ is a left invariant total ordering on $\mathrm{Sym}(X,\leq)$. It follows that $\mathrm{Sym}(X,\leq)$ is an orderable group. As any subgroup of an orderable group is orderable, we conclude that (b) implies (a). □

**Corollary 8.16.7.** *The group $\mathrm{Homeo}^+(\mathbb{R})$ of orientation-preserving homeomorphisms of $\mathbb{R}$ is orderable.* □

**Proposition 8.16.8.** *Every orderable group is a unique-product group.*

*Proof.* Let $G$ be an orderable group and let $\leq$ be a left-invariant total ordering on $G$. Let $A$ and $B$ be nonempty finite subsets of $G$. Set $b_m = \min B$ and let $a_m \in A$ be the unique element such that $a_m b_m = \min Ab_m$. Consider the element $g = a_m b_m$. For all $a \in A$ and $b \in B$, we have $a_m b_m \leq ab_m \leq ab$. Thus, if $g = ab$ for some $a \in A$ and $b \in B$, then $a_m b_m = ab_m = ab$, so that we get $a_m = a$ and $b_m = b$ after right and left cancellation. This shows that $G$ is a unique-product group. □

**Proposition 8.16.9.** *Let $G$ be a unique-product group and let $R$ be a ring with no zero-divisors. Then the group ring $R[G]$ has no zero-divisors.*

*Proof.* Let $\alpha$ and $\beta$ be nonzero elements in $R[G]$, and denote by $A$ and $B$ their supports. As $G$ is a unique-product group, there is an element $g \in AB$ such that there exists a unique element $(a, b) \in A \times B$ such that $g = ab$. Then we have

$$(\alpha\beta)(g) = \sum_{\substack{h_1 \in A, h_2 \in B \\ h_1 h_2 = g}} \alpha(h_1)\beta(h_2) = \alpha(a)\beta(b).$$

As $R$ has no zero-divisors, this implies $(\alpha\beta)(g) \neq 0$. Thus, we have $\alpha\beta \neq 0$. This shows that $R[G]$ has no zero-divisors. □

**Corollary 8.16.10.** *Let $G$ be a unique-product group and let $\mathbb{K}$ be a field. Then the group algebra $\mathbb{K}[G]$ has no zero-divisors.* □

Let $G$ be a group and let $\mathbb{K}$ be a field. We have seen in Corollary 8.5.3 that the map $\Psi\colon \mathbb{K}[G] \to \mathrm{LCA}(G; \mathbb{K})$ defined by

$$\Psi(\alpha)(x)(g) = \sum_{h \in G} \alpha(h)x(gh)$$

for all $\alpha \in \mathbb{K}[G]$, $x \in \mathbb{K}^G$, and $g \in G$, is a $\mathbb{K}$-algebra isomorphism.

**Proposition 8.16.11.** *Let $G$ be a group and let $\mathbb{K}$ be a field. Let $\alpha$ be a nonzero element in $\mathbb{K}[G]$. Then the following conditions are equivalent:*

(a) *$\alpha$ is not a left zero-divisor in $\mathbb{K}[G]$;*
(b) *the linear cellular automaton $\Psi(\alpha)\colon \mathbb{K}^G \to \mathbb{K}^G$ is pre-injective.*

*Proof.* By Proposition 8.2.5, the linear cellular automaton $\Psi(\alpha)$ is not pre-injective if and only if there exists a nonzero element $\beta \in \mathbb{K}[G]$ such that $\Psi(\alpha)(\beta) = 0$. As

$$\Psi(\alpha)(\beta)(g) = \sum_{h \in G} \alpha(h)\beta(gh) = \sum_{h \in G} \alpha(h)\beta^*(h^{-1}g^{-1}) = (\alpha\beta^*)(g^{-1}),$$

for all $\beta \in \mathbb{K}[G]$ and $g \in G$, we deduce that $\Psi(\alpha)$ is not pre-injective if and only if there exists a nonzero element $\beta \in \mathbb{K}[G]$ such that $\alpha\beta^* = 0$, that is, if and only if $\alpha$ is a left zero-divisor in $\mathbb{K}[G]$. This shows that conditions (a) and (b) are equivalent. □

**Corollary 8.16.12.** *Let $G$ be a group and let $\mathbb{K}$ be a field. Then the following conditions are equivalent:*

(a) *the group algebra $\mathbb{K}[G]$ has no zero-divisors;*
(b) *every non-identically-zero linear cellular automaton $\tau\colon \mathbb{K}^G \to \mathbb{K}^G$ is pre-injective.* □

From Corollary 8.16.12 and Theorem 8.9.6 we deduce the following.

**Corollary 8.16.13.** *Let $G$ be an amenable group and let $\mathbb{K}$ be a field such that $\mathbb{K}[G]$ has no zero-divisors. Then every non-identically-zero linear cellular automaton $\tau\colon \mathbb{K}^G \to \mathbb{K}^G$ is surjective.* □

Observe that it follows from Theorem 4.6.1, Example 8.16.5(e), Proposition 8.16.8, and Corollary 8.16.10 that torsion-free abelian groups satisfy the hypotheses of Corollary 8.16.13 for any field $\mathbb{K}$. This implies in particular that if $G$ is a torsion-free abelian group and $S$ is a nonempty subset of $G$ which is not reduced to the identity element, then the discrete Laplacian $\Delta_S\colon \mathbb{K}^G \to \mathbb{K}^G$ is surjective for any field $\mathbb{K}$.

## Notes

In the literature, the term *linear* is used by some authors with a different meaning, namely to designate a cellular automaton $\tau\colon A^G \to A^G$ for which the alphabet $A$ is a finite abelian group and $\tau$ is a group endomorphism of $A^G$. Such cellular automata are also called *additive cellular automata*.

Linear cellular automata with vector spaces as alphabets were considered by the authors in a series of papers starting with [CeC1]. The fact that the algebra of linear cellular automata over a group $G$ whose alphabet is a $d$-dimensional vector space over a field $\mathbb{K}$ is isomorphic to the algebra of $d \times d$ matrices with coefficients in $\mathbb{K}[G]$ (Corollary 8.7.8) was proved in Sect. 6 of [CeC1] for $d = 1$ and in Sect. 4 of [CeC2] for all $d \geq 1$. The representations of linear cellular automata both as elements in $\mathrm{End}_{\mathbb{K}}(V)[G]$ and as elements in $\mathrm{End}_{\mathbb{K}[G]}(V[G])$ were given in Sect. 4 of [CeC5].

Recall that a Laurent polynomial over a field $\mathbb{K}$ is a polynomial in the variable $t$ and its inverse $t^{-1}$ with coefficients in $\mathbb{K}$. The $\mathbb{K}$-algebra of Laurent polynomials is thus denoted by $\mathbb{K}[t, t^{-1}]$. Also, a Laurent polynomial matrix is a matrix whose entries are Laurent polynomials. For $d \geq 1$, denote by $\mathrm{Mat}_d(\mathbb{K}[t, t^{-1}])$ the $\mathbb{K}$-algebra of $d \times d$ Laurent polynomial matrices over $\mathbb{K}$. When $G = \mathbb{Z}$, there are canonical isomorphisms of $\mathbb{K}$-algebras $\mathbb{K}[\mathbb{Z}] \cong \mathbb{K}(t, t^{-1}]$ and $\mathrm{LCA}(\mathbb{Z}; \mathbb{K}^d) \cong \mathrm{Mat}_d(\mathbb{K}[t, t^{-1}])$. In [LiM, Sect. 1.6] cellular automata $\tau \in \mathrm{LCA}(\mathbb{Z}; \mathbb{K}^d)$ are called $(d \times d)$ *convolutional encoders*.

The notion of mean dimension for vector subspaces of $V^G$, where $G$ is an amenable group and $V$ is a finite-dimensional vector space, was introduced by Gromov in [Gro5] and [Gro6]. Mean dimension was used by Elek [Ele] to prove that, given any amenable group $G$ and any field $\mathbb{K}$, there exists a non-trivial homomorphism from the Grothendieck group of finitely generated modules over the group algebra $\mathbb{K}[G]$ into the additive group of real numbers. As for entropy, it can be shown, as an application of the Ornstein-Weiss convergence

theorem (see [OrW], [LiW], [Gro6], [Ele], [Kri]) that the lim sup in (8.18) is in fact a true limit and does not depend on the Følner net $\mathcal{F}$.

The fact that the image of a linear cellular automaton with finite dimensional alphabet is closed (Theorem 8.8.1) was proved in Sect. 3 of [CeC1]. The linear version of the Garden of Eden theorem (Theorem 8.9.6) was first proved for countable amenable groups in [CeC1] and then extended to all amenable groups in [CeC8] using induction and restriction for linear cellular automata. The linear version of the Garden of Eden theorem was generalized to $R$-linear cellular automata with coefficients in semisimple left $R$-modules of finite length over a ring $R$ and over amenable groups in [CeC4]. Invertibility of linear cellular automata with finite dimensional alphabet $V$ was proved in Sect. 3 of [CeC2] for bijective linear cellular automata $\tau\colon X \to Y$ between closed linear subshifts $X, Y \subset V^G$ over countable groups and in [CeC8] for bijective linear cellular automata $\tau\colon V^G \to V^G$ over any group $G$.

In [CeC11] it is shown that if $G$ is a non-periodic group, then for every infinite-dimensional vector space $V$ over a field $\mathbb{K}$ there exist a bijective cellular automaton $\tau\colon V^G \to V^G$ which is not invertible (cf. Theorem 8.12.1 and Remark 8.12.3) and a cellular automaton $\tau'\colon V^G \to V^G$ whose image $\tau'(V^G)$ is not closed in $V^G$ with respect to the prodiscrete topology (cf. Theorem 8.8.1 and Example 8.8.3).

Theorem 8.13.2 and Corollary 8.13.4 were proved in [CeC6] and [CeC9].

Directly finite rings are sometimes called *Dedekind finite rings* or *von Neumann finite rings*. In [Coh], P.M. Cohn constructed, for each integer $d \geq 1$, a ring $R$ such that $\mathrm{Mat}_d(R)$ is directly finite but $\mathrm{Mat}_{d+1}(R)$ is not directly finite. I. Kaplansky [Kap2, p. 122], [Kap3, Problem 23] observed that techniques from the theory of operator algebras could be used to prove that, for any group $G$ and any field $\mathbb{K}$ of characteristic 0, the group algebra $\mathbb{K}[G]$ is stably finite and asked whether this property remains true for fields of characteristic $p > 0$. The stable finiteness of $\mathbb{K}[G]$ in arbitrary characteristic was established for free-by-abelian groups $G$ by P. Ara, K.C. O'Meara and F. Perera in [AOP]. This was extended to all sofic groups by Elek and Szabó [ES1], using the notion of von Neumann dimension for continuous regular rings. The proof of Elek and Szabó's result via linear cellular automata which is presented in this chapter (cf. Corollary 8.15.8) was given in [CeC2]. The notion of L-surjunctivity was also introduced in [CeC2] and the equivalence between L-surjunctivity and stable finiteness was established in Corollary 4.3 therein. In [CeC3], the authors proved that if $R$ is a ring and $G$ is a residually finite group, then every injective $R$-linear cellular automaton over $G$ whose alphabet is an Artinian left $R$-module is surjunctive. In [CeC5], it was shown that if $R$ is a ring, then $R$-linear cellular automata with coefficients in left $R$-modules of finite length (thus a stronger condition than being Artinian) over sofic groups (thus a weaker condition than being residually finite) are surjunctive. This last result was used to show in [CeC5] that the group ring $R[G]$ is stably finite whenever $R$ is a left (or right) Artinian ring and $G$ is a

sofic group. This yields an extension of the Elek and Szabó stable finiteness result since any division ring is Artinian.

Unique-product groups were introduced by W. Rudin and H. Schneider in [RuS] under the name of *Ω-groups*. The question of the existence of a torsion-free group which is not unique-product was raised by Rudin and Schneider [RuS, p. 592]. This question was answered in the affirmative by E. Rips and Y. Segev [RiS] (see also [Pro]). More information on orderable groups may be found for example in [BoR], [Pas], and [Gla]. A group $G$ is said to be *locally indicable* if any nontrivial finitely generated subgroup of $G$ admits an infinite cyclic quotient. Every locally indicable group is orderable (see for example [BoR, Theorem 7.3.1] or [Gla, Lemma 6.9.1]). This implies in particular that locally nilpotent torsion-free groups, free groups, and fundamental groups of surfaces not homeomorphic to the real projective plane are all orderable groups. It was observed by G.M. Bergman [Ber] that the universal covering group $\widetilde{\mathrm{SL}_2(\mathbb{R})}$ of $\mathrm{SL}_2(\mathbb{R})$ is orderable but not locally indicable. The group $\widetilde{\mathrm{SL}_2(\mathbb{R})}$ is orderable since it has a natural faithful action by orientation-preserving homeomorphisms of the real line but it is not locally indicable since it contains nontrivial finitely generated perfect groups. By a recent result due to D.W. Morris [Morr, Theorem B], every amenable orderable group is locally indicable. The orderability of the braid groups $B_n$ was established independently by P. Dehornoy [Deh] and W. Thurston. It follows from a result of E.A. Gorin and V.Ja. Lin [GoL] that the group $B_n$ is not locally indicable for $n \geq 5$.

A group is called *bi-orderable* if it admits a total ordering which is both left and right invariant. All free groups and all locally nilpotent torsion-free groups are bi-orderable. For $n \geq 3$, the braid group $B_n$ is not bi-orderable. However, the pure braid groups $P_n$ (i.e., the kernel of the natural epimorphism of $B_n$ onto $\mathrm{Sym}_n$) is bi-orderable for all $n$. A theorem due to A.I. Mal'cev [Mal2] and B.H. Neumann [Neu1] says that, given a group $G$ equipped with a total ordering which is both left and right invariant and a field $\mathbb{K}$, the vector subspace of $\mathbb{K}^G$ consisting of all maps $x\colon G \to \mathbb{K}$ whose support is a well-ordered subset of $G$ is a division $\mathbb{K}$-algebra for the convolution product (see [Pas, Theorem 2.11 in Chap. 13]). When $G = \mathbb{Z}$, this division algebra is the field of Laurent series with coefficients in $\mathbb{K}$. The Mal'cev-Neumann theorem implies in particular that, for any bi-orderable group $G$ and any field $\mathbb{K}$, the group algebra $\mathbb{K}[G]$ embeds in a division $\mathbb{K}$-algebra and is therefore stably finite.

A famous conjecture attributed to Kaplansky is the *zero-divisor conjecture* which states that if $G$ is a torsion-free group then the group algebra $\mathbb{K}[G]$ has no zero-divisors for any field $\mathbb{K}$ (see [Kap1], [Kap2, p. 122], [Pas, Chap. 13]). The observation that unique-product groups (and hence orderable groups) satisfy the Kaplansky zero-divisor conjecture (see Corollary 8.16.10) was made by Rudin and Schneider [RuS, Theorem 3.2]. The class of *elementary amenable* groups is the smallest class of groups containing all finite

and all abelian groups that is closed under taking subgroups, quotients, extensions, and directed unions. It is known (see [KLM, Theorem 1.4]) that torsion-free elementary amenable groups satisfy the Kaplansky zero-divisors conjecture. This implies in particular that all torsion-free virtually solvable groups satisfy the Kaplansky zero-divisors conjecture. According to Corollary 8.16.12, the zero-divisor conjecture is equivalent to saying that if $G$ is a torsion-free group and $\mathbb{K}$ is a field then every non-identically-zero linear cellular automaton $\tau\colon \mathbb{K}^G \to \mathbb{K}^G$ is pre-injective. This reformulation of the Kaplansky zero-divisor conjecture in terms of linear cellular automata was given in Sect. 6 of [CeC1].

## Exercises

**8.1.** Let $G$ be a group and let $R$ be a nonzero ring. Show that the ring $R[G]$ is commutative if and only if both $G$ and $R$ are commutative.

**8.2.** Recall that the *center* of a ring $A$ is the subring $B$ of $A$ consisting of the elements $x \in A$ which satisfy $ax = xa$ for all $a \in A$. Let $G$ be a group and let $R$ be a commutative ring. Let $\alpha \in R[G]$. Show that $\alpha$ is in the center of $R[G]$ if and only if $\alpha$ is constant on each conjugacy class of $G$.

**8.3.** One says that a group $G$ has the ICC-*property* if every conjugacy class of $G$ except $\{1_G\}$ is infinite.

(a) Show that if a group $G$ has the ICC-property then the center of $G$ is reduced to the identity element.

(b) Show that if a group $G$ has the ICC-property then every normal subgroup of $G$ which is not reduced to the identity element is infinite.

(c) Let $X$ be an infinite set. Show that the group $\mathrm{Sym}_0(X)$, which consists of all permutations of $X$ with finite support, has the ICC-property.

(d) Show that every nonabelian free group has the ICC-property.

(e) Let $G$ be a group with the ICC-property and let $R$ be a commutative ring. Show that the center of $R[G]$ consists of the maps $x\colon G \to R$ which satisfy $x(g) = 0_R$ for all $g \in G \setminus \{1_G\}$. Hint: Use the result of Exercise 8.2.

(f) Let $G$ be a group and let $R$ be a nonzero commutative ring. Show that $G$ has the ICC-property if and only if the center of $R[G]$ is isomorphic to the ring $R$.

**8.4.** Let $G$ be a group, $\mathbb{K}$ a field, and $d \geq 1$ an integer. Show that the $\mathbb{K}$-algebras $\mathrm{Mat}_d(\mathbb{K})[G]$ and $\mathrm{Mat}_d(\mathbb{K}[G])$ are isomorphic.

**8.5.** Show that Theorem 8.13.2 remains valid if the field $\mathbb{R}$ is replaced by the field $\mathbb{C}$ of complex numbers. Hint: Write every $f \in \mathbb{C}^G$ in the form $f = f_0 + if_1$, where $f_0, f_1 \in \mathbb{R}^G$.

**8.6.** Let $G$ be a group and let $S \subset G$ be a non-empty finite subset. Denote by $\langle \cdot, \cdot \rangle$ the standard scalar product on $\mathbb{R}[G] \subset \ell^2(G)$ and let $\| \cdot \|$ denote the associated norm.

(a) Show that for all $x \in \mathbb{R}[G]$ and $\lambda \in \mathbb{R}$, one has

$$\langle x, \Delta_S(x) + \lambda x \rangle = \frac{1}{2} \sum_{g \in G} \sum_{s \in S} |x(g) - x(gs)|^2 + \lambda \|x\|^2.$$

(b) Show that if the subgroup generated by $S$ is infinite then for any non–empty finite subset $F \subset G$ there exists $s \in S$ such that $Fs \not\subset F$.

(c) Deduce from (a) and (b) that if $\lambda \geq 0$ and the subgroup generated by $S$ is infinite, then the linear cellular automaton $\Delta_S + \lambda \, \mathrm{Id}_{\mathbb{R}^G} \colon \mathbb{R}^G \to \mathbb{R}^G$ is pre-injective.

**8.7.** Show that if $(R_i)_{i \in I}$ is a family of directly finite rings then the product ring $P = \prod_{i \in I} R_i$ is directly finite.

**8.8.** Show that every subring of a directly finite (resp. stably finite) ring is directly finite (resp. stably finite).

**8.9.** Let $\mathbb{K}$ be a field and let $V$ be an infinite-dimensional vector space over $\mathbb{K}$. Show that the $\mathbb{K}$-algebra $\mathrm{End}_{\mathbb{K}}(V)$ is not directly finite.

**8.10.** Let $R$ be a ring and let $M$ be a left $R$-module. One says that the module $M$ is *Hopfian* if every surjective endomorphism of $M$ is injective. Show that if $M$ is Hopfian, then the ring $\mathrm{End}_R(M)$ is directly finite.

**8.11.** Let $R$ be a ring and let $M$ be a left $R$-module. One says that the module $M$ is *Noetherian* if its submodules satisfy the *ascending chain condition*, i.e., every increasing sequence

$$N_1 \subset N_2 \subset \dots$$

of submodules of $M$ stabilizes (there is an integer $i_0 \geq 1$ such that $N_i = N_{i_0}$ for all $i \geq i_0$). Show that if $M$ is Noetherian, then $M$ is Hopfian. Hint: Suppose that $f$ is a surjective endomorphism of $M$ and consider the sequence of submodules $N_i = \mathrm{Ker}(f^i)$, $i \geq 1$.

**8.12.** Let $R$ be a ring and let $P$ be a left $R$-module. One says that the module $P$ is *projective* if for every homomorphism $f \colon P \to M$ and every surjective homomorphism $g \colon \widetilde{M} \to M$ of left $R$-modules, there exists a homomorphism $h \colon P \to \widetilde{M}$ such that $f = g \circ h$. Show that if the module $P$ is projective, then $P$ is Hopfian if and only if the ring $\mathrm{End}_R(P)$ is directly finite.

**8.13.** Let $R$ be a ring and let $d \geq 1$ be an integer. Equip $R^d$ with its natural structure of left $R$-module. Show that $R^d$ is Hopfian if and only if the ring $\mathrm{Mat}_d(R)$ is directly finite.

**8.14.** One says that a ring $R$ is *left Noetherian* if $R$ is Noetherian as a left module over itself.

(a) Let $R$ be a left Noetherian ring. Show by induction that $R^d$ is Noetherian as a left $R$-module for each integer $d \geq 1$.

(b) Show that every left Noetherian ring is stably finite. Hint: Use Exercises 8.11 and 8.13.

**8.15.** One says that a ring $R$ has the *unique rank property* if it satisfies the following condition: if $m$ and $n$ are positive integers such that $R^m$ and $R^n$ are isomorphic as left $R$-modules, then one has $m = n$. Show that every stably finite ring has the unique rank property.

**8.16.** Let $\mathbb{K}$ be a field and let $V$ be an infinite-dimensional vector space over $\mathbb{K}$. Show that the ring $R = \mathrm{End}_{\mathbb{K}}(V)$ does not have the unique rank property. Hint: Observe that the vector spaces $V$ and $V \oplus V$ are isomorphic and then prove that the set consisting of all $\mathbb{K}$-linear maps $f \colon V \oplus V \to V$, with its natural structure of left $R$-module, is isomorphic to both $R$ and $R^2$.

**8.17.** Show that every division ring is stably finite.

**8.18.** A ring $R$ is said to be *unit-regular* if for any $a \in R$ there exists an invertible element $u \in R$ such that $a = aua$.

(a) Show that every division ring is unit-regular.

(b) Show that if $\mathbb{K}$ is a field then the ring $\mathrm{Mat}_d(\mathbb{K})$ is unit-regular for every $d \geq 1$. Hint: Prove that if $A \in \mathrm{Mat}_d(\mathbb{K})$ has rank $r$ then there exist invertible matrices $U, V \in \mathrm{GL}_d(\mathbb{K})$ such that $A = UD_rV$, where $D_r \in \mathrm{Mat}_d(\mathbb{K})$ is the diagonal matrix defined by $\delta_{ij} = 1$ if $1 \leq i = j \leq r$ and $\delta_{ij} = 0$ otherwise, and then observe that $A = AXA$, where $X = (UV)^{-1}$.

(c) Show that every unit-regular ring is directly finite.

(d) Prove that the ring $\mathbb{Z}$ is not unit-regular.

(e) A ring $R$ is called a *Boolean ring* if $a^2 = a$ for all $a \in R$. Show that every Boolean ring is commutative and unit-regular.

**8.19.** One says that a ring $R$ is a *right Ore ring* if $R$ has no zero-divisors and if for any pair $a, b$ of nonzero elements in $R$ there exist nonzero elements $u, v \in R$ such that $au = bv$. Left Ore rings are defined similarly. Show that any right (or left) Ore ring is directly finite.

Hint: Prove that $ab = 1_R$ implies $ba = 1_R$ by a direct argument, or show that $R$ can be embedded as a subring of a division ring by adapting the construction of the field of fractions of an integral domain.

**8.20.** Let $\mathbb{K}$ be a field. Show that every finite-dimensional $\mathbb{K}$-algebra is stably finite.

**8.21.** Let $G$ be an amenable group and let $\mathcal{F}$ be a Følner net for $G$. Let $V$ be a finite-dimensional vector space. Suppose that $X$ (resp. $Y$) is a vector subspace of $V^G$. Show that $\mathrm{mdim}_{\mathcal{F}}(X \cup Y) \leq \mathrm{mdim}_{\mathcal{F}}(X) + \mathrm{mdim}_{\mathcal{F}}(Y)$.

**8.22.** Use Corollary 8.15.7 and the result of Exercise 8.20 to recover the fact that every finite group is L-surjunctive.

**8.23.** Use Proposition 8.15.2 and Corollary 8.15.7 to recover the fact that every abelian group is L-surjunctive.

**8.24.** It follows from Theorem 8.14.4 that every residually finite group is L-surjunctive. The goal of this exercise is to present an alternative proof of this result. Let $G$ be a residually finite group and let $V$ be a finite-dimensional vector space over a field $\mathbb{K}$. Let $\tau\colon V^G \to V^G$ be an injective linear cellular automaton. Fix a family $(\Gamma_i)_{i\in I}$ of subgroups of finite index of $G$ such that $\bigcap_{i\in I}\Gamma_i = \{1_G\}$ (the existence of such a family follows from the residual finiteness of $G$).

(a) Show that, for each $i \in I$, the set $\mathrm{Fix}(\Gamma_i) = \{x \in V^G : gx = x \text{ for all } g \in \Gamma_i\}$ is a finite-dimensional vector subspace of $V^G$.

(b) Show that $\tau(\mathrm{Fix}(\Gamma_i)) = \mathrm{Fix}(\Gamma_i)$ for all $i \in I$.

(c) Prove that $\bigcup_{i\in I}\mathrm{Fix}(\Gamma_i)$ is dense in $V^G$ and conclude.

**8.25.** Let $R$ be a ring and let $x \in R$. Show that $x$ is an idempotent if and only if $1_R - x$ is an idempotent.

**8.26.** Show that any subgroup of a unique-product group is a unique-product group.

**8.27.** Let $G$ be a group. Suppose that $G$ contains a normal subgroup $N$ such that $N$ and $G/N$ are both unique-product groups. Show that $G$ is a unique-product group. Hint: See for example [RuS, Theorem 6.1].

**8.28.** Show that every residually orderable group is orderable.

**8.29.** Show that the limit of a projective system of orderable groups is orderable.

**8.30.** A group $G$ is called *bi-orderable* if it admits a total ordering $\leq$ which is both left and right invariant, i.e., such that $g_1 \leq g_2$ implies $gg_1 \leq gg_2$ and $g_1 g \leq g_2 g$ for all $g, g_1, g_2 \in G$. Let $G$ be a bi-orderable group. Suppose that an element $g \in G$ satisfies the following property: there exist an integer $n \geq 1$ and elements $h_1, h_2, \cdots, h_n \in G$ such that

$$h_1 g h_1^{-1} h_2 g h_2^{-1} \cdots h_n g h_n^{-1} = 1_G.$$

Show that $g = 1_G$.

**8.31.** Let $G$ be a bi-orderable group. Show that if $g, h \in G$ are such that $g^n = h^n$ for some integer $n \geq 1$, then $g = h$.

**8.32.** The *Klein bottle group* is the group $K$ given by the presentation $K = \langle x, y; xyx^{-1}y\rangle$. Thus, $K$ is the quotient group $K = F/N$, where $F$ is the free

group based on $x$ and $y$, and $N$ is the normal closure of $xyx^{-1}y$ in $F$. Let $\rho\colon F \to K$ denote the quotient homomorphism.

(a) Let $H$ denote the subgroup of $K$ generated by $\rho(y)$. Show that $H$ is normal in $K$ and that $H$ and $K/H$ are both infinite cyclic.

(b) Show that $K$ is an orderable group.

(c) Show that the group $K$ is not bi-orderable.

**8.33.** Let $G$ be an amenable group, $\mathcal{F}$ a right Følner net for $G$, and $V$ a finite-dimensional vector space over some field $\mathbb{K}$. Let $X$ be a vector subspace of $V^G$ and let $\overline{X}$ denote the closure of $X$ in $V^G$ for the prodiscrete topology. Show that $\overline{X}$ is a vector subspace of $V^G$ and that one has $\mathrm{mdim}_{\mathcal{F}}(\overline{X}) = \mathrm{mdim}_{\mathcal{F}}(X)$.

# Appendix A
# Nets and the Tychonoff Product Theorem

## A.1 Directed Sets

Recall that a *partially ordered* set is a set $I$ equipped with a binary relation $\leq$ which is both reflexive ($i \leq i$ for all $i \in I$) and transitive ($i \leq j$ and $j \leq k$ implies $i \leq k$ for all $i, j, k \in I$).

A *directed set* is a partially ordered set $I$ which satisfies the following condition: for all $i, j \in I$, there exists an element $k \in I$ such that $i \leq k$ and $j \leq k$.

*Examples A.1.1.* (a) The set $\mathbb{Z}$ equipped with the relation $\leq$ defined by

$$i \leq j \iff i \text{ divides } j$$

is a directed set.

(b) If $E$ is an arbitrary set, then the set $\mathcal{P}(E)$ of all subsets of $E$ is a directed set for inclusion.

(c)) If $X$ is a topological space and $x$ is a point of $X$, then the set of neighborhoods of $x$, equipped with the relation $\leq$ defined by

$$V \leq W \iff W \subset V,$$

is a directed set. Indeed, if $V$ and $W$ are neighborhoods of $x$, then $V \cap W$ is a neighborhood of $x$ satisfying $V \leq V \cap W$ and $W \leq V \cap W$.

## A.2 Nets in Topological Spaces

Let $X$ be a set. A *net* of points of $X$ (or *net* in $X$) is a family $(x_i)_{i \in I}$ of points of $X$ indexed by some directed set $I$.

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_9, 

Let $(x_i)_{i\in I}$ and $(y_j)_{j\in J}$ be nets in a set $X$ indexed by directed sets $I$ and $J$ respectively. One says that the net $(y_j)$ is a *subnet* of the net $(x_i)$ if there is a map $\varphi\colon J \to I$ which satisfies the following conditions:

(SN-1) $y_j = x_{\varphi(j)}$ for all $j \in J$;
(SN-2) for each $i \in I$, there exists $j \in J$ such that if $k \in J$ and $j \leq k$ then $i \leq \varphi(k)$.

Let $X$ be a topological space. Let $(x_i)_{i\in I}$ be a net in $X$ and let $a \in$. One says that the net $(x_i)$ *converges* to $a$, or that $a$ is a *limit point* of the net $(x_i)$, if the following condition is satisfied: for each neighborhood $V$ of $a$ in $X$, there exists an element $i_0 \in I$ such that $x_i \in V$ for all $i \geq i_0$. One says that the net $(x_i)$ is *convergent* if there exists a point $a$ in $X$ such that $(x_i)$ converges to $a$.

It is clear that if the net $(x_i)$ converges to $a$, then every subnet of $(x_i)$ also converges to $a$.

**Proposition A.2.1.** *Let $X$ be a topological space, $Y \subset X$ and $a \in X$. Let $\overline{Y}$ denote the closure of $Y$ in $X$. Then the following conditions are equivalent:*

(a) $a \in \overline{Y}$;
(b) *there exists a net $(y_i)_{i\in I}$ of points of $Y$ which converges to $a$ in $X$.*

*Proof.* Suppose that $(y_i)_{i\in I}$ is a net of points of $Y$ which converges to $a$. Then, for each neighborhood $V$ of $a$, there is an element $i_0 \in I$ such that $y_i \in V$ for all $i \geq i_0$. Thus, every neighborhood of $a$ meets $Y$. This shows that $a \in \overline{Y}$.

Conversely, suppose that $a \in \overline{Y}$. Let $I$ denote the directed set consisting of all neighborhoods of $a$ in $X$ partially ordered by reverse inclusion, that is,

$$V \leq W \iff W \subset V.$$

Since $a$ is in the closure of $Y$, we can find for each neighborhood $V \in I$ a point $y_V$ in $Y$ such that $y_V \in V$. It is clear that the net $(y_V)_{V\in I}$ converges to $a$. □

**Proposition A.2.2.** *A topological space $X$ is Hausdorff if and only if every convergent net in $X$ admits a unique limit.*

*Proof.* Suppose that $X$ is Hausdorff. Consider a net $(x_i)_{i\in I}$ in $X$ which converges to some point $a \in X$. Let $b$ be a point in $X$ with $a \neq b$. Since $X$ is Hausdorff, there exist a neighborhood $V$ of $a$ and a neighborhood $W$ of $b$ such that $V \cap W$. For $i$ large enough, the point $x_i$ is in $V$ and therefore not in $W$. Therefore the net $(x_i)_{i\in I}$ does not converge to $b$. This shows that every convergent net in $X$ has a unique limit.

Suppose now that $X$ is not Hausdorff. Then there exist distinct points $a$ and $b$ in $X$ such that each neighborhood of $a$ meets each neighborhood of $b$. Consider the set $I$ consisting of all pairs $(V, W)$, where $V$ is a neighborhood of

$a$ and $W$ is a neighborhood of $b$. partially ordered by declaring that $(V', W') \leq (V, W)$ if and only if $V \subset V'$ and $W \subset W'$. Clearly $I$ is a directed set. If we choose, for each $i = (V, W) \in I$ a point $x_i \in V \cap W$, then the net $(x_i)_{i \in I}$ admits both points $a$ and $b$ as limits. This proves the converse implication. □

Let $(x_i)_{i \in I}$ be a net in a topological space $X$. One says that a point $a \in X$ is a *cluster point* of the net $(x_i)$ if it satisfies the following condition: for each neighborhood $V$ of $a$ in $X$ and each $i \in I$, there exists an element $j \in I$ such that $i \leq j$ and $x_j \in V$.

**Proposition A.2.3.** *Let $X$ be a topological space, $(x_i)_{i \in I}$ a net in $X$, and $a \in X$. Then the following conditions are equivalent:*

(a) *the point $a$ is a cluster point of the net $(x_i)$;*
(b) *the net $(x_i)$ admits a subnet converging to $a$.*

*Proof.* Suppose that $(y_j)_{j \in J}$ is a subnet of the net $(x_i)_{i \in I}$ converging to $a$. Let $\varphi \colon J \to I$ be a map satisfying conditions (SN-1) and (SN-2) above. Consider a neighborhood $V$ of $a$ and an element $i_0 \in I$. By (SN-2), we may find $j_0 \in J$ such that $i_0 \leq \varphi(k)$ for all $k \in J$ satisfying $j_0 \leq k$. Since the net $(y_j)$ converges to $a$, there exists $k_0 \in J$ such that $j_0 \leq k_0$ and $y_{k_0} \in V$. Then we have $i_0 \leq \varphi(k_0)$ and $x_{\varphi(k_0)} = y_{k_0} \in V$. This shows that $a$ is a cluster point for the net $(x_i)$. Thus (b) implies (a).

Conversely, suppose that $a$ is a cluster point for the net $(x_i)$. Denote by $N_a$ the set of all neighborhoods of $a$, partially ordered by reverse inclusion. Let $J$ be the subset of the Cartesian product $I \times N_a$ consisting of all pairs $(i, V) \in I \times N_a$ such that $x_i \in V$. The fact that $a$ is a cluster point of the net $(x_i)$ implies that $J$ is a directed set for the partial ordering $\leq$ defined by

$$(i, V) \leq (i', V') \overset{def}{\iff} (i \leq i' \text{ and } V \leq V').$$

Consider the non decreasing map $\varphi \colon J \to I$ given by $\varphi((i, V)) = i$. If we set $y_j = x_{\varphi(j)}$ for all $j \in J$, it is clear that $\varphi$ satisfies conditions (SN-1) and (SN-2) above and that the net $(y_j)_{j \in J}$ converges to $a$. This shows that (a) implies (b). □

Note that if $f \colon X \to Y$ is a continuous map between topological spaces and $a \in X$ is a limit (resp. cluster) point of the net $(x_i)$, then $f(a)$ is a limit (resp. cluster) point of the net $(f(x_i)$. This immediately follows from the definition and the fact that $f^{-1}(W)$ is a neighborhood of $a$ for each neighborhood $W$ of $f(a)$.

## A.3 Initial Topology

Let $X$ be a set and let $(Y_\lambda)_{\lambda\in\Lambda}$ be a family of topological spaces indexed by an arbitrary set $\Lambda$. Suppose that we are given, for each $\lambda \in \Lambda$, a map $f_\lambda\colon X \to Y_\lambda$. Then one constructs a topology on $X$ in the following way.

Let $\mathcal{F}$ denote the set of all subsets of $X$ of the form $f_\lambda^{-1}(U_\lambda)$, where $\lambda \in \Lambda$ and $U_\lambda$ is an open subset of $Y_\lambda$. Let $\mathcal{B}$ be the set of all subsets of $X$ which may be written as a finite intersection of elements of $\mathcal{F}$. Finally, let $\mathcal{T}$ denote the set of all subsets of $X$ which may be written as a (finite or infinite) union of elements of $\mathcal{B}$. It is straightforward to verify that the set $\mathcal{T}$ is the set of open sets of a topology on $X$ which admits $\mathcal{B}$ as a base, and that this topology is the smallest topology on $X$ for which all maps $f_\lambda\colon X \to Y_\lambda$ are continuous. The topology on $X$ whose open sets are the elements of $\mathcal{T}$ is called the *initial topology* on $X$ associated with the topological spaces $(Y_\lambda)_{\lambda\in\Lambda}$ and the maps $(f_\lambda)_{\lambda\in\Lambda}$.

If $Z$ is a topological space and $g\colon Z \to X$ is a map, then $g$ is continuous with respect to the initial topology on $X$ if and only if all composite maps $f_\lambda \circ g\colon Z \to Y_\lambda$, $\lambda \in \Lambda$, are continuous. If $(x_i)_{i\in I}$ is a net in $X$ and $a$ is a point in $X$, then the net $(x_i)_{i\in I}$ converges to $a$ if and only if the net $(f_\lambda(x_i))_{i\in I}$ converges to $f_\lambda(a)$ for every $\lambda \in \Lambda$. Similarly, $a$ is a cluster point of the net $(x_i)_{i\in I}$ if and only if $f_\lambda(a)$ is a cluster point of the net $(f_\lambda(x_i))_{i\in I}$ for every $\lambda \in \Lambda$.

## A.4 Product Topology

Let $(X_\lambda)_{\lambda\in\Lambda}$ be a family of topological spaces indexed by a set $\Lambda$. The initial topology on the cartesian product $X = \prod_{\lambda\in\Lambda} X_\lambda$ associated with the projection maps $\pi_\lambda\colon X \to X_\lambda$ is called the *product topology* on $X$. A base for the product topology on $X$ consists of all subsets of the form $V = \prod_{\lambda\in\Lambda} U_\lambda$, where $U_\lambda$ is an open subset of $X_\lambda$ for each $\lambda \in \Lambda$ and $U_\lambda = X_\lambda$ for all but finitely many $\lambda \in \Lambda$.

In the case when each $X_\lambda$ is endowed with the discrete topology, the product topology on $X = \prod_{\lambda\in\Lambda} X_\lambda$ is called the *prodiscrete topology*.

**Proposition A.4.1.** *Let $(X_\lambda)_{\lambda\in\Lambda}$ be a family of Hausdorff topological spaces. Then $X = \prod_{\lambda\in\Lambda} X_\lambda$ is Hausdorff for the product topology.*

*Proof.* Let $x = (x_\lambda)$ and $y = (y_\lambda)$ be distinct points of $X$. Then there exists $\lambda_0 \in \Lambda$ such that $x_{\lambda_0} \neq y_{\lambda_0}$. Since $X_{\lambda_0}$ is Hausdorff, we may find disjoint open subsets $U$ and $V$ of $X_{\lambda_0}$ containing $x_{\lambda_0}$ and $y_{\lambda_0}$ respectively. The pullbacks of $U$ and $V$ by the projection map $\pi_{\lambda_0}\colon X \to X_{\lambda_0}$ are disjoint open subsets of $X$ containing $x$ and $y$ respectively. Therefore $X$ is Hausdorff. □

Recall that a topological space $X$ is called *totally disconnected* if any nonempty connected subset of $X$ is reduced to a single point.

**Proposition A.4.2.** *Let $(X_\lambda)_{\lambda\in\Lambda}$ be a family of totally disconnected topological spaces. Then $X = \prod_{\lambda\in\Lambda} X_\lambda$ is totally disconnected for the product topology.*

*Proof.* Let $C$ be a nonempty connected subset of $X$. Then, for each $\lambda \in \Lambda$, the image of $C$ by the projection map $\pi_\lambda \colon X \to X_\lambda$ is a nonempty connected subset of $X_\lambda$. Since $X_\lambda$ is totally disconnected, the set $\pi_\lambda(C)$ is reduced to a single point for every $\lambda \in \Lambda$. This implies that $C$ is reduced to a single point. Therefore, $X$ is totally disconnected. □

**Proposition A.4.3.** *Let $(X_\lambda)_{\lambda\in\Lambda}$ be a family of topological spaces. Suppose that $F_\lambda$ is a closed subset of $X_\lambda$ for each $\lambda \in \Lambda$. Then $F = \prod_{\lambda\in\Lambda} F_\lambda$ is a closed subset of $X = \prod_{\lambda\in\Lambda} X_\lambda$ for the product topology.*

*Proof.* We have

$$F = \bigcap_{\lambda\in\Lambda} \pi_\lambda^{-1}(F_\lambda).$$

Thus $F$ is closed in $X$ as it is the intersection of a family of closed subsets of $X$. □

## A.5 The Tychonoff Product Theorem

Recall that a topological space $X$ is called *compact* if every open cover of $X$ admits a finite subcover. This means that if $(U_\alpha)_{\alpha\in A}$ is a family of open subsets of $X$ with $X = \bigcup_{\alpha\in A} U_\alpha$, then there exists a finite subset $B \subset A$ such that $X = \bigcup_{\alpha\in B} U_\alpha$. By taking complements, one sees that the compactness of $X$ is equivalent to the fact that every family $(F_\alpha)_{\alpha\in A}$ of closed subsets of $X$ with the *finite intersection property*,, that is, $\bigcap_{\alpha\in B} F_\alpha \neq \varnothing$ for every finite subset $B \subset A$, has a nonempty intersection.

It is well known that a metric space $X$ is compact if and only if every sequence in $X$ admits a convergent subsequence. There is an analogous characterization of compactness for general topological spaces using nets:

**Theorem A.5.1.** *Let $X$ be a topological space. Then the following conditions are equivalent:*

(a) *$X$ is compact;*
(b) *every net in $X$ admits a cluster point;*
(c) *every net in $X$ admits a convergent subnet.*

*Proof.* The equivalence between conditions (b) and (c) follows from Proposition A.2.3. Thus, it suffices to prove that conditions (a) and (b) are equivalent.

Suppose first that $X$ is compact. Let $(x_i)_{i\in I}$ be a net in $X$. For each $i \in I$, define the subset $Y_i \subset X$ by

$$Y_i = \{x_j : j \in I \text{ and } i \leq j\}.$$

Let $J$ be a finite subset of $I$. Since $I$ is a directed set, we may find an element $k \in I$ such that $i \leq k$ for all $i \in J$. This implies $x_k \in Y_i$ for all $i \in J$ and hence $\bigcap_{i\in J} \overline{Y_i} \supset \bigcap_{i\in J} Y_i \neq \varnothing$. It follows that the family of closed sets $(\overline{Y_i})_{i\in I}$ has the finite intersection property. By compactness of $X$, we deduce that $\bigcap_{i\in I} \overline{Y_i} \neq \varnothing$. Let $a$ be a point in $\bigcap_{i\in I} \overline{Y_i}$. Then, any neighborhood of $a$ meets $Y_i$ for all $i \in I$. This means that $a$ is a cluster point of the net $(x_i)$. Therefore, (a) implies (b).

Conversely, suppose that every net in $X$ admits a cluster point. Let $(F_\alpha)_{\alpha\in A}$ be a family of closed subsets of $X$ with the finite intersection property. Consider the directed set $\mathcal{E}$ consisting of all finite subsets of $A$ partially ordered by inclusion. Choose, for each $E \in \mathcal{E}$, an element $x_E \in \bigcap_{\alpha\in E} F_\alpha$. By our hypothesis, the net $(x_E)_{E\in\mathcal{E}}$ admits a cluster point $a \in X$. Let $\alpha_0 \in A$. If $V$ is a neighborhood of $a$, then there exists a finite set $E_0 \subset A$ such that $\{\alpha_0\} \subset E_0$ and $x_{E_0} \in V$. Since $x_{E_0} \in \bigcap_{\alpha\in E_0} F_\alpha \subset F_{\alpha_0}$, it follows that any neighborhood of $a$ meets $F_{\alpha_0}$. Since the set $F_{\alpha_0}$ is closed, we deduce that $a \in F_{\alpha_0}$. As $\alpha_0$ was an arbitrary element in $A$, we conclude that $a \in \bigcap_{\alpha\in A} F_\alpha$. This shows that $\bigcap_{i\in I} F_i \neq \varnothing$. Consequently, the space $X$ is compact. This proves that (b) implies (a). □

**Theorem A.5.2 (Tychonoff theorem).** *Let $(X_\lambda)_{\lambda\in\Lambda}$ be a family of compact topological spaces. Then $X = \prod_{\lambda\in\Lambda} X_\lambda$ is compact for the product topology.*

*Proof.* By Theorem A.5.1, it suffices to prove that every net in $X$ admits a cluster point. Let $(x_i)$ be net in $X$. We shall prove that $(x_i)$ admits a cluster point by applying Zorn's Lemma. Let us first introduce some notation. Given a subset $A \subset \Lambda$, we set $X(A) = \prod_{\lambda\in A} X_\lambda$ and equip $X(A)$ with the product topology. If $A$ and $B$ are subsets of $\Lambda$ such that $A \subset B$, we denote by $\pi_A^B$ the projection map $X(B) \to X(A)$. Consider the set $\mathcal{E}$ consisting of all pairs $(A, a)$, where $A$ is a subset of $\Lambda$ and $a \in X(A)$ is a cluster point of the net $(\pi_A^\Lambda(x_i))_{i\in I}$. We partially order $\mathcal{E}$ by declaring that two elements $(A, a)$ and $(B, b)$ satisfy $(A, a) \leq (B, b)$ if and only if $A \subset B$ and $\pi_A^B(b) = a$. The set $\mathcal{E}$ is not empty since it contains the pair $(A, a)$, where $A = \varnothing$ and $a$ is the unique element of $X(\varnothing)$. On the other hand, $\mathcal{E}$ is inductive. Indeed, suppose that $\mathcal{F}$ is a totally ordered subset of $\mathcal{E}$. Let $B = \bigcup_{(A,a)\in\mathcal{F}} A$ and consider the unique element $b \in X(B)$ such that $\pi_A^B(b) = a$ for all $(A, a) \in \mathcal{F}$. Clearly $(B, b) \in \mathcal{E}$ and $(B, b)$ is an upper bound for $\mathcal{F}$. By applying Zorn's Lemma, we deduce that $\mathcal{E}$ admits a maximal element $(M, m)$. To prove that the net $(x_i)$ admits a cluster point, it suffices to show that $M = \Lambda$. Suppose not and choose an element $\lambda_0 \in \Lambda\backslash M$. Since the space $X_{\lambda_0}$ is compact, the net $(\pi_{\lambda_0}(x_i))$ admits a cluster point $a_0 \in X_{\lambda_0}$. Let us set $M' = M \cup \{\lambda_0\}$ and consider the element $m' \in X(M')$ defined by $\pi_M^{M'}(m') = m$ and $\pi_{\{\lambda_0\}}^{M'}(m') = a_0$. Clearly $m'$ is a cluster point of the net $(\pi_{M'}^\Lambda(x_i))$ and $(M, m) \leq (M', m')$. This contradicts the maximality of $(M, m)$ and completes the proof. □

It follows from the definition of compactness that if a topological space $X$ has only finitely many open subsets, then $X$ is compact. In particular, every finite topological space is compact. Therefore, an immediate consequence of Tychonoff theorem is the following:

**Corollary A.5.3 (Tychonoff theorem).** *Let $(X_\lambda)_{\lambda\in\Lambda}$ be a family of finite topological spaces. Then $X = \prod_{\lambda\in\Lambda} X_\lambda$ is compact for the product topology.* □

## Notes

The original Tychonoff theorem was only stated for product of compact intervals. The proof of the general Tychonoff theorem we have presented here is based on the one given by P. Chernoff in [Che]. Three other proofs may be found in Kelley's book [Kel]: a proof using Alexander's subbase theorem, Bourbaki's proof [Bou] using ultrafilters, and a proof based on universal nets. There is also another proof using non-standard analysis in the book of A. Robinson [RobA].

# Appendix B
# Uniform Structures

## B.1 Uniform Spaces

Let $X$ be a set. We shall use the following notation. We denote by $\Delta_X$ the diagonal in $X \times X$, that is,

$$\Delta_X = \{(x, x) : x \in X\} \subset X \times X.$$

Suppose that $R$ is a subset of $X \times X$ (in other words, $R$ is a binary relation on $X$). For $y \in X$, we define the set $R[y] \subset X$ by

$$R[y] = \{x \in X : (x, y) \in R\}.$$

The *inverse* $\overset{-1}{R} \subset X \times X$ of $R$ is defined by

$$\overset{-1}{R} = \{(x, y) : (y, x) \in R\}.$$

One says that $R$ is *symmetric* if it satisfies $\overset{-1}{R} = R$.

If $R$ and $S$ are subsets of $X \times X$, we define their *composite* $R \circ S \subset X \times X$ by

$$R \circ S = \{(x, y) : \text{ there exists } z \in X \text{ such that } (x, z) \in R \text{ and } (z, y) \in S\}.$$

**Definition B.1.1.** Let $X$ be a set. A *uniform structure* on $X$ is a non–empty set $\mathcal{U}$ of subsets of $X \times X$ satisfying the following conditions:

(UN-1) if $V \in \mathcal{U}$, then $\Delta_X \subset V$;
(UN-2) if $V \in \mathcal{U}$ and $V \subset V' \subset X \times X$, then $V' \in \mathcal{U}$;
(UN-3) if $V \in \mathcal{U}$ and $W \in \mathcal{U}$, then $V \cap W \in \mathcal{U}$;
(UN-4) if $V \in \mathcal{U}$, then $\overset{-1}{V} \in \mathcal{U}$;
(UN-5) if $V \in \mathcal{U}$, then there exists $W \in \mathcal{U}$ such that $W \circ W \subset V$.

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_10, 

A set $X$ equipped with a uniform structure $\mathcal{U}$ is called a *uniform space* and the elements of $\mathcal{U}$ are called the *entourages* of $X$ (see Fig. B.1).

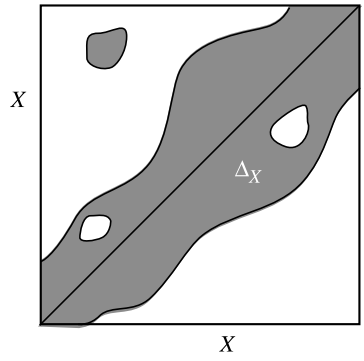



**Fig. B.1** An entourage in a uniform space $X$

*Examples B.1.2.* (a) Let $X$ be a set. Then $\mathcal{U} = \{X \times X\}$ is a uniform structure on $X$. This uniform structure is called the *trivial uniform structure* on $X$. It is the smallest uniform structure on $X$.

(b) The *discrete uniform structure* on a set $X$ is the uniform structure whose entourages consist of all subsets of $X \times X$ containing $\Delta_X$. This is the largest uniform structure on $X$. It follows from (UN-2) that the discrete uniform structure on $X$ is the only uniform structure on $X$ admitting the diagonal $\Delta_X \subset X \times X$ as an entourage.

(c) Suppose that $d$ is a metric on $X$. For each $\varepsilon > 0$ let $V_\varepsilon \subset X \times X$ denote the set of pairs $(x, y)$ such that $d(x, y) < \varepsilon$. Let $\mathcal{U}$ be the set of all subsets $W \subset X \times X$ such that one can find $\varepsilon > 0$ for which $V_\varepsilon \subset W$. Then $\mathcal{U}$ is a uniform structure on $X$ which is called the uniform structure *associated* with the metric $d$.

A uniform structure $\mathcal{U}$ on a set $X$ is said to be *metrizable* if $\mathcal{U}$ is the uniform structure on $X$ associated with some metric on $X$.

*Example B.1.3.* If $d$ is the *discrete* metric on a set $X$, that is, the metric given by $d(x, y) = 0$ if $x = y$ and $d(x, y) = 1$ otherwise, then the uniform structure defined by $d$ is the discrete uniform structure on $X$. Thus the discrete uniform structure on $X$ is metrizable.

Let $X$ be a uniform space. One easily verifies that is possible to define a topology on $X$ by taking as open sets the subsets $\Omega \subset X$ which satisfy the following property: for each $x \in \Omega$, there exists an entourage $V \subset X \times X$ such that $V[x] = \Omega$. One says that this topology is the topology *associated* with the uniform structure on $X$. A subset $N \subset X$ is a neighborhood of a point $x \in X$ for this topology if and only if there exists an entourage $V$ such

that $N = V[x]$. This topology is Hausdorff if and only if the intersection of the entourages of $X$ coincides with the diagonal $\Delta_X \subset X \times X$.

*Examples B.1.4.* (a) The topology associated with the discrete uniform structure on a set $X$ is the discrete topology on $X$ (every subset is open).

(b) If $\mathcal{U}$ is the uniform structure associated with a metric $d$ on a set $X$, then the topology defined by $\mathcal{U}$ coincides with the topology defined by $d$.

Let $\mathcal{U}$ be a uniform structure on a set $X$.

If $Y$ is a subset of $X$, then $\mathcal{U}_Y = \{V \cap (Y \times Y) : V \in \mathcal{U}\}$ is a uniform structure on $Y$, which is said to be *induced* by $\mathcal{U}$. The topology on $Y$ associated with $\mathcal{U}_Y$ is the topology induced by the topology on $X$ associated with $\mathcal{U}$.

A subset $\mathcal{B} \subset \mathcal{U}$ is called a *base* of $\mathcal{U}$ if for each $W \in \mathcal{U}$ there exists $V \in \mathcal{B}$ such that $V \subset W$.

*Example B.1.5.* If $d$ is a metric on $X$, then $\mathcal{B} = \{V_\varepsilon : \varepsilon > 0\}$, where $V_\varepsilon = \{(x, y) \in X \times X : d(x, y) < \varepsilon\}$, is a base for the uniform structure on $X$ associated with $d$.

The proof of the following statement is straightforward.

**Proposition B.1.6.** *Let $X$ be a set and let $\mathcal{B}$ be a nonempty set of subsets of $X \times X$. Then $\mathcal{B}$ is a base for some (necessarily unique) uniform structure on $X$ if and only if it satisfies the following properties:*

(BU-1) *if $V \in \mathcal{B}$, then $\Delta_X \subset V$;*
(BU-2) *if $V \in \mathcal{B}$ and $W \in \mathcal{B}$, then there exists $U \in \mathcal{B}$ such that $U \subset V \cap W$;*
(BU-3) *if $V \in \mathcal{B}$, then there exists $W \in \mathcal{B}$ such that $W \subset V^{-1}$;*
(BU-4) *if $V \in \mathcal{B}$, then there exists $W \in \mathcal{B}$ such that $W \circ W \subset V$.*

□

## B.2 Uniformly Continuous Maps

Let $X$ and $Y$ be uniform spaces. A map $f\colon X \to Y$ is called *uniformly continuous* if it satisfies the following condition: for each entourage $W$ of $Y$, there exists an entourage $V$ of $X$ such that $(f \times f)(V) \subset W$. Here $f \times f$ denotes the map from $X \times X$ into $Y \times Y$ defined by $(f \times f)(x_1, x_2) = (f(x_1), f(x_2))$ for all $(x_1, x_2) \in X \times X$.

If $\mathcal{B}$ (resp. $\mathcal{B}'$) is a base of the uniform structure on $X$ (resp. $Y$), then a map $f\colon X \to Y$ is uniformly continuous if and only if it satisfies the following condition: for each $W \in \mathcal{B}'$, there exists $V \in \mathcal{B}$ such that $(f \times f)(V) \subset W$. Note that this condition is equivalent to the fact that $(f \times f)^{-1}(W)$ is an entourage of $Y$ for each entourage $W$ of $X$.

*Example B.2.1.* Let $(X, d_X)$ and $(Y, d_Y)$ be metric spaces. Then a map $f\colon X \to Y$ is uniformly continuous if and only if it satisfies the following condition: for each $\varepsilon > 0$, there exists $\delta > 0$ such that $d_X(x_1, x_2) < \delta$ implies $d_Y(f(x_1), f(x_2)) < \varepsilon$.

**Proposition B.2.2.** *Let $X$ and $Y$ be uniform spaces. Then every uniformly continuous map $f\colon X \to Y$ is continuous (with respect to the topologies on $X$ and $Y$ associated with the uniform structures).*

*Proof.* Suppose that $f\colon X \to Y$ is uniformly continuous. Let $x \in X$ and let $N \subset Y$ be a neighborhood of $f(x)$. Then there exists an entourage $W$ of $Y$ such that $W[f(x)] = N$. Since $f$ is uniformly continuous, the set $V = (f \times f)^{-1}(W)$ is an entourage of $X$. The set $V[x]$ is a neighborhood of $x$ and satisfies $f(V[x]) \subset W[f(x)] = N$. This shows that $f$ is continuous. □

A continuous map between uniform spaces may fail to be uniformly continuous. For example, the map $x \mapsto x^2$ is not uniformly continuous on $\mathbb{R}$ (equipped with the uniform structure associated with its usual metric). However, this is true when the source space is compact:

**Theorem B.2.3.** *Let $X$ and $Y$ be uniform spaces and suppose that $X$ is compact. Then every continuous map $f\colon X \to Y$ is uniformly continuous.*

Let us first establish the following:

**Lemma B.2.4 (Lebesgue lemma).** *Let $(\Omega_i)_{i\in I}$ be an open cover of a compact uniform space $X$. Then there exists an entourage $\Lambda$ of $X$ satisfying the following property: for each $x \in X$, there exists an index $i \in I$ such that $\Lambda[x] \subset \Omega_i$.*

*Proof.* Let us choose, for each $x \in X$, an index $i(x) \in I$ such that $x \in \Omega_{i(x)}$. Since $\Omega_{i(x)}$ is a neighborhood of $x$, there is an entourage $V_x$ such that $V_x[x] = \Omega_{i(x)}$. By (UN-5), we may find an entourage $W_x$ such that $W_x \circ W_x \subset V_x$. The set $W_x[x]$ is a neighborhood of $x$ for each $x \in X$. By compactness of $X$, there exists a finite subset $A \subset X$ such that $X = \bigcup_{a\in A} W_a[a]$. Let us show that the entourage

$$\Lambda = \bigcap_{a\in A} W_a$$

has the required property. Let $x \in X$. Choose a point $a \in A$ such that $x \in W_a[a]$. Suppose that $y \in \Lambda[x]$. Since $(x, a) \in W_a$ and $(y, x) \in \Lambda \subset W_a$, we have $(y, a) \in W_a \circ W_a \subset V_a$. Thus $y \in V_a[a]$. Since $V_a[a] \subset \Omega_{i(a)}$, this shows that $\Lambda[x] \subset \Omega_{i(a)}$. □

*Proof of Theorem B.2.3.* Let $f\colon X \to Y$ be a continuous map and let $W$ be an entourage of $Y$. By (UN-3), (UN-4), and (UN-5), we may find a symmetric entourage $S$ of $Y$ such that $S \circ S \subset W$. Since $f$ is continuous, there exists, for each $x \in X$, an open neighborhood $\Omega_x$ of $X$ such that $f(\Omega_x) \subset S[f(x)]$.

By Lemma B.2.4, we may find an entourage $\Lambda$ of $X$ such that, for each $y \in X$, there exists $x \in X$ such that $\Lambda[y] \subset \Omega_x$. Suppose that $(x_1, x_2) \in \Lambda$. Choose $a \in X$ such that $\Lambda[x_2] \subset \Omega_a$. Since $x_1$ and $x_2$ are in $\Lambda[x_2]$, we deduce that the points $f(x_1)$ and $f(x_2)$ are in $S[f(a)]$. It follows that $(f(x_1), f(a))$ and $(f(a), f(x_2))$ are in $S$, and hence $(f(x_1), f(x_2)) \in S \circ S \subset W$. Thus $(f \times f)(\Lambda) \subset W$. This shows that $f$ is uniformly continuous. □

Let $X$ and $Y$ be uniform spaces.

One says that a map $f\colon X \to Y$ is a *uniform isomorphism* if $f$ is bijective and both $f$ and $f^{-1}$ are uniformly continuous.

One says that a map $f\colon X \to Y$ is a *uniform embedding* if $f$ is injective and induces a uniform isomorphism between $X$ and $f(X) \subset Y$.

**Proposition B.2.5.** *Let $X$ and $Y$ be uniform spaces with $X$ compact and $Y$ Hausdorff. Suppose that $f\colon X \to Y$ is a continuous injective map. Then $f$ is a uniform embedding.*

*Proof.* As $X$ is compact and $Y$ is Hausdorff, $f$ induces a homeomorphism from $X$ onto $f(X)$. This homeomorphism is a uniform isomorphism by Theorem B.2.3. □

## B.3 Product of Uniform Spaces

Let $X$ be a set. Suppose that we are given a family $(X_\lambda)_{\lambda \in \Lambda}$ of uniform spaces and a family $(f_\lambda)_{\lambda \in \Lambda}$ of maps $f_\lambda\colon X \to X_\lambda$. Then the *initial uniform structure* associated with these data is the smallest uniform structure on $X$ such that all maps $f_\lambda\colon X \to X_\lambda$, $\lambda \in \Lambda$, are uniformly continuous.

In the particular case when $X = \prod_{\lambda \in \Lambda} X_\lambda$ and $f_\lambda\colon X \to X_\lambda$ is the projection map, the associated initial uniform structure on $X$ is called the *product uniform structure*. A base of entourages for the product uniform structure on $X$ is obtained by taking all subsets of $X \times X$ which are of the form

$$\begin{aligned} \prod_{\lambda \in \Lambda} V_\lambda \subset \prod_{\lambda \in \Lambda} X_\lambda \times X_\lambda \\ = \left( \prod_{\lambda \in \Lambda} X_\lambda \right) \times \left( \prod_{\lambda \in \Lambda} X_\lambda \right) \\ = X \times X, \end{aligned}$$

where $V_\lambda \subset X_\lambda \times X_\lambda$ is an entourage of $X_\lambda$ and $V_\lambda = X_\lambda \times X_\lambda$ for all but finitely many $\lambda \in \Lambda$.

When each $X_\lambda$ is endowed with the discrete uniform structure, the product uniform structure on $X = \prod_{\lambda \in \Lambda} X_\lambda$ is called the *prodiscrete uniform structure*.

## B.4 The Hausdorff-Bourbaki Uniform Structure on Subsets

Let $X$ be a uniform space with uniform structure $\mathcal{U}$. In this section, we construct a uniform structure on the set $\mathcal{P}(X)$ of all subsets of $X$.

We shall use the following notation. Suppose that $R$ is a subset of $X \times X$. Given a subset $Y \subset X$, we set

$$R[Y] = \bigcup_{y \in Y} R[y] = \{x \in X : (x, y) \in R \text{ for some } y \in Y\}, \tag{B.1}$$

and we define the subset $\widehat{R} \subset \mathcal{P}(X) \times \mathcal{P}(X)$ by

$$\widehat{R} = \{(Y, Z) \in \mathcal{P}(X) \times \mathcal{P}(X) : Y \subset R[Z] \text{ and } Z \subset R[Y]\}. \tag{B.2}$$

**Proposition B.4.1.** *The set $\{\widehat{V} : V \in \mathcal{U}\}$ is a base for a uniform structure on $\mathcal{P}(X)$.*

*Proof.* Let us check that the conditions of Proposition B.1.6 are satisfied. Property (BU-1) follows from the fact that we have $Y \subset V[Y]$ for all $Y \in \mathcal{P}(X)$ and $V \in \mathcal{U}$ since $\Delta_X \subset V$. Then observe that

$$\begin{aligned}(V \cap W)[Y] &= \{x \in X : (x, y) \in V \cap W \text{ for some } y \in Y\} \\ &\subset \{x \in X : (x, y_V) \in V \text{ and } (x, y_W) \in W \text{ for some } y_V, y_W \in Y\} \\ &= V[Y] \cap W[Y],\end{aligned}$$

for all $V, W \in \mathcal{U}$ and $Y \subset X$. Therefore we have $\widehat{V \cap W} \subset \widehat{V} \cap \widehat{W}$ for all $V, W \in \mathcal{U}$, so that (BU-2) is satisfied. Property (BU-3) follows from the fact that the set $\widehat{V}$ is a symmetric subset of $\mathcal{P}(X) \times \mathcal{P}(X)$ for each $V \in \mathcal{U}$. Finally, let us verify (BU-4). Let $V \in \mathcal{U}$ and take $W \in \mathcal{U}$ such that $W \circ W \subset V$. We claim that $\widehat{W} \circ \widehat{W} \subset \widehat{V}$. To see this, let $(Y, Z) \in \widehat{W} \circ \widehat{W}$. This means that there exists $T \in \mathcal{P}(X)$ such that $(Y, T) \in \widehat{W}$ and $(T, Z) \in \widehat{W}$. In particular we have $Y \subset W[T]$ and $T \subset W[Z]$. Thus, given $y \in Y$, there exist $t \in T$ and $z \in Z$ such that $(y, t) \in W$ and $(t, z) \in W$. We have $(y, z) \in W \circ W \subset V$. This shows that $Y \subset V[Z]$. Similarly, we get $Z \subset V[Y]$ by using $Z \subset W[T]$ and $T \subset W[Y]$. We deduce that $(Y, Z) \in \widehat{V}$. This proves the claim. Consequently, (BU-4) is satisfied. □

The uniform structure on $\mathcal{P}(X)$ admitting $\{\widehat{V} : V \in \mathcal{U}\}$ as a base is called the *Hausdorff-Bourbaki uniform structure* on $\mathcal{P}(X)$ and the topology associated with this uniform structure is called the *Hausdorff-Bourbaki topology* on $\mathcal{P}(X)$.

*Remarks B.4.2.* (a) The empty set is an isolated point in $\mathcal{P}(X)$.

(b) One easily checks that the map $i\colon X \to \mathcal{P}(X)$ defined by $i(x) = \{x\}$ is a uniform embedding.

**Proposition B.4.3.** *Let $X$ be a uniform space. Let $Y$ and $Z$ be closed subsets of $X$. Suppose that there is a net $(T_i)_{i\in I}$ of subsets of $X$ which converges to both $Y$ and $Z$ with respect to the Hausdorff-Bourbaki topology on $\mathcal{P}(X)$. Then one has $Y = Z$.*

*Proof.* Let $y \in Y$ and let $\Omega$ be a neighborhood of $y$ in $X$. Then there is a symmetric entourage $V$ of $X$ such that $V[y] \subset \Omega$. Choose an entourage $W$ of $X$ such that $W \circ W \subset V$. Since the net $(T_i)_{i\in I}$ converges to both $Y$ and $Z$, we can find an element $i_0 \in I$ such that $\subset W[T_{i_0}]$ and $T_{i_0} \subset W[Z]$. Thus, there exist $t \in T_{i_0}$ and $z \in Z$ such that $(y,t) \in W$ and $(t,z) \in W$. This implies $(y,z) \in W \circ W \subset V$. As $V$ is symmetric, it follows that $(z,y) \in V$ and hence $z \in V[y] \subset \Omega$. This shows that $y$ is in the closure of $Z$. Since $Z$ is closed in $X$, we deduce that $Y \subset Z$. By symmetry, we also have $Z \subset Y$. Consequently, $Y = Z$. □

By using Proposition A.2.2, we immediately deduce from Proposition B.4.3 the following:

**Corollary B.4.4.** *Let $X$ be a uniform space. Then the topology induced by the Hausdorff-Bourbaki topology on the set of closed subsets of $X$ is Hausdorff.* □

*Remark B.4.5.* Suppose that $(X,d)$ is a metric space and let $\mathcal{C}_b(X)$ denote the set consisting of all closed bounded subsets of $X$. For $x \in X$ and $r > 0$, denote by $B(x,r)$ the open ball of radius $r$ centered at $x$. Then it is not difficult to verify that the map $\delta\colon \mathcal{C}_b(X) \times \mathcal{C}_b(X) \to \mathbb{R}$ defined by

$$\delta(Y,Z) = \inf\{r > 0 : Z \subset \bigcup_{y\in Y} B(y,r) \text{ and } Y \subset \bigcup_{z\in Z} B(z,r)\}$$

is a metric on $\mathcal{C}_b(X)$ and that the uniform structure associated with $\delta$ is the uniform structure induced by the Hausdorff-Bourbaki structure on the set of subsets of $X$. The metric $\delta$ is called the *Hausdorff metric* on $\mathcal{C}_b(X)$.

**Proposition B.4.6.** *Let $X$ and $Y$ be uniform spaces and let $f\colon X \to Y$ be a uniformly continuous map. Then the map $f_*\colon \mathcal{P}(X) \to \mathcal{P}(Y)$ which sends each subset $A \subset X$ to its image $f(A) \subset Y$ is uniformly continuous with respect to the Hausdorff-Bourbaki uniform structures on $\mathcal{P}(X)$ and $\mathcal{P}(Y)$.*

*Proof.* Let $W$ be an entourage of $Y$ and let

$$\widehat{W} = \{(B_1,B_2) \in \mathcal{P}(Y) \times \mathcal{P}(Y) : B_2 \subset W[B_1] \text{ and } B_1 \subset W[B_2]\}$$

be the associated entourage of $\mathcal{P}(Y)$. Since $f$ is uniformly continuous, there is an entourage $V$ of $X$ such that

$$(f \times f)(V) \subset W.$$

Suppose that $(A_1, A_2) \in \widehat{V}$, that is, $A_2 \subset V[A_1]$ and $A_1 \subset V[A_2]$. If $a_1 \in A_1$, then there exists $a_2 \in A_2$ such that $(a_1, a_2) \in V$ and hence $(f(a_1), f(a_2)) \in W$. Therefore, we have $f_*(A_1) \subset W[f_*(A_2)]$. Similarly, we get $f_*(A_2) \subset W[f_*(A_1)]$. It follows that $(f_*(A_1), f_*(A_2)) \in \widehat{W}$. This shows that $(f_* \times f_*)(\widehat{V}) \subset \widehat{W}$. Consequently, $f_*$ is uniformly continuous. □

## Notes

Uniform structures were introduced by André Weil [Weil]. The reader is referred to [Bou, Ch. 2], [Kel, Ch. 6], and [Jam] for a detailed exposition of the general theory of uniform spaces. The Hausdorff-Bourbaki uniform structure on the set of subsets of a uniform space was introduced in exercises by Bourbaki (see [Bou, ch. II exerc. 5 p. 34 and exerc. 6 p. 36]).

# Appendix C
# Symmetric Groups

## C.1 The Symmetric Group

Let $X$ be a set. A *permutation* of $X$ is a bijective map $\sigma\colon X \to X$. Let $\mathrm{Sym}(X)$ denote the set of all permutations of $X$. We equip $\mathrm{Sym}(X)$ with a group structure by defining the product $\sigma_1\sigma_2$ of two elements $\sigma_1, \sigma_2 \in \mathrm{Sym}(X)$ as the composite map $\sigma_1 \circ \sigma_2$. The associative property follows from the associativity of the composition of maps. The identity map $\mathrm{Id}_X\colon X \to X$ is the identity element and the inverse of $\sigma \in \mathrm{Sym}(X)$ is the inverse map $\sigma^{-1}$. The group $\mathrm{Sym}(X)$ is called the *symmetric group* on $X$.

*Remark C.1.1.* Suppose that $f\colon X \to Y$ is a bijection from a set $X$ onto a set $Y$. Then the map $f_*\colon \mathrm{Sym}(X) \to \mathrm{Sym}(Y)$ defined by $f_*(\sigma) = f \circ \sigma \circ f^{-1}$ is a group isomorphism. As a consequence, symmetric groups on equipotent sets are isomorphic.

**Theorem C.1.2 (Cayley's theorem).** *Every group $G$ is isomorphic to a subgroup of* $\mathrm{Sym}(G)$.

*Proof.* Let $G$ be a group. Given $g \in G$ denote by $L_g\colon G \to G$ the left multiplication by $g$, that is, the map defined by $L_g(h) = gh$ for all $h \in G$. Observe that $L_g \in \mathrm{Sym}(G)$. Indeed, $L_g$ is bijective since, given $h, h' \in G$, we have $L_g(h) = h'$ if and only if $h = g^{-1}h'$. Let us show that the map

$$\begin{aligned} L\colon G &\to \mathrm{Sym}(G) \\ g &\mapsto L_g \end{aligned}$$

is a group homomorphism. Given $g, g', h \in G$ we have $L_{gg'}(h) = gg'h = L_g(g'h) = L_g(L_{g'}(h))$ which shows that $L_{gg'} = L_gL_{g'}$. Moreover, $L$ is injective. Indeed, if $g \in \ker(L)$, that is, $L_g = \mathrm{Id}_G$, we have $g = g \cdot 1_G = L_g(1_G) = 1_G$. This shows that $\ker(L) = \{1_G\}$, and therefore $L$ is injective. It follows that $G$ is isomorphic to $L(G) \subset \mathrm{Sym}(G)$. □

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_11, 

## C.2 Permutations with Finite Support

Let $X$ be a set. The *support* of a permutation $\sigma \in \mathrm{Sym}(X)$ is the set $S(\sigma) \subset X$ consisting of all $x \in X$ such that $\sigma(x) \neq x$.

**Proposition C.2.1.** *Let $\sigma, \tau \in \mathrm{Sym}(X)$. Then*

(i) $\sigma(S(\sigma)) = S(\sigma)$;
(ii) $S(\sigma) = S(\sigma^{-1})$;
(iii) $S(\sigma\tau) \subset S(\sigma) \cup S(\tau)$;
(iv) *if* $S(\sigma) \cap S(\tau) = \varnothing$, *then* $\sigma\tau = \tau\sigma$;
(v) $S(\tau\sigma\tau^{-1}) = \tau(S(\sigma))$.

*Proof.* (i) Let $x \in X$. Then, $x \in S(\sigma)$, that is, $\sigma(x) \neq x$, if and only if $\sigma(\sigma(x)) \neq \sigma(x)$, that is, $\sigma(x) \in S(\sigma)$.

(ii) This follows from the fact that $\sigma(x) \neq x$ if and only if $x \neq \sigma^{-1}(x)$.

(iii) Suppose that $x \in X \setminus (S(\sigma) \cup S(\tau))$. Then $\sigma(x) = x = \tau(x)$ and therefore $(\sigma\tau)(x) = \sigma(\tau(x)) = \sigma(x) = x$. It follows that $x \notin S(\sigma\tau)$.

(iv) Let $x \in X$. Suppose first that $x \in X \setminus (S(\sigma) \cup S(\tau))$. Then, by (iii), $x \notin S(\sigma\tau)$ and $x \notin S(\tau\sigma)$, so that $\sigma(\tau(x)) = x = \tau(\sigma(x))$. Suppose now that $x$ is in the support of one of the two permutations, say $x \in S(\sigma)$. It then follows from our assumptions that $x \notin S(\tau)$ and therefore $\tau(x) = x$. Also, by (i), $\sigma(x) \in S(\sigma)$ and therefore, again by our assumptions, $\sigma(x) \notin S(\tau)$, so that $\tau(\sigma(x)) = \sigma(x)$. We thus have $\tau(\sigma(x)) = \sigma(x) = \sigma(\tau(x))$. It follows that $\sigma\tau = \tau\sigma$.

(v) Let $x \in X$. Then, $x \in S(\sigma)$, that is $\sigma(x) \neq x$, if and only if, $(\tau\sigma\tau^{-1})(\tau(x)) = \tau(\sigma(x))$ is not equal to $\tau(x)$. Thus, $x \in S(\sigma)$ if and only if $\tau(x) \in S(\tau\sigma\tau^{-1})$. □

Let $\mathrm{Sym}_0(X)$ denote the subset of $\mathrm{Sym}(X)$ consisting of all permutations of $X$ with finite support.

**Proposition C.2.2.** *Let $X$ be a set. Then the set $\mathrm{Sym}_0(X)$ is a normal subgroup of $\mathrm{Sym}(X)$.*

*Proof.* The support of the identity map $\mathrm{Id}_X$ is the empty set and therefore $\mathrm{Id}_X \in \mathrm{Sym}_0(X)$. By Proposition C.2.1(ii), if $\sigma \in \mathrm{Sym}_0(X)$ then $\sigma^{-1} \in \mathrm{Sym}_0(X)$. On the other hand, by Proposition C.2.1(iii), the set $\mathrm{Sym}_0(X)$ is closed under multiplication. Thus $\mathrm{Sym}_0(X)$ is a subgroup of $\mathrm{Sym}(X)$. Finally, from Proposition C.2.1(v) we deduce that if $\sigma \in \mathrm{Sym}_0(X)$ then $\tau\sigma\tau^{-1} \in \mathrm{Sym}_0(X)$ for all $\tau \in \mathrm{Sym}(X)$. It follows that $\mathrm{Sym}_0(X)$ is a normal subgroup of $\mathrm{Sym}(X)$. □

Let $r \geq 2$ be an integer and let $x_1, x_2, \ldots, x_r$ be distinct elements in $X$. We denote by

$$\gamma = (x_1\ x_2\ \cdots\ x_r)$$

the permutation of $X$ that maps $x_1$ to $x_2, x_2$ to $x_3, \ldots, x_{r-1}$ to $x_r$, $x_r$ to $x_1$, and maps each element of $X \setminus \{x_1, x_2, \ldots, x_r\}$ to itself. Thus, the support of $\gamma$ is the set $\{x_1, x_2, \ldots, x_r\}$. One says that $\gamma$ is a *cycle* of *length* $r$, or an *$r$-cycle*. A 2-cycle is called a *transposition*. Observe that

$$\gamma = (x\ \gamma(x)\ \gamma^2(x)\ \cdots\ \gamma^{r-1}(x))$$

for all $x \in \{x_1, x_2, \ldots, x_r\}$ and that the inverse of $\gamma$ is the $r$-cycle

$$\gamma^{-1} = (x_r\ x_{r-1}\ \cdots\ x_2\ x_1).$$

**Proposition C.2.3.** *Let $X$ be a set and let $\sigma \in \mathrm{Sym}_0(X)$. Then there exists an integer $n \geq 0$ and cycles $\gamma_1, \gamma_2, \ldots, \gamma_n$ with pairwise disjoint supports such that*

$$\sigma = \gamma_1 \gamma_2 \cdots \gamma_n. \tag{C.1}$$

*Moreover, such a factorization is unique up to a permutation of the factors.*

*Proof.* If $\sigma = \mathrm{Id}_X$ then $n = 0$ and there is nothing to prove.

Suppose now that $\sigma \neq \mathrm{Id}_X$. Let $S = S(\sigma) \subset X$ be the support of $\sigma$. We introduce an equivalence relation on $S$ by setting $x \sim y$ if and only if there exists $k \in \mathbb{Z}$ such that $y = \sigma^k(x)$. Let $S = X_1 \coprod X_2 \coprod \cdots \coprod X_n$ be the partition of $S$ into the equivalence classes of $\sim$ and let us set $r_i = |X_i|$ for $1 \leq i \leq n$. Note that $r_i \geq 2$ for all $i$. For each $i = 1, 2, \ldots, n$, choose a representative $x_i \in X_i$. Observe that the elements $x_i, \sigma(x_i), \sigma^2(x_i), \ldots, \sigma^{r_i - 1}(x_i)$ are all distinct since otherwise the class of $x_i$ would have less that $r_i$ elements. Consider the cycle

$$\gamma_i = (x_i\ \sigma(x_i)\ \sigma^2(x_i)\ \cdots\ \sigma^{r_i-1}(x_i)).$$

The support of $\gamma_i$ is $X_i$. Thus, the cycles $\gamma_1, \gamma_2, \ldots, \gamma_n$ have pairwise disjoint supports. Clearly $\sigma = \gamma_1 \gamma_2 \cdots \gamma_n$.

Suppose now that $\sigma = \delta_1 \delta_2 \cdots \delta_s$, where $\delta_1, \delta_2, \ldots, \delta_s$ are cycles with pairwise disjoint supports. The supports of the cycles $\delta_i$ are the equivalence classes of $\sim$. We deduce that $s = n$. Moreover, up to a permutation of the factors, we may suppose that the support of $\gamma_i$ equals the support of $\delta_i$ for all $i = 1, 2, \ldots, n$. We have

$$\begin{aligned}
\gamma_i &= (x_i\ \gamma_i(x_i)\ \gamma_i^2(x_i)\ \cdots\ \gamma_i^{r_i-1}(x_i)) \\
&= (x_i\ \sigma(x_i)\ \sigma^2(x_i)\ \cdots\ \sigma^{r_i-1}(x_i)) \\
&= (x_i\ \delta_i(x_i)\ \delta_i^2(x_i)\ \cdots\ \delta_i^{r_i-1}(x_i)) \\
&= \delta_i
\end{aligned}$$

and this completes the proof. □

**Corollary C.2.4.** *Every permutation in $\mathrm{Sym}_0(X)$ can be expressed as a product of transpositions.*

*Proof.* First observe that every cycle is a product of transpositions. Indeed, for all distinct $x_1, x_2, \ldots, x_r \in X$, we have

$$(x_1\ x_2\ \cdots\ x_r) = (x_1\ x_r)(x_1\ x_{r-1})\cdots(x_1\ x_3)(x_1\ x_2). \tag{C.2}$$

By applying Proposition C.2.3 we deduce that every $\sigma \in \mathrm{Sym}_0(X)$ is a product of transpositions. □

## C.3 Conjugacy Classes in $\mathrm{Sym}_0(X)$

Let $G$ be a group and let $H \subset G$ be a subgroup. We recall that two elements $h$ and $h'$ in $H$ are said to be *conjugate* in $G$ (resp. in $H$) if there exists an element $g \in G$ (resp. $g \in H$) such that $h' = ghg^{-1}$. Clearly conjugacy in $G$ (resp. in $H$) defines an equivalence relation on $H$.

**Proposition C.3.1.** *Let $X$ be a set. Let $\gamma \in \mathrm{Sym}_0(X)$ be a cycle of length $r$ and let $\sigma \in \mathrm{Sym}(X)$. Then $\sigma\gamma\sigma^{-1}$ is also a cycle of length $r$. More precisely, if $\gamma = (x_1\ x_2\ \cdots\ x_r)$, then $\sigma\gamma\sigma^{-1}$ equals the cycle*

$$(\sigma(x_1)\ \sigma(x_2)\ \cdots\ \sigma(x_r)). \tag{C.3}$$

*Proof.* First observe that, by Proposition C.2.1(v) the support of $\sigma\gamma\sigma^{-1}$ is the set $\{\sigma(x_1), \sigma(x_2), \ldots, \sigma(x_n)\}$. Given $1 \le i \le r$, we have

$$(\sigma\gamma\sigma^{-1})(\sigma(x_i)) = (\sigma\gamma)(x_i) = \sigma(x_{i+1}),$$

where $r + 1 = 1$. It follows that $\sigma\gamma\sigma^{-1} = (\sigma(x_1)\ \sigma(x_2)\ \cdots\ \sigma(x_r))$. □

Let $\sigma \in \mathrm{Sym}_0(X)$. The *type* of $\sigma$ is the sequence $t(\sigma) = (t_r)_{r\ge 2}$ where $t_r$ is the number of cycles of length $r$ in the factorization of $\sigma$ as a product of cycles with pairwise disjoint supports (cf. Proposition C.2.3).

**Proposition C.3.2.** *Let $X$ be a set and let $\sigma$ and $\sigma'$ in $\mathrm{Sym}_0(X)$. Then the following conditions are equivalent:*

(a) *$\sigma$ and $\sigma'$ are conjugate in $\mathrm{Sym}_0(X)$;*
(b) *$\sigma$ and $\sigma'$ are conjugate in $\mathrm{Sym}(X)$;*
(c) *$\sigma$ and $\sigma'$ have the same type.*

*Proof.* The implication (a) ⇒ (b) is obvious. Suppose (b). let

$$\sigma = \gamma_1\gamma_2\cdots\gamma_r \tag{C.4}$$

be the factorization of $\sigma$ as a product of cycles with disjoint supports and let $\alpha \in \mathrm{Sym}(X)$ be a permutation such that $\sigma' = \alpha\sigma\alpha^{-1}$. From (C.4) we deduce that $\sigma' = (\alpha\gamma_1\alpha^{-1})(\alpha\gamma_2\alpha^{-1})\cdots(\alpha\gamma_r\alpha^{-1})$. It follows from Proposition C.3.1 that $t(\sigma') = t(\sigma)$. This shows (b) ⇒ (c).

Finally, suppose that

$$\sigma = (x_1\ x_2\ \cdots\ x_{r_1})(y_1\ y_2\ \cdots\ y_{r_2})\cdots(z_1\ z_2\ \cdots\ z_{r_\ell})$$

and

$$\sigma' = (x'_1\ x'_2\ \cdots\ x'_{r_1})(y'_1\ y'_2\ \cdots y'_{r_2})\cdots(z'_1\ z'_2\ \cdots\ z'_{r_\ell})$$

are two permutations of the same type. Consider a permutation $\alpha$, with support the union of the supports of $\sigma$ and $\sigma'$, which maps $x_i$ to $x'_i$ for all $i = 1, 2, \ldots, r_1$, $y_j$ to $y'_j$, for all $j = 1, 2, \ldots, r_2$, ..., and $z_k$ to $z'_k$ for all $k = 1, 2, \ldots, r_\ell$. Note that $\alpha \in \mathrm{Sym}_0(X)$. Then (cf. the proof of Proposition C.3.1) $\alpha\sigma\alpha^{-1} = \alpha'$. It follows that $\sigma$ and $\sigma'$ are conjugate in $\mathrm{Sym}_0(X)$. □

## C.4 The Alternating Group

**Proposition C.4.1.** *Let $X$ be a set and let $\sigma \in \mathrm{Sym}_0(X)$. Suppose that $\sigma$ can be expressed as a product of $n$ transpositions. Then the parity of $n$ only depends on $\sigma$.*

*Proof.* Suppose that $\sigma$ can be expressed both as a product of an even and as a product of an odd number of transpositions. Then, the same holds for $\sigma^{-1}$. It follows that choosing an even writing for $\sigma$ and an odd one for $\sigma^{-1}$, we can write the identity element $\mathrm{Id}_X = \sigma\sigma^{-1}$ as a product of an odd number of transpositions, say

$$\mathrm{Id}_X = \tau_1\tau_2\cdots\tau_{2m+1}. \tag{C.5}$$

Let $x$ be an element in $X$ appearing in the support of one of the transpositions $\tau_i$ in (C.5).

As transpositions with disjoint support commute (cf. Proposition C.2.1(iv)) and $(y\ z)(x\ z) = (x\ y)(y\ z)$ for all distinct elements $y, z$ in $X \setminus \{x\}$, we can move all transpositions of the form $(x\ y)$ to the left in (C.5). In other words, we can write the identity $\mathrm{Id}_X$ as a product of $2m+1$ transpositions

$$\mathrm{Id}_X = \tau'_1\tau'_2\cdots\tau'_r\tau'_{r+1}\cdots\tau'_{2m+1} \tag{C.6}$$

where $1 \le r \le 2m+1$ and $x$ belongs to the support of $\tau'_i$ if and only if $1 \le i \le r$. Let $\tau'_r = (x\ y)$ and observe that it cannot appear only once in the product (C.6), otherwise the element $x$ would be mapped onto $y$, while it has to remain fixed, since that product is the identity. It follows that there exists $1 \le j \le r-1$ such that $\tau'_j = \tau'_r$ and $\tau'_i \ne \tau'_r$ for all $i = j+1, j+2, \ldots, r-1$. Now, for all $j+1 \le i \le r-1$, if $\tau'_i = (x\ z)$, we have $\tau'_i\tau'_r = (x\ z)(x\ y) = (x\ y)(y\ z) = \tau'_r(y\ z)$. Thus, we can move to the left the transposition $\tau'_r$ next to $\tau'_j$ and cancel them out, without changing the value of the product. Repeating this operation, we reduce every time by 2 the

number of transpositions in (C.6). Eventually, we reach a single transposition. This clearly yields a contradiction. □

Consider the map

$$\varepsilon\colon \operatorname{Sym}_0(X) \to \{-1, 1\} \tag{C.7}$$

defined by setting $\varepsilon(\sigma) = 1$ if $\sigma$ is a product of an even number of permutations and $\varepsilon(\sigma) = -1$ otherwise. Note that $\varepsilon$ is well defined by virtue of Proposition C.4.1. It is obvious that $\varepsilon$ is a group homomorphism. Moreover, $\varepsilon$ is surjective if $X$ has at least two elements.

The normal subgroup $\ker(\varepsilon) \subset \operatorname{Sym}_0(X)$ is called the *alternating group* on $X$ and it is denoted by $\operatorname{Sym}_0^+(X)$.

**Proposition C.4.2.** *Let $X$ be a set. An $r$-cycle is in $\operatorname{Sym}_0^+(X)$ if and only if $r$ is odd. In particular, every 3-cycle belongs to $\operatorname{Sym}_0^+(X)$.*

*Proof.* We have seen (cf. (C.2)) that and $r$-cycle is a product of $r - 1$ transpositions. □

Recall that a nontrivial group $G$ is simple if the only normal subgroups of $G$ are the trivial subgroup $\{1_G\}$ and $G$ itself.

**Theorem C.4.3.** *Let $X$ be a set having at least five distinct elements. Then the group $\operatorname{Sym}_0^+(X)$ is simple.*

*Proof.* Every element of $\operatorname{Sym}_0^+(X)$ is a product of permutations of the form $(s\ t)(u\ v)$ or $(s\ t)(s\ u)$, where $s, t, u$ and $v$ are distinct elements of $X$. Since $(s\ t)(u\ v) = (s\ u\ t)(s\ u\ v)$ and $(s\ t)(s\ u) = (s\ u\ t)$, it follows that $\operatorname{Sym}_0^+(X)$ is generated by the set of all 3-cycles.

Let $x$ and $y$ be two distinct elements in $X$. Clearly, any 3-cycle is of one of the forms $(x\ y\ s), (x\ s\ y), (x\ s\ t), (y\ t\ u)$, or $(s\ t\ u)$, where $s, t, u$ are distinct elements in $X \setminus \{x, y\}$. We have

$$\begin{aligned}
(x\ s\ y) &= (x\ y\ s)^2,\\
(x\ s\ t) &= (x\ y\ t)(x\ s\ y) = (x\ y\ t)(x\ y\ s)^2,\\
(y\ s\ t) &= (x\ t\ y)(x\ y\ s) = (x\ y\ t)^2(x\ y\ s),\\
(s\ t\ u) &= (x\ s\ y)(x\ y\ u)(x\ t\ y)(x\ y\ s) = (x\ y\ s)^2(x\ y\ u)(x\ y\ t)^2(x\ y\ s).
\end{aligned}$$

This shows that $\operatorname{Sym}_0^+(X)$ is generated by the 3-cycles $(x\ y\ z)$, where $z \in X \setminus \{x, y\}$.

Let now $N \subset \operatorname{Sym}_0^+(X)$ be a nontrivial normal subgroup. Let us show that $N = \operatorname{Sym}_0^+(X)$. We distinguish a few cases (corresponding to the different possible cycle structures of a nontrivial element in $N$).

**Case 1.** $N$ contains a 3-cycle $(x\ y\ s)$. Then, for any $z \in X \setminus \{x, y, s\}$ we have that the 3-cycle

$$(x\ y\ z) = (x\ y)(s\ z)(x\ y\ s)^2(s\ z)(x\ y) = [(x\ y)(s\ z)]^{-1}(x\ y\ z)^2[(s\ z)(x\ y)]$$

belongs to $N$. From the preceding part of the proof, we deduce that $N = \mathrm{Sym}_0^+(x)$.

**Case 2.** $N$ contains an element $\sigma$ whose factorization as product of cycles with disjoint supports contains a cycle $(x_1\ x_2\ \cdots\ x_r)$ of length $r \geq 4$. Write $\sigma = (x_1\ x_2\ \cdots\ x_r)\rho$, where $\rho \in \mathrm{Sym}_0(X)$ is the product of the remaining cycles. Consider the cycle $\gamma = (x_1\ x_2\ x_3)$. Then, $\sigma(\gamma\sigma^{-1}\gamma^{-1}) \in N$, as $N$ is a normal subgroup. By Proposition C.3.1, we have

$$\sigma(\gamma\sigma^{-1}\gamma^{-1}) = (\sigma\gamma\sigma^{-1})\gamma^{-1} = (x_2\ x_3\ x_4)(x_3\ x_2\ x_1) = (x_1\ x_4\ x_2).$$

Thus, $N$ contains a 3-cycle and, by Case 1, $N = \mathrm{Sym}_0^+(X)$.

**Case 3.** $N$ contains an element $\sigma$ whose factorization as a product of cycles with disjoint supports contains at least two cycles $(x_1\ x_2\ x_3)$ and $(x_4\ x_5\ x_6)$ of length 3. Write $\sigma = (x_1\ x_2\ x_3)(x_4\ x_5\ x_6)\rho$, where $\rho \in \mathrm{Sym}_0(X)$ is the product of the remaining cycles. Consider the cycle $\gamma = (x_1\ x_2\ x_4)$. Then, as above, $\sigma\gamma\sigma^{-1}\gamma^{-1} \in N$. But, again by Proposition C.3.1, we have

$$\sigma\gamma\sigma^{-1}\gamma^{-1} = (x_2\ x_3\ x_5)(x_4\ x_2\ x_1) = (x_1\ x_4\ x_3\ x_5\ x_2).$$

Thus, $N$ contains a 5-cycle and by Case 2, $N = \mathrm{Sym}_0^+(X)$.

**Case 4.** $N$ contains an element $\sigma$ which factorizes as a product of one single 3-cycle $(x_1\ x_2\ x_3)$ and transpositions with disjoint supports. Write $\sigma = (x_1\ x_2\ x_3)\rho$, where $\rho \in \mathrm{Sym}_0(X)$ is the product of the transpositions. Note that $\rho^2 = \mathrm{Id}_X$. Then $\sigma^2 \in N$ and

$$\sigma^2 = (x_1\ x_2\ x_3)\rho(x_1\ x_2\ x_3)\rho = (x_1\ x_2\ x_3)^2\rho^2 = (x_1\ x_3\ x_2).$$

Thus, again as in Case 1, $N = \mathrm{Sym}_0^+(X)$.

**Case 5.** $N$ contains an element $\sigma$ which is the product of an (even) number of transpositions with disjoint supports. We can write $\sigma = (x_1\ x_2)(x_3\ x_4)\rho$ where $\rho \in \mathrm{Sym}_0(X)$ satisfies $\rho^2 = \mathrm{Id}_X$. Consider the cycle $\gamma = (x_1\ x_2\ x_3)$. Then, the element $\pi = \sigma\gamma\sigma^{-1}\gamma^{-1}$ belongs to $N$. By Proposition C.3.1,

$$\pi = \sigma\gamma\sigma^{-1}\gamma^{-1} = (x_2\ x_1\ x_4)(x_1\ x_3\ x_2) = (x_1\ x_3)(x_2\ x_4).$$

Since $X$ has at least five distinct elements, there exists an element $y \in X \setminus \{x_1, x_2, x_3, x_4\}$. Set $\delta = (x_1\ x_3\ y)$. Then, $\pi(\delta\pi^{-1}\delta^{-1}) \in N$. But, one more time by Proposition C.3.1,

$$\pi(\delta\pi^{-1}\delta^{-1}) = (\pi\delta\pi^{-1})\delta^{-1} = (x_3\ x_1\ y)(y\ x_3\ x_1) = (x_1\ x_3\ y).$$

We are again in Case 1, and therefore $N = \mathrm{Sym}_0^+(X)$.

In all cases, $N = \mathrm{Sym}_0^+(X)$, and this shows that $\mathrm{Sym}_0^+(X)$ is simple. □

Note that if $X$ is a finite set, then $\mathrm{Sym}_0(X) = \mathrm{Sym}(X)$. Given an integer $n \geq 1$, we denote by $\mathrm{Sym}_n$ the symmetric group of the set $\{1, 2, \ldots, n\}$. The group $\mathrm{Sym}_n$ is called the *symmetric group* of *degree* $n$. By Remark C.1.1, if $X$ is a finite set with $|X| = n$, we have $\mathrm{Sym}_n \cong \mathrm{Sym}(X)$.

The subgroup $\mathrm{Sym}_0^+(\{1, 2, \ldots, n\})$ is called the *alternating group* of *degree* $n$ and it is denoted by $\mathrm{Sym}_n^+$.

*Remark C.4.4.* The groups $\mathrm{Sym}_1^+ (= \mathrm{Sym}_1)$ and $\mathrm{Sym}_2^+$ are trivial groups.

The group $\mathrm{Sym}_3^+ = \{\mathrm{Id}_{\{1,2,3\}}, (1\ 2\ 3), (1\ 3\ 2)\}$ is cyclic of order 3 and therefore it is a simple group.

On the other hand, the subgroup

$$K = \{\mathrm{Id}_{\{1,2,3,4\}}, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\} \subset \mathrm{Sym}_4^+$$

has index two in $\mathrm{Sym}_4^+$. It follows that $K$ is a normal subgroup of $\mathrm{Sym}_4^+$. Therefore, $\mathrm{Sym}_4^+$ is not a simple group.

From Theorem C.4.3 we deduce that the group $\mathrm{Sym}_n^+$ is simple for all $n \geq 5$.

# Appendix D
# Free Groups

## D.1 Concatenation of Words

Let $A$ be a set.

A *word* on the alphabet set $A$ is an element of the set

$$A^* = \bigcup_{n\in\mathbb{N}} A^n,$$

where $A^n$ is the Cartesian product of $A$ with itself $n$ times, that is, the set consisting of all $n$-tuples $(a_1, a_2, \ldots, a_n)$ with $a_k \in A$ for $1 \leq k \leq n$. The unique element of $A^0$ is denoted by $\epsilon$ and is called the *empty word*.

The *concatenation* of two words $w = (a_1, a_2, \ldots, a_m) \in A^m$ and $w' = (a'_1, a'_2, \ldots, a'_n) \in A^n$ is the word $ww' \in A^{m+n}$ defined by

$$ww' = (a_1, a_2, \ldots, a_m, a'_1, a'_2, \ldots, a'_n).$$

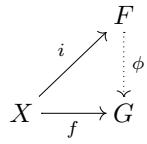
We have $\epsilon w = w\epsilon = w$ and $(ww')w'' = w(w'w'')$ for all $w, w', w'' \in A^*$. Thus, $A^*$ is a monoid for the concatenation product whose identity element is the empty word $\epsilon$.

Observe that each word $w = (a_1, a_2, \ldots, a_n) \in A^n$ may be uniquely written as a product of elements of $A = A^1$, namely $w = a_1 a_2 \cdots a_n$.

## D.2 Definition and Construction of Free Groups

**Definition D.2.1.** A *based free group* is a triple $(F, X, i)$, where $F$ is a group, $X$ is a set, and $i\colon X \to F$ is a map from $X$ to $F$ satisfying the following universal property: for every group $G$ and any map $f\colon X \to G$, there exists a unique homomorphism $\phi\colon F \to G$ such that $f = \phi \circ i$.

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_12, 




A group $F$ is called *free* if there exist a set $X$ and a map $i\colon X \to F$ such that the triple $(F, X, i)$ is a based free group. One then says that $(X, i)$ is a *free base* for $F$ and that $F$ is a free group *based* on $(X, i)$.

*Remarks D.2.2.* (a) If $(F, X, i)$ is a based free group and $\alpha\colon Y \to X$ is a bijective map from a set $Y$ onto $X$, then the triple $(F, Y, i \circ \alpha)$ is also a based free group. Indeed, if $f\colon Y \to G$ is a map from $Y$ into a group $G$, then there is a unique homomorphism $\phi\colon F \to G$ such that $f = \phi \circ i \circ \alpha$, namely the unique homomorphism $\phi\colon F \to G$ satisfying $f \circ \alpha^{-1} = \phi \circ i$.

(b) If $(F, X, i)$ is a based free group and $\psi\colon F \to F'$ is an isomorphism from $F$ onto a group $F'$, then the triple $(F', X, \psi \circ i)$ is a based free group. Indeed, if $f\colon X \to G$ is a map from $X$ into a group $G$, then there exists a unique homomorphism $\phi'\colon F' \to G$ such that $f = \phi' \circ \psi \circ i$, namely the homomorphism given by $\phi' = \phi \circ \psi^{-1}$, where $\phi\colon F \to G$ is the unique homomorphism satisfying $f = \phi \circ i$.

**Proposition D.2.3.** *Let $(F, X, i)$ be a based free group. Then the following hold:*

(i) *the map $i$ is injective;*
(ii) *the set $i(X)$ generates the group $F$;*
(iii) *the triple $(F, X', i')$, where $X' = i(X)$ and $i'\colon X' \to F$ is the inclusion map, is also a based free group.*

*Proof.* (i) Let $x_1$ and $x_2$ be two distinct elements in $X$. Consider the map $f\colon X \to \mathbb{Z}/2\mathbb{Z}$ defined by $f(x) = \overline{0}$ if $x \neq x_2$ and $f(x_2) = \overline{1}$. Since $(F, X, i)$ is a based free group, there exists a homomorphism $\phi\colon F \to \mathbb{Z}/2\mathbb{Z}$ such that $f = \phi \circ i$. As $f(x_1) \neq f(x_2)$, this implies $i(x_1) \neq i(x_2)$. Therefore, the map $i$ is injective.

(ii) Denote by $H$ the subgroup of $F$ generated by $i(X)$. Consider the map $i_*\colon X \to H$ defined by $i_*(x) = i(x)$ for all $x \in X$. Since $(F, X, i)$ is a based free group, there exists a homomorphism $\phi\colon F \to H$ such that $i_* = \phi \circ i$. Consider now the inclusion map $\rho\colon H \to F$. The homomorphisms $\mathrm{Id}_F$ and $\rho \circ \phi$ satisfy $\mathrm{Id}_F \circ i = \rho \circ \phi \circ i$. By uniqueness, we get $\mathrm{Id}_F = \rho \circ \phi$. This implies that $\rho$ is surjective, that is, $H = F$. Therefore, $i(X)$ generates $F$.

(iii) The fact that $(F, X', i')$ is a based free group immediately follows from Remark D.2.2(a) since $i' = i \circ j^{-1}$ where $j\colon X \to X'$ is the bijective map defined by $j(x) = i(x)$ for all $x \in X$. □

From Proposition D.2.3(iii), we deduce that if $F$ is a free group then there exists a subset $X \subset F$ such that the triple $(F, X, i)$, where $i\colon X \to F$ is the inclusion map, is a based free group. Such a subset $X \subset F$ is then called a *free base subset*, or simply a *base*, for $F$.

**Proposition D.2.4.** *Let $(F, X, i)$ be a based free group and let $Y \subset X$. Let $K$ denote the subgroup of $F$ generated by $Y$ and let $j\colon Y \to K$ be the map defined by $j(y) = i(y)$ for all $y \in Y$. Then $(K, Y, j)$ is a based free group.*

*Proof.* Let $G$ be a group and let $f\colon Y \to G$ be a map. Let us show that there exists a unique homomorphism $\phi\colon K \to G$ satisfying $f = \phi \circ j$. Uniqueness follows from the fact that $j(Y) = i(Y)$ generates $K$. Choose a map $f'\colon X \to G$ extending $f$. As $(F, X, i)$ is a based free group, there exists a homomorphism $\phi'\colon F \to G$ such that $f' = \phi' \circ i$. Then $\phi = \phi'|_K\colon K \to G$ satisfies $f = \phi \circ j$. This proves that $(K, Y, j)$ is a based free group. □

In the case when $i$ is an inclusion map, this gives us the following:

**Corollary D.2.5.** *Let $F$ be a free group with base $X \subset F$. Let $Y \subset X$ and let $K$ denote the subgroup of $F$ generated by $Y$. Then $K$ is a free group with base $Y$.* □

**Proposition D.2.6.** *Let $(F_1, X_1, i_1)$ and $(F_2, X_2, i_2)$ be two based free groups. Suppose that there is a bijective map $u\colon X_1 \to X_2$. Then there exists a unique group isomorphism $\varphi\colon F_1 \to F_2$ satisfying $\varphi \circ i_1 = i_2 \circ u$.*

$$\begin{array}{ccc} F_1 & \xrightarrow{\varphi} & F_2 \\ {\scriptstyle i_1}\uparrow & & \uparrow{\scriptstyle i_2} \\ X_1 & \xrightarrow[u]{} & X_2 \end{array}$$

*Proof.* Since $(F_1, X_1, i_1)$ is a based free group, there exists a unique homomorphism $\varphi\colon F_1 \to F_2$ such that $i_2 \circ u = \varphi \circ i_1$. It suffices to show that $\varphi$ is bijective. To see this, we now use the fact that $(F_2, X_2, i_2)$ is a based free group. This implies that there exists a homomorphism $\varphi'\colon F_2 \to F_1$ such that $i_1 \circ u^{-1} = \varphi' \circ i_2$. The maps $\mathrm{Id}_{F_1}$ and $\varphi' \circ \varphi$ are endomorphisms of $F_1$ satisfying $\mathrm{Id}_{F_1} \circ i_1 = i_1$ and $(\varphi' \circ \varphi) \circ i_1 = i_1$. Since $(F_1, X_1, i_1)$ is a based free group, it follows that $\mathrm{Id}_{F_1} = \varphi' \circ \varphi$ by uniqueness. Similarly, we get $\varphi \circ \varphi' = \mathrm{Id}_{F_2}$. This shows that $\varphi$ is bijective. □

**Theorem D.2.7.** *Let $X$ be a set. Then there exist a group $F$ and a map $i\colon X \to F$ such that the triple $(F, X, i)$ is a based free group.*

*Proof.* Let $X'$ be a disjoint copy of $X$, that is, a set $X'$ such that $X \cap X' = \varnothing$ together with a bijective map $\gamma\colon X \to X'$. Let $A = X \cup X'$. For each $a \in A$, define the element $\widetilde{a} \in A$ by setting

$$\widetilde{a} = \begin{cases} \gamma(a) & \text{if } a \in X, \\ \gamma^{-1}(a) & \text{if } a \in X'. \end{cases}$$

Observe that the map $a \mapsto \widetilde{a}$ is an involution of $A$, that is, it satisfies $\widetilde{\widetilde{a}} = a$ for all $a \in A$.

Consider now the set $A^*$ consisting of all words on the alphabet set $A$ (see Sect. D.1). Recall that $A^*$ is a monoid for the concatenation product whose identity element is the empty word $\epsilon$.

We say that a word $w \in A^*$ may be obtained from a word $w' \in A^*$ by an *elementary reduction* if there exist an element $a \in A$ and words $u, v \in A^*$ such that $w = uv$ and $w' = ua\widetilde{a}v$. Given words $w, w' \in A^*$, we write $w \sim w'$ if either $w$ may be obtained from $w'$ by an elementary reduction or $w'$ may be obtained from $w$ by an elementary reduction. Finally, we define a relation $\equiv$ on $A^*$ by writing $w \equiv w'$ for $w, w' \in A^*$ if and only if there exist an integer $n \geq 1$ and a sequence of words $w_1, w_2, \ldots, w_n \in A^*$ with $w_1 = w$ and $w_n = w'$ such that $w_i \sim w_{i+1}$ for each $i = 1, 2, \ldots, n-1$. It is immediate to check that $\equiv$ is an equivalence relation on $A^*$. Denote by $[w]$ the equivalence class of an element $w \in A^*$ and consider the quotient set $F = A^*/\equiv$, that is, the set consisting of all equivalence classes $[w]$, $w \in A^*$. Observe that if $u, u', v \in A^*$ and $u \sim u'$ then $uv \sim u'v$ and $vu \sim vu'$. It follows from this observation that if the words $u, v, u', v' \in A^*$ satisfy $u \equiv u'$ and $v \equiv v'$ then $uv \equiv u'v'$, so that we can define the product of $[u]$ and $[v]$ in $F$ by setting

$$[u][v] = [uv].$$

Let us show that this product gives a group structure on $F$. The associativity immediately follows from the associativity of the concatenation product in $A^*$. Indeed, for all $u, v, w \in A^*$, we have

$$([u][v])[w] = [uv][w] = [(uv)w] = [u(vw)] = [u][vw] = [u]([v][w]).$$

On the other hand, for all $w \in A^*$, we have

$$[\epsilon][w] = [\epsilon w] = [w] \quad \text{and} \quad [w][\epsilon] = [w\epsilon] = [w]$$

which shows that $[\epsilon]$ is an identity element. Finally, let us show that every element in $F$ admits an inverse. For $w = a_1 a_2 \cdots a_n \in A^*$, where $n \geq 0$ and $a_k \in A$ for $1 \leq k \leq n$, define the word $\widetilde{w} \in A^*$ by

$$\widetilde{w} = \widetilde{a_n}\widetilde{a_{n-1}} \cdots \widetilde{a_1}.$$

Observe that

$$\begin{aligned} w\widetilde{w} &= a_1 a_2 \cdots a_n \widetilde{a_n} \cdots \widetilde{a_2}\widetilde{a_1} \\ &\sim a_1 \cdots a_{n-1}\widetilde{a_{n-1}} \cdots \widetilde{a_1} \\ &\vdots \\ &\sim a_1 a_2 \widetilde{a_2}\widetilde{a_1} \\ &\sim a_1 \widetilde{a_1} \\ &\sim \epsilon. \end{aligned}$$

Thus, we have $w\widetilde{w} \equiv \epsilon$. This gives us $[w][\widetilde{w}] = [w\widetilde{w}] = [\epsilon]$. Similarly, we get $[\widetilde{w}][w] = [\epsilon]$. It follows that $[\widetilde{w}]$ is an inverse of $[w]$ for the product operation in $F$. This proves that $F$ is a group.

Consider the map $i\colon X \to F$ defined by $i(x) = [x]$ for all $x \in X$. If $w = a_1 a_2 \cdots a_n \in A^*$, where $a_k \in A$ for $1 \leq k \leq n$, then

$$[w] = [a_1 a_2 \cdots a_n] = [a_1][a_2] \cdots [a_n].$$

Since $[a] = i(a)$ if $a \in X$ and $[a] = [\widetilde{a}]^{-1} = i(\widetilde{a})^{-1}$ if $a \in X'$, we deduce that $i(X)$ generates the group $F$.

Let us show that the triple $(F, X, i)$ is a based free group. Suppose that $f\colon X \to G$ is a map from $X$ into a group $G$. We have to prove that there exists a unique homomorphism $\phi\colon F \to G$ such that $f = \phi \circ i$. Uniqueness follows from the fact that $i(X)$ generates $F$. To construct $\phi$, we first extend $f$ to a map $g\colon A \to G$ by setting

$$g(a) = \begin{cases} f(a) & \text{if } a \in X, \\ f(\widetilde{a})^{-1} & \text{if } a \in X'. \end{cases}$$

Observe that $g(\widetilde{a}) = g(a)^{-1}$ for all $a \in A$. Define now a map $\Phi\colon A^* \to G$ by setting

$$\Phi(w) = g(a_1)g(a_2) \cdots g(a_n)$$

for all $w = a_1 a_2 \cdots a_n \in A^*$. Note that we have

$$\Phi(ww') = \Phi(w)\Phi(w') \tag{D.1}$$

for all $w, w' \in A^*$. Moreover, if $w$ may be obtained from $w'$ by an elementary reduction, that is, $w = uv$ and $w' = ua\widetilde{a}v$ for some $a \in A$ and $u, v \in A^*$, then

$$\Phi(w') = \Phi(u)g(a)g(\widetilde{a})\Phi(v) = \Phi(u)g(a)g(a)^{-1}\Phi(v) = \Phi(u)\Phi(v) = \Phi(w).$$

It follows that $\Phi(w) = \Phi(w')$ whenever $w, w' \in A^*$ satisfy $w \sim w'$. By induction, we deduce that $\Phi(w) = \Phi(w')$ for all $w, w' \in A^*$ such that $w \equiv w'$. Thus, we can define a map $\phi\colon F \to G$ by setting $\phi([w]) = \Phi(w)$ for all $w \in A^*$. By using (D.1), we get

$$\phi([w][w']) = \phi([ww']) = \Phi(ww') = \Phi(w)\Phi(w') = \phi([w])\phi([w']).$$

Therefore, $\phi$ is a group homomorphism. On the other hand, for all $x \in X$, we have

$$\phi \circ i(x) = \phi([x]) = g(x) = f(x),$$

which shows that $\phi \circ i = f$. Consequently, the triple $(F, X, i)$ is a based free group. □

Given an arbitrary set $X$, it follows from Theorem D.2.7 that there exist a based free group $(F, X, i)$. Then one often says that $F = F(X)$ is *the* free

group based on $X$ (this is a minor abuse of language since if $(F', X, i')$ is another based free group then there is a unique isomorphism $\varphi\colon F \to F'$ such that $\varphi \circ i = i'$, by Proposition D.2.6). For $k \in \mathbb{N}$, the free group based on $\{1, 2, \ldots, k\}$ is denoted by $F_k$.

We recall that one says that two sets $X_1$ and $X_2$ are *equipotent*, or that they have the same *cardinality*, if there exists a bijective map $u\colon X_1 \to X_2$.

**Theorem D.2.8.** *Let $F_1$ and $F_2$ be two free groups based on $X_1 \subset F_1$ and $X_2 \subset F_2$. Then the following conditions are equivalent:*

(a) *the groups $F_1$ and $F_2$ are isomorphic;*
(b) *the sets $X_1$ and $X_2$ are equipotent.*

For the proof, we shall need a few preliminary results. We use the following notation. If $G_1$ and $G_2$ are groups, we denote by $\mathrm{Hom}(G_1, G_2)$ the set consisting of all homomorphisms $\phi\colon G_1 \to G_2$. We denote by $\mathcal{P}(X)$ the set of all subsets of a set $X$.

**Lemma D.2.9.** *Let $F$ be a free group based on $X \subset F$. Then the sets $\mathrm{Hom}(F, \mathbb{Z}/2\mathbb{Z})$ and $\mathcal{P}(X)$ are equipotent.*

*Proof.* Since $F$ is free with base $X$, each map $f\colon X \to \mathbb{Z}/2\mathbb{Z}$ can be uniquely extended to a homomorphism $\phi\colon F \to \mathbb{Z}/2\mathbb{Z}$. Therefore, the restriction map yields a bijection from $\mathrm{Hom}(F, \mathbb{Z}/2\mathbb{Z})$ onto the set of maps from $X$ to $\mathbb{Z}/2\mathbb{Z}$. Consequently, the sets $\mathrm{Hom}(F, \mathbb{Z}/2\mathbb{Z})$ and $\mathcal{P}(X)$ are equipotent. □

**Lemma D.2.10.** *Let $F$ be a free group based on $X \subset F$. Suppose that the set $X$ is infinite. Then the sets $X$ and $F$ are equipotent.*

*Proof.* Let $A = X \cup X^{-1}$. Since $X$ generates $F$, the map $\rho\colon A^* \to F$ defined by $\rho(a_1, a_2, \ldots, a_n) = a_1 a_2 \ldots a_n$ is surjective. As $X$ is infinite, the sets $X$, $A$, and $A^*$ are all equipotent. It follows that there is a surjective map from $X$ onto $F$ and therefore an injective map from $F$ into $X$. By applying the Cantor-Bernstein theorem (cf. Corollary H.3.5), we deduce that $X$ and $F$ are equipotent. □

*Proof of Theorem D.2.8.* The fact that (b) implies (a) immediately follows from Proposition D.2.6.

To prove the converse implication, suppose that there exists an isomorphism $\alpha\colon F_1 \to F_2$. Then $\alpha$ induces a bijective map between the sets $\mathrm{Hom}(F_1, \mathbb{Z}/2\mathbb{Z})$ and $\mathrm{Hom}(F_2, \mathbb{Z}/2\mathbb{Z})$. It follows that $\mathcal{P}(X_1)$ and $\mathcal{P}(X_2)$ are equipotent by Lemma D.2.9. If $X_1$ is infinite, this implies that $X_2$ is also infinite, and we conclude that $X_1$ and $X_2$ are equipotent by applying Lemma D.2.10. On the other hand, if $X_1$ is finite, the fact that $\mathcal{P}(X_1)$ and $\mathcal{P}(X_2)$ are equipotent implies that $X_2$ is also finite and that $2^{|X_1|} = 2^{|X_2|}$. This gives us $|X_1| = |X_2|$ and we conclude that $X_1$ and $X_2$ are also equipotent in this case. □

**Corollary D.2.11.** *Let $F$ be a free group and let $X_1, X_2 \subset F$ be two bases of $F$. Then $X_1$ and $X_2$ are equipotent.* □

Let $F$ be a free group. The cardinality of a base $X \subset F$ depends only on $F$ by Corollary D.2.11. This cardinality is called the *rank* of $F$. Two free groups are isomorphic if and only if they have the same rank by Theorem D.2.8. A group is free of finite rank $k \in \mathbb{N}$ if and only if it is isomorphic to $F_k$.

## D.3 Reduced Forms

In order to state the first result of this section, we use the notation introduced in the proof of Theorem D.2.7. A word $w \in A^*$ is said to be *reduced* if it contains no subword of the form $a\widetilde{a}$ with $a \in A$, that is, if there is no word $w' \in A^*$ which can be obtained from $w$ by applying an elementary reduction. Note that the empty word $\epsilon$ is reduced and that every subword of a reduced word is itself reduced.

**Theorem D.3.1.** *Every equivalence class for $\equiv$ contains a unique reduced word.*

*Proof.* It is clear that any word $w \in A^*$ can be transformed into a reduced word by applying a suitable finite sequence of elementary reductions. This shows the existence of a reduced word in any equivalence class for $\equiv$.

Let us prove uniqueness. Consider the subset $R \subset A^*$ consisting of all reduced words. For $a \in A$ and $r \in R$, define the word $\alpha_a(r)$ by

$$\alpha_a(r) = \begin{cases} w & \text{if } r = \widetilde{a}w \text{ for some } w \in A^*, \\ ar & \text{otherwise.} \end{cases}$$

Note that the word $\alpha_a(r)$ is always reduced. Thus, we get a map $\alpha_a \colon R \to R$ defined for each $a \in A$. Let us show that

$$\alpha_a \circ \alpha_{\widetilde{a}} = \alpha_{\widetilde{a}} \circ \alpha_a = \mathrm{Id}_R\,. \tag{D.2}$$

Let $a \in A$ and consider an arbitrary element $r \in R$. If $r = aw$ for some $w \in A^*$, then $\alpha_{\widetilde{a}}(r) = w$ and hence $\alpha_a(\alpha_{\widetilde{a}}(r)) = \alpha_a(w) = aw = r$ (observe that $w$ cannot start by $\widetilde{a}$ since $r$ is reduced). Otherwise, we have $\alpha_{\widetilde{a}}(r) = \widetilde{a}r$ and therefore $\alpha_a(\alpha_{\widetilde{a}}(r)) = \alpha_a(\widetilde{a}r) = r$. It follows that $\alpha_a \circ \alpha_{\widetilde{A}} = \mathrm{Id}_R$. By replacing $a$ by $\widetilde{a}$ in this equality, we get $\alpha_{\widetilde{a}} \circ \alpha_a = \mathrm{Id}_R$ since $\widetilde{\cdot}$ is an involution on $A$. This proves (D.2).

From (D.2), we deduce that $\alpha_a \in \mathrm{Sym}(R)$ for all $a \in A$. By setting

$$\rho(w) = \alpha_{a_1} \circ \alpha_{a_2} \circ \cdots \circ \alpha_{a_n}$$

for every word $w = a_1 a_2 \cdots a_n \in A^*$, we get a monoid homomorphism $\rho\colon A^* \to \mathrm{Sym}(R)$. Note that

$$\rho(r)(\epsilon) = r \quad \text{for all } r \in R. \tag{D.3}$$

On the other hand, it immediately follows from (D.2) that $\rho(w) = \rho(w')$ whenever $w, w' \in A^*$ satisfy $w \equiv w'$. Thus, if $r_1$ an $r_2$ are reduced words in the same equivalence class for $\equiv$, we have $\rho(r_1) = \rho(r_2)$ and therefore $r_1 = r_2$ by applying (D.3). □

**Corollary D.3.2.** *Let $(F, X, i)$ be a based free group. Then every element $f \in F$ can be uniquely written in the form*

$$f = i(x_1)^{h_1} i(x_2)^{h_2} \cdots i(x_n)^{h_n} \tag{D.4}$$

*with $n \geq 0$, $x_i \in X$ and $h_i \in \mathbb{Z} \setminus \{0\}$ for $1 \leq i \leq n$, and $x_i \neq x_{i+1}$ for $1 \leq i \leq n-1$.*

**Definition D.3.3.** The expression (D.4) is called the *reduced form* of the element $f$ in the based free group $(F, X, i)$.

*Proof of Corollary D.3.2.* We can assume that $(F, X, i)$ is the based free group constructed in the proof of Theorem D.2.7. By construction, the element $f \in F$ is an equivalence class for $\equiv$. This equivalence class contains a unique reduced word $r$ by Theorem D.3.1. The word $r$ can be uniquely written in the form

$$r = a_1^{k_1} a_2^{k_2} \cdots a_n^{k_n},$$

where $n \geq 0$, $a_i \in A$ and $k_i \in \mathbb{N} \setminus \{0\}$ for $1 \leq i \leq n$, and $a_i \neq a_{i+1}$ for $1 \leq i \leq n-1$. This gives us an expression of the form (D.4) with $x_i = a_i$ and $h_i = k_i$ if $a_i \in X$, and $x_i = \widetilde{a_i}$ and $h_i = -k_i$ otherwise. Uniqueness of such an expression for $f$ follows from the uniqueness of the reduced word $r \in f$. □

**Corollary D.3.4.** *Let $G$ be a group. Let $U \subset G$ be a subset such that*

$$u_1^{k_1} u_2^{k_2} \cdots u_n^{k_n} \neq 1_G, \tag{D.5}$$

*for all $u_1, u_2, \ldots, u_n \in U$ and $k_1, k_2, \ldots, k_n \in \mathbb{Z} \setminus \{0\}$ with $u_i \neq u_{i+1}$ for $1 \leq i \leq n-1$ and $n \geq 1$. Then the subgroup of $G$ generated by $U$ is free with base $U$.*

*Proof.* Denote by $H$ the subgroup of $G$ generated by $U$ and by $\iota\colon U \to H$ the inclusion map. Let $(F, U, i)$ be a based free group (cf. Theorem D.2.7). Then there exists a unique homomorphism $\phi\colon F \to H$ such that $\iota = \phi \circ i$. Note that $\phi$ is surjective since $\phi(i(U)) = \iota(U) = U$ generates $H$. Let us show that $\phi$ is also injective. Consider an element $f \in F$ written in reduced form, that is, in the form $f = i(u_1)^{h_1} i(u_2)^{h_2} \cdots i(u_n)^{h_n}$ with $n \geq 0$, $u_i \in U$ and $h_i \in \mathbb{Z} \setminus \{0\}$ for $1 \leq i \leq n$, and $u_i \neq u_{i+1}$ for $1 \leq i \leq n-1$. Then we

have $\phi(f) = u_1^{h_1} u_2^{h_2} \cdots u_n^{h_n}$ since $\phi(i(u_i)) = \iota(u_i) = u_i$ for $1 \leq i \leq n$. It follows from (D.5) that $\phi(f) = 1_H$ if and only if $n = 0$, that is, if and only if $f = 1_F$. This shows that $\phi$ is an isomorphism. It follows from Remark D.2.2 that $(H, U, \iota)$ is a based free group. □

## D.4 Presentations of Groups

**Proposition D.4.1.** *Let $G$ be a group and let $S$ be a generating subset of $G$. Let $F$ denote the free group based on $S$. Then, the group $G$ is isomorphic to a quotient of $F$.*

*Proof.* Let $i\colon S \to F$ and $f\colon S \to G$ denote the inclusion maps. Since $F$ is free with base $S$, there exists a homomorphism $\phi\colon F \to G$ such that $f = \phi \circ i$. This implies that $S$ is contained in the image of $\phi$. Since $S$ generates $G$, we deduce that $\phi$ is surjective. Therefore, the group $G$ is isomorphic to the quotient group $F/\operatorname{Ker}(\phi)$. □

As every group admits a generating subset (e.g., the group itself), we deduce the following:

**Corollary D.4.2.** *Every group is isomorphic to a quotient of a free group.* □

A group is said to be *finitely generated* if it admits a finite generating subset. Proposition D.4.1 gives us:

**Corollary D.4.3.** *Every finitely generated group is isomorphic to a quotient of a free group of finite rank.* □

Let $A$ be a subset of a group $G$. The intersection of all normal subgroups of $G$ containing $A$ is a normal subgroup of $G$ which is called the *normal closure* of $A$ in $G$.

Let $G$ be a group. By Corollary D.4.2, there exist a free group $F$ and an epimorphism $\phi\colon F \to G$. Let $X$ be a free base for $F$. If $R$ is a subset of $F$ whose normal closure is the kernel of $\phi$, then one says that $G$ admits the *presentation*

$$G = \langle X; R \rangle. \tag{D.6}$$

Note that as $X$ generates $F$ (cf. Proposition D.2.3(ii)), we have that $\phi(X)$ generates $G$. The elements $x \in X$ (rather than the $\phi(x) \in G, x \in X$) are called, by abuse of language, the *generators* of the presentation (D.6). The elements $r \in R$ are called the *relators* of the presentation (D.6). For every $r \in R$ let $u_r, v_r$ be elements in $F$ such that $r = u_r(v_r)^{-1}$. Then (D.6) is often expressed as $G = \langle X : r = 1, r \in R \rangle$ or $G = \langle X : u_r = v_r, r \in R \rangle$.

Note that $G$ admits a presentation (D.6) with $X$ finite if and only if $G$ is finitely generated (cf. Corollary D.4.3).

If a group $G$ admits a presentation (D.6) with both $X$ and $R$ finite, then $G$ is said to be *finitely presented*.

## D.5 The Klein Ping-Pong Theorem

The following theorem is often used to prove that a group is free:

**Theorem D.5.1.** *Let $G$ be a group. Let $X$ be a generating subset of $G$ having at least two distinct elements. Suppose that $G$ acts on a set $E$ and that there is a family $(A_x)_{x\in X}$ of nonempty pairwise disjoint subsets of $E$ such that*

$$x^k \left( \bigcup_{y\in X\setminus\{x\}} A_y \right) \subset A_x \text{ for all } x \in X \text{ and } k \in \mathbb{Z}\setminus\{0\}. \tag{D.7}$$

*Then $G$ is a free group with base $X$.*

*Proof.* Consider an element $g \in G$ written as a nontrivial reduced word on the generating subset $X$, that is, in the form

$$g = x_1^{k_1} x_2^{k_2} \dots x_n^{k_n},$$

where $n \geq 1$, $x_i \in X$ and $k_i \in \mathbb{Z}\setminus\{0\}$ for $1 \leq i \leq n$, and $x_i \neq x_{i+1}$ for $1 \leq i \leq n-1$. By Corollary D.3.4, we have to show that $g \neq 1_G$.

Suppose first that either $X$ contains at least three distinct elements or $X$ has exactly two elements and $x_1 = x_n$. In this case, we can find an element $y \in X$ such that $y \neq x_1$ and $y \neq x_n$. By successive applications of D.7, we get

$$\begin{aligned}
gA_y &= x_1^{k_1} x_2^{k_2} \dots x_{n-2}^{k_{n-2}} x_{n-1}^{k_{n-1}} x_n^{k_n} A_y \\
&\subset x_1^{k_1} x_2^{k_2} \dots x_{n-2}^{k_{n-2}} x_{n-1}^{k_{n-1}} A_{x_n} \\
&\subset x_1^{k_1} x_2^{k_2} \dots x_{n-2}^{k_{n-2}} A_{x_{n-1}} \\
&\subset x_1^{k_1} x_2^{k_2} \dots A_{x_{n-2}} \\
&\dots \\
&\subset x_1^{k_1} x_2^{k_2} A_{x_3} \\
&\subset x_1^{k_1} A_{x_2} \\
&\subset A_{x_1}.
\end{aligned}$$

As the sets $A_y$ and $A_{x_1}$ are disjoint and $A_y \neq \varnothing$ by our hypotheses, we deduce that $g \neq 1_G$.

It remains to treat the case when $X$ has exactly two elements and $x_1 \neq x_n$. Then

$$x_1^{k_1} g x_1^{-k_1} = x_1^{2k_1} x_2^{k_2} x_3^{k_3} \dots x_n^{k_n} x_1^{-k_1}$$

is a reduced form of $x_1^{k_1} g x_1^{-k_1}$. We have $x_1^{k_1} g x_1^{-k_1} \neq 1_G$ by the first case. We deduce that we also have $g \neq 1_G$ in this case. □

*Remark D.5.2.* The proof of Theorem D.5.1 shows that the hypotheses on the family $(A_x)_{x \in X}$ may be relaxed: in fact, it suffices that the subsets $A_x \subset X$, $x \in X$, satisfy D.7 and $A_y \not\subset A_x$ for all distinct elements $x, y \in X$.

**Corollary D.5.3.** *Let $F$ be a free group of rank 2 and let $n \geq 2$ be an integer. Then $F$ contains a free subgroup of rank $n$.*

*Proof.* Suppose that $F$ is based on the elements $a$ and $b$. Let $G$ be the subgroup of $F$ generated by the subset

$$X = \{a^i b a^{-i} : 0 \leq i \leq n-1\}.$$

Consider the action of $G$ on $F$ given by left multiplication. Define, for each $x = a^i b a^{-i} \in X$, the subset $A_x \subset F$ as being the set of elements of $F$ whose reduced form starts by $a^i b^k$ for some $k \in \mathbb{Z} \setminus \{0\}$. The subsets $A_x$, $x \in X$, clearly satisfy the hypotheses of Theorem D.5.1. Therefore $G$ is free with base $X$. □

# Appendix E
# Inductive Limits and Projective Limits of Groups

## E.1 Inductive Limits of Groups

Let $I$ be a directed set. An *inductive system of groups* over $I$ consists of the following data: (1) a family of groups $(G_i)_{i\in I}$ indexed by $I$, (2) for each pair $i, j \in I$ such that $i \leq j$, a homomorphism $\psi_{ji}\colon G_i \to G_j$ satisfying the following conditions:

$$\psi_{ii} = \mathrm{Id}_{G_i} \text{ (identity map on } G_i) \text{ for all } i \in I,$$

$$\psi_{kj} \circ \psi_{ji} = \psi_{ki} \text{ for all } i, j, k \in I \text{ such that } i \leq j \leq k.$$

Then one speaks of the inductive system $(G_i, \psi_{ji})$ or simply of the inductive system $(G_i)$ if the homomorphisms $\psi_{ji}$ are understood.

Let $(G_i, \psi_{ji})$ be an inductive system of groups over $I$. Consider the relation $\sim$ on the disjoint union $E = \coprod_{i\in I} G_i$ defined as follows. If $x_i \in G_i$ and $x_j \in G_j$ are elements of $E$, then $x_i \sim x_j$ if and only if there is an element $k \in I$ such that $i \leq k$, $j \leq k$ and $\psi_{ki}(x_i) = \psi_{kj}(x_j)$. It easy to check that $\sim$ is a equivalence relation on the set $E$. Let $G = E/\sim$ be the set of equivalence classes of $\sim$. For $x_i \in G_i$, let $[x_i] \in G$ denote the class of $x_i$. One defines a binary operation on $G$ in the following way. Given $x_i \in G_i$ and $x_j \in G_j$, one defines the class $[x_i][x_j]$ by setting

$$[x_i][x_j] = [\psi_{ki}(x_i)\psi_{kj}(x_j)],$$

where $k \in I$ is such that $i \leq k$ and $j \leq k$. One checks that $[x_i][x_j]$ depends neither on the choice of the representatives $x_i$ and $x_j$, nor on the choice of $k$. Moreover, this operation gives a group structure on $G$. The group $G$ is called the *inductive limit* (or the *direct limit*) of the inductive system $(G_i)$ and one writes $G = \varinjlim G_i$. For each $i \in I$, there is a canonical homomorphism $h_i\colon G_i \to G$ defined by $h_i(x_i) = [x_i]$. Note that it immediately follows from the construction of $G$ that $G = \bigcup_{i\in I} h_i(G_i)$. Moreover, one has

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_13, 

$$h_j \circ \psi_{ji} = h_i$$

for all $i, j \in I$ such that $i \leq j$.

*Examples E.1.1.* (a) Let $(G_i)_{i\in I}$ be a family of groups and, for $i, j \in I$ such that $i \leq j$, let $\psi_{ji} \colon G_i \to G_j$ be the trivial homomorphism, namely, $\psi_{ji}(g) = 1_{G_j}$ for all $g \in G_i$. Then, $(G_i, \psi_{ji})$ is an inductive system whose inductive limit is a trivial group.

(b) Let $G$ be a group and denote by $\mathcal{H}$ the set of all finitely generated subgroups of $G$. Then $(\mathcal{H}, \subset)$ is a directed set and $\bigcup_{H\in\mathcal{H}} H = G$. Moreover, the set $\mathcal{H}$ together with the inclusion maps $\psi_{K,H} \colon H \to K$, for all $H, K \in \mathcal{H}$ with $H \subset K$, is an inductive system whose limit is canonically isomorphic to $G$.

## E.2 Projective Limits of Groups

Let $I$ be a directed set. A *projective system of groups* over $I$ consists of the following data: (1) a family of groups $(G_i)_{i\in I}$ indexed by $I$, (2) for each pair $i, j \in I$ such that $i \leq j$, a homomorphism $\varphi_{ij} \colon G_j \to G_i$ satisfying the following conditions:

$$\varphi_{ii} = \mathrm{Id}_{G_i} \text{ (identity map on } G_i) \text{ for all } i \in I,$$

$$\varphi_{ij} \circ \varphi_{jk} = \varphi_{ik} \text{ for all } i, j, k \in I \text{ such that } i \leq j \leq k.$$

Then one speaks of the projective system $(G_i, \varphi_{ij})$ or simply of the projective system $(G_i)$ if the homomorphisms $\varphi_{ij}$ are understood.

Let $(G_i, \varphi_{ij})$ be a projective system of groups over $I$. Let $P = \prod_{i\in I} G_i$ denote the direct product of the groups $G_i$. One immediately checks that

$$G = \{(x_i) \in P : \varphi_{ij}(x_j) = x_i \text{ for all } i, j \in I \text{ such that } i \leq j\}$$

is a subgroup of $P$. The group $G$ is called the *projective limit* of the projective system $(G_i)$, and one writes $G = \varprojlim G_i$. For each $i \in I$, there is a canonical homomorphism $f_i \colon G \to G_i$ obtained by restriction of the projection map $\pi_i \colon P \to G_i$. One has

$$\varphi_{ij} \circ f_j = f_i$$

for all $i, j \in I$ such that $i \leq j$.

*Examples E.2.1.* (a) Let $(G_i)_{i\in I}$ be a family of groups and, for $i, j \in I$ such that $i \leq j$, let $\varphi_{ij} \colon G_j \to G_i$ be the trivial homomorphism. Then, $(G_i, \varphi_{ij})$ is a projective system whose projective limit is a trivial group.

(b) Let $G$ be a group. Denote by $\mathcal{N}$ the set of all normal subgroups of $G$. Then $(\mathcal{N}, \supset)$ is a directed set. The family of groups $(G/H)_{H\in\mathcal{N}}$, when

equipped with the canonical quotient homomorphisms $\varphi_{H,K} \colon G/K \to G/H$ for all $H, K \in \mathcal{N}$ with $H \supset K$, gives rise to a projective system whose projective limit is canonically isomorphic to $G$.

# Appendix F
# The Banach-Alaoglu Theorem

All vector spaces considered in this appendix are vector spaces over the field $\mathbb{R}$ of real numbers.

## F.1 Topological Vector Spaces

A *topological vector space* is a real vector space $X$ endowed with a topology such that the maps

$$\begin{aligned} X \times X &\to X \\ (x, y) &\mapsto x + y \end{aligned}$$

and

$$\begin{aligned} \mathbb{R} \times X &\to X \\ (\lambda, x) &\mapsto \lambda x \end{aligned}$$

are continuous.

*Example F.1.1.* Let $\|\cdot\|$ be a norm on a real vector space $X$ and let $d$ denote the metric on $X$ defined by $d(x, y) = \|x - y\|$ for all $x, y \in X$. Then the topology defined by $d$ yields a structure of topological vector space on $X$.

Recall that a subset $C$ of a real vector space $X$ is said to be *convex* if

$$(1 - \lambda)x + \lambda y \in C$$

for all $\lambda \in [0, 1]$ and $x, y \in C$.

A topological real vector space $X$ is said to be *locally convex* if there is a base of neighborhoods of 0 consisting of convex subsets of $X$.

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_14, 

## F.2 The Weak-$*$ Topology

Let $X$ be a real vector space equipped with a norm $\|\cdot\|$. Recall that a linear map $u\colon X \to \mathbb{R}$ is continuous if and only if $u$ is bounded on the unit ball $B(X) = \{x \in X : \|x\| \leq 1\}$, that is, if and only if there is a constant $C \geq 0$ such that $|u(x)| \leq C$ for all $x \in B(X)$. The *topological dual* of $X$ is the vector space $X^*$ consisting of all continuous linear maps $u\colon X \to \mathbb{R}$. The *operator norm* on $X^*$ is the norm defined by

$$\|u\| = \sup_{x \in B(X)} |u(x)|.$$

The topology defined by the operator norm is called the *strong topology* on $X^*$.

Given $x \in X$, let $\psi_x\colon X^* \to \mathbb{R}$ denote the evaluation map $u \mapsto u(x)$ at $x$. The *weak-$*$ topology* on $X^*$ is the initial topology associated with the family of all evaluation maps $\psi_x\colon X^* \to \mathbb{R}$, $x \in X$. Thus, the weak-$*$ topology is the smallest topology on $X^*$ for which all evaluation maps $\psi_x$ are continuous. Observe that every subset of $X^*$ which is open for the weak-$*$ topology is also open for the strong topology since all the evaluation maps are continuous for the strong topology. It follows that convergence with respect to the strong topology implies convergence with respect to the weak-$*$ topology.

The weak-$*$ topology provides a topological vector space structure on $X^*$. A base of open neighborhoods of 0 for the weak-$*$ topology is given by all subsets of the form

$$V(F, \varepsilon) = \{u \in X^* : |u(x)| < \varepsilon \text{ for all } x \in F\},$$

where $F$ is a finite subset of $X$ and $\varepsilon > 0$. Since all the sets $V(F, \varepsilon)$ are convex, the topological vector space structure associated with the weak-$*$ topology on $X^*$ is locally convex.

The weak-$*$ topology on $X^*$ is Hausdorff. Indeed, if $u_1$ and $u_2$ are two distinct elements in $X^*$, then there exists $x \in X$ such that $u_1(x) \neq u_2(x)$. If $U_1$ and $U_2$ are disjoint open subsets of $\mathbb{R}$ containing $u_1(x)$ and $u_2(x)$ respectively, then $\psi_x^{-1}(U_1)$ and $\psi_x^{-1}(U_2)$ are disjoint open subsets of $X^*$ containing $u_1$ and $u_2$ respectively.

## F.3 The Banach-Alaoglu Theorem

In general, the unit ball in $X^*$ is not compact for the strong topology (this follows from the fact that the unit ball in a normed space is never compact unless the space is finite-dimensional). However, this ball is always compact for the weak-$*$ topology. This result, which is known as the Banach-Alaoglu

theorem is one of the central results of classical functional analysis and may be easily deduced from the Tychonoff theorem.

**Theorem F.3.1 (Banach-Alaoglu theorem).** *Let $X$ be a real normed vector space and let $\|\cdot\|$ denote the operator norm on its topological dual $X^*$. Then the unit ball $B(X^*) = \{u \in X^* : \|u\| \leq 1\}$ is compact for the weak-$*$ topology on $X^*$.*

*Proof.* Observe first that $X^*$ is a vector subspace of the vector space $\mathbb{R}^X$ consisting of all real-valued functions on $X$. On the other hand, setting $I_x = [-\|x\|, \|x\|] \subset \mathbb{R}$ for each $x \in X$. We have $B(X^*) \subset \prod_{x\in X} I_x$ by definition of the operator norm. Let us equip $\mathbb{R}^X = \prod_{x\in X} \mathbb{R}$ with the product topology. Then it is clear that the topology induced on $\prod_{x\in X} I_x$ is the product topology and that the topology induced on $X^*$ is the weak-$*$ topology. Let $f$ be an element of $\mathbb{R}^X$ which is the limit of a net $(u_i)$ of elements of $B(x^*)$. For all $x, y \in X$ and $\lambda \in \mathbb{R}$, we have $u_i(\lambda x) = \lambda u_i(x)$, $u_i(x + y) = u_i(x) + u_i(y)$ and $|u_i(x)| \leq \|x\|$ since $u_i \in B(X^*)$. By taking limits, we get $f(\lambda x) = \lambda f(x)$, $f(x + y) = f(x) + f(y)$ and $|f(x)| \leq \|x\|$. This shows that $f \in B(X^*)$. Consequently, $B(X^*)$ is closed in $\mathbb{R}^X$. Since $\prod_{x\in X} I_x$ is compact by Tychonoff theorem (Theorem A.5.2), we deduce that $B(X^*)$ is compact. □

# Appendix G
# The Markov-Kakutani Fixed Point Theorem

All vector spaces considered in this appendix are vector spaces over the field $\mathbb{R}$ of real numbers.

## G.1 Statement of the Theorem

Let $C$ be a convex subset of a real vector space $X$. A map $f\colon C \to C$ is called *affine* if

$$f((1-\lambda)x + \lambda y) = (1-\lambda)f(x) + \lambda f(y)$$

for all $\lambda \in [0,1]$ and $x, y \in C$.

**Theorem G.1.1 (Markov-Kakutani).** *Let $K$ be a nonempty convex compact subset of a Hausdorff topological vector space $X$. Let $\mathcal{F}$ be a set of continuous affine maps $f\colon K \to K$. Suppose that all elements of $\mathcal{F}$ commute, that is, $f_1 \circ f_2 = f_2 \circ f_1$ for all $f_1, f_2 \in \mathcal{F}$. Then there exists a point in $K$ which is fixed by all the elements of $\mathcal{F}$.*

## G.2 Proof of the Theorem

In the proof of the Markov-Kakutani theorem, we shall use the following lemmas.

**Lemma G.2.1.** *Let $K$ be a compact subset of a topological vector space $X$ and let $V$ be a neighborhood of $0$ in $X$. Then there exists a real number $\alpha > 0$ such that $\lambda K \subset V$ for every real number $\lambda$ such that $|\lambda| < \alpha$.*

*Proof.* Since the multiplication by a scalar $\mathbb{R} \times X \to X$ is continuous, we can find, for each $x \in X$, a real number $\alpha_x$ and an open neighborhood $\Omega_x \subset X$ of $x$ such that

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_15, 

$$|\lambda| < \alpha_x \Rightarrow \lambda_x \Omega_x \subset V. \tag{G.1}$$

The sets $\Omega_x$, $x \in K$, form an open cover of $K$. As $K$ is compact, there is a finite subset $F \subset K$ such that

$$K \subset \bigcup_{x \in F} \Omega_x. \tag{G.2}$$

If we take $\alpha = \min_{x \in F} \alpha_x$, then $\alpha > 0$ and

$$|\lambda| < \alpha \Rightarrow \lambda K \subset V$$

by (G.1) and (G.2). □

**Lemma G.2.2.** *Let $K$ be a compact subset of a topological vector space $X$. Let $(x_i)_{i \in I}$ be a net of points in $K$ and let $(\lambda_i)_{i \in I}$ be a net of real numbers converging to 0 in $\mathbb{R}$. Then the net $(\lambda_i x_i)_{i \in I}$ converges to 0 in $X$.*

*Proof.* Let $V$ be a neighborhood of 0 in $X$. By Lemma G.2.1, we can find $\alpha > 0$ such that $\lambda K \subset V$ for every $\lambda$ such that $|\lambda| < \alpha$. As the net $(\lambda_i)$ converges to 0, there exists $i_0 \in I$ such that $i \geq i_0$ implies $|\lambda_i| < \alpha$. Thus we have $\lambda_i x_i \in V$ for all $i \geq i_0$. This shows that the net $(\lambda_i x_i)$ converges to 0. □

The following lemma is the theorem of Markov-Kakutani in the particular case when the set $\mathcal{F}$ is reduced to a single element.

**Lemma G.2.3.** *Let $K$ be a nonempty convex compact subset of a Hausdorff topological vector space $X$ and let $f\colon K \to K$ be an affine continuous map. Then $f$ has a fixed point in $K$.*

*Proof.* Let us set $C = \{y - f(y) : y \in K\}$. The fact that $f$ admits a fixed point in $K$ is equivalent to the fact that $0 \in C$. Choose an arbitrary point $x \in K$ and consider the sequence $(x_n)_{n \geq 1}$ of points of $X$ defined by

$$x_n = \frac{1}{n} \sum_{k=0}^{n-1} (f^k(x) - f^{k+1}(x)).$$

We have $f^k(x) - f^{k+1}(x) = f^k(x) - f(f^k(x)) \in C$ for $0 \leq k \leq n-1$. On the other hand, the set $C$ is convex, since $K$ is convex and $f$ is affine. Thus $x_n \in C$ for every $n \geq 1$.

As

$$x_n = \frac{1}{n} x - \frac{1}{n} f^n(x).$$

and $f^n(x) \in K$ for every $n \geq 1$, it follows from Lemma G.2.2 that the sequence $(x_n)_{n \geq 1}$ converges to 0. The set $C$ is compact since it is the image of the compact set $K$ by the continuous map $y \mapsto y - f(y)$. As every compact subset of a Hausdorff space is closed, we deduce that $C$ is closed in $X$. Thus $0 \in C$. □

*Proof of Theorem G.1.1.* Let $f \in \mathcal{F}$ and consider the set

$$\mathrm{Fix}(f) = \{x \in K : f(x) = x\}$$

of its fixed points. The set $\mathrm{Fix}(f)$ is not empty by Lemma G.2.3 and it is compact since it is a closed subset of the compact set $K$. On the other hand, $\mathrm{Fix}(f)$ is convex since $K$ is convex and $f$ is affine. If $g \in \mathcal{F}$ and $x \in \mathrm{Fix}(f)$, then the fact that $f$ and $g$ commute implies that $g(x) \in \mathrm{Fix}(f)$ since

$$f(g(x)) = g(f(x)) = g(x).$$

Therefore we can apply Lemma G.2.3 to the restriction of $g$ to $\mathrm{Fix}(f)$. It follows that $g$ fixes a point in $\mathrm{Fix}(f)$, that is,

$$\mathrm{Fix}(f) \cap \mathrm{Fix}(g) \neq \varnothing.$$

By induction on $n$, we get

$$\mathrm{Fix}(f_1) \cap \mathrm{Fix}(f_2) \cap \cdots \cap \mathrm{Fix}(f_n) \neq \varnothing$$

for all $f_1, f_2, \ldots, f_n \in \mathcal{F}$. Since $K$ is compact, from the finite intersection property (see Sect. A.5) we deduce that

$$\bigcap_{f \in \mathcal{F}} \mathrm{Fix}(f) \neq \varnothing.$$

This shows that there is a point in $K$ which is fixed by all elements of $\mathcal{F}$. □

## Notes

The proof presented here is due to S. Kakutani [Kak] (see [Jac]). In the proof of A. Markov [Mar], the local convexity of $X$ is needed.

# Appendix H
# The Hall Harem Theorem

## H.1 Bipartite Graphs

A *bipartite graph* is a triple $\mathcal{G} = (X, Y, E)$, where $X$ and $Y$ are arbitrary sets, and $E$ is a subset of the Cartesian product $X \times Y$. The set $X$ (resp. $Y$) is called the set of *left* (resp. *right*) *vertices* and $E$ is called the set of *edges* of the bipartite graph $\mathcal{G}$ (see Fig. H.1).

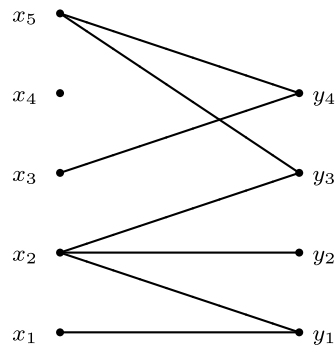



**Fig. H.1** The bipartite graph $\mathcal{G} = (X, Y, E)$ with $X = \{x_1, x_2, x_3, x_4, x_5\}$, $Y = \{y_1, y_2, y_3, y_4\}$ and $E = \{(x_1, y_1), (x_2, y_1), (x_2, y_2), (x_2, y_3), (x_3, y_4), (x_5, y_3), (x_5, y_4)\}$

A *bipartite subgraph* of a bipartite graph $\mathcal{G} = (X, Y, E)$ is a bipartite graph $\mathcal{G}' = (X', Y', E')$ with $X' \subset X$, $Y' \subset Y$ and $E' \subset E$ (see Fig. H.2).

Let $\mathcal{G} = (X, Y, E)$ be a bipartite graph.

Two edges $(x, y), (x', y') \in E$ are said to be *adjacent* if $x = x'$ or $y = y'$.

Given a vertex $x \in X$ (resp. $y \in Y$) the *right-neighborhood* of $x$ (resp. the *left-neighborhood* of $y$) is the subset $\mathcal{N}_R(x) \subset Y$ (resp. $\mathcal{N}_L(y) \subset X$) defined by (see Figs. H.3–H.4):

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_16, 

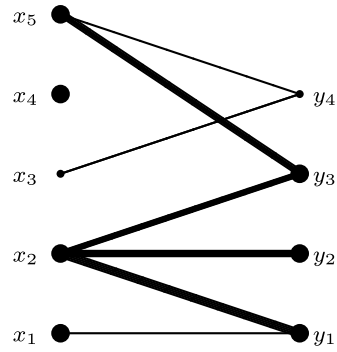



**Fig. H.2** The bipartite subgraph $\mathcal{G}' = (X', Y', E')$ of the bipartite graph $\mathcal{G}$ in Fig. H.1 with $X' = \{x_1, x_2, x_4, x_5\}$, $Y' = \{y_1, y_2, y_3\}$ and $E' = \{(x_2, y_1), (x_2, y_2), (x_2, y_3), (x_5, y_3)\}$

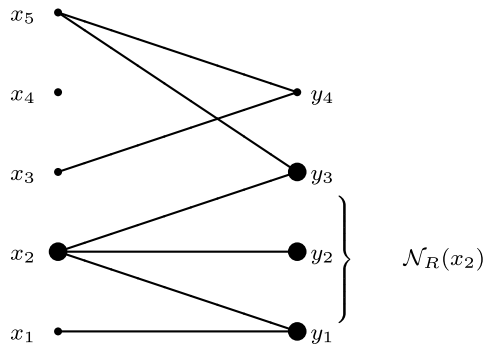



**Fig. H.3** The right-neighborhood $\mathcal{N}_R(x_2) \subset Y$ of the vertex $x_2 \in X$ in the bipartite graph $\mathcal{G}$ of Fig. H.1

$$\mathcal{N}_R(x) = \mathcal{N}_R^{\mathcal{G}}(x) = \{y \in Y : (x, y) \in E\}$$
$$(\text{resp. } \mathcal{N}_L(y) = \mathcal{N}_L^{\mathcal{G}}(y) = \{x \in X : (x, y) \in E\}).$$

For subsets $A \subset X$ and $B \subset Y$, we define the *right-neighborhood* $\mathcal{N}_R(A)$ of $A$ and the *left-neighborhood* $\mathcal{N}_L(B)$ of $B$ by

$$\mathcal{N}_R(A) = \mathcal{N}_R^{\mathcal{G}}(A) = \bigcup_{a \in A} \mathcal{N}_R(a) \quad \text{and} \quad \mathcal{N}_L(B) = \mathcal{N}_L^{\mathcal{G}}(B) = \bigcup_{b \in B} \mathcal{N}_L(b).$$

One says that the bipartite graph $\mathcal{G} = (X, Y, E)$ is *finite* if the sets $X$ and $Y$ are finite. One says that $\mathcal{G}$ is *locally finite* if the sets $\mathcal{N}_R(x)$ and $\mathcal{N}_R(y)$ are finite for all $x \in X$ and $y \in Y$. Note that if $\mathcal{G}$ is locally finite then the sets $\mathcal{N}_R(A)$ and $\mathcal{N}_L(B)$ are finite for all finite subsets $A \subset X$ and $B \subset Y$.

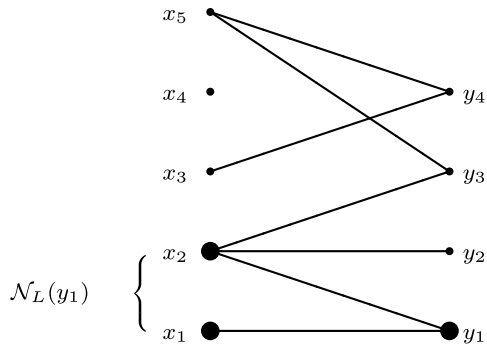



**Fig. H.4** The left-neighborhood $\mathcal{N}_L(y_1) \subset X$ of the vertex $y_1 \in Y$ in the bipartite graph $\mathcal{G}$ of Fig. H.1

## H.2 Matchings

Let $\mathcal{G} = (X, Y, E)$ be a bipartite graph.

A *matching* in $\mathcal{G}$ is a subset $M \subset E$ of pairwise nonadjacent edges. In other words, a subset $M \subset E$ is a matching if and only if both projection maps $p\colon M \to X$ and $q\colon M \to Y$ are injective.

A matching $M$ is called *left-perfect* (resp. *right-perfect*) if for each $x \in X$ (resp. $y \in Y$), there exists $y \in Y$ (resp. $x \in X$) such that $(x, y) \in M$ (see Fig. H.5). Thus, a matching $M$ is left-perfect (resp. right-perfect) if and only if the projection map $p\colon M \to X$ (resp. $q\colon M \to Y$) is surjective (and therefore bijective).

A matching $M$ is called *perfect* if it is both left-perfect and right-perfect.

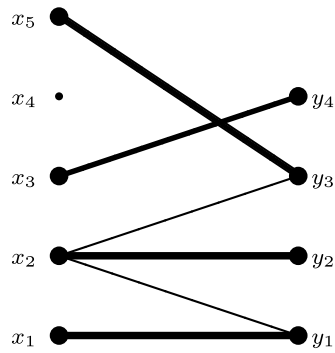



**Fig. H.5** A right-perfect matching $M \subset E$ in the bipartite graph $\mathcal{G}$ of Fig. H.1. Note that there is no left-perfect matching (and therefore no perfect matching) in $\mathcal{G}$ since $|X| > |Y|$

*Remarks H.2.1.* (a) A subset $M \subset E$ is a left-perfect (resp. right-perfect) matching if and only if there is an injective map $\varphi\colon X \to Y$ (resp. an injective

map $\psi\colon Y \to X$) such that $M = \{(x, \varphi(x)) : x \in X\}$ (resp. $M = \{(\psi(y), y) : y \in Y\}$).

(b) Similarly, a subset $M \subset E$ is a perfect matching if and only if there is a bijective map $\varphi\colon X \to Y$ such that $M = \{(x, \varphi(x)) : x \in X\}$.

Given a bipartite graph $\mathcal{G} = (X, Y, E)$, one often regards the set $X$ as a set of *boys* and $Y$ as a set of *girls*. One interprets $(x, y) \in E$ as the condition that $x$ and $y$ *know* each other. In this context, a matching $M \subset E$ is a process of getting boys and girls that know each other married (no polygamy is allowed here). The matching is left-perfect (resp. right-perfect) if and only if every boy (resp. girl) gets married. Finally, $M$ is perfect if and only if there remain no singles.

## H.3 The Hall Marriage Theorem

We use the notation $|\cdot|$ to denote cardinality of sets.

**Definition H.3.1 (Hall conditions).** Let $\mathcal{G} = (X, Y, E)$ be a bipartite graph. One says that $\mathcal{G}$ satisfies the *left* (resp. *right*) *Hall condition* if

$$|\mathcal{N}_R(A)| \geq |A| \tag{H.1}$$

(resp. $|\mathcal{N}_L(B)| \geq |B|$) for every finite subset $A \subset X$ (resp. for every finite subset $B \subset Y$).

One says that $\mathcal{G}$ satisfies the *Hall marriage conditions* if $\mathcal{G}$ satisfies both the left and the right Hall conditions.

**Theorem H.3.2.** *Let $\mathcal{G} = (X, Y, E)$ be a locally finite bipartite graph. Then the following conditions are equivalent.*

(a) *$\mathcal{G}$ satisfies the left (resp. right) Hall condition;*
(b) *$\mathcal{G}$ admits a left (resp. right) perfect matching.*

*Proof.* It is obvious that (b) implies (a). Let us prove that (a) implies (b). By symmetry, it suffices to show that if $\mathcal{G}$ satisfies the left Hall condition then it admits a left perfect matching.

We fist treat the case when the set $X$ is finite by induction on $n = |X|$. In the case $|X| = 1$, the statement is trivially satisfied. Suppose that we have proved the statement whenever $|X| \leq n - 1$ and let us prove it for $|X| = n$. We distinguish two cases.

Case (i): Suppose that

$$|\mathcal{N}_R(A)| \geq |A| + 1 \tag{H.2}$$

for all nonempty proper subsets $A \subset X$. Then fix $x_0 \in X$ and $y_0 \in Y$ such that $(x_0, y_0) \in E$. Consider the bipartite subgraph $\mathcal{G}' = (X', Y', E')$,

where $X' = X \setminus \{x_0\}$, $Y' = Y \setminus \{y_0\}$, and $E' = E \cap (X' \times Y')$. Then, for every nonempty subset $A' \subset X'$, we have $|\mathcal{N}_R(A')| \geq |A'| + 1$ by (H.2), and hence $|\mathcal{N}_R^{\mathcal{G}'}(A')| \geq |A'|$ since $\mathcal{N}_R^{\mathcal{G}'}(A') = \mathcal{N}_R(A') \setminus \{y_0\}$. As $|X'| = n - 1$, it follows from our induction hypothesis that $\mathcal{G}'$ admits a left-perfect matching $M' \subset E'$. Then $M = M' \cup \{(x_0, y_0)\}$ is a left-perfect matching for $\mathcal{G}$.

Case (ii): Suppose that we are not in Case (i). This means that there exists a nonempty proper subset $X' \subset X$ such that

$$|\mathcal{N}_R(X')| = |X'|. \tag{H.3}$$

Then the bipartite subgraph $\mathcal{G}' = (X', Y', E')$, where $Y' = \mathcal{N}_R(X')$ and $E' = E \cap (X' \times Y')$ clearly satisfies the left Hall condition. As $|X'| \leq n - 1$, there exists, by our induction hypothesis, a left-perfect matching $M' \subset E'$ for $\mathcal{G}'$. Consider now the bipartite subgraph $\mathcal{G}'' = (X'', Y'', E'')$, where $X'' = X \setminus X'$, $Y'' = Y \setminus Y'$, and $E'' = E \cap (X'' \times Y'')$. We claim that the left Hall condition also holds for $\mathcal{G}''$. Otherwise, there would be some subset $A'' \subset X''$ such that,

$$|\mathcal{N}_R^{\mathcal{G}''}(A'')| < |A''|, \tag{H.4}$$

and then the left Hall condition for $\mathcal{G}$ would be violated by $A = X' \cup A'' \subset X$ since

$$\begin{aligned}
|\mathcal{N}_R(A)| &= |\mathcal{N}_R(X' \cup A'')| \\
&= |\mathcal{N}_R(X') \cup \mathcal{N}_R(A'')| \\
&= |\mathcal{N}_R(X') \cup \mathcal{N}_R^{\mathcal{G}''}(A'')| \\
&\leq |\mathcal{N}_R(X')| + |\mathcal{N}_R^{\mathcal{G}''}(A'')| \\
&< |X'| + |A''| \quad \text{(by (H.3) and (H.4))} \\
&= |X' \cup A''| \\
&= |A|.
\end{aligned}$$

Therefore, as $|X''| < |X| = n$, induction applies again yielding a left-perfect matching $M'' \subset E''$ for $\mathcal{G}''$. It then follows that $M = M' \cup M''$ is a left-perfect matching for $\mathcal{G}$. This completes the proof that (a) implies (b) in the case when $X$ is finite.

To treat the general case, we shall apply the Tychonoff product theorem. Suppose that $\mathcal{G} = (X, Y, E)$ is a (possibly infinite) locally finite bipartite graph satisfying the left Hall condition, that is, $|\mathcal{N}_R(A)| \geq |A|$ for every finite subset $A \subset X$. Let us equip the Cartesian product $K = \prod_{x \in X} \mathcal{N}_R(x)$ with its prodiscrete topology, that is, with the product topology obtained by taking the discrete topology on each factor $\mathcal{N}_R(x)$. As each set $\mathcal{N}_R(x)$ is finite, the space $K$ is compact by the Tychonoff product theorem (Corollary A.5.3).

For each $x \in X$, let $\pi_x \colon K \to \mathcal{N}_R(x)$ denote the projection map. Let $\mathcal{F}$ be the set consisting of all nonempty finite subsets of $X$. Consider, for each $F \in \mathcal{F}$, the set $C(F) \subset K$ consisting of all $z \in K$ which satisfy $\pi_{x_1}(z) \neq \pi_{x_2}(z)$ for all distinct elements $x_1$ and $x_2$ in $F$. It follows from this definition and the continuity of the projection maps $\pi_x$, that $C(F)$ is an intersection of closed subsets of $X$ and hence closed in $K$. On the other hand, $C(F)$ is not empty. Indeed, choose, for each $x \in X$, an element $\psi(x) \in \mathcal{N}_R(x)$ (observe that $\mathcal{N}_R(x)$ is not empty by the left Hall condition). Also set $\mathcal{G}' = (X', Y', E')$ where $X' = F$, $Y' = \mathcal{N}_R(F)$ and $E' = E \cap (X' \times Y')$. Then, the left Hall condition for $\mathcal{G}$ implies the left Hall condition for the finite bipartite graph $\mathcal{G}'$ and therefore, by Theorem H.3.2, there exists a perfect matching for $\mathcal{G}'$. By Remark H.2.1, there exists an injective mapping $\varphi \colon F = X' \to Y' = \mathcal{N}_R(F)$. Then, the element $(\Phi(x))_{x \in X} \in K$, defined by $\Phi(x) = \varphi(x)$ if $x \in F$ and $\Phi(x) = \psi(x)$ if $x \in X \setminus F$, clearly belongs to $C(F)$.

Since

$$C(F_1) \cap C(F_2) \cap \cdots \cap C(F_n) \supset C(F_1 \cup F_2 \cup \cdots \cup F_n)$$

for all $F_1, F_2, \ldots, F_n \in \mathcal{F}$, we deduce that the family $\{C(F) : F \in \mathcal{F}\}$ of subsets of $K$ has the finite intersection property. By compactness of $K$, there is a point $z_0 \in \bigcap_{F \in \mathcal{F}} F$. Then $M = \{(x, \pi_x(z_0)) : x \in X\}$ is a left-perfect matching for $\mathcal{G}$. □

*Remark H.3.3.* The hypothesis of local finiteness for the bipartite graph can not be removed from the statement of the previous theorem. To see this, consider the bipartite graph $\mathcal{G} = (X, Y, E)$, where $X = Y = \mathbb{N}$ and $E = \{(0, n+1) : n \in \mathbb{N}\} \cup \{(n+1, n) : n \in \mathbb{N}\}$. Observe that $\mathcal{G}$ is not locally finite as $|\mathcal{N}_R(0)| = \infty$. Also, it satisfies the left Hall condition. Indeed, given a finite subset $A \subset X$, the set $\mathcal{N}_R(A)$ is infinite if $0 \in A$, while $|\mathcal{N}_R(A)| = |\{n - 1 : n \in A\}| = |A|$ if $0 \notin A$. However, $\mathcal{G}$ admits no left perfect matching. Indeed, for any matching $M \subset E$, either $0 \in X$ remains unmatched, or, if $(0, n) \in M$ for some $n \in \mathbb{N}$, then $n + 1 \in X$ remains unmatched (see Fig. H.6).

**Theorem H.3.4.** *Let $\mathcal{G} = (X, Y, E)$ be a bipartite graph. Suppose that $\mathcal{G}$ admits both a left perfect matching and a right perfect matching. Then $\mathcal{G}$ admits a perfect matching.*

*Proof.* Let $M_X$ (resp. $M_Y$) be a left perfect (resp. right perfect) matching for $\mathcal{G}$. Consider the equivalence relation in $\overline{M} = M_X \cup M_Y$ defined by declaring two edges $e$ and $e'$ in relation if there exists a finite sequence $e = e_0, e_1, \ldots, e_n = e'$ in $\overline{M}$ such that $e_i$ and $e_{i+1}$ are adjacent for all $0 \leq i \leq n - 1$. Then, each equivalence class consists of either (see Fig. H.7):

– a single edge, or
– a cycle of even length $2n \geq 4$, or
– an infinite chain which can be bi-infinite, or infinite only in one direction.

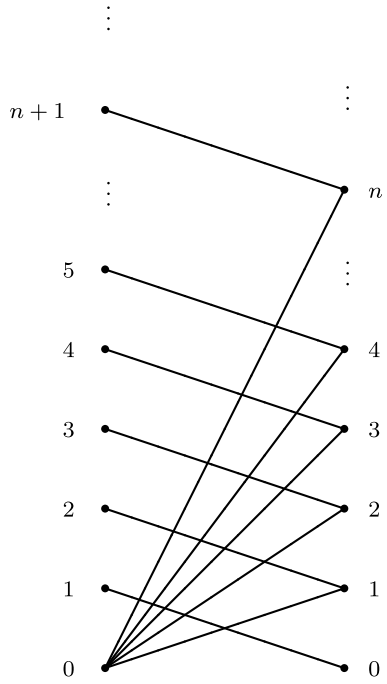



**Fig. H.6** The bipartite graph $\mathcal{G} = (X, Y, E)$, where $X = Y = \mathbb{N}$ and $E = \{(0, n+1) : n \in \mathbb{N}\} \cup \{(n+1, n) : n \in \mathbb{N}\}$

Note that an equivalence class is reduced to a single edge $(x, y)$ if and only if $(x, y) \in M_X \cap M_Y$.

On the other hand, an equivalence class is a cycle of length $2n$ if and only if it is of the form

$$\begin{aligned}\mathcal{C} =&\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\} \cup \{(x_1, y_2), (x_2, y_3),\\ &\ldots, (x_{n-1}, y_n), (x_n, y_1)\}\end{aligned}$$

with $x_i \in X$, all distinct, and $y_j \in Y$, all distinct, $1 \leq i, j \leq n$. In this case, we then set

$$M(\mathcal{C}) = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}.$$

Also, a bi-infinite chain is an equivalence class of the form

$$\mathcal{C} = \{(x_n, y_n) : n \in \mathbb{Z}\} \cup \{(x_n, y_{n+1}) : n \in \mathbb{Z}\}$$

with $x_i \in X$, all distinct, and $y_j \in Y$, all distinct, $i, j \in \mathbb{Z}$. In this case, we then set

$$M(\mathcal{C}) = \{(x_n, y_n) : n \in \mathbb{Z}\}.$$

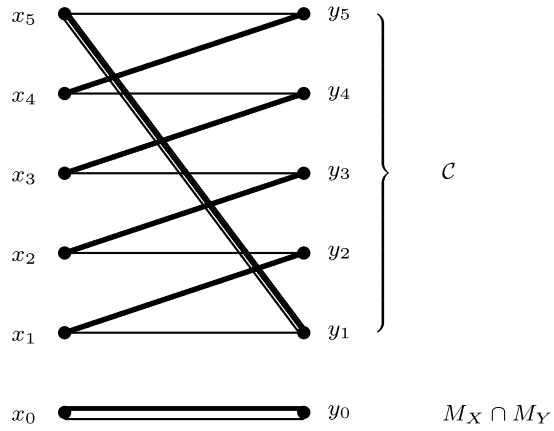



**Fig. H.7** In the bipartite graph $\mathcal{G} = (X, Y, E)$, where $X = \{x_0, x_1, x_2, x_3, x_4, x_5\}$, $Y = \{y_0, y_1, y_2, y_3, y_4, y_5\}$ and $E = \{(x_i, y_i) : i = 0, 1, \ldots, 5\} \cup \{(x_i, y_{i+1}) : i = 1, 2, \ldots, 4\} \cup \{(x_5, y_1)\}$, we have a left-perfect matching $M_X = \{(x_i, y_i) : i = 0, 1, \ldots, 5\}$ (which is indeed perfect) and a right-perfect matching $M_Y = \{(x_0, y_0)\} \cup \{(x_i, y_{i+1}) : i = 1, 2, \ldots, 4\} \cup \{(x_5, y_1)\}$ (which is also perfect). There are two equivalence classes in $\overline{M} = M_X \cup M_Y$, namely, $M_X \cap M_Y = \{(x_0, y_0)\}$, which consists of a single edge, and $\mathcal{C} = \{(x_i, y_i) : i = 1, 2 \ldots, 5\} \cup \{(x_i, y_{i+1}) : i = 1, 2, \ldots, 4\} \cup \{(x_5, y_1)\}$, which is a cycle of length 10

Finally, if the equivalence class is an infinite chain, which is infinite only in one direction, then it is either of the form

$$\mathcal{C} = \{(x_n, y_n) : n \in \mathbb{N}\} \cup \{(x_n, y_{n+1}) : n \in \mathbb{N}\}$$

or

$$\mathcal{C} = \{(x_n, y_n) : n \in \mathbb{N}\} \cup \{(x_{n+1}, y_n) : n \in \mathbb{N}\}$$

with $x_i \in X$, all distinct, and $y_j \in Y$, all distinct, $i, j \in \mathbb{N}$. In both cases, we then set

$$M(\mathcal{C}) = \{(x_n, y_n) : n \in \mathbb{N}\}.$$

It is then clear that the set $M \subset E$ defined by

$$M = \bigcup_{\mathcal{C}} M(\mathcal{C}),$$

where $\mathcal{C}$ runs over all equivalence classes in $\overline{M}$, is the required perfect matching for $\mathcal{G}$. □

**Corollary H.3.5 (Cantor–Bernstein Theorem).** *Let $X$ and $Y$ be two sets. Suppose that there exist injective maps $f\colon X \to Y$ and $g\colon Y \to X$. Then there exists a bijective map $h\colon X \to Y$.*

*Proof.* Consider the bipartite graph $\mathcal{G} = (X, Y, E)$, where $E = \{(x, f(x)) : x \in X\} \cup \{(g(y), y) : y \in Y\}$. Now, $M_X = \{(x, f(x)) : x \in X\}$ and

$M_Y = \{(g(y), y) : y \in Y\}$ are left perfect and right perfect matchings in $\mathcal{G}$, respectively (cf. Remark H.2.1(a)). By the previous theorem, there exists a perfect matching $M \subset E$ for $\mathcal{G}$. Then $M = \{(x, h(x)) : x \in X\}$, where $h\colon X \to Y$ is bijective (cf. Remark H.2.1(b)). □

**Theorem H.3.6 (The Hall marriage Theorem).** *Let $\mathcal{G} = (X, Y, E)$ be a locally finite bipartite graph. Then the following conditions are equivalent:*

(a) *$\mathcal{G}$ satisfies the Hall-marriage conditions;*
(b) *$\mathcal{G}$ admits a perfect matching.*

*Proof.* The left (resp. right) Hall condition implies, by Theorem H.3.2, the existence of a left- (resp. right-) perfect matching for $\mathcal{G}$. Then, Theorem H.3.4 guarantees the existence of a perfect matching for $\mathcal{G}$. The converse implication is trivial. □

## H.4 The Hall Harem Theorem

Let $\mathcal{G} = (X, Y, E)$ be a bipartite graph and let $k \geq 1$ be an integer.

A subset $M \subset E$ is called a *perfect $(1, k)$-matching* if it satisfies the following conditions: (1) for each $x \in X$, there are exactly $k$ elements in $y \in Y$ such that $(x, y) \in M$, (2) for each $y \in Y$, there is a unique element $x \in X$ such that $(x, y) \in M$ (see Fig. H.8). Thus, a subset $M \subset E$ is a perfect $(1, k)$-matching if and only if there exists a $k$-to-one surjective map $\psi\colon Y \to X$ such that $M = \{(\psi(y), y) : y \in Y\}$ (recall that a surjective map $f\colon S \to T$ from a set $S$ onto a set $T$ is said to be *$k$-to-one* if each element in $T$ has exactly $k$ preimages in $S$). Note that when $k = 1$, a perfect $(1, k)$-matching is the same thing as a perfect matching.

In the language of boys, girls, and marriages, a perfect $(1, k)$-matching is a process for marrying each boy with exactly $k$ girls (among the girls he knows) in such a way that each girl is married with exactly one boy (among the boys she knows). The girls that are married with a given boy constitute his *harem*.

**Definition H.4.1.** Let $\mathcal{G} = (X, Y, E)$ be a locally finite bipartite graph and let $k \geq 1$ be an integer. One says that $\mathcal{G}$ satisfies the *Hall $k$-harem conditions* if

$$\begin{aligned} &|\mathcal{N}_R(A)| \geq k|A| \quad \text{and} \\ &|\mathcal{N}_L(B)| \geq \tfrac{1}{k}|B| \quad \text{for all finite subsets } A \subset X \text{ and } B \subset Y. \end{aligned} \tag{H.5}$$

**Theorem H.4.2 (The Hall harem Theorem).** *Let $\mathcal{G} = (X, Y, E)$ be a locally finite bipartite graph and let $k \geq 1$ be an integer. Then, the following conditions are equivalent.*

(a) *$\mathcal{G}$ satisfies the Hall $k$-harem conditions;*
(b) *$\mathcal{G}$ admits a perfect $(1, k)$-matching.*

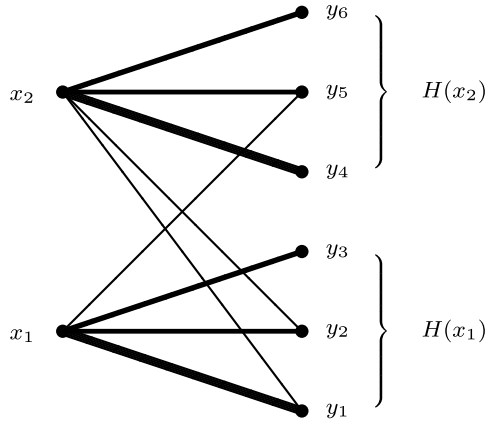



**Fig. H.8** The bipartite graph $\mathcal{G} = (X, Y, E)$, where $X = \{x_1, x_2\}$, $Y = \{y_1, y_2, y_3, y_4\}$ and $E = \{(x_1, y_i) : i = 1, 2, 3, 5\} \cup \{(x_2, y_j) : j = 1, 2, 4, 5, 6\}$, with a perfect $(1, 3)$-matching $M = \{(x_1, y_i) : i = 1, 2, 3\} \cup \{(x_2, y_j) : j = 4, 5, 6\}$ and the harems $H(x_1)$ and $H(x_2)$ of $x_1$ and $x_2$ respectively

*Proof.* The fact that (b) implies (a) is trivial.

Let us prove that (a) implies (b). Suppose that the Hall $k$-harem conditions are satisfied by $\mathcal{G}$. Let $X_1, X_2, \ldots, X_k$ be disjoint copies of $X$ and let $\phi_i : X \to X_i$, $i = 1, 2, \ldots, k$, denote the copy maps. It will be helpful to think of $\phi_i(x)$ as the *clone* of $x \in X$ in $X_i$.

Consider the new bipartite graph $\mathcal{G}' = (X', Y', E')$, where $X' = \coprod_{i=1}^{k} X_i$, $Y' = Y$, and

$$E' = \{(\phi_i(x), y) : (x, y) \in E, i = 1, 2, \ldots, k\} \subset X' \times Y'.$$

Let $A'$ be a finite subset of $X'$ and denote by $\overline{A'}$ the set of $x \in X$ such that some clone of $x$ belongs to $A'$. Observe that $|A'| \leq k|\overline{A'}|$ and $\mathcal{N}_R^{\mathcal{G}'}(A') = \mathcal{N}_R^{\mathcal{G}}(\overline{A'})$. Thus, using (a), we get

$$|\mathcal{N}_R^{\mathcal{G}'}(A')| = |\mathcal{N}_R^{\mathcal{G}}(\overline{A'})| \geq k|\overline{A'}| \geq |A'|. \tag{H.6}$$

On the other hand, if $B'$ is a finite subset of $Y' = Y$, then $\mathcal{N}_L^{\mathcal{G}'}(B')$ is the set consisting of all clones of elements of $\mathcal{N}_L^{\mathcal{G}}(B')$ so that

$$|\mathcal{N}_L^{\mathcal{G}'}(B')| = k|\mathcal{N}_L^{\mathcal{G}}(B')| \geq k\left(\frac{1}{k}|B'|\right) = |B'| \tag{H.7}$$

by (a). Inequalities H.6 and H.7 say that $\mathcal{G}'$ satisfies the Hall marriage conditions. Therefore, $\mathcal{G}'$ admits a perfect matching $M' \subset E'$ by Theorem H.3.6. Then the set $M$, consisting of all pairs $(x, y) \in E$ such that $(x', y) \in M'$ for some clone $x'$ of $x$, is clearly a perfect $(1, k)$-matching for $\mathcal{G}$. □

## Notes

The Hall marriage theorem was first established for finite bipartite graphs by P. Hall [Hall1] and then extended to infinite locally finite bipartite graphs by M. Hall [Hall-M-1]. The proof of Theorem H.3.2 which is given in this appendix is based on [Halm].

# Appendix I
# Complements of Functional Analysis

All vector spaces considered in this appendix are vector spaces over the field $\mathbb{R}$ of real numbers.

## I.1 The Baire Theorem

Let $(X, d)$ be a metric space. Given $x \in X$ and $r > 0$, we denote by

$$B_X(x, r) = \{y \in X : d(x, y) \leq r\}$$

the closed ball of radius $r$ centered at $x$ and by

$$OB_X(x, r) = \{y \in X : d(x, y) < r\}$$

the open ball of radius $r$ centered at $x$. For a subset $A \subset X$ we denote by $\overline{A}$ (resp. $\operatorname{Int} A$) the closure (resp. the interior) of $A$.

**Theorem I.1.1 (Baire's Theorem).** *Let $(X, d)$ be a complete metric space. Let $(X_n)_{n\geq 1}$ be a sequence of closed subsets such that*

$$\operatorname{Int} X_n = \varnothing \tag{I.1}$$

*for all $n \geq 1$. Then*

$$\operatorname{Int}\left(\bigcup_{n\geq 1} X_n\right) = \varnothing. \tag{I.2}$$

*Proof.* Let us set $A_n = X \setminus X_n$, so that $A_n$ is an open dense subset of $X$. Let us show that $\bigcap_{n\geq 1} A_n$ is dense in $X$. Let $x_0 \in X$ and $r_0 > 0$. We have to show that

$$B_X(x_0, r_0) \bigcap \left(\bigcap_{n\geq 1} A_n\right) \neq \varnothing. \tag{I.3}$$

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_17, 

As $A_1$ is dense and open we can find $x_1 \in B_X(x_0, r_0) \bigcap A_1$ and $r_1 > 0$ such that

$$\begin{cases} B_X(x_1, r_1) \subset OB_X(x_0, r_0) \bigcap A_1 \\ 0 < r_1 < \frac{r_0}{2}. \end{cases}$$

Continuing this way, we produce by induction a sequence $(x_n)_{n\geq 1}$ in $X$ and a sequence $(r_n)_{n\geq 1}$ in $\mathbb{R}$ such that

$$\begin{cases} B_X(x_{n+1}, r_{n+1}) \subset OB_X(x_n, r_n) \bigcap A_{n+1} \\ 0 < r_{n+1} < \frac{r_n}{2}. \end{cases} \tag{I.4}$$

Given $0 \leq n \leq m$ we have

$$x_m \in B_X(x_n, r_n) \tag{I.5}$$

so that $d(x_n, x_m) < \frac{r_0}{2^n}$. We deduce that $(x_n)_{n\geq 1}$ is a Cauchy sequence. As $X$ is complete, $(x_n)_{n\geq 1}$ is convergent so that there exists $x \in X$ such that $\lim_{n\to\infty} x_n = x$. Passing to the limit (for $m \to \infty$) in (I.5) we obtain $x \in B_X(x_n, r_n) \subset B_X(x_0, r_0)$ and, by (I.4), $x \in B_X(x_n, r_n) \subset A_n$ for all $n \geq 1$. As a consequence, $x \in B_{x_0,r_0} \bigcap (\bigcap_{n\geq 1} A_n)$ and (I.3) follows. This shows that $\bigcap_{n\geq 1} A_n$ is dense in $X$. Taking complements, we deduce that $\bigcup_{n\geq 1} X_n = \bigcup_{n\geq 1} (X \setminus A_n) = X \setminus \bigcap_{n\geq 1} A_n$ has empty interior, and (I.2) follows. □

*Remark I.1.2.* We can restate Baire's theorem in the following two ways.

(i) Let $(X, d)$ be a nonempty complete metric space. Let $(X_n)_{n\geq 1}$ be a sequence of closed subsets such that $\bigcup_{n\geq 1} X_n = X$. Then there exists $n_0$ such that $\operatorname{Int} X_{n_0} \neq \varnothing$.

(ii) (By passing to complements) Let $(X, d)$ be a nonempty complete metric space. Let $(\Omega_n)_{n\geq 1}$ be a sequence of open dense subsets in $X$. Then $\bigcap_{n\geq 1} \Omega_n$ is dense in $X$.

## I.2 The Open Mapping Theorem

Recall that if $X$ and $Y$ are two topological spaces, then a map $T\colon X \to Y$ is said to be *open* if for any open set $A$ in $X$ the image $T(A)$ is open in $Y$.

**Theorem I.2.1 (Open mapping theorem).** *Let $X$ and $Y$ be two Banach spaces. Then every surjective continuous linear map $T\colon X \to Y$ is open.*

*Proof.* Let $T\colon X \to Y$ be a surjective continuous linear map. We need to show that for every $x \in X$ and every neighborhood $N$ of $x$, the image $T(N)$ is a neighborhood of $T(x)$. Since, by linearity, $T(x + N) = T(x) + T(N)$, we can reduce to the case when $x = 0$. Moreover, since neighborhoods contain open balls, it is sufficient to show that for every $r' > 0$ there exists $r'' > 0$ such that

$T(OB_X(0,r')) \supset OB_Y(0,r'')$. Finally, since $T(OB_X(0,r')) = r'T(OB_X(0,1))$ and $OB_Y(0,r'') = r''OB_Y(0,1)$, we are only left to show that there exists $r > 0$ such that

$$T(OB_X(0,1)) \supset OB_Y(0,r). \tag{I.6}$$

For all $n \geq 1$ set $Y_n = n\overline{T(OB_X(0,1))} = \overline{T(OB_X(0,n))}$. As $T$ is surjective we have $\bigcup_{n\geq 1} Y_n = Y$. It follows from Theorem I.1.1 that there exists $n_0 \geq 1$ such that $\operatorname{Int} Y_{n_0} \neq \varnothing$. Since $\operatorname{Int}(n_0\overline{T(OB_X(0,1))}) = n_0 \operatorname{Int}(\overline{T(OB_X(0,1))})$, we deduce that $\operatorname{Int}(\overline{T(OB_X(0,1))}) \neq \varnothing$. Thus we can find $r > 0$ and $y \in Y$ such that

$$OB_Y(y,4r) \subset \overline{T(OB_X(0,1))}. \tag{I.7}$$

It follows that $y \in \overline{T(OB_X(0,1))}$ and, by symmetry,

$$-y \in \overline{T(OB_X(0,1))}. \tag{I.8}$$

Summing (I.7) and (I.8) we obtain

$$\begin{aligned} 2OB_Y(0,2r) &= OB_Y(0,4r) \\ &= OB_Y(y,4r) - y \\ &\subset \overline{T(OB_X(0,1))} + \overline{T(OB_X(0,1))} \\ &\subset 2\overline{T(OB_X(0,1))}. \end{aligned}$$

We deduce that $OB_Y(0,2r) \subset \overline{T(OB_X(0,1))}$ so that, for all $n \geq 0$,

$$OB_Y\left(0, \frac{r}{2^n}\right) \subset \overline{T(OB_X(0,1/2^{n+1}))}. \tag{I.9}$$

Let $y_0 \in OB_Y(0,r)$ and let us show that there exists $x' \in OB_X(0,1)$ such that $y_0 = T(x')$. We deduce from (I.9) (with $n = 0$) that there exists $x_0 \in OB_X(0,\frac{1}{2})$ such that $\|y_0 - T(x_0)\| < \frac{r}{2}$. Set $y_1 = y_0 - T(x_0)$ and observe that $y_1 \in OB_Y(0,\frac{r}{2})$. Continuing this way, we find a sequence $(y_n)_{n\geq 1}$ in $Y$ and a sequence $(x_n)_{n\geq 1}$ in $X$ such that

$$y_{n+1} = y_n - T(x_n) \in OB_Y\left(0, \frac{r}{2^{n+1}}\right), \tag{I.10}$$

$$x_n \in OB_X\left(0, \frac{1}{2^{n+1}}\right) \tag{I.11}$$

and

$$\|y_n - T(x_n)\| < \frac{r}{2^{n+1}}. \tag{I.12}$$

Now, (I.11) is equivalent to $\|x_n\| < \frac{1}{2^{n+1}}$ so that setting $x'_n = x_0 + x_1 + \cdots + x_n$, the sequence $(x'_n)_{n\geq 1}$ is convergent in $X$. It follows that there exists $x' \in X$ such that $\lim_{n\to\infty} x'_n = x'$ and, moreover $\|x'\| < 1$. On the other hand,

(I.10) and (I.12) give $\|y_0 - T(x'_n)\| = \|y_0 - T(x_0 + x_1 + \cdots + x_n)\| = \|y_1 - T(x_1 + x_2 + \cdots + x_n)\| = \cdots = \|y_n - T(x_n)\| \leq \frac{r}{2^{n+1}}$ so that, passing to the limit, we deduce that $T(x') = y_0$, since $T$ is continuous. This shows that $OB_Y(0, r) \subset T(OB_X(0, 1))$. It follows that $T$ is an open map. □

**Corollary I.2.2.** *Let $X$ and $Y$ be two Banach spaces. Let $T\colon X \to Y$ be a bijective continuous linear map. Then the inverse linear map $T^{-1}\colon Y \to X$ is continuous.*

*Proof.* This follows immediately from Theorem I.2.1 since the inverse map $T^{-1}\colon Y \to X$ is continuous if and only if $T\colon X \to Y$ is an open map. □

**Corollary I.2.3.** *Let $X$ and $Y$ be two Banach spaces. Let $T\colon X \to Y$ be a continuous linear map. Suppose that for every $\varepsilon > 0$ there exists $x \in X$ such that $\|x\| = 1$ and $\|T(x)\| < \varepsilon$. Then $T$ is not bijective.*

*Proof.* Suppose by contradiction that $T$ is bijective. Then, by Corollary I.2.2, the inverse linear map $T^{-1}\colon Y \to X$ is continuous. Thus we can find a constant $M > 0$ such that $\|T^{-1}(y)\| \leq M\|y\|$ for all $y \in Y$. Since $T$ is bijective, this is equivalent to $\|x\| \leq M\|T(x)\|$ for all $x \in X$. This clearly contradicts the hypotheses. Thus, $T$ is not bijective. □

## I.3 Spectra of Linear Maps

Let $(X, \|\cdot\|)$ be a Banach space. We denote by $\mathcal{L}(X)$ the space of all continuous linear maps $T\colon X \to X$ endowed with the norm

$$\|T\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{\|T(x)\|}{\|x\|}.$$

We denote by $\mathrm{Id}_X\colon X \to X$ the identity map.

**Definition I.3.1.** Let $T \in \mathcal{L}(X)$. The set

$$\sigma(T) = \{\lambda \in \mathbb{R} : (T - \lambda \,\mathrm{Id}_X) \text{ is not bijective}\} \tag{I.13}$$

is called the *real spectrum* of $T$.

**Proposition I.3.2.** *Let $T \in \mathcal{L}(X)$. Then the spectrum $\sigma(T)$ is a compact set and*

$$\sigma(T) \subset [-\|T\|, \|T\|]. \tag{I.14}$$

*Proof.* Let $\lambda \in \mathbb{R}$ and suppose that $|\lambda| > \|T\|$. For $y \in X$ the equation

$$(T - \lambda \,\mathrm{Id}_X)(x) = y \tag{I.15}$$

admits a unique solution $x \in X$. Indeed (I.15) is equivalent to

$$x = \frac{1}{\lambda}(T(x) - y). \tag{I.16}$$

Moreover, $\|\frac{1}{\lambda}(T(x_1)-y)-\frac{1}{\lambda}(T(x_2)-y)\| \leq \frac{1}{|\lambda|}\|T\|\cdot\|x_1-x_2\|$ for all $x_1, x_2 \in X$ and $\frac{1}{|\lambda|}\|T\| < 1$. It then follows from the Banach fixed point theorem that there exists a unique $x \in X$ satisfying (I.16). It follows that $T - \lambda \operatorname{Id}_X$ is bijective and therefore $\lambda \notin \sigma(T)$. This shows (I.14).

Let us show that $\mathbb{R} \setminus \sigma(T)$ is open. Let $\lambda_0 \in \mathbb{R} \setminus \sigma(T)$ so that $T - \lambda_0 \operatorname{Id}_X$ is bijective. By Corollary I.2.2, we have that the inverse map $(T - \lambda_0 \operatorname{Id}_X)^{-1}$ is also continuous so that $0 \neq \|(T - \lambda_0 \operatorname{Id}_X)^{-1}\| < \infty$. Let $\lambda \in \mathbb{R}$ such that $|\lambda - \lambda_0| < \frac{1}{\|(T-\lambda_0 \operatorname{Id}_X)^{-1}\|}$. For $y \in X$ we have that the linear equation (I.15) can be written as $T(x) - \lambda_0 x = y + (\lambda - \lambda_0)x$, that is,

$$x = (T - \lambda_0 \operatorname{Id}_X)^{-1}[y + (\lambda - \lambda_0)x]. \tag{I.17}$$

By applying again the Banach fixed point theorem, we deduce that there exists a unique $x \in X$ satisfying (I.17). This shows that $T-\lambda \operatorname{Id}_X$ is bijective, that is, $\lambda \in \mathbb{R}\setminus\sigma(T)$. It follows that $\mathbb{R}\setminus\sigma(T)$ is open. Therefore $\sigma(T)$ is closed. □

## I.4 Uniform Convexity

**Definition I.4.1.** A normed space $(X, \|\cdot\|)$ is said to be *uniformly convex* if, for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\|x - y\| > \varepsilon \quad \text{implies} \quad \left\|\frac{x+y}{2}\right\| < 1 - \delta$$

for all $x, y \in X$ with $\|x\|, \|y\| \leq 1$.

Let $Z$ be a nonempty set not reduced to a single point. Then the Banach space $\ell^1(Z)$ is not uniformly convex. For instance, if $z, z' \in Z$ are distinct, setting $x = \delta_z$ and $y = \delta_{z'}$, one has $\|x\|_1 = \|y\|_1 = 1$, $\|x - y\|_1 = 2$, but $\|\frac{x+y}{2}\|_1 = 1$.

On the other hand, we have the following:

**Proposition I.4.2.** *Let $X$ be a vector space equipped with a scalar product $\langle\cdot,\cdot\rangle$ and denote by $\|\cdot\|$ the associated norm. Then the normed space $(X, \|\cdot\|)$ is uniformly convex. In particular, every Hilbert space is uniformly convex.*

*Proof.* Let $x, y \in X$ such that $\|x\|, \|y\| \leq 1$. Then, we have

$$\begin{aligned}\left\|\frac{x+y}{2}\right\|^2 + \frac{1}{4}\|x-y\|^2 &= \frac{1}{4}\left(\langle x+y, x+y\rangle + \langle x-y, x-y\rangle\right)\\ &= \frac{1}{4}\left(\langle x,x\rangle + \langle y,y\rangle + 2\langle x,y\rangle + \langle x,x\rangle + \langle y,y\rangle - 2\langle x,y\rangle\right)\\ &= \frac{1}{4}\left(2\|x\|^2 + 2\|y\|^2\right)\\ &\leq 1.\end{aligned}$$

Therefore,

$$\left\|\frac{x+y}{2}\right\|^2 \leq 1 - \frac{1}{4}\|x-y\|^2. \tag{I.18}$$

Let now $\varepsilon > 0$ and set $\delta = 1 - \frac{1}{2}\sqrt{4-\varepsilon^2}$. Suppose that $\|x-y\| > \varepsilon$. This implies $1 - \frac{1}{4}\|x-y\|^2 < 1 - \frac{\varepsilon^2}{4} = (1-\delta)^2$. From (I.18) we then deduce that $\|\frac{x+y}{2}\| < 1-\delta$. This shows that $X$ is uniformly convex. □

# Appendix J
# Ultrafilters

## J.1 Filters and Ultrafilters

Let $X$ be a set. We denote by $\mathcal{P}(X)$ the set of all subsets of $X$.

**Definition J.1.1.** A *filter* on $X$ is a nonempty set $\mathcal{F} \subset \mathcal{P}(X)$ satisfying the following conditions:

(F-1) $\varnothing \notin \mathcal{F}$;
(F-2) if $A \in \mathcal{F}$ and $A \subset B \subset X$, then $B \in \mathcal{F}$;
(F-3) if $A \in \mathcal{F}$ and $B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$.

*Examples J.1.2.* (a) Let $X$ be a set and let $A_0$ be a nonempty subset of $X$. Then the set $\{A \in \mathcal{P}(X) : A_0 \subset A\}$ is a filter on $X$. A filter $\mathcal{F}$ on $X$ is said to be a *principal filter* if there exists a nonempty subset $A_0$ of $X$ such that $\mathcal{F} = \{A \in \mathcal{P}(X) : A_0 \subset A\}$. One then says that $\mathcal{F}$ is the principal filter *based* on $A_0$.

(b) Let $X$ be a set and $\Omega \subset \mathcal{P}(X)$. Suppose that $\varnothing \notin \Omega$ and that, given $A'_1, A'_2 \in \Omega$, there exists $A' \in \Omega$ such that $A' \subset A'_1 \cap A'_2$. Then the set

$$\mathcal{F}(\Omega) = \{A \in \mathcal{P}(X) : \text{ there exists } A' \in \Omega \text{ such that } A' \subset A\} \tag{J.1}$$

is a filter on $X$ and one has $\Omega \subset \mathcal{F}(\Omega)$. The filter $\mathcal{F}(\Omega)$ is called the filter *generated* by $\Omega$.

(c) Let $(I, \leq)$ be a nonempty directed set. A subset $A \subset I$ is called *residual* in $I$ if there exists $i \in I$ such that $A \supset \{j \in I : i \leq j\}$. Clearly, the set $\mathcal{F}_r(I)$ of all residual subsets of $I$ is a filter on $I$. It is the filter generated by the sets $\{j \in I : i \leq j\}$, $i \in I$. The filter $\mathcal{F}_r(I)$ is called the *residual filter* on $I$.

(d) Let $X$ be an infinite set. Then the set $\mathcal{F} = \{A \in \mathcal{P}(X) : X \setminus A \text{ is finite}\}$ is a filter. It is called the *Fréchet filter* on $X$. If we consider the directed set $(\mathbb{N}, \leq)$, a subset $A \subset \mathbb{N}$ is residual if and only if its complement $\mathbb{N} \setminus A$ is finite. Therefore, the residual filter on $\mathbb{N}$ equals the Fréchet filter on $\mathbb{N}$.

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1_18, 

(e) Let $X$ be a topological space and let $x$ be a point in $X$. Then, the set $\mathcal{F}_x$ of all neighborhoods of $x$ is a filter on $X$.

(f) Let $X$ be a nonempty uniform space (cf. Appendix B). Then the set $\mathcal{U} \subset \mathcal{P}(X \times X)$ of all entourages of $X$ is a filter on $X \times X$.

Note that for a filter $\mathcal{F} \subset \mathcal{P}(X)$ the following also holds true:

(F-4) if $A \subset X$, then $A$ and $X \setminus A$ cannot both belong to $\mathcal{F}$;
(F-5) if $A_1, A_2, \ldots, A_n \in \mathcal{F}$, $n \geq 1$, then $\bigcap_{i=1}^{n} A_i \in \mathcal{F}$;
(F-6) $X \in \mathcal{F}$.

Indeed, (F-4) follows from (F-1) and (F-3) with $B = X \setminus A$. From (F-3) we deduce by induction condition (F-5). Finally, (F-6) follows from the fact that $\mathcal{F} \neq \varnothing$ and (F-2).

**Proposition J.1.3.** *Let $X$ be a set and $\Omega_0 \subset \mathcal{P}(X)$. Then, there exists a filter on $X$ containing $\Omega_0$ if an only if $\Omega_0$ has the finite intersection property, that is, every finite family of elements in $\Omega_0$ has nonempty intersection.*

*Proof.* Let $\Omega_0 \subset \mathcal{P}(X)$ be a set and suppose that there exists a filter $\mathcal{F}$ such that $\Omega_0 \subset \mathcal{F}$. By (F-1) and (F-5) we have that $\Omega_0$ has the finite intersection property. Conversely, suppose that $\Omega_0$ has the finite intersection property. Consider the set $\Omega$ consisting of all finite intersections of elements of $\Omega_0$. Then, $\Omega$ satisfies the conditions in Example J.1.2(b), and therefore the filter generated by $\Omega$ is a filter containing $\Omega_0$. □

**Definition J.1.4.** A filter $\omega \subset \mathcal{P}(X)$ is called an *ultrafilter* on $X$ if it satisfies the condition

(UF) if $A \subset X$, then $A \in \omega$ or $(X \setminus A) \in \omega$.

*Example J.1.5.* Let $X$ be a set and $x \in X$. Then the principal filter based on $\{x\}$ is an ultrafilter. An ultrafilter $\omega$ on $X$ is called *principal* if there exists $x \in X$ such that $\omega$ is the principal filter based on $\{x\}$. Note that a principal filter is an ultrafilter if and only if it is based on a singleton. If $X$ is finite, then every ultrafilter is, clearly, principal.

An ultrafilter which is not principal is called *non-principal* (or *free*).

**Theorem J.1.6.** *Let $X$ be a nonempty set. Let $\mathcal{F}_0$ be a filter on $X$. Then there exists an ultrafilter $\omega$ on $X$ containing $\mathcal{F}_0$.*

Let us first prove the following:

**Lemma J.1.7.** *Let $X$ be a set. Let $\mathcal{F}$ be a filter on $X$. Suppose there exists a set $A_0 \in \mathcal{P}(X)$ such that $A_0 \notin \mathcal{F}$ and $(X \setminus A_0) \notin \mathcal{F}$. Then there exists a filter $\mathcal{F}'$ on $X$ such that*

(1) $\mathcal{F} \subset \mathcal{F}'$;
(2) $A_0 \in \mathcal{F}'$.

*Proof.* Consider the principal filter $\mathcal{F}_0$ based on $A_0$ and set

$$\mathcal{F}' = \{B \in \mathcal{P}(X) : B \supset A \cap A' \text{ for some } A \in \mathcal{F} \text{ and } A' \in \mathcal{F}_0\}. \tag{J.2}$$

Let us show that $\mathcal{F}'$ satisfies the required conditions. First of all, taking $A' = X$ we have $A = A \cap X \in \mathcal{F}'$ for all $A \in \mathcal{F}$. This shows (1). On the other hand, taking $A = X$ and $A' = A_0$ we have $A_0 = X \cap A_0 \in \mathcal{F}$, and this shows (2).

We are only left to show that $\mathcal{F}'$ is a filter. Let $B \in \mathcal{F}'$ and denote by $A \in \mathcal{F}$ and $A' \in \mathcal{F}_0$ two sets such that $A \cap A' \subset B$.

Let us first show that $A \cap A' \neq \varnothing$. Suppose the contrary. Then $A \subset (X \setminus A') \subset (X \setminus A_0)$. As $A$ belongs to the filter $\mathcal{F}$, it follows from (F-2) that $(X \setminus A_0) \in \mathcal{F}$, contradicting our assumptions. It follows that $A \cap A' \neq \varnothing$ and therefore $B \neq \varnothing$. This shows (F-1). Suppose now that $B' \in \mathcal{P}(X)$ contains $B$. Then $(A \cap A') \subset B'$ so that $B' \in \mathcal{F}'$. This shows (F-2). Finally, let $B_1, B_2 \in \mathcal{F}'$. For $i = 1, 2$ denote by $A_i \in \mathcal{F}$ and $A_i' \in \mathcal{F}_0$ two sets such that $A_i \cap A_i' \subset B_i$. We have

$$B_1 \cap B_2 \supset (A_1 \cap A_1') \bigcap (A_2 \cap A_2') = (A_1 \cap A_2) \bigcap (A_1' \cap A_2').$$

But $A_1 \cap A_2$ belongs to the filter $\mathcal{F}$, and $A_1' \cap A_2' \in \mathcal{F}_0$ as both $A_1'$ and $A_2'$ contain $A_0$. It follows that $B_1 \cap B_2 \in \mathcal{F}'$. This shows (F-3). It follows that $\mathcal{F}'$ is a filter. □

*Proof of Theorem J.1.6.* Consider the set $\Phi_0$ consisting of all filters on $X$ containing $\mathcal{F}_0$. This is a nonempty set, partially ordered by inclusion. Let $\Phi$ be a totally ordered subset of $\Phi_0$. We claim that $\widetilde{\mathcal{F}} = \bigcup_{\mathcal{F} \in \Phi} \mathcal{F}$ is an upper bound for $\Phi$. We only have to show that $\widetilde{\mathcal{F}}$ belongs to $\Phi_0$. As $\varnothing \notin \mathcal{F}$ for all $\mathcal{F} \in \Phi$ we also have $\varnothing \notin \widetilde{\mathcal{F}}$. This shows that $\widetilde{\mathcal{F}}$ satisfies condition (F-1). Let now $A \in \widetilde{\mathcal{F}}$ and $B \subset X$ be such that $A \subset B$. Then there exists $\mathcal{F} \in \Phi$ such that $A \in \mathcal{F}$. As $\mathcal{F}$ is a filter, we have $B \in \mathcal{F}$, by (F-2). It then follows that $B \in \widetilde{\mathcal{F}}$. This shows that $\widetilde{\mathcal{F}}$ satisfies condition (F-2). Finally, suppose that $A, B \in \widetilde{\mathcal{F}}$. Then there exist $\mathcal{F}_A, \mathcal{F}_B \in \Phi$ such that $A \in \mathcal{F}_A$ and $B \in \mathcal{F}_B$. As $\Phi$ is totally ordered, up to exchanging $A$ and $B$ we can suppose that $\mathcal{F}_A \subset \mathcal{F}_B$. We then have $A, B \in \mathcal{F}_B$ and therefore $A \cap B \in \mathcal{F}_B$, by (F-3). It follows that $A \cap B \in \widetilde{\mathcal{F}}$. Therefore $\widetilde{\mathcal{F}}$ satisfies condition (F-3) as well. This shows that $\Phi_0$ is inductive.

By Zorn's lemma, $\Phi_0$ contains a maximal element $\omega$. Let us show that $\omega$ is an ultrafilter on $X$. Let $A_0 \subset X$. Suppose that $A_0, (X \setminus A_0) \notin \omega$. Then, by Lemma J.1.7, there exists a filter $\omega'$ containing $A_0$ and $\omega$. As $\omega'$ is in $\Phi_0$ and properly contains $\omega$, this contradicts the maximality of $\omega$. It follows that either $A_0$ or $X \setminus A_0$ belongs to $\omega$. This shows that $\omega$ satisfies condition (UF), and therefore it is an ultrafilter. □

**Corollary J.1.8.** *Let $X$ be a set. Let $\mathcal{F}$ be a filter on $X$. Then the following conditions are equivalent.*

(a) *$\mathcal{F}$ is an ultrafilter;*
(b) *$\mathcal{F}$ is a maximal filter, that is, if $\mathcal{F}'$ is a filter containing $\mathcal{F}$, then $\mathcal{F}' = \mathcal{F}$.*

*Proof.* Suppose (a) and let $\mathcal{F}'$ be a filter properly containing $\mathcal{F}$. Let $A \in \mathcal{F}' \setminus \mathcal{F}$. As $\mathcal{F}$ is an ultrafilter, $(X \setminus A) \in \mathcal{F} \subset \mathcal{F}'$. Thus, both $A$ and $(X \setminus A)$ belong to $\mathcal{F}'$ contradicting (F-4). This shows that $\mathcal{F}' = \mathcal{F}$ and the implication (a) $\Rightarrow$ (b) follows. Suppose now that $\mathcal{F}$ is a maximal filter. By Theorem J.1.6 there exists an ultrafilter $\omega$ containing $\mathcal{F}$. By maximality we have $\mathcal{F} = \omega$, that is, $\mathcal{F}$ is an ultrafilter. This shows (b) $\Rightarrow$ (a). □

## J.2 Limits Along Filters

**Definition J.2.1.** Let $X$ be a topological space. A filter $\mathcal{F}$ on $X$ is said to be *convergent* if there exists a point $x_0 \in X$ such that all neighborhoods of $x_0$ belong to $\mathcal{F}$. One then says that $x_0$ is a *limit* of $\mathcal{F}$ and that $\mathcal{F}$ converges to $x_0$.

*Remark J.2.2.* A topological space $X$ is Hausdorff if and only if every convergent filter on $X$ has a unique limit point in $X$. Indeed, suppose that $X$ is Hausdorff and let $\mathcal{F}$ be a filter converging to two distinct points $x, y \in X$. Let $U \in \mathcal{F}_x \subset \mathcal{F}$ and $V \in \mathcal{F}_y \subset \mathcal{F}$ be such that $U \cap V = \varnothing$. As $U, V \in \mathcal{F}$ we also have $U \cap V \in \mathcal{F}$ and this contradicts (F-1). Conversely, suppose that $X$ is not Hausdorff. Then there exist two distinct points $x, y \in X$ such that for all $U \in \mathcal{F}_x$ and $V \in \mathcal{F}_y$ one has $U \cap V \neq \varnothing$. It follows that the set $\Omega = \mathcal{F}_x \cup \mathcal{F}_y \subset \mathcal{P}(X)$ has the finite intersection property. Thus, by Proposition J.1.3, there exists a filter $\mathcal{F}$ containing both $\mathcal{F}_x$ and $\mathcal{F}_y$. It follows that $\mathcal{F}$ converges to both $x$ and $y$.

**Theorem J.2.3.** *Let $X$ be a topological space. Then the following conditions are equivalent:*

(a) *$X$ is compact;*
(b) *every ultrafilter on $X$ is convergent.*

*Proof.* Suppose (a) and let $\omega$ be an ultrafilter on $X$. Let $A_1, A_2, \ldots, A_n$, $n \geq 1$, be elements in $\omega$ and denote by $\overline{A}_1, \overline{A}_2, \ldots, \overline{A}_n$ their closures. As $\bigcap_{i=1}^n \overline{A_i} \supset \bigcap_{i=1}^n A_i \neq \varnothing$ (by (F-1) and (F-5)), we have that the family $(\overline{A})_{A \in \omega}$ in $X$ has the finite intersection property and therefore, by compactness of $X$, it has a nonempty intersection. Let $x \in \bigcap_{A \in \omega} \overline{A}$. As $x \in \overline{A}$ for all $A \in \omega$, it follows that for every $V \in \mathcal{F}_x$ we have $A \cap V \neq \varnothing$. Consider the set $\Omega = \{A \cap V : A \in \omega, V \in \mathcal{F}_x\} \subset \mathcal{P}(X)$. Observe that $\Omega$ has the finite intersection property and denote by $\mathcal{F}(\Omega)$ the filter generated by $\Omega$

(cf. Example J.1.2(b)). We have $\omega \subset \Omega \subset \mathcal{F}(\Omega)$ and, as $\omega$ is an ultrafilter, it follows from Corollary J.1.8 that $\omega = \mathcal{F}(\Omega)$. As $\mathcal{F}_x \subset \Omega$, we deduce that $\mathcal{F}_x \subset \omega$, that is, $\omega$ converges to $x$.

Conversely, suppose (b). Let $(C_i)_{i\in I}$ be a family of closed subsets of $X$ with the finite intersection property. To prove that $X$ is compact we have to show that

$$\bigcap_{i\in I} C_i \neq \varnothing. \tag{J.3}$$

By Proposition J.1.3 there exists a filter $\mathcal{F}$ such that $C_i \in \mathcal{F}$ for all $i \in I$. By Theorem J.1.6 there exists an ultrafilter $\omega$ such that $\mathcal{F} \subset \omega$. By our assumptions, there exists $x \in X$ such that $\omega$ converges to $x$, equivalently, $\mathcal{F}_x \subset \omega$. Let $i \in I$. by (F-1) and (F-3), we have $C_i \cap V \neq \varnothing$ for all $V \in \mathcal{F}_x$. As $C_i$ is closed, we have $x \in C_i$. Thus $x \in \bigcap_{i\in I} C_i$, and (J.3) follows. □

**Definition J.2.4.** Let $X$ be a set, $Y$ a topological space, $y_0$ a point of $Y$, $f\colon X \to Y$ a map and $\mathcal{F}$ a filter on $X$. One says that $y_0$ is a *limit* of $f$ along $\mathcal{F}$ (or that $f(x)$ *converges* to $y_0$ along $\mathcal{F}$), and one writes

$$f(x) \underset{x\to\mathcal{F}}{\longrightarrow} y_0,$$

if $f^{-1}(V) = \{x \in X : f(x) \in V\}$ belongs to $\mathcal{F}$ for all neighborhoods $V$ of $y_0$. If such a limit point $y_0$ is unique one writes

$$\lim_{x\to\mathcal{F}} f(x) = y_0.$$

*Examples J.2.5.* (a) Let $X$ and $Y$ be two topological spaces, $f\colon X \to Y$ a map, $x_0 \in X$ and $y_0 \in Y$. One has that $f(x)$ converges to $y_0$ in $Y$ for $x$ tending to $x_0$ if and only if $f(x) \underset{x\to\mathcal{F}_{x_0}}{\longrightarrow} y_0$.

(b) Let $X$ be a topological space and $f\colon \mathbb{N} \to X$ a map. The sequence $(f(n))_{n\in\mathbb{N}}$ converges to $x \in X$, if and only if $x$ is a limit of $f$ along the Fréchet filter on $\mathbb{N}$.

(c) Let $(I, \leq)$ be a directed set and let $\mathcal{F}$ be the residual filter on $I$. Let $Y$ be a topological space and $(y_i)_{i\in I}$ a net in $Y$. Then $(y_i)_{i\in I}$ converges to a point $y_0 \in Y$ if and only if $y_i \underset{i\to\mathcal{F}}{\longrightarrow} y_0$. Note that (b) is a particular case of the present example.

**Corollary J.2.6.** *Let $X$ be a set, $Y$ a compact topological space, $f\colon X \to Y$ a map, and $\omega$ an ultrafilter on $X$. Then there exists $y_0 \in Y$ such that $f(x) \underset{x\to\mathcal{F}}{\longrightarrow} y_0$. Moreover, if $X$ is Hausdorff such an $y_0$ is unique.*

*Proof.* The set $f(\omega) = \{f(A) : A \in \omega\} \subset \mathcal{P}(Y)$ has the finite intersection property since $f(A)\cap f(B) \supset f(A\cap B) \neq \varnothing$ for all $A, B \in \omega$. Let $\mathcal{F}$ denote the filter generated by $f(\omega)$. By Theorem J.1.6, there exists an ultrafilter $\omega'$ on $Y$ which contains $\mathcal{F}$. As $Y$ is compact, by Theorem J.2.3 there exists $y_0 \in Y$ such that $\omega'$ converges to $y_0$. Let us show that if $V$ is a neighborhood of $y_0$, then

$f^{-1}(V)$ belongs to $\omega$. Suppose the contrary. As $\omega$ is an ultrafilter, by (UF) we have $X \setminus f^{-1}(V) = \{x \in X : f(x) \notin V\} \in \omega$. Setting $U = f(X \setminus f^{-1}(V))$, we have $U \in f(\omega) \subset \mathcal{F} \subset \omega'$. But $V \in \mathcal{F}_{y_0} \subset \omega'$ and, by construction, $U \cap V = \varnothing$. As $U \cap V \in \omega'$, this contradicts (F-1). This shows that $y_0$ is a limit of $f$ along $\omega$. If $X$ is Hausdorff, then uniqueness of $y_0$ follows from Remark J.2.2. □

**Proposition J.2.7.** *Let $X$ be a set and let $\mathcal{F}$ be a filter on $X$. Let $Y$, $Z$ be two topological spaces and $f\colon X \to Y$ and $g\colon X \to Z$ be two maps. Equip $Y \times Z$ with the product topology and let $F\colon X \to Y \times Z$ be the map defined by $F(x) = (f(x), g(x))$ for all $x \in X$. Suppose that there exist $y_0 \in Y$ and $z_0 \in Z$ such that $f(x) \underset{x\to\mathcal{F}}{\longrightarrow} y_0$ and $g(x) \underset{x\to\mathcal{F}}{\longrightarrow} z_0$. Then $F(x) \underset{x\to\mathcal{F}}{\longrightarrow} (y_0, z_0)$.*

*Proof.* Let $W \subset Y \times Z$ be a neighborhood of the point $(y_0, z_0)$. By definition of the product topology, there exist neighborhoods $U \subset Y$ of $y_0$ and $V \subset Z$ of $z_0$ such that $U \times V \subset W$. As $f(x) \underset{x\to\mathcal{F}}{\longrightarrow} y_0$, we have $f^{-1}(U) \in \mathcal{F}$. Similarly, as $g(x) \underset{x\to\mathcal{F}}{\longrightarrow} z_0$, we have $g^{-1}(V) \in \mathcal{F}$. It follows that $f^{-1}(U) \cap g^{-1}(V)$ belongs to $\mathcal{F}$ and as $F^{-1}(W) \supset F^{-1}(U \times V) = f^{-1}(U) \cap g^{-1}(V)$, we deduce from (F-2) that $F^{-1}(W) \in \mathcal{F}$. This shows that $F(x) \underset{x\to\mathcal{F}}{\longrightarrow} (y_0, z_0)$. □

**Proposition J.2.8.** *Let $X$ be a set and let $\mathcal{F}$ be a filter on $X$. Let $Y$, $Z$ be two topological spaces and $f\colon X \to Y$ and $g\colon Y \to Z$ be two maps. Suppose that there exists $y_0 \in Y$ such that $f(x) \underset{x\to\mathcal{F}}{\longrightarrow} y_0$ and that $g$ is continuous at $y_0$. Then $(g \circ f)(x) \underset{x\to\mathcal{F}}{\longrightarrow} g(y_0)$.*

*Proof.* Let $W \subset Z$ be a neighborhood of $g(y_0)$. By continuity of $g$ there exists a neighborhood $V \subset Y$ of $y_0$ such that $g^{-1}(W) \supset V$. By definition of limit, we have $f^{-1}(V) \in \mathcal{F}$. As $(g \circ f)^{-1}(W) = f^{-1}(g^{-1}(W)) \supset f^{-1}(V)$, it follows from (F-2) that $(g \circ f)^{-1}(W) \in \mathcal{F}$. This shows that $(g \circ f)(x) \underset{x\to\mathcal{F}}{\longrightarrow} g(y_0)$. □

**Proposition J.2.9.** *Let $X$ be a set and let $\mathcal{F}$ be a filter on $X$. Let $f_1, f_2\colon X \to \mathbb{R}$ be two maps such that $f_1 \leq f_2$ (i.e. $f_1(x) \leq f_2(x)$ for all $x \in X$). Suppose that there exists $y_1 \in \mathbb{R}$ (resp. $y_2 \in \mathbb{R}$) such that $y_1 = \lim_{x\to\mathcal{F}} f_1(x)$ (resp. $y_2 = \lim_{x\to\mathcal{F}} f_2(x)$). Then $y_1 \leq y_2$.*

*Proof.* Suppose, by contradiction, that $y_1 > y_2$. Set $r = \frac{y_1 - y_2}{3}$ and let $V_1 = (y_1 - r, y_1 + r)$ and $V_2 = (y_2 - r, y_2 + r)$. Note that $V_1 \cap V_2 = \varnothing$. By definition of limit we have $f_1^{-1}(V_1) \in \mathcal{F}$ and $f_2^{-1}(V_2) \in \mathcal{F}$. As $\mathcal{F}$ is a filter we deduce that $f_1^{-1}(V_1) \cap f_2^{-1}(V_2) \in \mathcal{F}$. On the other hand we have $f_1^{-1}(V_1) \cap f_2^{-1}(V_2) = \varnothing$ since $f_1 \leq f_2$. This contradicts (F-2). It follows that $y_1 \leq y_2$. □

Recall that given a set $X$, we denote by $\ell^\infty(X)$ the Banach space consisting of all bounded real maps $f\colon X \to \mathbb{R}$ equipped with the norm $\|f\|_\infty = \sup\{|f(x)| : x \in X\}$. Moreover, a linear map $m\colon \ell^\infty(X) \to \mathbb{R}$ satisfying $m(1) = 1$ and $m(x) \geq 0$ for all $x \in \ell^\infty(E)$ such that $x \geq 0$, is called a mean on $X$ (cf. Definition 4.1.4).

**Corollary J.2.10.** *Let $X$ be a set and let $\omega$ be an ultrafilter on $X$. For every $f \in \ell^\infty(X)$ there exists a unique $y_0 \in \mathbb{R}$ such that $f(x) \underset{x\to\omega}{\longrightarrow} y_0$. Moreover, the map $m_\omega \colon \ell^\infty(X) \to \mathbb{R}$ defined by $m_\omega(f) = \lim_{x\to\omega} f(x)$ is a mean on $X$, in particular $m_\omega$ is continuous and $\|m_\omega\| = 1$.*

*Proof.* Let $f \in \ell^\infty(X)$. Using the fact that $f(x) \in Y = [-\|f\|_\infty, \|f\|_\infty]$ for all $x \in X$ and that $Y$ is compact Hausdorff, we deduce from Corollary J.2.6 that there exists a unique $y_0 \in \mathbb{R}$ such that $f(x) \underset{x\to\omega}{\longrightarrow} y_0$.

Let us show that the map $m_\omega$ is linear. Let $a \in \mathbb{R}$. Consider the continuous map $g \colon \mathbb{R} \to \mathbb{R}$ defined by $g(y) = ay$ for all $y \in \mathbb{R}$. Then $af = g \circ f$ and from Proposition J.2.8 we deduce

$$m_\omega(af) = \lim_{x\to\omega} (af)(x) = a \lim_{i\to\omega} f(x) = a m_\omega(f).$$

Let now $f, g \in \ell^\infty(X)$. Consider the map $F \colon X \to \mathbb{R}^2$ defined by setting $F(x) = (f(x), g(x))$ for all $x \in X$ and the continuous map $G \colon \mathbb{R}^2 \to \mathbb{R}$ defined by $G(y_1, y_2) = y_1 + y_2$. Then $f + g = G \circ F$ and from Proposition J.2.7 and Proposition J.2.8 we deduce

$$m_\omega(f+g) = \lim_{x\to\omega} (f+g)(x) = \Big(\lim_{x\to\omega} f(x)\Big) + \Big(\lim_{x\to\omega} g(x)\Big) = m_\omega(f) + m_\omega(g).$$

This shows that $m_\omega$ is linear.

Let now $f(x) = 1$ for all $x \in X$. For every neighborhood $V$ of $1 \in \mathbb{R}$ we have $f^{-1}(V) = \{x \in X : f(x) \in V\} = X \in \omega$, by (F-6). This shows that $m_\omega(1) = m_\omega(f) = \lim_{x\to\omega} f(x) = 1$.

Finally, it follows immediately from Proposition J.2.9 that if $f \geq 0$ then $m_\omega(f) = \lim_{x\to\omega} f(x) \geq 0$.

We have shown that $m_\omega$ is a mean. The last properties of $m_\omega$ follow from Proposition 4.1.7. □

## Notes

The definition of filter is due to H. Cartan (1937). The full treatment of convergence along filters is given in Bourbaki [Bou] as an alternative to the similar notion of a net developed in 1922 by E. H. Moore and H. L. Smith. Given a Hausdorff topological space $X$ the set $\beta X$ of all ultrafilters on $X$ can be given the structure of a compact Hausdorff space, called the *Stone-Čech compactification* of $X$. This is the largest compact Hausdorff space "generated" by $X$, in the sense that any map from $X$ to a compact Hausdorff space factors through $\beta X$ in a unique way. The elements $x$ of $X$ correspond to the principal ultrafilters $\omega(\{x\})$ on $X$.

Let now $X$ be any set. With every ultrafilter $\omega$ on $X$ one associates the $\{0,1\}$-valued finitely additive probability measure $\mu_\omega$ on $X$ defined

by $\mu_\omega(A) = 1$ if $A \in \omega$ and $\mu_\omega(A) = 0$ if $A \in \mathcal{P}(X) \setminus \omega$. Conversely, given a $\{0,1\}$-valued finitely additive probability measure $\mu$ on $X$, the set $\omega_\mu = \{A \in \mathcal{P}(X) : \mu(A) = 1\}$ is an ultrafilter on $X$. This establishes a one-to-one correspondence between the ultrafilters $\omega$ on $X$ and the $\{0,1\}$-valued finitely additive probability measures on $X$. In Sect. 4.1 we considered the set $\mathcal{MP}(X)$ (resp. $\mathcal{M}(X)$) of all finitely additive probability measures (resp. of all means) on $X$ and we showed that there exists a natural bijective map $\Phi\colon \mathcal{MP}(X) \to \mathcal{M}(X)$. Then one has $\Phi^{-1}(\mu_\omega) = m_\omega$ for all ultrafilters $\omega$ on $X$.

# Open Problems

In the list below we collect some open problems related to the topics treated in this book.

(OP-1) Let $G$ be an amenable periodic group which is not locally finite. Does there exist a finite set $A$ and a cellular automaton $\tau\colon A^G \to A^G$ which is surjective but not injective?

(OP-2) Let $G$ be a periodic group which is not locally finite and let $A$ be an infinite set. Does there exist a bijective cellular automaton $\tau\colon A^G \to A^G$ which is not invertible?

(OP-3) Let $G$ be a periodic group which is not locally finite and let $V$ be an infinite-dimensional vector space over a field $\mathbb{K}$. Does there exist a bijective linear cellular automaton $\tau\colon V^G \to V^G$ which is not invertible?

(OP-4) Is every Gromov-hyperbolic group residually finite (resp. residually amenable, resp. sofic, resp. surjunctive)?

(OP-5) (*Gottschalk's conjecture*) Is every group surjunctive?

(OP-6) Let $G$ be a periodic group which is not locally finite and let $A$ be an infinite set. Does there exist a cellular automaton $\tau\colon A^G \to A^G$ whose image $\tau(A^G)$ is not closed in $A^G$ with respect to the prodiscrete topology?

(OP-7) Let $G$ be a periodic group which is not locally finite and let $V$ be an infinite-dimensional vector space over a field $\mathbb{K}$. Does there exist a linear cellular automaton $\tau\colon V^G \to V^G$ whose image $\tau(V^G)$ is not closed in $V^G$ with respect to the prodiscrete topology?

(OP-8) Let $G$ and $H$ be two quasi-isometric groups. Suppose that $G$ is surjunctive. Is it true that $H$ is surjunctive?

(OP-9) Let $G$ be a non-amenable group. Does there exist a finite set $A$ and a cellular automaton $\tau\colon A^G \to A^G$ which is pre-injective but not surjective?

(OP-10) Does there exist a non-sofic group?

(OP-11) Does there exist a surjunctive group which is non-sofic?

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1, 

(OP-12) Let $G$ and $H$ be two quasi-isometric groups. Suppose that $G$ is sofic. Is it true that $H$ is sofic?

(OP-13) Let $G$ be a non-amenable group and let $\mathbb{K}$ be a field. Does there exist a finite-dimensional $\mathbb{K}$-vector space $V$ and a linear cellular automaton $\tau\colon V^G \to V^G$ which is pre-injective but not surjective?

(OP-14) Let $G$ be a non-amenable group and let $\mathbb{K}$ be a field. Does there exist a finite-dimensional $\mathbb{K}$-vector space $V$ and a linear cellular automaton $\tau\colon V^G \to V^G$ which is surjective but not pre-injective?

(OP-15) (*Kaplanski's stable finiteness conjecture*) Is the group algebra $\mathbb{K}[G]$ stably finite for any group $G$ and any field $\mathbb{K}$? Equivalently, is every group L-surjunctive, that is, is it true that, for any group $G$, any field $\mathbb{K}$, and any finite-dimensional $\mathbb{K}$-vector space $V$, every injective linear cellular a automaton $\tau\colon V^G \to V^G$ is surjective?

(OP-16) (*Kaplanski's zero-divisors conjecture*) Is it true that the group algebra $\mathbb{K}[G]$ has no zero-divisors for any torsion-free group $G$ and any field $\mathbb{K}$? Equivalently, is it true that, for any torsion-free group $G$ and any field $\mathbb{K}$, every non-identically-zero linear cellular automaton $\tau\colon \mathbb{K}^G \to \mathbb{K}^G$ is pre-injective?

(OP-17) Is every unique-product group orderable?

## Comments

(OP-1) The answer to this question is affirmative if $G$ is non-periodic, i.e., it contains an element of infinite order (see Exercise 3.23), or if $G$ is non-amenable (Theorem 5.12.1). On the other hand, if $G$ is a locally finite group and $A$ is a finite set, then every surjective cellular automaton $\tau\colon A^G \to A^G$ is injective (see Exercise 3.21). An example of an amenable periodic group which is not locally finite is provided by the Grigorchuck group described in Sect. 6.9.

(OP-2) The answer is affirmative if $G$ is not periodic (cf. [CeC11, Corollary 1.2]). On the other hand, if $G$ is locally finite and $A$ is an arbitrary set, then every bijective cellular automaton $\tau\colon A^G \to A^G$ is invertible (cf. Exercise 3.20 or [CeC11, Proposition 4.1]).

(OP-3) The answer is affirmative if $G$ is not periodic (cf. [CeC11, Theorem 1.1]). On the other hand, if $G$ is locally finite and $V$ is an arbitrary vector space, then every bijective linear cellular automaton $\tau\colon V^G \to V^G$ is invertible (cf. [CeC11, Proposition 4.1]).

(OP-5) Every sofic group is surjunctive (cf. Theorem 7.8.1).

(OP-6) When $A$ is a finite set and $G$ is an arbitrary group, it follows from Lemma 3.3.2 that the image of every cellular automaton $\tau\colon A^G \to A^G$ is closed in $A^G$. When $A$ is an infinite set and $G$ is a non-periodic group, it is shown in [CeC11, Corollary 1.4] that there exists a cellular automaton $\tau\colon A^G \to A^G$ whose image is not closed in $A^G$. On the other hand, when $G$

is locally finite, then, for any set $A$, the image of every cellular automaton $\tau\colon A^G \to A^G$ is closed in $A^G$ (cf. Exercise 3.22 or [CeC11, Proposition 4.1]).

(OP-6) When $V$ is a finite-dimensional vector space over a field $\mathbb{K}$ and $G$ is an arbitrary group, it follows from Theorem 8.8.1 that the image of every linear cellular automaton $\tau\colon V^G \to V^G$ is closed in $V^G$. When $V$ is an infinite-dimensional vector space and $G$ is a non-periodic group, it is shown in [CeC11, Theorem 1.3] that there exists a linear cellular automaton $\tau\colon V^G \to V^G$ whose image is not closed in $V^G$. On the other hand, when $G$ is locally finite, then, for any vector space $V$, the image of every linear cellular automaton $\tau\colon V^G \to V^G$ is closed in $V^G$ (cf. [CeC11, Proposition 4.1]).

(OP-9) The answer is affirmative if $G$ contains a nonabelian free subgroup (cf. Proposition 5.11.1).

(OP-13) The answer is affirmative if $G$ contains a nonabelian free subgroup (cf. Corollary 8.10.2).

(OP-14) The answer is affirmative if $G$ contains a nonabelian free subgroup (cf. Corollary 8.11.2).

(OP-15) See the discussion in the notes at the end of Chap. 8.

(OP-16) See the discussion in the notes at the end of Chap. 8.

# References


[Abe] Abels, H.: An example of a finitely presented solvable group. In: Homological Group Theory Proc. Sympos., Durham, 1977. London Math. Soc. Lecture Note Ser., vol. 36, pp. 205–211. Cambridge University Press, Cambridge (1979)

[Ady] Adyan, S.I.: Random walks on free periodic groups. Math. USSR, Izv. **21**, 425–434 (1983)

[AlS] Allouche, J.-P., Shallit, J.: Automatic Sequences: Theory, Applications, Generalizations. Cambridge University Press, Cambridge (2003)

[Amo] Amoroso, S., Patt, Y.N.: Decision procedures for surjectivity and injectivity of parallel maps for tessellation structures. J. Comput. Syst. Sci. **6**, 448–464 (1972)

[AOP] Ara, P., O'Meara, K.C., Perera, F.: Stable finiteness of group rings in arbitrary characteristic. Adv. Math. **170**, 224–238 (2002)

[Aus] Auslander, L.: On a problem of Philip Hall. Ann. Math. **86**, 112–116 (1967)

[BanMS] Bandini, S., Mauri, G., Serra, R.: Cellular automata: from a theoreticel parallel computational model to its application to complex systems. Parallel Comput. **27**(5), 539–553 (2001). Cellular Automata: From Modeling to Applications (Trieste, 1998)

[Bar] Bartholdi, L.: Gardens of Eden and amenability on cellular automata. J. Eur. Math. Soc. **12**, 241–248 (2010)

[Bas] Bass, H.: The degree of polynomial growth of finitely generated nilpotent groups. Proc. Lond. Math. Soc. **25**, 603–614 (1972)

[Bau] Baumslag, G.: Automorphism groups of residually finite groups. J. Lond. Math. Soc. **38**, 117–118 (1963)

[BaS] Baumslag, G., Solitar, D.: Some two-generator one-relator non-Hopfian groups. Bull. Am. Math. Soc. **68**, 199–201 (1962)

[BDV] Bekka, B., de la Harpe, P., Valette, A.: Kazhdan's Property ($T$). New Mathematical Monographs, vol. 11. Cambridge University Press, Cambridge (2008)

[Ber] Bergman, G.M.: Right orderable groups that are not locally indicable. Pac. J. Math. **147**, 243–248 (1991)

[BCG] Berlekamp, E.R., Conway, J.H., Guy, R.K.: Winning Ways for Your Mathematical Plays, vol. 2, 2nd edn. AK Peters, Natick (2003)

[BKM] Blanchard, F., Kůrka, P., Maass, A.: Topological and measure-theoretic properties of one-dimensional cellular automata. Physica D **103**(1–4), 86–99 (1997). Lattice Dynamics (Paris, 1995)

[BoR] Botto Mura, R., Rhemtulla, A.: Orderable Groups. Lecture Notes in Pure and Applied Mathematics, vol. 27. Dekker, New York (1977)

[Bou] Bourbaki, N.: Topologie Générale, Chapitres 1 à 4. Hermann, Paris (1971)



T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1,

[Bow] Bowen, L.: Measure conjugacy invariants for actions of countable sofic groups. J. Amer. Math. Soc. **23**, 217–245 (2010)

[BuM] Burger, M., Mozes, S.: Lattices in product of trees. Publ. Math. IHÉS **92**, 151–194 (2000)

[Bur] Burks, A.W.: von Neumann's self-reproducing automata. In: Burks, A.W. (ed.) Essays on Cellular Automata, pp. 3–64. University of Illinois Press, Champaign (1971)

[CFP] Cannon, J.W., Floyd, W.J., Parry, W.R.: Introductory notes on Richard Thompsons groups. Enseign. Math. **42**, 215–256 (1996)

[CeC1] Ceccherini-Silberstein, T., Coornaert, M.: The Garden of Eden theorem for linear cellular automata. Ergod. Theory Dyn. Syst. **26**, 53–68 (2006)

[CeC2] Ceccherini-Silberstein, T., Coornaert, M.: Injective linear cellular automata and sofic groups. Isr. J. Math. **161**, 1–15 (2007)

[CeC3] Ceccherini-Silberstein, T., Coornaert, M.: On the surjunctivity of Artinian linear cellular automata over residually finite groups. In: Geometric Group Theory. Trends in Mathematics, pp. 37–44. Birkhäuser, Basel (2007)

[CeC4] Ceccherini-Silberstein, T., Coornaert, M.: Amenability and linear cellular automata over semisimple modules of finite length. Commun. Algebra **36**, 1320–1335 (2008)

[CeC5] Ceccherini-Silberstein, T., Coornaert, M.: Linear cellular automata over modules of finite length and stable finiteness of group rings. J. Algebra **317**, 743–758 (2007)

[CeC6] Ceccherini-Silberstein, T., Coornaert, M.: A note on Laplace operators on groups. In: Limits of Graphs in Group Theory and Computer Science, pp. 37–40. EPFL Press, Lausanne (2009)

[CeC7] Ceccherini-Silberstein, T., Coornaert, M.: A generalization of the Curtis-Hedlund theorem. Theor. Comput. Sci. **400**, 225–229 (2008)

[CeC8] Ceccherini-Silberstein, T., Coornaert, M.: Induction and restriction of cellular automata. Ergod. Theory Dyn. Syst. **29**, 371–380 (2009)

[CeC9] Ceccherini-Silberstein, T., Coornaert, M.: On a characterization of locally finite groups in terms of linear cellular automata, J. Cell. Autom., to appear

[CeC10] Ceccherini-Silberstein, T., Coornaert, M.: *Expansive actions on uniform spaces and surjunctive maps*, arXiv:0810.1295

[CeC11] Ceccherini-Silberstein, T., Coornaert, M.: *On the reversibility and the closed image property of linear cellular automata*, arXiv:0910.0863

[CFS] Ceccherini-Silberstein, T., Fiorenzi, F., Scarabotti, F.: The Garden of Eden theorem for cellular automata and for symbolic dynamical systems. In: Kaimanovich, V.A. (ed.) Random Walks and Geometry, pp. 73–108. de Gruyter, Berlin (2004)

[CGH1] Ceccherini-Silberstein, T., Grigorchuk, R.I., de la Harpe, P.: Amenability and paradoxical decompositions for pseudogroups and discrete metric spaces. Proc. Steklov Inst. Math. **224**, 57–97 (1999)

[CGH2] Ceccherini-Silberstein, T., Grigorchuk, R.I., de la Harpe, P.: Décompositions paradoxales des groupes de Burnside [Paradoxical decompositions of free Burnside groups]. C. R. Acad. Sci., Sér. 1 Math. **327**, 127–132 (1998)

[CMS1] Ceccherini-Silberstein, T.G., Machì, A., Scarabotti, F.: Amenable groups and cellular automata. Ann. Inst. Fourier **49**, 673–685 (1999)

[CMS2] Ceccherini-Silberstein, T.G., Machì, A., Scarabotti, F.: The Grigorchuk group of intermediate growth. Rend. Circ. Mat. Palermo **50**(1), 67–102 (2001)

[Cha] Champetier, C.: L'espace des groupes de type fini. Topology **39**, 657–680 (2000)

[Che] Chernoff, P.R.: A simple proof of Tychonoff's theorem via nets. Am. Math. Mon. **99**, 932–934 (1992)

[Coh] Cohn, P.M.: Some remarks on the invariant basis property. Topology **5**, 215–228 (1966)
[Con] Connell, I.G.: On the group ring. Can. J. Math. **15**, 650–685 (1963)
[Cor] Cornulier, Y.: A sofic group away from amenable groups. arXiv:0906.3374
[CGP] de Cornulier, Y., Guyot, L., Pitsch, W.: On the isolated points in the space of groups. J. Algebra **307**, 254–277 (2007)
[CovN] Coven, E.M., Nitecki, Z.W.: On the genesis of symbolic dynamics as we know it. Colloq. Math. **110**, 227–242 (2008)
[CovP] Coven, E., Paul, M.: Endomorphisms of irreducible shifts of finite type. Math. Syst. Theory **8**, 167–175 (1974)
[Day1] Day, M.M.: Amenable semigroups. Ill. J. Math. **1**, 509–544 (1957)
[Day2] Day, M.M.: Fixed-point theorems for compact convex sets. Ill. J. Math. **5**, 585–590 (1961)
[Day3] Day, M.M.: Convolutions, means and spectra. Ill. J. Math. **8**, 100–111 (1964)
[Deh] Dehornoy, P.: Braid groups and left distributive operations. Trans. Am. Math. Soc. **345**, 115–151 (1994)
[DS] Dicks, W., Schick, T.: The spectral measure of certain elements of the complex group ring of a wreath product. Geom. Dedic. **93**, 121–137 (2002)
[DrS] Druţu, C., Sapir, M.: Non-linear residually finite groups. J. Algebra **284**, 174–178 (2005)
[Efr] Efremovich, V.A.: The geometry of proximity. I. Mat. Sbornik N. S. **31**(73), 189–200 (1952)
[Ele] Elek, G.: The rank of finitely generated modules over group algebras. Proc. Am. Math. Soc. **131**, 3477–3485 (2003)
[ES1] Elek, G., Szabó, E.: Sofic groups and direct finiteness. J. Algebra **280**, 426–434 (2004)
[ES2] Elek, G., Szabó, E.: Hyperlinearity, essentially free actions and $L^2$-invariants. The sofic property. Math. Ann. **332**, 421–441 (2005)
[ES3] Elek, G., Szabó, E.: On sofic groups. J. Group Theory **9**, 161–171 (2006)
[Fio1] Fiorenzi, F.: The Garden of Eden theorem for sofic shifts. Pure Math. Appl. **11**, 471–484 (2000)
[Fio2] Fiorenzi, F.: Cellular automata and strongly irreducible shifts of finite type. Theor. Comput. Sci. **299**, 477–493 (2003)
[Føl] Følner, E.: On groups with full Banach mean value. Math. Scand. **3**, 245–254 (1955)
[Gar-1] Gardner, M.: The game of life. Mathematical games. Sci. Am. **223**, 120–123 (1970)
[Gar-2] Gardner, M.: Wheels, Life, and Other Mathematical Amusements. W. H. Freeman & Co., San Francisco (1983)
[Gla] Glass, A.M.W.: Partially Ordered Groups. Series in Algebra, vol. 7. World Scientific, River Edge (1999)
[GlG] Glebsky, L.Yu., Gordon, E.I.: On surjunctivity of the transition functions of cellular automata on groups. Taiwan. J. Math. **9**, 511–520 (2005)
[GlR] Glebsky, L., Rivera, L.M.: Sofic groups and profinite topology on free groups. J. Algebra **320**, 3512–3518 (2008)
[GolS] Golod, E.S., Shafarevich, I.R.: On the class field tower. Izv. Akad. Nauk SSSR, Ser. Mat. **28**, 261–272 (1964)
[GoL] Gorin, E.A., Lin, V.Ja.: Algebraic equations with continuous coefficients, and certain questions of the algebraic theory of braids. Math. USSR Sb. **7**, 569–596 (1969)
[Got] Gottschalk, W.H.: Some general dynamical systems. In: Recent Advances in Topological Dynamics. Lecture Notes in Mathematics, vol. 318, pp. 120–125. Springer, Berlin (1973)

[GoH] Gottschalk, W.H., Hedlund, G.A.: Topological Dynamics. American Mathematical Society Colloquium Publications, vol. 36. Am. Math. Soc., Providence (1955)
[Gre] Greenleaf, F.P.: Invariant Means on Topological Groups and their Applications. Van Nostrand Mathematical Studies, vol. 16. Van Nostrand-Reinhold, New York (1969)
[Gri1] Grigorchuk, R.I.: Symmetrical random walks on discrete groups. In: Multicomponent Random Systems. Adv. Probab. Related Topics, vol. 6, pp. 285–325. Dekker, New York (1980)
[Gri2] Grigorchuk, R.I.: On Burnside's problem on periodic groups. Funct. Anal. Appl. **14**, 41–43 (1980)
[Gri3] Grigorchuk, R.I.: On the Milnor problem of group growth. Sov. Math. Dokl. **28**, 23–26 (1983)
[Gri4] Grigorchuk, R.I.: Degrees of growth of finitely generated groups and the theory of invariant means. Izv. Akad. Nauk SSSR, Ser. Mat. **48**, 939–985 (1984)
[Gri5] Grigorchuk, R.I.: Solved and unsolved problems around one group. In: Infinite Groups: Geometric, Combinatorial and Dynamical Aspects. Progr. Math., vol. 248, pp. 117–218. Birkhäuser, Basel (2005)
[GriNS] Grigorchuk, R.I., Nekrashevych, V.V., Sushchanskii, V.I.: Automata, dynamical systems, and groups. Proc. Steklov Inst. Math. **231**, 128–203 (2000)
[GriP] Grigorchuk, R.I., Pak, I.: Groups of intermediate growth: an introduction. Enseign. Math. **54**, 251–272 (2008)
[Gro1] Gromov, M.: Infinite groups as geometric objects. In: Proceedings of the International Congress of Mathematicians, vol. 1, 2 (Warsaw, 1983), pp. 385–392. PWN, Warsaw (1984)
[Gro2] Gromov, M.: Groups of polynomial growth and expanding maps. Publ. Math. IHÉS **53**, 53–73 (1981)
[Gro3] Gromov, M.: Hyperbolic groups. In: Essays in Group Theory. Math. Sci. Res. Inst. Publ., vol. 8, pp. 75–263. Springer, New York (1987)
[Gro4] Gromov, M.: Asymptotic invariants of infinite groups. In: Niblo, G.A., Roller, M.A. (eds.) Geometry Group Theory, Sussex 1991, vol. 2. London Math. Soc. Lecture Note Ser., vol. 182. Cambridge University Press, Cambridge (1993)
[Gro5] Gromov, M.: Endomorphisms of symbolic algebraic varieties. J. Eur. Math. Soc. **1**, 109–197 (1999)
[Gro6] Gromov, M.: Topological invariants of dynamical systems and spaces of holomorphic maps, Part I. Math. Phys. Anal. Geom. **2**, 323–415 (1999)
[Had] Hadamard, J.: Les surfaces à courbures opposées et leurs lignes géodésiques. J. Math. Pures Appl. **4**, 27–73 (1898)
[Hall-M-1] Hall, M.: Distinct representatives of subsets. Bull. Am. Math. Soc. **54**, 922–926 (1948)
[Hal-M-2] Hall, M.: Subgroups of finite index in free groups. Can. J. Math. **1**, 187–190 (1949)
[Hall1] Hall, P.: On representatives of subsets. J. Lond. Math. Soc. **10**, 26–30 (1935)
[Hall2] Hall, P.: On the finiteness of certain soluble groups. Proc. Lond. Math. Soc. **9**, 595–622 (1959)
[Halm] Halmos, P.R., Vaughan, H.E.: The marriage problem. Am. J. Math. **72**, 214–215 (1950)
[Har1] de la Harpe, P.: Free groups in linear groups. Enseign. Math. **29**, 129–144 (1983)
[Har2] de la Harpe, P.: Topics in Geometric Group Theory. Chicago Lectures in Mathematics. University of Chicago Press, Chicago (2000)
[Har3] de la Harpe, P.: Mesures finiment additives et paradoxes. In: Autour du Centenaire Lebesgue. Panor. Synthèses, vol. 18, pp. 39–61. Soc. Math. France, Paris (2004)

[Hed-1] Hedlund, G.A.: Sturmian minimal sets. Am. J. Math. **66**, 605–620 (1944)
[Hed-2] Hedlund, G.A.: Transformations commuting with the shift. In: Gottschalk, W.H., Auslander, J. (eds) Topological Dynamics. Benjamin, New York (1968)
[Hed-3] Hedlund, G.A.: Endormorphisms and automorphisms of the shift dynamical system. Math. Syst. Theory **3**, 320–375 (1969)
[Hig] Higman, G.: A finitely related group with an isomorphic proper factor group. J. Lond. Math. Soc. **26**, 59–61 (1951)
[Hir] Hirsch, K.A.: On infinite soluble groups. IV. J. Lond. Math. Soc. **27**, 81–85 (1952)
[Hun] Hungerford, T.W.: Algebra. Graduate Texts in Mathematics. Springer, New York (1987)
[Isb] Isbell, J.R.: Uniform Spaces. Mathematical Surveys, vol. 12. Am. Math. Soc., Providence (1964)
[Jac] Jachymski, J.R.: Another proof of the Markov-Kakutani theorem, and an extension. Math. Jpn. **47**, 19–20 (1998)
[Jam] James, I.M.: Introduction to Uniform Spaces. London Mathematical Society Lecture Note Series, vol. 144. Cambridge University Press, Cambridge (1990)
[KaV] Kaimanovich, V.A., Vershik, A.M.: Random walks on discrete groups: boundary and entropy. Ann. Probab. **11**, 457–490 (1983)
[Kak] Kakutani, S.: Two fixed-point theorems concerning bicompact convex sets. Proc. Imp. Acad. (Tokyo) **14**, 242–245 (1938)
[Kap1] Kaplansky, I.: Problems in the theory of rings. Report of a conference on linear algebras, June 1956, 1–3, National Academy of Sciences-National Research Council, Washington, Publ. 502 (1957)
[Kap2] Kaplansky, I.: Fields and Rings. Chicago Lectures in Math. University of Chicago Press, Chicago (1969)
[Kap3] Kaplansky, I.: "Problems in the theory of rings" revisited. Am. Math. Mon. **77**, 445–454 (1970)
[Kar1] Kari, J.: Reversibility of 2D cellular automata is undecidable. Physica D **45**, 379–385 (1990)
[Kar2] Kari, J.: Reversibility and surjectivity problems of cellular automata. J. Comput. Syst. Sci. **48**, 149–182 (1994)
[Kar3] Kari, J.: Theory of cellular automata: a survey. Theor. Comput. Sci. **334**, 3–33 (2005)
[Kel] Kelley, J.L.: General Topology. Graduate Texts in Mathematics, vol. 27. Springer, New York (1975)
[Kes1] Kesten, H.: Symmetric random walks on groups. Trans. Am. Math. Soc. **92**, 336–354 (1959)
[Kes2] Kesten, H.: Full Banach mean values on countable groups. Math. Scand. **7**, 146–156 (1959)
[Kit] Kitchens, B.P.: Symbolic dynamics. One-sided, two-sided and countable state Markov shifts. Universitext. Springer, Berlin (1998)
[Ko1] Kolmogorov, A.N.: A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces. Dokl. Akad. Nauk SSSR **119**, 861–864 (1958)
[Ko2] Kolmogorov, A.N.: Entropy per unit time as a metric invariant of automorphisms. Dokl. Akad. Nauk SSSR **124**, 754–755 (1959)
[Kri] Krieger, F.: Le lemme d'Ornstein-Weiss d'après Gromov. In: Dynamics, Ergodic Theory, and Geometry. Math. Sci. Res. Inst. Publ., vol. 54, pp. 99–111. Cambridge University Press, Cambridge (2007)
[KLM] Kropoholler, P.H., Linnell, P.A., Moody, J.A.: Applications of a new K-theoretic theorem to soluble group rings. Proc. Am. Math. Soc. **104**, 675–684 (1988)

[Kur] Kůrka, P.: Topological and Symbolic Dynamics. Cours Spécialisés, vol. 11. Société Mathématique de France, Paris (2003)
[Law] Lawton, W.: Note on symbolic transformation groups. Not. Am. Math. Soc. **19**, A-375 (1972) (abstract)
[Lev] Levi, F.: Über die Untergruppen der freien Gruppen. II. Math. Z. **37**, 90–97 (1933)
[LiM] Lind, D., Marcus, B.: An Introduction to Symbolic Dynamics and Coding. Cambridge University Press, Cambridge (1995)
[LiW] Lindenstrauss, E., Weiss, B.: Mean topological dimension. Isr. J. Math. **115**, 1–24 (2000)
[LS] Lyndon, R.C., Schupp, P.E., Combinatorial Group Theory. Reprint of the 1977 edition. Classics in Mathematics. Springer, Berlin (2001)
[Lys] Lysënok, I.G.: A set of defining relations for the Grigorchuk group. Mat. Zametki **38**, 503–516, 634 (1985)
[MaM] Machì, A., Mignosi, F.: Garden of Eden configurations for cellular automata on Cayley graphs of groups. SIAM J. Discrete Math. **6**, 44–56 (1993)
[Mag] Magnus, W.: Residually finite groups. Bull. Am. Math. Soc. **75**, 305–316 (1969)
[MaKS] Magnus, W., Karras, A., Solitar, D., Combinatorial Group Theory, Presentations of Groups in Terms of Generators and Relations. Reprint of the 1976 second edition. Dover, New York (2004)
[Mal1] Mal'cev, A.I.: On isomorphic matrix representations of infinite groups. Rec. Math. [Mat. Sb.] **8**(50), 405–422 (1940)
[Mal2] Mal'cev, A.I.: On the embedding of group algebras in division algebras. Dokl. Akad. Nauk SSSR **60**, 1499–1501 (1948)
[Mal3] Mal'cev, A.I.: On homomorphisms onto finite groups. In: A.I. Mal'cev, Selected Works, Vol. 1. Classical Algebra, pp. 450–462. Nauka, Moscow (1976)
[Mar] Markov, A.: Quelques théorèmes sur les ensembles abéliens. C.R. (Doklady) Acad. Sci. URSS **1**, 311–313 (1936)
[Mes] Meskin, S.: Nonresidually finite one-relator groups. Trans. Am. Math. Soc. **164**, 105–114 (1972)
[Mil1] Milnor, J.: A note on curvature and fundamental group. J. Differ. Geom. **2**, 1–7 (1968)
[Mil2] Milnor, J.: Growth of finitely generated solvable groups. J. Differ. Geom. **2**, 447–449 (1968)
[Moo] Moore, E.F.: Machine Models of Self-Reproduction. In: Proc. Symp. Appl. Math., vol. 14, pp. 17–34. Am. Math. Soc., Providence (1963)
[Morr] Morris, D.W.: Amenable groups that act on the line. Algebraic Geom. Topol. **6**, 2509–2518 (2006)
[Mors] Morse, M.: Recurrent geodesics on a surface of negative curvature. Trans. Am. Math. Soc. **22**, 84–100 (1921)
[Myh] Myhill, J.: The converse of Moore's Garden of Eden Theorem. Proc. Am. Math. Soc. **14**, 685–686 (1963)
[Nam] Namioka, I.: Følner's conditions for amenable semi-groups. Math. Scand. **15**, 18–28 (1964)
[NaB] Narici, L., Beckenstein, E.: Topological Vector Spaces. Monographs and Textbooks in Pure and Applied Mathematics, vol. **95**. Dekker, New York (1985)
[Neu1] Neumann, B.H.: On ordered division rings. Trans. Am. Math. Soc. **66**, 202–252 (1949)
[Neu2] Neumann, B.H.: A two-generator group isomorphic to a proper factor group. J. Lond. Math. Soc. **25**, 247–248 (1950)
[vNeu1] von Neumann, J.: Zur allgemeine theorie des masses. Fundam. Math. **13**, 73–116 (1929)
[vNeu2] von Neumann, J.: In: Burks, A.W. (ed.) The Theory of Self-reproducing Automata. University of Illinois Press, Urbana (1966)

[Ols] Ol'shanskii, A.Yu.: On the question of the existence of an invariant mean on a group. Usp. Mat. Nauk **35**, 199–200 (1980)
[OlS] Ol'shanskii, A.Yu., Sapir, M.: Non-amenable finitely presented torsion-by-cyclic groups. Publ. Math. IHÉS **96**, 43–169 (2002) (2003)
[OrW] Ornstein, D.S., Weiss, B.: Entropy and isomorphism theorems for actions of amenable groups. J. Anal. Math. **48**, 1–141 (1987)
[Pas] Passman, D.S.: The Algebraic Structure of Group Rings. Reprint of the 1977 original. Krieger, Melbourne (1985)
[Pat] Paterson, A.: Amenability. AMS Mathematical Surveys and Monographs, vol. 29. Am. Math. Soc., Providence (1988)
[Pes] Pestov, V.G.: Hyperlinear and sofic groups: a brief guide. Bull. Symb. Log. **14**, 449–480 (2008)
[PeK] Pestov, V.G., Kwiatkowska, A.: An introduction to hyperlinear and sofic groups. arXiv:0911.4266
[Pro] Promislow, S.D.: A simple example of a torsion-free, nonunique product group. Bull. Lond. Math. Soc. **20**, 302–304 (1988)
[RhR] Rhemtulla, A., Rolfsen, D.: Local indicability in ordered groups: braids and elementary amenable groups. Proc. Am. Math. Soc. **130**, 2569–2577 (2002)
[RiS] Rips, E., Segev, Y.: Torsion-free group without unique product property. J. Algebra **108**, 116–126 (1987)
[RobA] Robinson, A.: Non-standard analysis. North-Holland, Amsterdam (1966)
[RobD] Robinson, D.J.S.: A Course in the Theory of Groups, 2nd edn. Graduate Texts in Mathematics, vol. 80. Springer, New York (1996)
[Rot] Rotman, J.J.: An Introduction to the Theory of Groups, 4th edn. Graduate Texts in Mathematics, vol. 148. Springer, New York (1995)
[Rud] Rudin, W.: Functional Analysis. McGraw-Hill Series in Higher Mathematics. McGraw-Hill, New York (1973)
[RuS] Rudin, W., Schneider, H.: Idempotents in group rings. Duke Math. J. **31**, 585–602 (1964)
[San] Sanov, I.N.: A property of a representation of a free group. Dokl. Akad. Nauk SSSR **57**, 657–659 (1947)
[Sca] Scarabotti, F.: On a lemma of Gromov and the entropy of a graph. Eur. J. Comb. **23**, 631–633 (2002)
[Sha] Shalom, Y.: Harmonic analysis, cohomology, and the large-scale geometry of amenable groups. Acta Math. **192**, 119–185 (2004)
[Stë1] Stëpin, A.M.: Approximability of groups and group actions. Russ. Math. Surv. **38**, 131–132 (1983)
[Stë2] Stëpin, A.M.: Approximations of groups and group actions, the Cayley topology. In: Ergodic Theory of $\mathbb{Z}^d$-Actions. London Math. Soc. Lecture Note Ser., vol. 228, pp. 475–484. Cambridge University Press, Cambridge (1996)
[Sva] Švarc, A.S.: A volume invariant of coverings. Dokl. Akad. Nauk SSSR **105**, 32–34 (1955)
[Tar1] Tarski, A.: Sur les fonctions additives dans les classes abstraites et leur application au problème de la mesure. Comptes Rendus des Séances de la Société des Sciences et des Lettres de Varsovie, Classe III **22**, 243–248 (1929)
[Tar2] Tarski, A.: Algebraische fassung des massproblems. Fundam. Math. **31**, 47–66 (1938)
[Tho] Thom, A.: Examples of hyperlinear groups without factorization property. Groups Geom. Dyn. **4**, 195–208 (2010)
[Tits] Tits, J.: Free subgroups in linear groups. J. Algebra **20**, 250–270 (1972)
[Ver] Vershik, A.M.: Amenability and approximation of infinite groups. Sel. Math. Sov. **2**, 311–330 (1982)
[VeG] Vershik, A.M., Gordon, E.I.: Groups that are locally embeddable in the class of finite groups. St. Petersburg Math. J. **9**, 49–67 (1998)

[Wag] Wagon, S.: The Banach-Tarski paradox. Encyclopedia of Mathematics and its Applications, vol. 24. Cambridge University Press, Cambridge (1985)
[Weil] Weil, A.: Sur les Espaces à Structure Uniforme et sur la Topologie Générale. Hermann, Paris (1938)
[Weiss] Weiss, B.: Sofic groups and dynamical systems (Ergodic theory and harmonic analysis, Mumbai, 1999). Sankhya Ser. A **62**, 350–359 (2000)
[Woe1] Woess, W.: Random walks on infinite graphs and groups – a survey on selected topics. Bull. Lond. Math. Soc. **26**, 1–60 (1994)
[Woe2] Woess, W.: Random Walks on Infinite Graphs and Groups. Cambridge Tracts in Mathematics **138**. Cambridge University Press, Cambridge (2000)
[Wol] Wolf, J.A.: Growth of finitely generated solvable groups and curvature of Riemannian manifolds. J. Differ. Geom. **2**, 421–446 (1968)
[Wolfr1] Wolfram, S.: Statistical mechanics of cellular automata. Rev. Mod. Phys. **55**, 601–644 (1983)
[Wolfr2] Wolfram, S.: Universality and complexity in cellular automata. Physica D **10**, 1–35 (1984)
[Wolfr3] Wolfram, S.: A New Kind of Science. Wolfram Media, Inc., Champaign, IL (2002)

# List of Symbols

| Symbol | Definition | Page |
|---|---|---|
| $\epsilon$ | the empty word | 367 |
| $\preceq$ | the dominance relation in the set of growth functions $\gamma\colon \mathbb{N} \to [0, +\infty)$ | 162 |
| $\sim$ | the equivalence relation in the set of growth functions $\gamma\colon \mathbb{N} \to [0, +\infty)$ | 162 |
| $\|T\|_{p\to p}$ | the $\ell^p$-norm of a linear map $T\colon \ell^p(E) \to \ell^p(E)$ | 194 |
| $0_R, 0$ | the zero element of the ring $R$ | 291 |
| $1_G$ | the identity element of the group $G$ | 2 |
| $1_R, 1$ | the unity element of the ring $R$ | 291 |
| $\gamma_S^G$, $\gamma_S$ | the growth function of the group $G$ relative to the finite symmetric generating subset $S \subset G$ | 160 |
| $\Delta_S^G$, $\Delta_S$ | the discrete laplacian on the group $G$ associated with the subset $S \subset G$ | 9 |
| $\Delta_S^{(2)}$ | the restriction of $\Delta_S$ to the Hilbert space $\ell^2(G)$ | 201 |
| $\partial_E(\Omega)$ | the $E$-boundary of the subset $\Omega \subset G$ | 116 |
| $\iota_S(G)$ | the isoperimetric constant of the group $G$ with respect to the finite symmetric generating subset $S \subset G$ | 191 |
| $\lambda_S^G$, $\lambda_S$ | the growth rate of the group $G$ with respect to the finite symmetric generating subset $S \subset G$ | 169 |
| $\lambda(e)$ | the label of the edge $e \in E$ in a labeled graph $\mathcal{G} = (Q, E)$ | 153 |
| $\lambda(\pi)$ | the label of the path $\pi$ in a labeled graph $\mathcal{G} = (Q, E)$ | 154, 223 |
| $\pi^-$ | the initial vertex of the path $\pi$ in an $S$-labeled graph | 154 |

T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1, 

| Symbol | Definition | Page |
|---|---|---|
| $gx$ | the configuration defined by $gx(h) = x(g^{-1}h)$ | 2 |
| $\mathcal{G} = (X, Y, E)$ | the bipartite graph with left (resp. right) vertex set $X$ (resp. $Y$) and set of edges $E$ | 391 |
| $H_R$ | the Heisenberg group with coefficients in the ring $R$ | 94 |
| $\mathrm{ICA}(G; A)$ | the group consisting of all invertible cellular automata $\tau\colon A^G \to A^G$ | 24 |
| $\mathrm{Id}_X$ | the identity map on the set $X$ | 2 |
| $\ell(w)$ | the length of the word $w \in A^*$ | 35 |
| $\ell^p(E)$ | the Banach space of all $p$-summable functions $x\colon E \to \mathbb{R}$ | 193 |
| $\ell^\infty(E)$ | the Banach space of all bounded functions $x\colon E \to \mathbb{R}$ | 78 |
| $\ell_S^G(g)$, $\ell_S(g)$ | the word-length of the element $g \in G$ with respect to the finite symmetric generating subset $S \subset G$ | 152 |
| $\mathrm{LCA}(G; V)$ | the algebra of all linear cellular automata $\tau\colon V^G \to V^G$ | 287 |
| $\mathrm{LCA}(G, H; V)$ | the subalgebra of $\mathrm{LCA}(G; V)$ consisting of all linear cellular automata $\tau\colon V^G \to V^G$ admitting a memory set $S$ such that $S \subset H$ | 289 |
| $L(X)$ | the language associated with the subshift $X$ | 35 |
| $L_n(X)$ | the set of admissible words of length $n$ of the subshift $X$ | 35 |
| $\mathcal{L}(X)$ | the space of all continous endomorphisms of the Banach space $X$ | 406 |
| $M_S^{(p)}$ | the $\ell^p$-Markov operator associated with the finite subset $S \subset G$ | 195 |
| $\mathrm{Mat}_d(R)$ | the ring consisting of all $d \times d$ matrices with entries in the ring $R$ | 305 |
| $\mathrm{mdim}_{\mathcal{F}}(X)$ | the mean dimension of the vector subspace $X \subset V^G$ with respect to the right Følner net $\mathcal{F}$ | 308 |
| $\mathcal{M}(E)$ | the set of all means on the set $E$ | 79 |
| $\mathcal{PM}(E)$ | the set of all finitely additive probability measures on the set $E$ | 79 |
| $\mathcal{N}(\Gamma)$ | the space of all normal subgroups of the group $\Gamma$ or, equivalently, the space of all $\Gamma$-marked groups | 61 |
| $\mathcal{N}_L(B) \subset X$ | the left-neighborhood of the subset $B \subset Y$ in the bipartite graph $(X, Y, E)$ | 392 |
| $\mathcal{N}_L(y) \subset X$ | the left-neighborhood of the vertex $y \in Y$ in the bipartite graph $(X, Y, E)$ | 391 |

| Symbol | Definition | Page |
|---|---|---|
| $\mathcal{N}_R(A) \subset Y$ | the right-neighborhood of the subset $A \subset X$ in the bipartite graph $(X, Y, E)$ | 392 |
| $\mathcal{N}_R(x) \subset Y$ | the right-neighborhood of the vertex $x \in X$ in the bipartite graph $(X, Y, E)$ | 391 |
| $\mathcal{P}(E)$ | the set of all subsets of the set $E$ | 77 |
| $\mathrm{per}(B)$ | the period of the matrix $B$ | 227 |
| $\mathrm{per}(\mathcal{G})$ | the period of the labeled graph $\mathcal{G}$ | 227 |
| $\mathrm{per}(X)$ | the period of the irreducible sofic subshift $X$ | 227 |
| $\mathrm{per}_n(X)$ | the number of $n\mathbb{Z}$-periodic configurations in the subshift $X$ | 227 |
| $Q(r)$ | the set of all vertices of the $S$-labeled graph $\mathcal{Q} = (Q, E)$ for which there exists an $S$-labeled graph isomorphism $\psi_{q,r} \colon B_S(r) \to B(q, r)$ satisfying $\psi_{q,r}(1_G) = q$ | 265 |
| $\mathcal{Q} = (Q, E)$ | the $S$-labeled graph with vertex set $Q$ and edge set $E \subset Q \times S \times Q$ | 153 |
| $R[G]$ | the group ring of the group $G$ with coefficients in the ring $R$ | 292 |
| $R^{op}$ | the opposite ring of the ring $R$ | 293 |
| $\mathrm{Sym}(X)$ | the symmetric group of the set $X$ | 359 |
| $\mathrm{Sym}_0(X)$ | the subgroup of $\mathrm{Sym}(X)$ consisting of all permutations with finite support | 360 |
| $\mathrm{Sym}_0^+(X)$ | the alternating group on $X$ | 364 |
| $\mathrm{Sym}_n$ | the symmetric group of degree $n$ | 366 |
| $\mathrm{Sym}_n^+$ | the alternating group of degree $n$ | 366 |
| $\mathrm{Sym}(X, \preceq)$ | the subgroup of $\mathrm{Sym}(X)$ that preserve the partial order $\preceq$ of the set $X$ | 179, 333 |
| $U(R)$ | the multiplicative group consisting of all invertible elements in the ring $R$, | 292 |
| $V[G]$ | the vector subspace of $V^G$ consisting of all configurations $x \colon V \to G$ with finite support | 288 |
| $x\|_\Omega$ | the restriction of the configuration $x \in A^G$ to the subset $\Omega \subset G$ | 3 |
| $X^*$ | the topological dual of the real normed space $X$ | 384 |
| $X(\mathcal{A})$ | the subshift of finite type defined by the set of admissible patterns $\mathcal{A}$ | 32 |
| $X_f$ | the set of all configurations in the subshift $X$ whose $G$-orbit is finite | 71 |
| $X_\mathcal{P}$ | the subshift defined by the set of forbidden patterns $\mathcal{P}$ | 32 |
| $X^\mathcal{G}$ | the subshift defined by the labeled graph $\mathcal{G}$ | 223 |
| $Z(G)$ | the center of the group $G$ | 94 |

# Index




T. Ceccherini-Silberstein, M. Coornaert, *Cellular Automata and Groups*,
Springer Monographs in Mathematics,
DOI 10.1007/978-3-642-14034-1,