Chapter 1
Introduction

1.1 What Is This Book About?

This book deals with "products of random matrices". Let us describe in concrete terms the questions we will be studying throughout this book. Let $d \geq 1$ be a positive integer. We choose a sequence $g_1, \ldots, g_n, \ldots$ of $d \times d$ invertible matrices with real coefficients. These matrices are chosen independently and according to an identical law μ. We want to study the sequence of product matrices $p_n := g_n \cdots g_1$. In particular, we want to know:

$$\textit{Can one describe the asymptotic behavior of the matrices } p_n\textit{?} \tag{1.1}$$

A naive way to ask this question is to fix a Euclidean norm on the vector space $V = \mathbb{R}^d$, to fix a nonzero vector v on V and a nonzero linear functional f on V and to ask

$$\textit{What is the asymptotic behavior of the norms } \|p_n\|\textit{?} \tag{1.2}$$

$$\textit{What is the asymptotic behavior of the coefficients } f(p_n v)\textit{?} \tag{1.3}$$

The first aim of this book is to explain the answer to these questions, which was guessed at a very early stage of the theory: under suitable irreducibility and moment assumptions, the real random variables $\log \|p_n\|$ and $\log |f(p_n v)|$ behave very much like a "sum of independent identically distributed (iid) real random variables".

Indeed we will see that, under suitable assumptions, these variables satisfy many properties that are classical for "sums of iid random real numbers" like the Law of Large Numbers (LLN), the Central Limit Theorem (CLT), the Law of Iterated Logarithm (LIL), the Large Deviations Principle (LDP), and the Local Limit Theorem (LLT).

The answer to Questions (1.2) and (1.3) will be obtained by focusing first on the following two related questions:

$$\textit{What is the asymptotic distribution of the vectors } \frac{p_n v}{\|p_n v\|}\textit{?} \tag{1.4}$$

$$\textit{What is the asymptotic behavior of the norms } \|p_n v\|\textit{?} \tag{1.5}$$

Y. Benoist, J.-F. Quint, *Random Walks on Reductive Groups*,
Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of
Modern Surveys in Mathematics 62, DOI 10.1007/978-3-319-47721-3_1

1.2 When Did This Topic Emerge?

The theory of "products of random matrices" or more precisely "products of iid random matrices" is sometimes also called "random walks on linear groups". It began in the middle of the 20^{th} century. It finds its roots in the speculative work of Bellman in [8] who guessed that an analog of classical Probability Theory for "sums of random numbers" might be true for the coefficients of products of random matrices. The pioneers of this topic are Kesten, Furstenberg, Guivarc'h, ...

At that time, in 1960, Probability Theory was already based on very strong mathematical foundations, and the language of σ-algebras, measure theory and Fourier transforms was widely adopted among the specialists interested in probabilistic phenomena. A few textbooks on "sum of random numbers" were already available (like those of Kolmogorov [80] in the USSR, Lévy [85] in France and Cramér [35] in the UK, ...), and many more were about to appear, such as those of Loève [86], Spitzer [119], Breiman [29], Feller [44], ...

It took about half a century for the theory of "products of random matrices" to achieve its maturity. The reason may be the following. Even though some of the new characters who happen to play an important role in this new realm, like the "martingales and the Markov chains" and the "ergodic theory of cocycles" were very popular among specialists of this topic, some of them like the "semisimple algebraic groups" and the "highest weight representations" were less popular, moreover, some of them like the "spectral theory of transfer operators" and the "asymptotic properties of discrete linear groups" were not yet known.

This book is also an introduction to all of these tools.

The main contributors of the theorems we are going to explain in this book are not only Kesten, Furstenberg, Guivarc'h, but also Kifer, Le Page, Raugi, Margulis, Goldsheid, ...

The topic of this book is the same as the nice and very influential book written by Bougerol–Lacroix 30 years ago. We also recommend the surveys by Ledrappier [84] and Furman [48] on related topics. This theory has recently found nice applications to the study of subgroups of Lie groups (as in [58], [27] or [26, Sect. 12]). Beyond these applications, we were urged to write this book so that it could serve as a background reference for our joint work in [13], [15], and [16].

Even though our topic is very much related to the almost homonymous topic "random walks on countable groups", we will not discuss here this aspect of the theory and its ties with the "geometric group theory" and the "growth of groups".

1.3 Is This Topic Related to Sums of Random Numbers?

Yes. The classical theory of "sums of random numbers" or more precisely "sums of iid random numbers" is sometimes also called "random walks on $\mathbb{R}^d$". Let us describe in concrete terms the question studied in this classical theory.

We choose a sequence $t_1, \ldots, t_n, \ldots$ of real numbers. These real numbers are chosen independently and according to an identical law μ. This law μ is a Borel probability measure on the real line $\mathbb{R}$. We denote by A the support of μ. For instance, when $\mu = \frac{1}{2}(\delta_0 + \delta_1)$, the set A is $\{0, 1\}$, and we are choosing the t_k to be either 0 or 1 with equal probability and independently of the previous choices of t_j for $j < k$. We want to study the sequence of partial sums $s_n := t_1 + \cdots + t_n$. In particular, we want to know:

$$\textit{What is the asymptotic behavior of } s_n\textit{?} \tag{1.6}$$

We will explain in Sect. 1.4 various classical answers to this question.

On the one hand, some of these classical answers describe the behavior in law of this sequence. They tell us what we can expect at time n when n is large. These statements only involve the law of the random variable s_n which is nothing else than the n^{th}-convolution power μ^{*n} of μ, i.e.

$$\mu^{*n} = \mu * \cdots * \mu.$$

For instance, the Central Limit Theorem (CLT), the Large Deviations Principle (LDP) and the Local Limit Theorem (LLT) are statements in law. An important tool in this point of view is Fourier analysis.

On the other hand, some classical answers describe the behavior of the individual trajectories $s_1, s_2, \ldots, s_n, \ldots$ These statements are true for almost every trajectory. The trajectories are determined by elements of the Bernoulli space

$$B := A^{\mathbb{N}^*} := \{b = (t_1, \ldots, t_n, \ldots) \mid t_n \in A\}$$

of all possible sequences of random choices. Here "almost every" refers to the Bernoulli probability measure

$$\beta := \mu^{\otimes \mathbb{N}^*}$$

on this space B. This space B is also called the space of *forward trajectories*. For instance, the Law of Large Numbers (LLN) and the Law of the Iterated Logarithm (LIL) are statements about almost every trajectory. An important tool in this point of view is the conditional expectation.

The interplay between these two aspects is an important feature of Probability Theory. The Borel–Cantelli lemma sometimes allows one to transfer results in law into almost-sure results. Conversely, the point of view of trajectories gives us a much deeper level of analysis on the probabilistic phenomena that cannot be reached by the sole study of the laws μ^{*n}.

1.4 What Classical Results Should I Know?

This short book is as self-contained as possible. We will reprove many classical facts from Probability Theory. However we will take for granted basic facts from Linear Algebra, Integration Theory and Functional Analysis. A few results on real reductive algebraic groups, their representations and their discrete subgroups will be quoted without proof.

The reader will more easily appreciate the streamlining of this book if he or she knows classical Probability Theory. Indeed, the main objective of this book is to present for "products of iid random matrices" the analogs of the following five classical theorems for "sums of iid random numbers".

In these five classical theorems, we fix a probability measure μ on $\mathbb{R}$ and set $b = (t_1, \ldots, t_n, \ldots) \in B$ and $s_n = t_1 + \cdots + t_n$ for the partial sums. The sequence b is chosen according to the law β, which means that the coordinates t_k are iid random real numbers of law μ.

The first theorem is the Law of Large Numbers due to many authors from Bernoulli up to Kolmogorov. It tells us that, when μ has a finite first moment, i.e. when $\int_{\mathbb{R}} |t| \, \mathrm{d}\mu(t) < \infty$, almost every trajectory has a drift which is equal to the average of the law:

$$\lambda := \int_{\mathbb{R}} t \, \mathrm{d}\mu(t). \tag{1.7}$$

Theorem 1.1 (LLN) *Let μ be a Borel probability measure on $\mathbb{R}$ with a finite first moment. Then, for β-almost all b in B, one has*

$$\lim_{n\to\infty} \tfrac{1}{n} s_n = \lambda. \tag{1.8}$$

The second theorem is the Central Limit Theorem which is also due to many authors from Laplace up to Lindeberg and Lévy. It tells us that, when μ is non-degenerate, i.e. is not a Dirac mass, and when μ has a finite second moment, i.e. when $\int_{\mathbb{R}} t^2 \, \mathrm{d}\mu(t) < \infty$, the recentered law of μ^{*n} spreads at speed $\sqrt{n}$, more precisely, it tells us that the renormalized variables $\frac{s_n - n\lambda}{\sqrt{n}}$ converge in law to a Gaussian variable which has the same variance Φ as μ:

$$\Phi := \int_{\mathbb{R}} (t - \lambda)^2 \, \mathrm{d}\mu(t).$$

Theorem 1.2 (CLT) *Let μ be a non-degenerate Borel probability measure on $\mathbb{R}$ with a finite second moment. Then, for any bounded continuous function ψ on $\mathbb{R}$, one has*

$$\lim_{n\to\infty} \int_{\mathbb{R}} \psi\left(\tfrac{s - n\lambda}{\sqrt{n}}\right) \mathrm{d}\mu^{*n}(s) = \int_{\mathbb{R}} \psi(s) \frac{e^{-\frac{s^2}{2\Phi}}}{\sqrt{2\pi\Phi}} \, \mathrm{d}s. \tag{1.9}$$

The third theorem is the Law of the Iterated Logarithm discovered by Khinchin. It tells us that almost all recentered trajectories spread at a slightly higher speed than $\sqrt{n}$. More precisely, it tells us that the precise scale at which almost all recentered trajectories fill a bounded interval is $\sqrt{n \log\log n}$.

Theorem 1.3 (LIL) *Let μ be a non-degenerate Borel probability measure on $\mathbb{R}$ with a finite second moment. Then, for β-almost all b in B, the set of cluster points of the sequence*

$$\frac{s_n - n\lambda}{\sqrt{2\Phi\, n \log\log n}}$$

is equal to the interval $[-1, 1]$.

The fourth theorem is the Large Deviations Principle due to Cramér. It tells us that when μ has a finite exponential moment, i.e. when $\int_{\mathbb{R}} e^{\alpha|t|}\,\mathrm{d}\mu(t) < \infty$, for some $\alpha > 0$, the probability of an excursion away from the average decays exponentially. We will just state below the upper bound in the large deviations principle.

Theorem 1.4 (LDP) *Let μ be a Borel probability measure on $\mathbb{R}$ with a finite exponential moment. Then, for any $t_0 > 0$, one has*

$$\limsup_{n\to\infty} \mu^{*n}(\{t \in \mathbb{R} \mid |t - n\lambda| \geq nt_0\})^{\frac{1}{n}} < 1. \tag{1.10}$$

The fifth theorem is the Local Limit Theorem due to many authors from de Moivre up to Stone. It tells us that the rate of decay for the probability that the recentered sum $s_n - n\lambda$ belongs to a fixed interval is $1/\sqrt{n}$. For the sake of simplicity, we will assume below that μ is aperiodic, i.e. μ is not supported by an arithmetic progression $m_0 + t\mathbb{Z}$ with $m_0 \in \mathbb{R}$ and $t > 0$. Indeed, the statement is just slightly different when μ is supported by an arithmetic progression.

Theorem 1.5 (LLT) *Let μ be an aperiodic Borel probability measure on $\mathbb{R}$ with a finite second moment. Then, for all $a_1 \leq a_2$, one has*

$$\lim_{n\to\infty} \sqrt{n}\ \mu^{*n}(n\lambda + [a_1, a_2]) = \frac{a_2 - a_1}{\sqrt{2\pi\Phi}}.$$

1.5 Can You Show Me Some Nice Sample Results from This Topic?

The five main results that we will explain in this book are the analogs of the five classical theorems that we just quoted in the previous section. We will state below special cases of these five results. We will explain in Sect. 1.9 what kind of generalizations of these special cases is needed for a better answer to Question 1.1.

In these five results, we fix a Borel probability measure μ on the special linear group $G := \mathrm{SL}(d, \mathbb{R})$, we set $V = \mathbb{R}^d$, and we fix a Euclidean norm $\|\cdot\|$ on V. We denote by A the support of μ, and by Γ_μ the closed subsemigroup of G spanned by A. For $n \geq 1$, we denote by μ^{*n} the n^{th}-convolution power

$$\mu^{*n} := \mu * \cdots * \mu.$$

The forward trajectories are determined by elements of the Bernoulli space

$$B := A^{\mathbb{N}^*} := \{b = (g_1, \dots, g_n, \dots) \mid g_n \in A\} \tag{1.11}$$

endowed with the Bernoulli probability measure

$$\beta := \mu^{\otimes \mathbb{N}^*}.$$

As in Sect. 1.4, the sequence b is chosen according to the law β which means that b is a sequence of iid random matrices g_k chosen with law μ, and we want to understand the asymptotic behavior of the products $p_n := g_n \cdots g_1$. We assume, to simplify this introduction, that

$$\begin{array}{l} - \mu \textit{ has a finite exponential moment,} \\ - \Gamma_\mu \textit{ is unbounded and acts strongly irreducibly on } V. \end{array} \tag{1.12}$$

Among these assumptions, *finite exponential moment* means that

$$\int_G \|g\|^\alpha \, \mathrm{d}\mu(g) < \infty \quad \text{for some } \alpha > 0.$$

Notice that the word *exponential* is natural in this context if one wants this terminology to be compatible with that introduced in Sect. 1.4 and if one keeps in mind the equality $\|g\|^\alpha = e^{\alpha \log \|g\|}$.

In these assumptions, *strongly irreducible* means that no proper finite union of vector subspaces of V is Γ_μ-invariant.

These conditions are satisfied, for instance, when

$$\mu = \tfrac{1}{2}(\delta_{a_0} + \delta_{a_1}), \quad \text{where } a_0 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \text{ and } a_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

or, more generally, where

$$a_0 = \begin{pmatrix} 2 & 1 & 0 & . & 0 \\ 1 & 1 & 0 & . & 0 \\ 0 & 0 & 1 & . & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & . & 1 \end{pmatrix} \text{ and } a_1 = \begin{pmatrix} 0 & -1 & 0 & . & 0 \\ 0 & 0 & -1 & . & 0 \\ 0 & 0 & 0 & . & 0 \\ . & . & . & . & -1 \\ 1 & 0 & 0 & . & 0 \end{pmatrix}. \tag{1.13}$$

In this example, one has $A = \{a_0, a_1\}$ and we are choosing the g_k to be either a_0 or a_1 with equal probability and independently of the previous choices of g_j for $j < k$. The partial products $p_n := g_n \cdots g_1$ can take 2^n values with equal probability. The precise value of these two matrices a_0 and a_1 are not very important: they have just been chosen to satisfy the condition (1.12). This kind of concrete example is very interesting to keep in mind. Indeed, the whole machinery we are going to explain in this book is necessary to understand the asymptotic behavior of p_n even for a law μ as simple as the one given by this example.

We denote by $\lambda_1 = \lambda_{1,\mu}$ the first Lyapunov exponent of μ, i.e.

$$\lambda_1 := \lim_{n\to\infty} \frac{1}{n} \int_G \log \|g\| \, \mathrm{d}\mu^{*n}(g). \tag{1.14}$$

The first result tells us that the variables $\log \|p_n v\|$ satisfy the Law of Large Numbers. It is due to Furstenberg.

Theorem 1.6 (LLN) *For all v in $V \smallsetminus \{0\}$, for β-almost all b in B, one has*

$$\lim_{n\to\infty} \tfrac{1}{n} \log \|g_n \cdots g_1 v\| = \lambda_1, \ \textit{ and one has } \ \lambda_1 > 0. \tag{1.15}$$

The second result tells us that the variables $\log \|p_n v\|$ satisfy the Central Limit Theorem, i.e. that the renormalized variables $\frac{\log \|p_n v\| - n\lambda_1}{\sqrt{n}}$ converge in law to a nondegenerate Gaussian variable.

Theorem 1.7 (CLT) *The limit*

$$\Phi := \lim_{n\to\infty} \frac{1}{n} \int_G (\log \|g\| - n\,\lambda_1)^2 \, \mathrm{d}\mu^{*n}(g)$$

exists and is positive $\Phi > 0$. For all v in $V \smallsetminus \{0\}$, for any bounded continuous function ψ on $\mathbb{R}$, one has

$$\lim_{n\to\infty} \int_G \psi\left(\tfrac{\log \|gv\| - n\lambda_1}{\sqrt{n}}\right) \mathrm{d}\mu^{*n}(g) = \int_{\mathbb{R}} \psi(s) \frac{e^{-\frac{s^2}{2\Phi}}}{\sqrt{2\pi\,\Phi}} \, \mathrm{d}s. \tag{1.16}$$

The third result tells us that the variables $\log \|p_n v\|$ satisfy a law of the iterated logarithm.

Theorem 1.8 (LIL) *For all v in $V \smallsetminus \{0\}$, for β-almost all b in B, the set of cluster points of the sequence*

$$\frac{\log \|g_n \cdots g_1 v\| - n\lambda_1}{\sqrt{2\Phi\, n \log\log n}}$$

is equal to the interval $[-1, 1]$.

The fourth result tells us that the variables $\log \|p_n v\|$ satisfy a Large Deviations Principle.

Theorem 1.9 (LDP) *For all v in $V \smallsetminus \{0\}$, for any $t_0 > 0$, one has*

$$\limsup_{n\to\infty} \mu^{*n}(\{g \in G \mid |\log \|gv\| - n\lambda_1| \geq n t_0\})^{\frac{1}{n}} < 1. \tag{1.17}$$

The fifth result tells us that the variables $\log \|p_n v\|$ satisfy a Local Limit Theorem.

Theorem 1.10 (LLT) *For all $a_1 \le a_2$, for all v in $V \smallsetminus \{0\}$, one has*

$$\lim_{n\to\infty} \sqrt{n}\, \mu^{*n}(\{g \in G \mid \log \|gv\| - n\lambda_1 \in [a_1, a_2]\}) = \frac{a_2 - a_1}{\sqrt{2\pi \Phi}}.$$

Theorems 1.7 up to 1.10 are in Le Page's thesis under technical assumptions. Since then, the statements have been extended and simplified by Guivarc'h, Raugi, Goldsheid, Margulis, and the authors.

1.6 How Does One Prove These Nice Results?

Thanks for your enthusiasm. As for sums of random numbers, we will use tools coming from Probability Theory like the Doob Martingale Theorem, tools coming from Ergodic Theory like the Birkhoff Ergodic Theorem and tools coming from Harmonic Analysis like the Fourier Inversion Theorem.

New tools will be needed. We will be able to understand the asymptotic behavior of the product p_n of iid random matrices by first studying the associated Markov chain on the projective space $\mathbb{P}(V)$ whose trajectories, starting from $x = \mathbb{R}v$, are $n \mapsto x_n := p_n x$. We will also study the ergodic properties along these trajectories of the cocycle σ_1 on $\mathbb{P}(V)$ given by

$$\sigma_1(g, x) = \log \tfrac{\|gv\|}{\|v\|}.$$

Indeed, for a vector v of norm $\|v\| = 1$, the quantity $s_n := \log \|p_n v\|$ that we want to study is nothing else than the sum

$$\log \|p_n v\| = \sum_{k=1}^{n} \sigma_1(g_k, x_{k-1}).$$

These random real variables $t_k := \sigma_1(g_k, x_{k-1})$, whose sum is s_n, are not always independent because the point x_{k-1} depends on what happened before. This is why we will need tools from Markov chains.

First we have to understand the statistics of the trajectories x_k, i.e. we have to answer Question (1.4). That is why we will study the invariant probability measures ν of this Markov chain, i.e. the probability measures ν on $\mathbb{P}(V)$ which satisfy $\mu * \nu = \nu$. Those probability measures ν are also called *μ-stationary*. This will allow us to prove the LLN and to give a formula for the drift analog to (1.7):

$$\lambda_1 = \int_{G\times\mathbb{P}(V)} \sigma_1(g, x)\, \mathrm{d}\mu(g)\, \mathrm{d}\nu(x). \tag{1.18}$$

This formula is due to Furstenberg.

We will see that, when the action of Γ_μ on V is *proximal* the invariant probability measure ν on $\mathbb{P}(V)$ is unique. The assumption *proximal* means that there exists a rank-one matrix which is a limit of matrices $\lambda_n \gamma_n$ with $\lambda_n > 0$ and γ_n in Γ_μ. In this case Furstenberg's formula (1.18) reflects the fact that, for all starting points x in $\mathbb{P}(V)$, the sequence $(x_n)_{n\geq 1}$ becomes equidistributed according to the law ν, for β-almost all b. When Γ_μ is not proximal, the asymptotic behavior of the sequence $(x_n)_{n\geq 1}$ is described in [17].

Second we have to understand the *transfer operator* P and its generalization the *complex transfer operator* $P_{i\theta}$ with $\theta \in \mathbb{R}$. This operator $P_{i\theta}$ is the bounded operator on $\mathscr{C}^0(\mathbb{P}(V))$, given by, for any φ in $\mathscr{C}^0(\mathbb{P}(V))$ and any x in $\mathbb{P}(V)$,

$$P_{i\theta}\varphi(x) = \int_G e^{i\theta\,\sigma_1(g,x)}\varphi(gx)\,\mathrm{d}\mu(g). \tag{1.19}$$

The CLT 1.7 describes the asymptotic behavior of the probability measures on $\mathbb{R}$

$$\mu_{n,x} := \text{image of } \mu^{*n} \text{ by the map } g \mapsto \log \tfrac{\|gv\|}{\|v\|}.$$

The Fourier transform of these measures is given by the classical and elegant formula with θ in $\mathbb{R}$,

$$\widehat{\mu_{n,x}}(\theta) = P_{i\theta}^n \mathbf{1}(x), \tag{1.20}$$

where $\mathbf{1}$ is the constant function on $\mathbb{P}(V)$ equal to 1. The behavior of the right-hand side of this formula will be controlled by the "largest" eigenvalue of $P_{i\theta}$. This formula (1.20) explains how spectral data from the complex transfer operator $P_{i\theta}$ can be used in combination with the Fourier Inversion Theorem to prove not only the CLT but also the LIL, the LDP and the LLT. We will be able to reduce our analysis to the case where the action of Γ_μ on V is proximal. We will see then that this operator $P_{i\theta}$ has a unique "largest" eigenvalue $\lambda_{i\theta}$ when θ is small, and that this eigenvalue $\lambda_{i\theta}$ varies analytically with θ.

1.7 Can You Answer Your Own Questions Now?

You are right, what took us so long are nothing but the answers to Questions (1.4) and (1.5). We will deduce the answers to Questions (1.2) and (1.3) from these.

Indeed, we will first check that, under assumption (1.12), the random variables $\log\|p_n\|$ satisfy the same LLN, CLT, LIL and LDP as $\log\|p_n v\|$. Technically, this will not be too difficult since these four limit laws involve a renormalization which will erase the difference between $\log\|p_n\|$ and $\log\|p_n v\|$.

We will also check that, when, moreover, Γ_μ is proximal, the random variables $\log|f(p_n v)|$ satisfy the same LLN, CLT, LIL and LDP as $\log\|p_n v\|$. This will be more delicate since we will have to control the excursions of the sequence $p_n x$ near the kernel of f. The key point will be to prove a Hölder regularity result for the stationary measure ν which is due to Guivarc'h.

1.8 Where Can I Find These Answers in This Book?

The LLN for $\log\|p_n v\|$ and $\log\|p_n\|$ are in Sects. 4.6 and 4.7.
The LLN for $\log|f(p_n v)|$ is in Sect. 14.4.
The CLT, LIL, LDP for $\log\|p_n v\|$ and $\log\|p_n\|$ are in Sect. 14.7.
The CLT, LIL, LDP for $\log|f(p_n v)|$ are in Sect. 14.8.
The LLT for $\log\|p_n v\|$ and $\log\|p_n\|$ are in Sects. 17.5.

1.9 Why Is This Book Less Simple than These Samples?

The quantity

$$\kappa_1(g) := \|g\|$$

gives us information on the size of a matrix g only "in one direction". It is much more useful in the applications to deal with all the singular values $\kappa_j(g) := \frac{\|\wedge^j(g)\|}{\|\wedge^{j-1}(g)\|}$ and to introduce the "multinorm"

$$\kappa_V(g) := (\log\kappa_1(g), \ldots, \log\kappa_d(g)). \tag{1.21}$$

A less naive way to ask our question (1.1) is:

$$\textit{Can one describe the asymptotic behavior of } \kappa_V(p_n)\textit{?} \tag{1.22}$$

The answer to this question is Yes! These random variables $\kappa_V(p_n)$ satisfy the LLN with average λ. However they do not exactly satisfy a CLT: the renormalized variable $\frac{\kappa_V(p_n)-n\lambda}{\sqrt{n}}$ converges in law but the limit law is only a "folded Gaussian law", i.e. the "image of a Gaussian law by a homogeneous continuous locally linear map"!

The support of this limit law depends only on λ and the "Zariski closure" G_μ of the semigroup Γ_μ. This Zariski closure G_μ is always a reductive algebraic group with compact center. The "folding" phenomenon occurs already when $d = 4$ and $G_\mu = \mathrm{SO}(2, 2)$!

The whole picture becomes much clearer when one adopts the following more intrinsic point of view.

We start with a connected real semisimple algebraic group, call it again G, and a Borel probability measure μ on G. We consider iid random variables $g_n \in G$ of law μ and want, again, to describe the asymptotic behavior of the products $p_n := g_n \cdots g_1$. In this point of view, we forget about the embedding ρ of G in $\mathrm{GL}(V)$ which was responsible for the folding of the Gaussian law. We replace the conditions (1.12) by

$$\begin{aligned} &\text{– } \mu \textit{ has a finite exponential moment,} \\ &\text{– } \textit{the semigroup } \Gamma_\mu \textit{ spanned by } A \textit{ is Zariski dense in } G, \end{aligned} \tag{1.23}$$

where A is the support of μ.

The projective space $\mathbb{P}(V)$ is replaced by the flag variety $\mathscr{P}$ of G, and the norm is replaced by the *Cartan projection* κ of G. Exactly as in Sect. 1.6, we will use a cocycle $\sigma(g,\eta)$ on the flag variety $\mathscr{P}$, called the Iwasawa or Busemann cocycle. The Iwasawa cocycle σ takes its values in a real vector space $\mathfrak{a}$ called the *Cartan subspace*, whose dimension is the *real rank* r of G. The Cartan projection κ takes its values in a simplicial cone $\mathfrak{a}^+$ of $\mathfrak{a}$ called the *Weyl chamber*. The precise definitions will be given later. For every η in $\mathscr{P}$, the asymptotic behavior of $\kappa(p_n)$ will be related to the asymptotic behavior of $\sigma(p_n,\eta)$. Our questions now become

$$\textit{What is the asymptotic behavior of } \kappa(p_n) \textit{ and } \sigma(p_n,\eta)\textit{?} \tag{1.24}$$

We will see that the random variables $\sigma(p_n,\eta)$ and $\kappa(p_n)$ satisfy the LLN, CLT, LIL and LDP. We will also check the LLT for the random variables $\sigma(p_n,\eta)$.

1.10 Can You State These More General Limit Theorems?

Here are the statements for the Iwasawa cocycle σ. The assumptions on μ are given in (1.23).

Theorem 1.11 (LLN) *There exists a unique μ-stationary probability measure ν on $\mathscr{P}$. The average*

$$\sigma_\mu := \int_{G\times\mathscr{P}} \sigma(g,\eta)\,\mathrm{d}\mu(g)\,\mathrm{d}\nu(\eta)$$

belongs to the interior of the Weyl chamber $\mathfrak{a}^+$.

For η in $\mathscr{P}$, for β-almost all b in B, one has

$$\lim_{n\to\infty} \tfrac{1}{n}\sigma(g_n\cdots g_1,\eta) = \sigma_\mu .$$

This multidimensional version of Theorem 1.6 is due to Guivarc'h–Raugi and Goldsheid–Margulis. An important new output there is the fact that the Lyapunov vector σ_μ belongs to the interior of the Weyl chamber $\mathfrak{a}^+$.

Theorem 1.12 (CLT) *There exists a Euclidean norm $\|\cdot\|_\mu$ on $\mathfrak{a}$ such that, for all η in $\mathscr{P}$, for any bounded continuous function ψ on $\mathfrak{a}$,*

$$\lim_{n\to\infty}\int_G \psi\left(\tfrac{\sigma(g,\eta)-n\sigma_\mu}{\sqrt{n}}\right)\mathrm{d}\mu^{*n}(g) = (2\pi)^{-r/2}\int_{\mathfrak{a}}\psi(v)e^{-\frac{\|v\|_\mu^2}{2}}\,\mathrm{d}\pi_\mu(v),$$

where $\mathrm{d}\pi_\mu(v) = \mathrm{d}v_1\cdots\mathrm{d}v_r$ in an orthonormal basis for $\|\cdot\|_\mu$.

This multidimensional version of Theorem 1.7 is due to Guivarc'h and Goldsheid. An important new output there is the fact that the support of the limit Gaussian law is the whole Cartan subspace $\mathfrak{a}$.

Here are the multidimensional versions of Theorems 1.8, 1.9 and 1.10.

Theorem 1.13 (LIL) *For all η in $\mathscr{P}$, for β-almost all b in B, the set of cluster points of the sequence*

$$\frac{\sigma(g_n \cdots g_1, \eta) - n\sigma_\mu}{\sqrt{2n \log\log n}}$$

is equal to the unit ball K_μ of $\|\cdot\|_\mu$.

Theorem 1.14 (LDP) *For any $t_0 > 0$, one has*

$$\limsup_{n\to\infty} \sup_{\eta\in\mathscr{P}} \mu^{*n}(\{g \in G \mid \|\sigma(g, \eta) - n\sigma_\mu\| \geq nt_0\})^{\frac{1}{n}} < 1.$$

Theorem 1.15 (LLT) *For all bounded open convex sets C of $\mathfrak{a}$, for all η in $\mathscr{P}$ belonging to the support of ν, one has*

$$\lim_{n\to\infty} (2\pi\, n)^{r/2}\, \mu^{*n}(\{g \in G \mid \sigma(g, \eta) - n\sigma_\mu \in C\}) = \pi_\mu(C).$$

It is remarkable that, in Theorem 1.15, no further "aperiodicity" assumptions have to be made as in Theorem 1.5. This will follow from a general fact for "Zariski dense subgroups of semisimple Lie groups" in [11].

We will also prove a version of this local limit theorem where we allow moderate deviation, i.e. where we allow the "window" C to be translated by a vector $v_n \in \mathfrak{a}$ as soon as $\|v_n\|$ do not grow faster than $\sqrt{n \log n}$. Indeed this version, which adapts Breuillard's LLT for sums of iid real numbers in [30], is the one which is needed in [15].

1.11 Are the Proofs as Simple as for the Simple Samples?

Well, ... at least the proofs of these five theorems follow the same lines as in Sect. 1.6.

First we study the associated Markov chain on the flag variety $\mathscr{P}$. Since this flag variety is equivariantly embedded in the product of projective spaces on which the action of Γ_μ is "proximal", we will be able to use results previously proven for these proximal actions.

Second, we study the spectral properties of the complex transfer operator. This operator $P_{i\theta}$ is defined for any $\theta \in \mathfrak{a}^*$. It is the bounded operator on $\mathscr{C}^0(\mathscr{P})$, given, for any φ in $\mathscr{C}^0(\mathscr{P})$ and η in $\mathscr{P}$, by the following formula similar to (1.19),

$$P_{i\theta}\varphi(\eta) = \int_G e^{i\theta(\sigma(g,\eta))}\varphi(g\eta)\, \mathrm{d}\mu(g).$$

Another consequence of the contraction property of the action on $\mathscr{P}$ will again be the existence of a unique "largest" eigenvalue $\lambda_{i\theta}$ for the operator $P_{i\theta}$ when θ is small, and the fact that this eigenvalue $\lambda_{i\theta}$ varies analytically with θ.

The CLT 1.12 for the Iwasawa cocycle σ describes the asymptotic behavior of the probability measures on $\mathfrak{a}$

$$\mu_{n,\eta} := \text{image of } \mu^{*n} \text{ by the map } g \mapsto \sigma(g,\eta).$$

The Fourier transform of these measures is given by the classical and elegant formula similar to (1.20), with θ in $\mathfrak{a}^*$,

$$\widehat{\mu_{n,\eta}}(\theta) = P_{i\theta}^n \mathbf{1}(\eta). \tag{1.25}$$

Thanks to this formula, we can use, as in Sect. 1.6, the uniqueness of the "largest" eigenvalue of the complex transfer operator $P_{i\theta}$, in combination with the Fourier Inversion Theorem, to prove the CLT for the Iwasawa cocycle σ.

This intrinsic approach allows us to answer Question (1.5) not only when the action of the semigroup Γ_μ on $\mathbb{R}^d$ is irreducible but also when this action is semisimple, i.e. when every Γ_μ-invariant vector subspace of $\mathbb{R}^d$ admits a Γ_μ-invariant complementary subspace.

1.12 Why Is the Iwasawa Cocycle so Important to You?

Both the Cartan projection and the Iwasawa cocycle are important to us. We recall that they are constructed thanks to the Cartan decomposition and the Iwasawa decomposition of a connected real reductive algebraic group

$$G = K \exp \mathfrak{a}^+ K \quad \text{and} \quad G = K \exp \mathfrak{a} N.$$

Here K is a maximal compact subgroup of G, exp is the exponential map of G, $\mathfrak{a}$ is a Cartan subspace of the Lie algebra $\mathfrak{g}$ of G that is orthogonal to the Lie algebra $\mathfrak{k}$ of K with respect to the Killing form, $\mathfrak{a}^+$ is a Weyl chamber in $\mathfrak{a}$, and N is the corresponding unipotent subgroup of G. Let M be the centralizer of $\mathfrak{a}$ in K. With these notations, the *flag variety* is the quotient space

$$\mathscr{P} = G/P, \quad \text{where} \quad P = M \exp \mathfrak{a} N$$

is the normalizer of N. This group P is called the *minimal parabolic subgroup* associated to $\mathfrak{a}^+$.

The precise formulae defining κ and σ are, for g in G and η in $\mathscr{P}$,

$$g \in K e^{\kappa(g)} K \quad \text{and} \quad gk \in K e^{\sigma(g,\eta)} N,$$

where k in K is chosen so that $k^{-1}\eta$ is N-invariant.

For instance, when $G = \mathrm{GL}(d,\mathbb{R})$, one can take $\mathfrak{a}$ to be the space of diagonal matrices, $\mathfrak{a}^+$ the subset of diagonal matrices with non-increasing coefficients, K to be $\mathrm{SO}(d,\mathbb{R})$, and N the group of upper triangular unipotent matrices. In this case the

Cartan decomposition is the "polar decomposition", the Cartan projection κ is the multinorm κ_V given by formula (1.21), and the Iwasawa decomposition is obtained by the "Gram-Schmidt orthonormalization process".

For g in G, the Cartan projection $\kappa(g)$ is important because it simultaneously controls for all representations ρ of G the norms of the matrices $\rho(g)$. Similarly, for g in G and η in $\mathscr{P}$, the Iwasawa cocycle $\sigma(g, \eta)$ is important because it controls simultaneously the norms of all vectors $\frac{1}{\|v\|}\rho(g)v$ when $\mathbb{R}v$ is a line invariant under the stabilizer of η. More precisely, one has the following fact:

When (V, ρ) is an irreducible algebraic representation of G, one has, for a suitable K-invariant norm on V, the equalities, for all g in G, η in $\mathscr{P}$, and every line $\mathbb{R}v$ in V which is invariant under the stabilizer of η,

$$\log \|\rho(g)\| = \chi(\kappa(g)) \ \text{ and } \ \log \frac{\|\rho(g)v\|}{\|v\|} = \chi(\sigma(g, \eta)),$$

where the linear functional $\chi \in \mathfrak{a}^*$ is the "highest weight" of V.

Because of this fact, the five theorems of Sect. 1.10 are multidimensional extensions of the five theorems of Sect. 1.5.

1.13 I Am Allergic to Local Fields. Is It Safe to Open This Book?

In this text we will not only study the asymptotic behavior of product of iid random real matrices, but we will allow the coefficients of these matrices to be in any local field $\mathbb{K}$. We recall that a local field $\mathbb{K}$ is a finite extension of either the field of p-adic numbers $\mathbb{Q}_p$, the field of Laurent series $\mathbb{F}_p((T))$ with coefficients in the finite field $\mathbb{F}_p$, where p is prime number, or the field $\mathbb{Q}_\infty = \mathbb{R}$.

For a first reading, you can assume that $\mathbb{K} = \mathbb{R}$. Except in very few places that we will point out, the proofs are no simpler over $\mathbb{R}$ than they are over any local field $\mathbb{K}$. A reader more familiar with local fields may assume that $\mathbb{K} = \mathbb{R}$ or $\mathbb{Q}_p$ since all the difficulties already occur in these cases.

So you may wonder in the first place why we want to state these results over local fields. The reason is that those extended results give new information of an arithmetic flavor. For instance when the support of the law μ consists of finitely many matrices in $\mathrm{SL}(d, \mathbb{Q})$, the coefficients of the random products p_n are rational numbers. The results over $\mathbb{K} = \mathbb{R}$ give information on the size of these coefficients while the extended results over $\mathbb{K} = \mathbb{Q}_p$ give information on the size of the denominators of these coefficients, and more precisely on the powers of the prime number p which occur in these denominators.

As a by-product of this point of view, we will be see that the five limit theorems we quoted in Sect. 1.5 can be adapted over any local field $\mathbb{K}$, even in positive characteristic, except that the variance Φ might be equal to 0 (see Sect. 14.7).

1.14 Why Are There so Many Chapters in This Book?

Sometimes chapters are related in pairs, the first one dealing with general cocycles over semigroup actions, the second one applying these general results to products of random matrices.

In Chap. 2, we recall basic facts on Markov chains.

In Chap. 3, we prove the LLN for cocycles over a semigroup action.

In Chap. 4, we prove the LLN for products of random matrices.

In Chap. 5, we explain how to induce a random walk to a finite index subsemigroup.

In Chap. 6, we check that Zariski dense semigroups in semisimple real Lie groups always contain loxodromic elements.

In Chap. 7, we focus on the Jordan projection of Zariski dense semigroups in semisimple real Lie groups.

In Chap. 8, we recall a few basic facts on reductive algebraic groups over local fields, their algebraic representations, their flag varieties, their Iwasawa cocycle and their Cartan projection.

In Chap. 9, we study the Zariski dense semigroups in algebraic reductive $\mathscr{S}$-adic Lie groups.

In Chap. 10, we reformulate the LLN for products of random matrices in the intrinsic language of Chap. 8.

In Chap. 11, we study the spectral properties of the complex transfer operator for a cocycle over a contracting semigroup action.

In Chap. 12, we prove the CLT, LIL and LDP for a cocycle over a contracting semigroup action.

In Chap. 13, we deduce the CLT, LIL and LDP for the Iwasawa cocycle and the Cartan projection.

In Chap. 14, we give a short proof of the Hölder regularity of the stationary measure on the flag variety. We apply it to prove the LLN, CLT, LIL and LDP for the coefficients and for the spectral radius.

In Chap. 15, we study more deeply the spectral properties of the complex transfer operator.

In Chap. 16, we prove the LLT for a cocycle over a contracting semigroup action.

In Chap. 17, we deduce the LLT for the Iwasawa cocycle. We apply it to prove the LLT for the Cartan projection, and for the norm of vectors.

In Appendix A, we recall basic facts on Martingales and their applications to the LLN for “sums of random numbers”.

In Appendix B, we recall basic facts on bounded operators in Banach spaces, their spectrum and their essential spectrum. These facts are used in the proof of the Local Limit Theorem.

In Appendix C, we quote our sources.

1.15 Whom Do You Thank?

Institutions, referees, colleagues, students, friends, and families who financed us, teased us, helped us, read us, encouraged us, and supported us.

Springer

http://www.springer.com/978-3-319-47719-0

Random Walks on Reductive Groups
Benoist, Y.; Quint, J.-F.
2016, XI, 323 p., Hardcover
ISBN: 978-3-319-47719-0