Some topics at the interface between probability theory, algebro-geometric structures, and mathematical physics

Reda Chhaibi

Manuscrit
présenté pour l'obtention de
L'Habilitation à Diriger des Recherches
par
Reda CHHAIBI
de l'Institut de Mathématiques de Toulouse

Some topics at the interface between probability theory, algebro-geometric structures, and mathematical physics.

Soutenue le 21 mars 2024, devant le jury suivant :

Rapporteurs :

Philippe BIANE	DR CNRS, Université Gustave-Eiffel
Patricia GONCALVES	Professeur, Técnico Lisboa
Alice GUIONNET	DR CNRS, ENS Lyon

Examinateurs :

Mireille CAPITAINE	DR CNRS, IMT, Toulouse III
Gersende FORT	DR CNRS, IMT, Toulouse III
Grégory SCHEHR	DR CNRS, LPTHE, Sorbonne

Mentor :

Michel LEDOUX	Professeur émérite, IMT, Toulouse III

Membres invités du jury:

Philippe BOUGEROL	Processeur émérite, LPSM, Sorbonne
Francesco COSTANTINO	Professeur, IMT, Toulouse III

Contents

Remerciements / Acknowledgements in French

Si le document est écrit en anglais, les remerciements sont dans ma langue de coeur.

Je dédie cette thèse à la ville de Toulouse, qui m'a accueilli il y a de cela bientôt 10 ans et qui m'a nourri de rencontres enrichissantes. Je considère que c'est une chance d'y avoir passé une grande partie de ma carrière académique. De ce point de vue, l'Institut de Mathématiques de Toulouse joue un role particulier pour m'avoir accueilli et soutenu dans mes recherches.

Je tiens à exprimer ma profonde gratitude envers les rapporteurs Philippe Biane, Patricia Goncalves et Alice Guionnet pour avoir pris la peine de lire ce document, in fine, bien plus long que j'aurai souhaité. Leurs rapports m'ont profondément touché.

Je souhaite également remercier chaleureusement les autres membres du jury pour leur implication dans l'évaluation de mon travail. Mireille Capitaine a toujours été un repère constant à l'institut. Michel Ledoux me fait l'honneur d'être mon "parrain" d'HDR, et a souvent été de bon conseil lors de moments de doute et de questionnement. Gersende Fort et Grégory Schehr me font le privilège d'incarner une ouverture vers les statistiques d'une part, et vers la physique théorique d'autre part.

Un grand merci s'adresse également à mes co-auteurs, dont les qualités scientifiques et humaines ont rendu la recherche non seulement stimulante, mais aussi agréable: Tristan Benoist, Cédric Bernardin, François Chapon, Raphaël Chetrite, Tariq Daouda, Ibrahim Ekren, Emma Hovhannisyan, Ezechiel Kahn, Pierre-Loic Méliot, Thomas Madaule, Joseph Najnudel, Ashkan Nikeghbali, Clément Pellegrini, Brad Rodgers.

Les membres invités sont Philippe Bougerol, mon directeur de thèse, qui me fait le plaisir de continuer de s'intérresser à mes travaux, et Francesco Costantino, avec qui j'encadre une thèse, et avec qui nous discutons si souvent de TQFT sans pour autant arriver à fixer nos idées.

À mesure que le temps passe, je prends conscience de plus en plus de la qualité des chercheurs qui m'entourent. J'espère un jour pouvoir contribuer, à hauteur de ne serait-ce qu'à une fraction, de ce qu'ils ont fait pour moi.

Enfin, je souhaite exprimer ma reconnaissance envers la famille Pinel, à qui cette HDR doit son écriture, à la montagne noire, dans le Cabardès.

CHAPTER 1

Introduction

1.1. Foreword

The goal of this document is to draw a coherent picture of multiple contributions to mathematics over the past 10 years.

Indeed, since my PhD defense in January 2013, I engaged in multiple questions with the desire to learn as much mathematics as possible. From various remarks from colleagues, it was more or less clear that my choices in mathematics were perceived as incoherent, which I do not really mind. I had decided that this habilitation thesis would be the occasion to lift such misunderstandings.

Beyond this rather serious goal, I also decided I wanted to have fun in writing this document. As such, I will detail amusing numerical experiments and write amusing proofs that I accumulated over the years, yet that never made into in papers. And I hope that a habilitation thesis with novel proofs is not too strange, as the underlying intention never was to be original. Perhaps the perspective of telling old stories felt unbearably boring. In fact, after finishing writing this thesis, I realized I spent more time writing down and polishing these proofs that never made it in papers. Hopefully these will be basis of subsequent work.

Finally, unlike mathematical papers where one avoids the using the pronoun "I", a habilitation thesis is more personal. Therefore I will not shy away from making personal statements using that pronoun. All such opinions are entirely mine, and it does not reflect what my co-authors think for example. That should not detract from the fact that I am immensely indebted to them, for the commitment, the hard work and the passion in sharing mathematics.

1.2. Structure of the document

In Chapter 2, we discuss our interests in Random Matrix Theory (RMT). The guiding principle has always been to understand the couplings for growing families of groups or homogenous spaces. Since these groups or homogenous spaces are realized as matrices or act on matrix spaces, understanding their geometry and how they are nested is essential for RMT. If n is the index growing to infinity, the coupling method gives particularly tractable descriptions of the large n limits. The main results are follow the contributions of [**CNN17**, **CHN$^+$19**, **CMN18**, **CN19**].

In Chapter 3, we discuss the Toda integrable system and its relationship to RMT. More specifically we explain how its scattering unveils the logarithmic interaction of Coulomb gases.

In Chapter 4, I go back to Pitman-type theorems and the Representation Theory (RT) of Lie groups, which was my "madeleine de Proust" during the PhD years. The relationship to RMT is actually straightforward for mathematicians familiar with

Kirillov's orbit method. It says that in a very precise sense, RT of Lie groups is the quantization of RMT.

Because that fact is often treated as folklore by the community, I chose to give an explicit exposition of that in Section 4.2. There we start from the spherical integrals often used in RMT to integrate out any unitary symmetry. And finishing with two precise statements about semi-classical limits of representations.

Then in Section 4.3, I explain the contributions of the paper [CC21]. It focuses on the group $G = SL_2(\mathbb{C})$, arguably the simplest, or rather it focuses on the quantum group version. I wanted to fully understand how the representation theory of the quantum group $\mathcal{U}_q(\mathfrak{s}l_2)$ and its semi-classical limit, while keeping track of a certain curvature limit.

Finally, in Section 4.4, I explain how the aforementioned representation theory allows to bridge RMT and certain integrable models in mathematical physics: directed last Passage percolation and directed polymers. In fact, a landmark result in that area due to Johannson [Joh00, Proposition 1.4] is reinterpreted as a Pitman-type theorem.

In the Chapter 5, we leave the topic of quantization while remaining the topic of quantum mechanics, this time with the problem of measurement and wavefunction collapse. More precisely we present work on the strong noise limits of SDEs following [BCC+23, BBC+21, BCNP22]. The community is in fact motivated by quantum mechanics and more precisely quantum open systems. In terms of personal motivation, I feel that this topic nicely complements the previous chapter. Indeed, we move from the quantization of (very algebraic) Hamiltonian systems, to the the mathematical challenges behind the measurement axiom. That was also a very natural endeavor as the required tool is the theory of stochastic processes – perhaps even more probabilistic than I initially thought!

Finally, I conclude with perspectives in Chapter 6. There I explain current and future projects. First, there is the use of the Toda integrable system in RMT and in order to start dabbling with hydrodynamical limits, following recent work of mathematical physicist Spohn. Second, there is some musing on the semi-classical limits of quantum groups at roots of unity. Third, I describe an ongoing project in Optimal Transport. Fourth, I detail ongoing projects in statistics and learning theory.

For convenience, the selection of papers treated in this habilitation is reproduced at the end of this introduction chapter. They are nevertheless present in the bibliography just like the other references as [CNN17, CMN18, CHN+19, CN19, CC21, BCC+23, BBC+21, BCNP22, CDK22]

1.3. What is new and what is not.

For the sake of full disclosure and transparency, I would like to make clear what is new and what is not in this document. By new, I mean what is not already published in a peer-reviewed journal nor available on the arxiv as preprints. The following content is new in that sense, and everything else appears in papers.

First, all of Chapter 3 is extracted from various notes that I accumulated over the years. While the RMT result it contains is well-established thanks to Dumitriu-Edelman [DE02], one could argue that the proof is original. Yet I never even posted this on the arxiv, as I had a different goal in mind – which does not work.

Second, there is the semi-classical limit form of Theorem 4.2.4 for one orbit and Theorem 4.2.5 for the convolution of two orbits. These are definitely known to specialists, at least in the form when observables $\mathcal{F}$ belong to the maximal commutative algebra. Yet they are the theorems I would have liked to see written and proved with care 10 years ago. To me they embody Kirillov's orbit method in a language that is both probabilistic and geometric. These theorems correspond to the flat curvature setting of my paper [**CC21**].

Third, Subsection 4.3.3 and Section 4.4 give a self-contained proof of Johansson's result entirely based on spherical transforms and our curvature deformation point of view. That is perhaps the proof I had the most fun writing. Parts of it did surprise me.

Finally, let me mention that the Chapter 6 of perspectives is mandatory for a habilitation thesis, by rules of the doctoral school in Toulouse. It contains a summary of ongoing projects, including works done with two of my PhD students:

- Anirban BOSE who is working on computational aspects of optimal tranport.
- Alexey LAZAREV who is working on the use of Ricci-type flows for the regularization of latent spaces in machines learning.

1.4. Ariane's threads

In order to flesh out some coherence in this body of research, it is fair to say that while my interest is always inherently probabilistic, I find probability all the more interesting that it has something to say about:

- Quantum mechanics: Quantization and the measurement problem in quantum mechanics.
- Harmonic analysis: Representation theory is nothing but the study of noncommutative Fourier transforms. Spherical transforms are exactly Fourier transforms under additionnal symmetries.
- Couplings of random variables: Whether it is the coupling of permutations, unitary matrices or stochastic processes, or the systematic study of couplings via optimal transport.

In Fig. 1.4.1 is a tentative graph of the topics I have been interested in, and how they relate to each other.

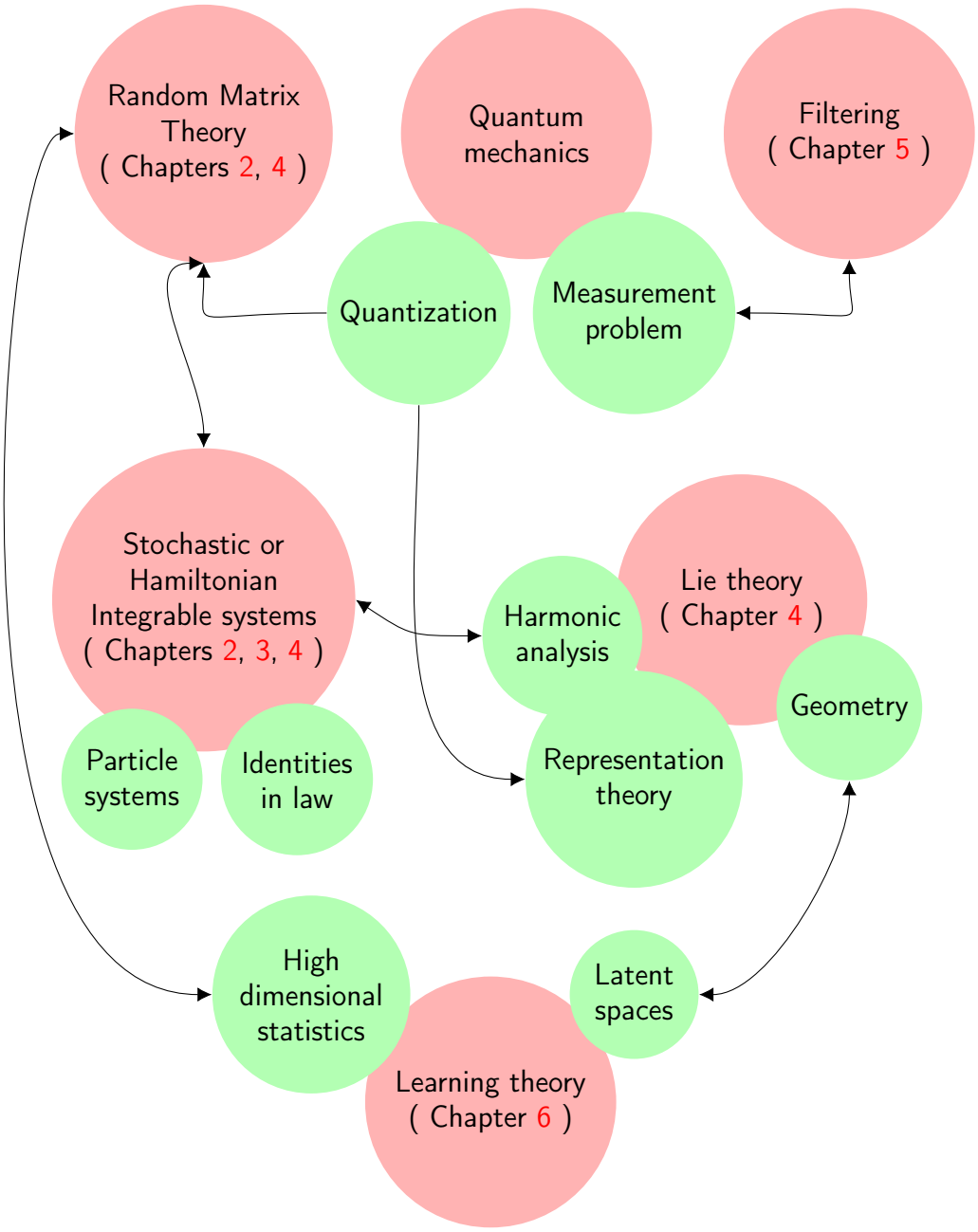

FIGURE 1.4.1. Sketch of my research interests and their interactions. In pink are large topics, which often correspond to chapters of this thesis. In green are relevant subtopics, which tend to interact with other fields.

Selected papers for this habilitation thesis

[Chh1] Reda Chhaibi, Joseph Najnudel, and Ashkan Nikeghbali. The circular unitary ensemble and the riemann zeta function: the microscopic landscape and a new approach to ratios. *Inventiones mathematicae*, 207:23–113, 2017.

[Chh2] Reda Chhaibi, Thomas Madaule, and Joseph Najnudel. On the maximum of the CβE field. 2018.

[Chh3] Reda Chhaibi, Emma Hovhannisyan, Joseph Najnudel, Ashkan Nikeghbali, and Brad Rodgers. The limiting characteristic polynomial of classical random matrix ensembles. In *Annales Henri Poincaré*, volume 20, pages 1093–1119. Springer, 2019.

[Chh4] Reda Chhaibi and Joseph Najnudel. On the circle, $GMC^\gamma = \varprojlim C\beta E_n$ for $\gamma = \sqrt{\frac{2}{\beta}}$, $(\gamma \leq 1)$. *arXiv preprint arXiv:1904.00578*, 2019.

[Chh5] François Chapon and Reda Chhaibi. Quantum sl_2, infinite curvature and Pitman's 2M-X theorem. *Probability Theory and Related Fields*, 179:835–888, 2021.

[Chh6] Cédric Bernardin, Raphaël Chetrite, Reda Chhaibi, Joseph Najnudel, and Clément Pellegrini. Spiking and collapsing in large noise limits of sdes. *The Annals of Applied Probability*, 33(1):417–446, 2023.

[Chh7] Tristan Benoist, Cédric Bernardin, Raphaël Chetrite, Reda Chhaibi, Joseph Najnudel, and Clément Pellegrini. Emergence of jumps in quantum trajectories via homogenization. *Communications in Mathematical Physics*, 387(3):1821–1867, 2021.

[Chh8] Cédric Bernardin, Reda Chhaibi, Joseph Najnudel, and Clément Pellegrini. To spike or not to spike: the whims of the wonham filter in the strong noise regime. *arXiv preprint arXiv:2211.02032*, 2022.

[Chh9] Reda Chhaibi, Tariq Daouda, and Ezechiel Kahn. Free probability for predicting the performance of feed-forward fully connected neural networks. *Advances in Neural Information Processing Systems*, 35:2439–2450, 2022.

CHAPTER 2

Random Matrix Theory (RMT)

Random Matrix Theory (RMT) was pioneered by Wigner in his seminal work [Wig58] proving the semi-circle law.

2.1. Introduction

Classical matrix ensembles with unitary symmetry: For $\Lambda \in \mathbb{R}^n$, we write the Vandermonde determinant as

$$\Delta(\Lambda) := \prod_{1 \le j < k \le n} (\Lambda_j - \Lambda_k) \ .$$

We define the Wishart ensemble first: Consider a matrix of complex i.i.d. Gaussians ($\mathbb{E}\mathcal{N}^{\mathbb{C}} = 0$, $\mathbb{E}|\mathcal{N}^{\mathbb{C}}|^2 = 1$):

$$\mathbb{G}_{M,N} := \left(\mathcal{N}^{\mathbb{C}}_{i,j}\right)_{\substack{1 \le i \le N \\ 1 \le j \le M}} = [\xi_1, \dots, \xi_M] \ ,$$

where the ξ_j's are the column vectors. A Wishart matrix is the sample covariance matrix of the ξ_j's:

$$\mathbb{W}_{M,N} := \mathbb{G}_{M,N}\mathbb{G}^*_{M,N} = \sum_{j=1}^{M} \xi_j \xi_j^* \ .$$

Notice that $(\mathbb{W}_{M,N} \ ; \ M \ge 0)$ is a random walk in M, with i.i.d. isotropic increments.

Second, we have the GUE (Gaussian Unitary Ensemble). A matrix from this ensemble $\mathrm{GUE}^{(N)}$ has Gaussian entries, following the Hermitian symmetry. In particular, a process with GUE marginals is the Hermitian Brownian motion $\left(\mathrm{GUE}_t^{(N)} \ ; t \ge 0\right)$. By Donkser's invariance principle, it can be for example obtained by rescaling the Wishart random walk

$$\forall t \ge 0, \ \mathrm{GUE}_t^{(N)} = \lim_{M \to \infty} \frac{\mathbb{W}_{\lfloor Mt \rfloor, N} - \lfloor Mt \rfloor I_N}{\sqrt{M}} \ .$$

The spectrum of $\mathbb{W}_{M,N}$ and $\mathrm{GUE}_t^{(N)}$ is given by [AGZ10, Proposition 4.1.3 and Theorem 2.5.2]

$$\mathbb{P}\left(\mathrm{Spec}(\mathbb{W}_{M,N}) \in d\Lambda\right) = \frac{\mathbb{1}_{\mathbb{R}^n_+}(\Lambda)}{Z_{M,N}} \Delta(\Lambda)^2 e^{-\frac{1}{2}\sum_{j=1}^N \Lambda_j} \prod_{j=1}^N \Lambda_j^{M-N} d\Lambda \ , \tag{2.1.1}$$

$$\mathbb{P}\left(\mathrm{Spec}(\mathrm{GUE}_t^{(N)}) \in d\Lambda\right) = \frac{1}{Z_t^{(N)}} \Delta(\Lambda)^2 e^{-\frac{\|\Lambda\|^2}{2t}} d\Lambda \ . \tag{2.1.2}$$

Remark 2.1.1 (These models have independent diagonals). *In the Wishart ensemble* $\mathrm{diag}(\xi_j \xi_j^*)$ *is a vector of i.i.d. exponential random variables (Box–Muller). The GUE also has independent diagonals in the form of real Gaussians.*

Fluctuations at the edge:

DEFINITION 2.1.2 (Definition-Theorem). *The Tracy-Widom distribution is the only law whose cumulative distribution function* $F(s) = \mathbb{P}(\mathrm{TW}_2 \leq s)$ *is equivalently given by*

- *a solution to a Painlevé II equation by setting* F *to be as follows.*

$$F(s) := \exp\left(-\int_s^\infty (x-s)u(x)^2 dx\right) ,$$

with

$$\begin{cases} u'' = 2u^3 + xu \\ u(x) \sim_{x\to\infty} Ai(x) . \end{cases}$$

- *a Fredholm determinant*

$$\begin{aligned} F(s) &= \det\left(Id - Ai\right)_{L^2([s,\infty))} \\ &:= \sum_{k=1}^\infty \frac{(-1)^k}{k!} \int_{[s,\infty)^k} \det_{k\times k}\left(Ai(x_i,x_j)\right) \prod_{j=1}^k dx_j , \end{aligned}$$

where the Airy kernel is $Ai(x,y) = \frac{Ai(x)Ai'(y) - Ai(y)Ai'(y)}{x-y}$.

A landmark result which envigorated RMT among mathematicians was the proof that fluctuations of the largest eigenvalue is the Tracy-Widom distribution. More precisely, let us write $\lambda_1(M)$ as the largest eigenvalue of a normal matrix M. Then as stated in [Joh00, Theorem 1.2] and [AGZ10, Theorem 3.1.4]

THEOREM 2.1.3 (Johansson [Joh00], Tracy-Widom [TW94]). *We have the following limits in law:*

$$(NM)^{\frac{1}{6}} \left(\sqrt{N}+\sqrt{M}\right)^{\frac{2}{3}} \left(\frac{\lambda_1(\mathbb{W}_{M,N})}{\left(\sqrt{N}+\sqrt{M}\right)^2} - 1\right) \overset{N,M\to\infty}{\longrightarrow} \mathrm{TW}_2 ,$$

$$N^{\frac{2}{3}} \left(\frac{1}{\sqrt{Nt}}\lambda_1(\mathrm{GUE}_t^{(N)}) - 2\right) \overset{N\to\infty}{\longrightarrow} \mathrm{TW}_2 .$$

In the first limit, one requires the ratio N/M *to remain bounded away from zero and infinity.*

2.2. Projective measures and couplings

The coupling argument is one of the most powerful methods in probability theory, to the point that there is an entire book devoted to it [Lin02]. In our case of interest, consider a sequence of homogenous spaces or groups G_n increasing in size, along with a natural inclusions $G_n \hookrightarrow G_{n+1}$. This allows the definition of the inductive limit

$$G_\infty = \varinjlim G_n .$$

This is also called the direct limit or the injective limit, as the object where all the G_n's can be injected in.

More interesting is the projective limit

$$G^\infty = \varprojlim G_n$$

provided we have a sequence of projections from the larger objects to the smaller ones. If the G_n's are endowed with natural measures, this naturally gives rise to couplings.

Here are three examples, the first two being guiding examples, while the last one being the one of interest in this Chapter.

2.2.1. The infinite symmetric group $\mathfrak{S}^\infty$. Any $\sigma_n \in \mathfrak{S}_n$ can also be seen as an element of $\mathfrak{S}_m$ for $m > n$ by considering the elements beyond n as fixed. This simple fact defines the inclusion maps $\iota_{m,n} : \mathfrak{S}_m \to \mathfrak{S}_n$. This way, any permutation $\sigma_n \in \mathfrak{S}_n$ can be seen as belonging to the "inductive limit" i.e. elements in

$$\mathfrak{S}_\infty := \varinjlim \mathfrak{S}_m \ .$$

The projective limit $\mathfrak{S}^\infty$ on the other hand is different. These are called virtual permutations in [BNN13] and $(\sigma_n \ ; \ n \geq 1) \in \mathfrak{S}^\infty$ can naturally be constructed by the so-called Chinese restaurant process (see e.g. [Pit06]), as follows:

- σ_1 is the unique permutation in $\mathfrak{S}_1$;
- for $n \geq 1$, σ_{n+1} is obtained from σ_n either by adding $n+1$ as a fixed point, or by inserting $n+1$ inside a cycle of σ_n.

More precisely the coupling between Haar measures is given by

THEOREM 2.2.1. *Given a Haar distributed element $\sigma_{n+1} \in \mathfrak{S}_{n+1}$, it decomposes uniquely as the product of σ_n is Haar distributed in $\mathfrak{S}_n$ and a transposition:*

$$\sigma_{n+1} = \sigma_n \circ (k_{n+1} \ (n+1)) \ ,$$

where k_{n+1} is uniform in $\{1, 2, \ldots, n+1\}$.

ELEMENTS OF PROOF. The decomposition requires $k_{n+1} = \sigma_{n+1}^{-1}(n+1)$. The fact that this decomposition induces the correct distribution is a simple couting argument. □

Notice that the mere fact of writing this coupling leaves implicit the inclusions $\iota_{n,n+1}$, as strictly speaking one should write

$$\sigma_{n+1} = \iota_{n,n+1}(\sigma_n) \circ (k_{n+1} \ (n+1)) \ .$$

2.2.2. Infinite Hermitian matrices H^∞. The vector space of Hermitian matrices is written $H_n := \{m \in M_n(\mathbb{C}) \mid m = m^*\}$. There is natural inclusion $H_n \hookrightarrow H_{n+1}$ which consists in identifying an $n \times n$ matrix with the top-left block of an $(n+1) \times (n+1)$ matrix. The inductive limit $H_\infty = \varinjlim H_n$ consists of infinite Hermitian matrices with finitely many non-zero coefficients.

The projective limit $H^\infty = \varprojlim H_n$, on the other hand, consists of infinite Hermitian matrices indexed by $\mathbb{N}$. The projection to H_n consists in simply cutting a top-left corner of size n. Notice that U_n acts by conjugation on H_n and it is possible to make sense of measures on H^∞ whose all restrictions are invariant under the unitary group. Restrictions are naturally to the top-left corners. Among the examples, we have classical ensembles:

- the infinite GUE i.e. the infinite matrix with i.i.d. complex Gaussian entries

$$[\mathrm{GUE}^\infty]_{i,j} = \frac{G_{i,j} + G_{j,i}}{\sqrt{2}} \ .$$

- the rank one matrix $\xi\xi^*$ where $\xi \in M_{\mathbb{N},1}(\mathbb{C})$ is an infinite vector of i.i.d. Gaussian entries.

A beautiful result by Olshanski and Vershik [OV96] consists in the classification of the ergodic (extremal) measures on H^∞ invariant by unitary conjugation.

THEOREM 2.2.2 ([OV96]). *The unitary invariant ergodic measures on H^∞ are exactly of the form*

$$\gamma \mathrm{GUE} + \sum_i \alpha_i \xi_i \xi_i^*$$

for a sequence of positive real numbers $\gamma > 0$, $\alpha_i > 0$ such that

$$\gamma^2 + \sum_i \alpha_i^2 < \infty \ .$$

Notice that the diagonal of such measures are necessarily exchangeable, because of the invariance under conjugation by the permutation group. Because De Finetti's theorem, such ergodic measures necessarily have i.i.d. diagonals. Actually this is sufficient.

2.2.3. The infinite unitary group U^∞. The unitary group U_n. This is the case that will interest us the most. We shall explain in this chapter that the coupling method is virtually at the core of the three contributions [CNN17, CMN18, CN19].

Bourgade, Najnudel and Nikeghbali introduced in [BNN13] the concept of virtual isometry. Following the idea explained above, for every integers $m \leq n$, we have a natural inclusion map $\iota_{m,n} : U_m \hookrightarrow U_n$ which maps a matrix $u_m \in U_m$ to the block matrix

$$\iota_{m,n}(u_m) := \begin{pmatrix} u_m & 0_{m\times(n-m)} \\ 0_{(n-m)\times m} & I_{n-m} \end{pmatrix}$$

This map depends on the choice of basis $\mathcal{B}_n = (e_1, e_2, \ldots, e_n)$. In order to define the dual operation and projective measures, we need to consider a natural way of coupling the Haar measure between the group U_n and U_{n+1}. The proposed solution in [BNN13] takes the form of the so-called virtual isometries. A virtual isometry $u \in U^\infty$ is a sequence of unitary matrices $u = (u_n \in U_n \ ; n \geq 1)$ with a coupling of the form:

$$u_n = u_{n-1} \circ r_n \ ,$$

and $r_n \in U_n$ is a complex reflection sending the canonical basis vector e_n to a vector on the sphere v_n. By complex reflection, we mean a unitary map with $\mathrm{rank}(r_n - I_n) = 1$. By setting $v_n = r_n(e_n)$, one finds the expression:

$$r_n = I_n - 2\frac{1 - \Re\langle v_n, e_n\rangle}{1 - \langle v_n, e_n\rangle} \frac{(v_n - e_n)\,(v_n - e_n)^*}{\|v_n - e_n\|^2} \ . \tag{2.2.1}$$

The analogous decomposition to Theorem 2.2.1 is:

THEOREM 2.2.3. *Given a Haar distributed element $u_n \in U_n$, it decomposes uniquely as a product of a Haar distributed $u_{n-1} \in U_{n-1}$ and a complex reflection r_n:*

$$u_n = u_{n-1} \circ r_n \ .$$

The complex reflection r_n has the form in Eq. (2.2.1) with v_n being uniform on the sphere $\{v \in \mathbb{C}^n \mid \|v\|^2 = 1\}$.

Now define the following two version of characteristic polynomials of a unitary matrix via

$$\Phi_n^*(z) := \det(I_n - zu_n^{-1}) = \prod_{j=1}^{n} \left(1 - ze^{-i\theta_j^{(n)}}\right), \tag{2.2.2}$$

$$\Phi_n(z) := \det(z - u_n) = \prod_{j=1}^{n} \left(z - e^{i\theta_j^{(n)}}\right). \tag{2.2.3}$$

The $\left(\theta_j^{(n)} \ ; \ 1 \leq j \leq n\right)$ are the eigenvalues of u_n.

Clearly we have the relationship

$$\Phi_n(z) = z^n \overline{\Phi_n(1/\bar{z})} \ . \tag{2.2.4}$$

The matrix determinant Lemma states that for a matrix $A \in \mathbb{C}^{n\times n}$ and $(u, v) \in \mathbb{C}^{n\times 1} \times \mathbb{C}^{n\times 1}$ column vectors, we have:

$$\det(A + uv^*) = \det(A)\left(1 + v^* A^{-1} u\right) \ .$$

Let us now apply this identity to $A = I_n - z\iota_{n-1,n}\left(u_{n-1}\right)^{-1}$ and

$$A + uv^* = I_n - zu_n^{-1} = I_n - zu_{n-1}^{-1} r_n^{-1} = I_n - zu_{n-1}^{-1} - z\left(r_n^{-1} - I_n\right) \ .$$

Notice that in the above computation, we identified $\iota_{n-1,n}\left(u_{n-1}\right)$ to u_{n-1} to the cost of adding an eigenvalue equal to one. As such $\det A = \Phi_{n-1}^*(z)(1-z)$. Moreover, from Eq. (2.2.1) we have

$$r_n^{-1} = I_n - 2\frac{1 - \Re\langle v_n, e_n\rangle}{1 - \overline{\langle v_n, e_n\rangle}} \frac{(v_n - e_n)(v_n - e_n)^*}{\|v_n - e_n\|^2} \ , \tag{2.2.5}$$

which is can be deduced without computation from reversing the roles of e_n and v_n. As such

$$uv^* = 2z\frac{1 - \Re\langle v_n, e_n\rangle}{1 - \overline{\langle v_n, e_n\rangle}} \frac{(v_n - e_n)(v_n - e_n)^*}{\|v_n - e_n\|^2} = z\frac{(v_n - e_n)(v_n - e_n)^*}{1 - \overline{\langle v_n, e_n\rangle}} \ .$$

In the end, we have

$$\begin{aligned}
\Phi_n^*(z) &= \det\left(I_n - zu_n^{-1}\right) \\
&= \det\left(A + uv^*\right) \\
&= \det A\left(1 + v^* A^{-1} u\right) \\
&= \Phi_{n-1}^*(z)(1-z)\left(1 + v^* A^{-1} u\right) \\
&= \Phi_{n-1}^*(z)(1-z)\left(1 + z\frac{(v_n - e_n)^*\left(I_n - zu_{n-1}^{-1}\right)^{-1}(v_n - e_n)}{1 - \overline{\langle v_n, e_n\rangle}}\right) \ .
\end{aligned}$$

The above equation yields a coupling between characteristic polynomials. It is the basis of a strong control on eigenvalues of unitary matrices, which leads to strong almost sure convergence results. For example, upon extending the eigenvalue indices to all of $\mathbb{Z}$ by 2π periodicity on the circle, we have the following.

THEOREM 2.2.4 ([MNN20]). *Almost surely, the point process*

$$\left(y_k^{(n)} := \frac{n}{2\pi}\theta_k^{(n)}, \ k \in \mathbb{Z}\right)$$

converges pointwise to a determinantal sine-kernel point process $(y_k, \ k \in \mathbb{Z})$.

And moreover, almost surely, the following estimate holds for all $\varepsilon > 0$:

$$\forall k \in [-n^{\frac{1}{4}}, n^{\frac{1}{4}}],\ y_k^{(n)} = y_k + O_\varepsilon\left((1+k^2)n^{-\frac{1}{3}+\varepsilon}\right)$$

Such estimates have been improved in [VV22].

2.3. The Stochastic Zeta function

Following a philosophy due to Keating-Snaith on the one hand and to Sarnak on the other hand, the characteristic polynomials of random matrices are models for the Riemann function ζ on the critical line $\Re s = \frac{1}{2}$. This has led to a large number of predictions, and some theorems, which would probably have remained out of reach without the direction given by random matrix theory. A conjecture raised by Virag at an AIM conference was the construction of a random zeta function as a limit of characteristic polynomials of random matrices. In [CNN17], we construct such a universal function through a fine control of the eigenvalues of distributed Haar matrices. This control is made possible by the virtual isometries described in the previous sections. Although we never named that object, more recent work by Valko and Virag [VV22] uses the name of "Stochastic Zeta Function". This random function allows to refine the correspondence between random matrices and number theory from a functional point of view. In [CHN+19], we construct the same function for sets of classical random matrices other than the unitary case, and where we have weaker controls on the eigenvalues.

Let us now give a precise statement. Consider the microscopic rescaling of the characteristic polynomial X_n of a Haar distributed matrix in U_n:

$$\xi_n(z) = \frac{X_n(e^{2iz\pi/n})}{X_n(1)}\ . \tag{2.3.1}$$

The first main theorem of [CNN17] is the following:

THEOREM 2.3.1 ([CNN17]). *Almost surely and uniformly on compact subsets of $\mathbb{C}$, we have the convergence:*

$$\xi_n(z) \overset{n\to\infty}{\longrightarrow} \xi_\infty(z) := e^{i\pi z}\prod_{k\in\mathbb{Z}}\left(1-\frac{z}{y_k}\right)$$

Here, the infinite product is not absolutely convergent. It has to be understood as the limit of the following product, obtained by regrouping the factors two by two:

$$\left(1-\frac{z}{y_0}\right)\prod_{k\geq 1}\left[\left(1-\frac{z}{y_k}\right)\left(1-\frac{z}{y_{-k}}\right)\right],$$

which is absolutely convergent.

The second main result consists in giving new compact formulas for moments of ratios, and a credible conjecture for the Riemann zeta function ζ.

THEOREM 2.3.2 ([CNN17]). *The following results on ratios hold:*

(1) For any $p > 0$ and any compact set $K \subset \mathbb{C}\backslash\mathbb{R}$, we have:

$$\sup_{n\in\mathbb{N}\sqcup\{\infty\}}\mathbb{E}\left(\sup_{(z,z')\in K^2}\left|\frac{\xi_n(z')}{\xi_n(z)}\right|^p\right) < \infty.$$

(2) For $z_1, \dots, z_k, z'_1, \dots, z'_k \in \mathbb{C}\backslash\mathbb{R}$, and for all $n \in \mathbb{N} \sqcup \{\infty\}$,

$$\mathbb{E}\left(\prod_{j=1}^{k}\left|\frac{\xi_n(z'_j)}{\xi_n(z_j)}\right|\right) < \infty$$

Moreover, for every compact set K in $\mathbb{C}\backslash\mathbb{R}$, we have the following convergence, uniformly in $z_1, z_2, \dots, z_k, z'_1, \dots, z'_k \in K$:

$$\mathbb{E}\left(\prod_{j=1}^{k}\frac{\xi_n(z'_j)}{\xi_n(z_j)}\right) \underset{n\to\infty}{\longrightarrow} \mathbb{E}\left(\prod_{j=1}^{k}\frac{\xi_\infty(z'_j)}{\xi_\infty(z_j)}\right).$$

(3) For all $z_1, \dots, z_k, z'_1, \dots, z'_k \in \mathbb{C}\backslash\mathbb{R}$ such that $z_i \neq z'_j$ for $1 \leq i, j \leq n$, we have

$$\det\left(\frac{1}{z_i - z'_j}\right)_{i,j=1}^{k} \mathbb{E}\left(\prod_{j=1}^{k}\frac{\xi_\infty(z'_j)}{\xi_\infty(z_j)}\right) = \det\left(\frac{1}{z_i - z'_j}\mathbb{E}\left(\frac{\xi_\infty(z'_j)}{\xi_\infty(z_i)}\right)\right)_{i,j=1}^{k}$$

and moreover:

$$\mathbb{E}\left(\frac{\xi_\infty(z')}{\xi_\infty(z)}\right) = \begin{cases} 1 & \text{if } \Im(z) > 0 \\ e^{i2\pi(z'-z)} & \text{if } \Im(z) < 0 \end{cases}$$

And we conjecture that if ω is a uniform random variable on $[0,1]$ and $T > 0$ a real parameter going to infinity, then, for all $z_1, \dots, z_k, z'_1, \dots, z'_k \in \mathbb{C}\backslash\mathbb{R}$, such that $z_i \neq z'_j$ for all i, j,

$$\mathbb{E}\left(\prod_{j=1}^{k}\frac{\zeta\left(\frac{1}{2} + iT\omega - \frac{i2\pi z'_j}{\log T}\right)}{\zeta\left(\frac{1}{2} + iT\omega - \frac{i2\pi z_j}{\log T}\right)}\right)$$
$$\overset{T\to\infty}{\longrightarrow} \det\left(\frac{1}{z_i - z'_j}\right)^{-1} \det\left(\frac{\mathbb{1}_{\Im(z_i)>0} + e^{2i\pi(z'_j - z_i)}\mathbb{1}_{\Im(z_i)<0}}{z_i - z'_j}\right)_{i,j=1}^{k},$$

where the last expression is well-defined where the z_i and the z'_j are all distinct, and is extended by continuity to the case where some of the z_i or some of the z'_j are equal.

2.4. Realization of couplings from OPUC

A different coupling can be realized for a model of random matrices more general than the CUE.

2.4.1. Orthogonal Polynomials on the Unit Circle (OPUC). Consider a probability measure μ on the unit circle $\partial\mathbb{D}$, $\mathbb{D}$ being the unit disc. By applying the Gram-Schmidt orthogonalization procedure to monomials $\{1, z, z^2, \dots\}$, one obtains a sequence $(\Phi_n)_{n\geq 0}$ of OPUC which satisfies the Szegö recurrence:

$$\begin{pmatrix}\Phi_{k+1}(z)\\ \Phi^*_{k+1}(z)\end{pmatrix} = \begin{pmatrix} z & -\overline{\alpha}_k \\ -\alpha_k z & 1\end{pmatrix}\begin{pmatrix}\Phi_k(z)\\ \Phi^*_k(z)\end{pmatrix}, \tag{2.4.1}$$

where

$$\Phi^*_n(z) := z^n\overline{\Phi_n(1/\bar{z})}.$$

The Szegö recurrence is the analogue of the three term recurrence for orthogonal polynomials on the line $\mathbb{R}$. The coefficients α_j belong to the closed disc, $\overline{\mathbb{D}}$, and are

called Verblunsky coefficients. If a measure μ determines the Verblunsky coefficients, the converse is also true (see [**Sim05a**, Theorem 1.7.11 p.97]):

THEOREM 2.4.1 (Verblunsky's theorem). *Let $\mathcal{M}_1(\partial\mathbb{D})$ be the simplex of probability measures on the circle, endowed with the weak topology, and let*

$$\mathcal{D} := \mathbb{D}^{\mathbb{N}} \sqcup \left(\sqcup_{n\in\mathbb{Z}_+} \mathbb{D}^n \times \partial\mathbb{D}\right)$$

be endowed with the topology related to the following notion of convergence: a sequence $(A_p)_{p\geq 1}$ in $\mathcal{D}$ converges to an element $A_\infty = (\alpha_j)_{0\leq j<K}$ with finitely many or infinitely many components (K finite or infinite) if and only if for all $j < K$, the coefficient of order j of A_p is well-defined for p large enough and converges to α_j. Then, the map

$$\mathbb{V}: \quad \mathcal{M}_1(\partial\mathbb{D}) \quad \to \quad \mathcal{D}$$

given by the sequence of Verblunsky coefficients is a homeomorphism. Atomic measures with n atoms have n Verblunsky coefficients, the last one being of modulus one, other measures have infinitely many Verblunsky coefficients.

If *Leb* is the Lebesgue measure on $\partial\mathbb{D}$, then $\mathbb{V}(Leb) = (0, 0, \dots)$. In fact, the tangent map of the Verblunsky map, at the point *Leb*, gives exactly the Fourier coefficients of the perturbation. Hence the Verblunsky map is inherently spectral in nature and Verblunsky coefficients can be seen as non-linear Fourier coefficients.

Now let us introduce the random matrix model of interest.

2.4.2. The Circular Beta Ensemble ($C\beta E$). For this paragraph, $\beta > 0$ plays the role of a coupling constant. Consider n points on the unit circle whose probability distribution is:

$$(C\beta E_n) \quad \frac{1}{Z_{n,\beta}} \prod_{1\leq k<l\leq n} \left|e^{i\theta_k} - e^{i\theta_l}\right|^\beta d\theta. \tag{2.4.2}$$

For $\beta = 2$, we recognize Weyl's integration formula for central functions on the unitary group $U(n)$. In this case the $C\beta E_n$ reduces to an ensemble known as the Circular Unitary Ensemble (CUE_n). It is nothing but the distribution of the eigenvalues of a Haar distributed random matrix. Naturally, the study of this case is very rich in the representation theory of unitary groups. See for example Diaconis-Shahshahani [**DS94**] and Bump-Gamburd [**BG06**], which are some of my favorite papers mixing RMT and representation theory.

For general $\beta > 0$, the representation-theoretic picture is more complicated. $C\beta E_n$ is the orthogonality measure for Jack polynomials in n variables [**Mac98**]. In turn, Jack polynomials are also intimately related to representation theory via rational Cherednik algebras [**DG10**]. Our point of view will be more direct. From the work of Killip and Nenciu [**KN04a**], the characteristic polynomial

$$X_n(z) := \det\left(\mathrm{id} - zU_n^*\right) = \prod_{1\leq j\leq n} \left(1 - ze^{-i\theta_j^{(n)}}\right)$$

can be realized as the last term of the Szegö recurrence, whose distribution of the Verblunsky coefficients is explicitly given. This distribution is described as follows: the coefficients are independent, the last one is uniform on the unit circle, and for

$0 \leq j \leq n-2$, α_j is rotationally invariant and $|\alpha_j|^2$ is a Beta random variable with parameters 1 and $\beta_j := \frac{\beta(j+1)}{2}$:

$$\mathbb{P}\left(|\alpha_j|^2 \in dx\right) = \beta_j \left(1-x\right)^{\beta_j - 1} \mathbb{1}_{\{0<x<1\}} dx \ . \tag{2.4.3}$$

In passing, let us record the following basic properties:

- Rotation invariance: if $|\lambda| = 1$, then

$$\left(\Phi_n^*(\lambda z)\right)_{z\in\mathbb{C}} \overset{\mathcal{L}}{=} \left(\Phi_n^*(z)\right)_{z\in\mathbb{C}}$$

- The equalities:

$$\mathbb{E}\left(|\alpha_j|^2\right) = \frac{1}{1+\beta_j} \tag{2.4.4}$$

and

$$\mathbb{E}\left[-\log\left(1-|\alpha_j|^2\right)\right] = \frac{1}{\beta_j} \tag{2.4.5}$$

hold.

Remark 2.4.2. *In fact, Killip and Nenciu first prove in their* [KN04a, Theorem 1] *and* [KN04a, Proposition 4.2] *that*

$$\mathbb{V}^{-1}\left(\alpha_{n-2}, \ldots, \alpha_1, \alpha_0, \eta\right) = \sum_{j=1}^{n} \widetilde{\pi}_j \delta_{\theta_j^{(n)}}(d\theta) \ ,$$

where the weights $(\widetilde{\pi}_j)_{1\leq j\leq n}$ *have a* β*-Dirichlet distribution and the support is independently distributed according to the* $C\beta E_n$ *given in* (2.4.2).

Moreover, thanks to [KN04a, Proposition B.2], *reversing the order of Verblunsky coefficients, except the last one* η*, changes the weights but preserves the support.*

Having in mind the previous remark, a fruitful idea consists in using the reversed order of Verblunksy coefficients and incorporating the weights in the definition of the $C\beta E_n$. Therefore, we *redefine* the Circular β Ensemble with n points as the random probability measure:

$$C\beta E_n := \mathbb{V}^{-1}\left(\alpha_0, \alpha_1, \ldots, \alpha_{n-2}, \eta\right) = \sum_{j=1}^{n} \pi_j \delta_{\theta_j^{(n)}}(d\theta) \ . \tag{2.4.6}$$

The support points $\left(\theta_j^{(n)}\right)_{1\leq j\leq n}$ are the zeroes of X_n and are still distributed as in (2.4.2). Nevertheless, the distribution of the weights $(\pi_j)_{1\leq j\leq n}$ is not known explicitly, with a tractable form.

From Eq. (2.4.6) a remarkable fact is that the sequence of Verblunsky coefficients is *consistent*: $C\beta E_n$ and $C\beta E_{n+1}$ have a priori no reason for living on the same probability space. However, it is possible to couple them in such a way that the $n-1$ first Verblunsky coefficients are exactly the same. This provides a way to couple the characteristic polynomial of $C\beta E_n$ for all values of $n \geq 1$: if $(\alpha_j)_{j\geq 0}$ is an *infinite* sequence of independent variables whose distribution is given as above, and if η is an independent variable, uniform on the unit circle, then the last orthogonal polynomial given by the sequence of Verblunsky coefficients $(\alpha_0, \ldots, \alpha_{n-2}, \eta)$ has the same law as the $C\beta E_n$ for all $n \geq 1$.

2.5. Maxima of characteristic polynomials

Still combining number theory and random matrices, Fyodoroff-Hiary-Keating formulate a very precise conjecture on the maximum of the function ζ on short intervals and on the maximum of the characteristic polynomial of a random matrix. This conjecture is informed by an analogy between branching models and the log-correlated nature of the random fields considered.

For shorter notation, the unit disc will sometimes be denoted $\mathbb{U} := \partial\mathbb{D}$. A standard result in RMT says that the log-characteristic polynomial of CβE converges to a Gaussian field. More precisely, the following holds:

$$(\log|X_n(z)| \ ; \ z \in \mathbb{D}) \overset{n\to\infty}{\longrightarrow} \left(\sqrt{\frac{2}{\beta}}G(z) \ ; \ z \in \mathbb{D}\right) \ ,$$

where for $z \in \mathbb{D}$,

$$G(z) := 2\Re \sum_{k=1}^{\infty} \frac{z^k}{\sqrt{k}} \mathcal{N}_k^{\mathbb{C}} \ . \tag{2.5.1}$$

Extremal statistics. Let us formulate and explain the intuition of the Fyodoroff-Hiary-Keating conjecture by restricting ourselves to $\beta = 2$ for now. By an explicit computation [**FHK12**], it is possible to prove that

$$\mathrm{Var}\,(\log|X_n(z)|) \sim \frac{1}{2}\log n$$

and that the correlation saturates at the scale $|\theta - \theta'| \sim \frac{1}{n}$. By correlation saturation, we simply mean that the order of magnitude of the correlation remains the same for $\theta - \theta'$ going to zero and for $|\theta - \theta'| \sim 1/n$. Thus, the naive analogy consists in approximating the function $\log|X_n(z)|$ on the circle by its values at $\mathcal{O}(n)$ points. Each point would be assigned an independent copy of a Gaussian with variance $\frac{1}{2}\log n$. It is classical that the maximum of such independent Gaussians is of order $\log n$, which intuitively explains the leading order. One hopes to show that the proof of this first order does not depend on the correlation structure. The story is different for the second order term. If not for the correlations, the asymptotic expansion would be $\log n - \frac{1}{4}\log\log n$ by approximating the field by $\mathcal{O}(n)$ independent Gaussians.

From this discussion, one sees that $(\log|X_n(z)|)_{z\in\mathbb{U}}$ is a complicated (yet integrable) regularization of a log-correlated Gaussian field $(G(z))_{z\in\mathbb{U}}$. In terms of global features, it is in every way similar to the "cone construction" (see Arguin, Zindy [**AZ14**, Fig. 1]): correlation is of logarithmic nature and saturates at the scale $\frac{1}{n}$. In that universality class, one expects:

$$\max_{z\in\mathbb{U}} \log|X_n(z)| \sim \log n - \frac{3}{4}\log\log n,$$

which is an established result in many cases. In the case of tree models such as branching Brownian motion and branching random walks, the result holds at fairly large level of generality (See [**HS09**, **AS10**, **Aïd13**]). By "tree model", we mean a model where a tree structure is apparent and explicit. Among non-tree models, where one needs to identify an approximate branching structure, the result holds for log-correlated Gaussian fields [**Mad15**, **DRZ15**], discrete GFF (Gaussian Free Fields) as described in [**BZ12**, **BDZ16**], and cover times [**BK14**]. The constant $\frac{3}{4}$ is strongly related to such an underlying hierarchical structure.

It is also worth mentioning that the field $\left(\Re G\left(e^{i\theta}\right)\right)_{\theta\in[0,2\pi)}$ can be regularized into a Gaussian field by evaluating the random field $z \mapsto \Re G(z)$ in the interior of the unit disk. The existing technology for Gaussian log-correlated fields is applicable to $\theta \mapsto \Re G\left(e^{-\frac{1}{n}+i\theta}\right)$, with mild modifications. It yields the expected results for this simple regularization where all the Random Matrix Theory is lost.

In two very insightful papers [**FK14**, **FHK12**], Fyodorov, Hiary and Keating formulate the following conjecture.

CONJECTURE 2.5.1 (The Fyodoroff-Hiary-Keating conjecture). *When $\beta = 2$, the following convergence in law holds:*

$$\sup_{z\in\mathbb{U}} \log|X_n(z)| - \left(\log n - \frac{3}{4}\log\log n\right) \underset{n\to\infty}{\longrightarrow} \frac{1}{2}(K_1 + K_2)$$

where K_1 and K_2 are two independent Gumbel random variables.

Indeed, in the notations of these papers, $-(K_1 + K_2)$ is a random variable with density

$$p(x) = 2e^x K_0(2e^{\frac{x}{2}}) = e^x \int_{\mathbb{R}} dy e^{-e^{x/2}\cosh(y)} .$$

Here K_0 is the modified Bessel function of the second kind. A quick computation of moment generating functions allows us to realize that we are dealing indeed with minus the sum of two independent Gumbel random variables.

It is a very challenging problem to prove (or disprove) such a precise conjecture. However, progress has recently been made in this direction. In a first breakthrough [**ABB16**], Arguin, Belius and Bourgade have proven that

$$\frac{\sup_{z\in\mathbb{U}} \log|X_n(z)|}{\log n} \underset{n\to\infty}{\longrightarrow} 1$$

in probability, and shortly afterwards, using different methods, Paquette and Zeitouni [**PZ16**] have refined this result by showing:

$$\frac{\sup_{z\in\mathbb{U}} \log|X_n(z)| - \log n}{\log\log n} \underset{n\to\infty}{\longrightarrow} -\frac{3}{4}$$

in probability. The refinement given by Paquette and Zeitouni is an important progress as the constant $\frac{3}{4}$ morally confirms the existence of hierarchical structures.

Our result. The main theorem of [**CMN18**] answers Conjecture 2.5.1 up to the third order, and in the setting of the Circular Beta Ensemble where $\beta > 0$ is not necessarily equal to 2.

For $\beta \neq 2$, the point process of the eigenvalue is not determinantal, and then it is more difficult to get exact formulas for this model. The tool we will use to deal with this problem is the theory of orthogonal polynomials on the unit circle, described for example in the book by Simon [**Sim05b**]. In [**KN04b**], Killip and Nenciu give the construction of an ensemble of random matrices whose eigenvalue distribution follows the CβE, and prove that the characteristic polynomial can be written as the last term of a sequence of orthogonal polynomials whose parameters, called Verblunsky coefficients, have a distribution which is explicitly given. In the beautiful paper [**KS09**], Killip and Stoiciu use this model in order to deduce the existence of a limiting point process for the microscopic behavior of the CβE. More details are given in the next section, along with the notions we will need.

The precise statement of our main result is the following.

THEOREM 2.5.2 (Main result of [CMN18]). *If* $\mathbb{U}' := \mathbb{U}\backslash\{\overline{\lambda_1}, \dots, \overline{\lambda_n}\}$, *the following family of random variables:*

$$\sqrt{\frac{\beta}{2}}\left(\sup_{z\in\mathbb{U}'} \Re \log X_n(z) - \left(\log n - \frac{3}{4}\log\log n\right)\right)_{n\geq 2}$$

and for $\sigma \in \{-1, 1\}$,

$$\sqrt{\frac{\beta}{2}}\left(\sup_{z\in\mathbb{U}'} (\sigma \Im \log X_n(z)) - \left(\log n - \frac{3}{4}\log\log n\right)\right)_{n\geq 2}$$

are tight.

It seems reasonable to expect that these families of random variables have a limiting distribution, however, we are not sure about what this distribution should be. It is interesting to state the previous result with the imaginary part of the characteristic polynomial, since this gives some information about the number of points among $(\lambda_j)_{1\leq j\leq n}$ which lie in a given arc of circle. In particular, we get the following corollary:

COROLLARY 2.5.3. *For* $z_1, z_2 \in \mathbb{U}$, *let* $N_n(z_1, z_2)$ *be the number of points of the* $C\beta E$ *lying in the arc coming counterclockwise from* z_1 *to* z_2, *and let* $N_n^{(0)}(z_1, z_2)$ *be the expectation of* $N_n(z_1, z_2)$, *which is equal to the length of the arc multiplied by* $n/2\pi$. *Then, the following family of random variables is tight:*

$$\left(\pi\sqrt{\frac{\beta}{8}} \sup_{z_1,z_2\in\mathbb{U}} |N_n(z_1, z_2) - N_n^{(0)}(z_1, z_2)| - \left(\log n - \frac{3}{4}\log\log n\right)\right)_{n\geq 2}.$$

The values of z_1 and z_2 maximizing $|N_n(z_1, z_2) - N_n^{(0)}(z_1, z_2)|$ correspond to the extreme values of the imaginary part of $\log X_n$ on $\mathbb{U}$.

Let us conclude this section by mentioning the latest refinement of Theorem 2.5.2 given in [PZ22]. There, Paquette and Zeitouni refine our method to prove that the fluctuations are indeed given by a sum of two independent random variables. First there is a Gumbel random variable which appears when aggregating decorrelated locations on the circle. The second random variable captures the extremal landscape and is the analogue of the derivative martingale in the branching Brownian motion case. The paper [PZ22] does not identify the limiting distribution of the second random variable, but it is conjectured to be a Gumbel random variable as well.

The limit is in fact expected to match (the log of) the total mass of a critical Gaussian Multiplicative Chaos (GMC) on the circle. It was when studying this object that we unveiled a stronger connection with RMT, which we will discuss in the next section.

2.6. From RMT to the Gaussian Multiplicative Chaos

The starting point is revisiting Eq. (2.4.6), which we look at through the lens of Verblunsky's Theorem 2.4.1. With this particular coupling, the Verblunsky coefficients provide a sequence of random measures indexed by n, supported by the points of the $C\beta E_n$, and tending to a limiting random measure μ^β, whose Verblunsky coefficients are $(\alpha_j)_{j\geq 0}$.

In light of Verblunsky's Theorem 2.4.1, this remark begs the question:

QUESTION 2.6.1. *Is there anything remarkable or canonical about the projective limit*

$$\varprojlim C\beta E_n := \mathbb{V}^{-1}(\alpha_0, \alpha_1, \alpha_2, \dots) = \mu^\beta \ ,$$

obtained from using all Verblunsky coefficients? Does this measure arise in other circumstances?

Before discussing this question, it is worth explaining why the points $C\beta E_n$ can be seen as quadrature points of the infinite random measure $\varprojlim C\beta E_n = \mu^\beta$. Any sequence of measures, indexed by n, whose $n-1$ first Verblunsky coefficients match the $n-1$ first elements of the sequence $(\alpha_0, \alpha_1, \dots)$ will converge to μ^β, in the topology of weak convergence. Moreover, if we assume that the Verblunsky coefficients are $(\alpha_0, \dots, \alpha_{n-2}, \eta)$ with $|\eta| = 1$, then the approximating measure is atomic, supported by n points. The general theory of orthogonal polynomials dictates that for all polynomials P of degree $\deg P \leq n-1$:

$$\int_{\partial\mathbb{D}} P \ \mu^\beta = \sum_{j=1}^{n} \pi_j P(e^{i\theta_j^{(n)}}) \ .$$

In the language of approximation theory, that is exactly to say that $(\pi_j)_{1\leq j\leq n}$ are (random) quadrature weights and that the n of points $C\beta E_n$ can be seen as the n (random) quadrature points for the (random) measure $\mu^\beta = \varprojlim C\beta E_n$.

It is in our paper [CN19] that we tackle Question 2.6.1. There we clarify the links between random matrices and log-correlated fields by proving a surprising connection. We show that, in a precise sense, the n-point CβE is the quadrature of the circle-invariant GMC (Gaussian Multiplicative Chaos). In a certain sense, these two models are in fact equal and not simply similar as suggested in a conjecture of Virag on the equality of the multifractal spectrum of the GMC and of the measure associated to the CβE. This of course solves Virag's question. But as a corollary we also obtain a new proof of the Fyodoroff-Bouchaud conjecture on the total mass of the GMC on the circle, proved one year before by Rémy by conformal field theory techniques. Our method also gives the law of the other trigonometric moments. For me, emphasizing the integrability of the GMC thanks to the CβE model of random matrices still corresponds to the philosophy of integrable probability. An attractive idea, still vague, would be to link the integrability of the conformal field theory, where the GMC is a central object which should be related to the (higher) representation theory associated to the CβE.

The Gaussian Multiplicative Chaos (GMC^γ). In this paragraph, $\gamma > 0$ plays the role of coupling constant in an a priori different context. Recall from Eq. (2.5.1) the Gaussian field on the unit disc:

$$G(z) := 2\Re \sum_{k=1}^{\infty} \frac{z^k}{\sqrt{k}} \mathcal{N}_k^{\mathbb{C}}$$

where $(\mathcal{N}_k^{\mathbb{C}})_{k\geq 0}$ denote i.i.d complex Gaussian variables, such that

$$\mathbb{E}\left[(\mathcal{N}_k^{\mathbb{C}})^2\right] = \mathbb{E}\left[\mathcal{N}_k^{\mathbb{C}}\right] = 0, \ \mathbb{E}\left[|\mathcal{N}_k^{\mathbb{C}}|^2\right] = 1 \ .$$

One can establish that:

- $\mathrm{Cov}(G(w), G(z)) = -2\log|1 - w\bar{z}|$.

- The field can be extended to the closed unit disc $\overline{\mathbb{D}}$ but its restriction to the circle is not a function. In fact, $G_{|\partial\mathbb{D}}$ is almost surely a random Schwartz distribution in $\cap_{\varepsilon>0}H^{-\varepsilon}(\partial\mathbb{D})$ where the Sobolev spaces are given for all $s\in\mathbb{R}$ by:

$$H^s(\partial\mathbb{D}) := \left\{ f \;\middle|\; \sum_{n\in\mathbb{Z}} |n|^s |\widehat{f}(n)|^2 < \infty \right\} .$$

- Because G is harmonic, $G(re^{i\theta}) = \left(G_{|\partial\mathbb{D}} * P_r\right)\left(e^{i\theta}\right)$ where $*$ denotes convolution and P_r is the Poisson kernel.

We can define the measure

$$GMC_r^\gamma(f) := \int_{\partial\mathbb{D}} \frac{d\theta}{2\pi} f(e^{i\theta}) \exp\left(\gamma G(re^{i\theta}) - \frac{1}{2}\gamma^2 \operatorname{Var}(G(re^{i\theta}))\right) = \int_{\partial\mathbb{D}} \frac{d\theta}{2\pi} f(e^{i\theta}) e^{\gamma G(re^{i\theta})} \left(1-r^2\right)^{\gamma^2} . \tag{2.6.1}$$

The Gaussian Multiplicative Chaos with coupling constant $\gamma > 0$ is the weak limit:

$$GMC^\gamma := \lim_{r\to 1} GMC_r^\gamma . \tag{2.6.2}$$

To be exact, the above limit holds in probability, upon integrating against continuous functions. The existence of such a limit for all $\gamma > 0$ is well-established via standard regularization techniques such as convolution or Karhunen-Loeve expansions of Gaussian processes [**RV13**, **B$^+$17**]. The literature treats higher dimensions and different geometries as well. Of course, this includes our particular case of convolution by the Poisson kernel. However, there are different regimes regarding the limit (2.6.2):

- $\gamma < 1$, Sub-critical phase. GMC^γ is a non-degenerate random measure, which can be seen from the following L^1 convergence.

  THEOREM 2.6.2 (Theorem 1.2 in [**B$^+$17**]). *For all nonnegative, smooth functions f, and for $\gamma < 1$, i.e. in the sub-critical regime:*

$$GMC_r^\gamma(f) \xrightarrow{r\to 1} GMC^\gamma(f) ,$$

  the convergence being in probability and in $L^1(\Omega, \mathcal{B}, \mathbb{P})$.

- $\gamma = 1$, Critical phase. The limit in (2.6.2) is the trivial zero measure, however one can perform different normalizations in order to obtain the so-called critical GMC. A random renormalization via the so-called derivative martingale has been implemented in [**DRS$^+$14**], while the Seneta-Heyde renormalization has been implemented in [**JS$^+$17**]. Both constructions agree [**Pow18**]. Moreover, Aru, Powell and Sepulveda [**APS18**, Section 4.1] have proven that the critical GMC can be written as the limit of the subcritical GMC when the parameter tends to 1 from below. This allows us to bootstrap the construction of the sub-critical GMC and obtain the critical GMC via the limit in probability:

$$GMC^{\gamma=1} = \lim_{\gamma\to 1^-} \frac{GMC^\gamma}{1-\gamma} , \tag{2.6.3}$$

when the random measures GMC^γ are constructed from the same field G for all values of $\gamma \in (0,1)$. The critical GMC is known to be non-atomic and it is conjectured to assign full measure to a random set of Hausdorff dimension zero (see the overview section of [DRS+14]).

As a corollary of our main result, we shall see that this latter conjecture holds, in the context of the circle.

- $\gamma > 1$, Supercritical phase. In this case, there are two constructions resulting in different measures.

  A first point of view consists in noticing that the renormalization of Eq. (2.6.1) by a factor $(1-r^2)^{\gamma^2}$ is too strong, and the limit (2.6.2) is the zero measure. One needs a different renormalization procedure so that a non-trivial limit holds. The correct normalization at the exponential scale is given by the precise asymptotic behavior of the maximum $\max_{\theta\in\mathbb{R}} G(re^{i\theta})$ as $r \to 1^-$. As such, one naturally expects the limit to be atomic, giving mass to the Gaussian field's maxima. This was done in [MRV+16]. With such a construction, the $\gamma > 1$ regime is called the glassy phase and the transition is referred to as a freezing transition. The term "freezing" comes from the fact that the logarithm of the total mass of the measure behaves linearly in γ because of the new renormalization. All in all, the result is that the limiting measure can be described as follows: one starts with the critical GMC, and conditionally on the corresponding random measure $GMC^{\gamma=1}$, one takes a strictly positive stable noise of scaling exponent $\frac{1}{\gamma}$ and intensity $GMC^{\gamma=1}$. In loose terms, in the supercritical regime, one only sees Dirac masses corresponding to the extrema of the underlying Gaussian field, and which are "sprinkled" on the circle with an intensity depending on the critical measure.

  Another version of the supercritical Gaussian multiplicative chaos has been previously constructed in [BJRV13] by taking a *subcritical* GMC with coupling constant $\gamma' = \frac{1}{\gamma}$, as the intensity of a stable noise of scaling exponent $\frac{1}{\gamma^2}$. We use a different normalization, hence extra factors 2 in [BJRV13]. The constructed measure is named the KPZ dual measure. As explained in that paper, the name stems from the relationship to the KPZ formula and its symmetry with respect to the transform $\gamma \mapsto \frac{1}{\gamma}$. This last construction cannot be naturally recovered from a logarithmically correlated Gaussian field on the circle without adding some extra randomness, contrarily to the construction of [MRV+16] with a freezing transition. Nevertheless, the KPZ dual measure seems to have better analyticity properties than the construction with a freezing transition. We will make further remarks on the topic at the end of the next section.

The result. The Main Theorem of [CN19] provides a direct link between the a priori unrelated objects introduced in the previous section: namely, it shows that up to a suitable normalization, the random measure $\varprojlim C\beta E_n$ and the Gaussian multiplicative chaos of parameter $\gamma := \sqrt{\frac{2}{\beta}}$ have the same distribution in the subcritical and the critical cases, i.e. for $\beta \geq 2$.

Notice that the construction of $\varprojlim C\beta E_n$ bypasses the phase transition involved in the definition of the GMC, since the description in terms of Verblunsky coefficients is uniform for all values of $\beta > 0$. However, we do not exactly know how the two random measures $C\beta E_\infty$ and GMC^γ are related in the supercritical case.

The precise statement is the following:

THEOREM 2.6.3 (Main Theorem of [CN19] - $GMC^\gamma = \varprojlim C\beta E_n$). *For $\beta \geq 2$, let $(\alpha_j)_{j\geq 0}$ be a sequence of independent, rotationally invariant complex-valued random variables, such that $|\alpha_j|^2$ is Beta-distributed with parameters 1 and $\beta_j = \frac{\beta}{2}(j+1)$. Let μ^β be the random probability measure whose Verblunsky coefficients are given by the sequence $(\alpha_j)_{j\geq 0}$, and let*

$$C_0 := \begin{cases} \prod_{j=0}^{\infty} \left(1-|\alpha_j|^2\right)^{-1} \left(1-\frac{2}{\beta(j+1)}\right) & \text{if } \beta > 2 \\ 2\left(1-|\alpha_0|^2\right)^{-1} \prod_{j=1}^{\infty} \left(1-|\alpha_j|^2\right)^{-1} \left(1-\frac{2}{\beta(j+1)}\right) & \text{if } \beta = 2\ . \end{cases}$$

Then, the product of C_0 by the measure μ^β has the same law as the measure corresponding to the Gaussian multiplicative chaos GMC^γ, with parameter $\gamma = \sqrt{\frac{2}{\beta}} \leq 1$. In particular, μ^β has the same law as GMC^γ, renormalized into a probability measure, and the total mass of GMC^γ has the same law as C_0.

CHAPTER 3

The Toda system, its scattering and tridiagonal models

If the previous chapter was focused on couplings for circular ensembles. Of course, one might wonder what are the analogues of the previous sections on the real line. Here we will revisit the matrix model of Dumitriu-Edelman [DE02], which yields a coupling which is arguably the oldest and the most well-known.

The goal of this chapter is to present an original description of tridiaonal models in terms of the Toda integrable system. The classical Toda flow is a well-known integrable Hamiltonian system that diagonalizes matrices. By keeping track of the distribution of entries and precise scattering asymptotics, one can exhibit matrix models for log-gases on the real line. These types of scattering asymptotics date back to fundamental work of Moser.

More precisely, using the classical Toda flow acting on symmetric real tridiagonal matrices, we give a "symplectic" proof of the fact that the Dumitriu-Edelman tridiagonal model has a spectrum following the Gaussian β-ensemble.

3.1. Gaussian β-ensembles and the Macdonald-Mehta-Opdam formula

The Gaussian β-ensemble is the probability distribution for an n-point configurations in the real line:

$$(G\beta E_n) \quad \mathbb{P}(\Lambda \in dx) := \frac{1}{Z_n^\beta} |\Delta(x)|^\beta e^{-\frac{1}{2}\sum_{j=1}^n x_i^2} \prod_{j=1}^n dx_i \ , \tag{3.1.1}$$

with Z_n^β being the normalization constant. It is a particular case of β-ensembles with general confining potential V:

$$\mathbb{P}(\Lambda \in dx) := \frac{1}{Z_n^\beta} |\Delta(x)|^\beta e^{-\sum_{j=1}^n V(x_i)} \prod_{j=1}^n dx_i \ . \tag{3.1.2}$$

From [DE02], the $G\beta E_n$is conveniently obtained as the spectrum of the tridiagonal matrix:

$$T_\beta = \begin{pmatrix} \mathcal{N}_1 & \chi_{\frac{1}{2}(n-1)\beta} & 0 & \dots & 0 & 0 \\ \chi_{\frac{1}{2}(n-1)\beta} & \mathcal{N}_2 & \chi_{\frac{1}{2}(n-2)\beta} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \mathcal{N}_{n-1} & \chi_{\frac{1}{2}\beta} \\ 0 & 0 & 0 & \dots & \chi_{\frac{1}{2}\beta} & \mathcal{N}_n \end{pmatrix} \tag{3.1.3}$$

where the variables with different symbols are independent. χ_k stands for a χ-distributed random variable with k degrees of freedom and $\mathcal{N}_i$ stands for a standard Gaussian random variable. We record this fact for future reference as

THEOREM 3.1.1 (Dumitriu and Edelman, [DE02]). *The spectrum of T_β, as given in Eq. (3.1.3), is the β-ensemble on the line, with quadratic confinement potential.*

Moreover, the Macdonald-Mehta-Opdam identity holds true:

$$Z^{\beta}(\mathbb{R}^n) = \int_{\mathbb{R}^n} dx \ \ |\Delta(x)|^{\beta} e^{-\frac{1}{2}\sum_{j=1}^n x_i^2} = (2\pi)^{\frac{1}{2}n} \prod_{j=1}^n \frac{\Gamma\left(1+\frac{\beta}{2}j\right)}{\Gamma\left(1+\frac{\beta}{2}\right)} \ .$$

Thanks to this tridiagonal model, the spectra of $G\beta E_n$ and $G\beta E_{n+1}$ are naturally coupled. This is again the same coupling in $\beta = 2$ of the projective measures of Olshanki-Vershik 2.2.2, after the so-called Trotter reduction.

From a physical perspective, the spectrum gives a log-gas with quadratic confinement potential. The fact that Dumitriu and Edelman's model has independent entries is miraculous and shows that β-ensembles are integrable in a sense.

The present work provides another proof that Dumitriu and Edelman's triadiagonal model has a spectrum distributed according to the β-ensemble on the line. While there is no new result per se, it is the approach that is novel. This derivation uses a Hamiltonian technique based on the scattering for the Toda flow. It is perhaps "the symplectic proof" that Dumitriu and Edelman mention in their paper [**DE02**] in the form of their Remark 2.10. Also, since this geometric proof splits the space into independent entries, it gives the change of variables which produces a proof of the Macdonald-Mehta-Opdam (MMO) integral.

The general approach goes as follows. We consider a random matrix with fixed distribution as starting point for the Toda flow. As the Toda flow is an integrable dynamical system that diagonalizes matrices, we are able to keep track of the matrix distribution throughout the flow.

Another interesting point regards the nature of the integrability of such models and we shed some light on the matter. Indeed, with initial measures expressed in terms of Casimirs, i.e. invariants of motion, the spectral distribution has a tractable and closed form expression. The particularity of quadratic potentials is that they are expressed using the first Casimir only, whose special structure leads to independence in the entries of the matrix model.

Summary. We start by developing the necessary results from the theory of the Toda lattice in Section 3.2.

Then we explicitly compute the scattering asymptotics, using orthogonal polynomials techniques. We will be particularly interested in precise scattering asymptotics which are originally due to Moser [**Mos75**] and which show the appearance of a logarithmic interaction via a Vandermonde. This allows to construct a scattering map between the generalised Toda flow and a free dynamic.

Before proving the main result, we find invariant volume forms under the Toda flow. These will play the role of reference measures. Putting everything together in Section 3.4 shows that, indeed, the spectrum is distributed as (3.1.1). The MMO formula is obtained by keeping track of the normalizing constant.

3.2. Definition of the Toda flow

Notations. Let $\mathfrak{k}$, $\mathfrak{a}$ and $\mathfrak{n}$ be respectively the subspaces in $M_n(\mathbb{C})$ of anti-Hermitian matrices, diagonal real matrices and upper triangular matrices. We have the direct sum decomposition:

$$M_n(\mathbb{C}) = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n} \tag{3.2.1}$$

Any matrix $X \in M_n(\mathbb{C})$ has a unique triangular decomposition into:

$$X = [X]_- + [X]_0 + [X]_+ \tag{3.2.2}$$

where $[X]_-$ (resp. $[X]_+$) are respectively lower triangular and upper triangular. $[X]_0$ is diagonal. For each subspace E in the direct sum decomposition (3.2.1), we denote the projection onto E by Π_E. These projection are given for $X \in M_n(\mathbb{C})$ by the expressions:

$$\Pi_{\mathfrak{k}}(X) = [X]_- - [X]_-^* + i\Im[X]_0\,; \quad \Pi_{\mathfrak{a}}(X) = \Re[X]_0\,; \quad \Pi_{\mathfrak{n}}(X) = [X + X^*]_+ \ .$$

Indeed, one can easily check that $\Pi_{\mathfrak{k}} + \Pi_{\mathfrak{a}} + \Pi_{\mathfrak{n}} = \mathrm{id}_{M_n(\mathbb{C})}$.

3.2.1. Flow definition. Let $\mathcal{T}$ be the space of symmetric tridiagonal matrix form:

$$X = \begin{pmatrix} a_1 & b_1 & 0 & \dots & 0 & 0 \\ b_1 & a_2 & b_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & \dots & b_{n-1} & a_n \end{pmatrix},$$

with $b_j > 0$.

The Toda flow acts on $\mathcal{T}$ via the differential equation:

$$\dot{X} = [X, \Pi_{\mathfrak{k}}(X)] \ . \tag{3.2.3}$$

Formally, define the vector field V at $x \in \mathcal{T}$ by:

$$V_x = [x, \Pi_{\mathfrak{k}}(x)] \ .$$

And the Toda flow is obtained by exponentiating the vector field V. In fact the flow is equivalent to the pair of equations:

$$\begin{cases} X_t & = Q_t \Lambda Q_t^* \\ \dot{Q} & = -\Pi_{\mathfrak{k}}(X)\, Q \end{cases}$$

by noticing that equation (3.2.3) can be rewritten:

$$\dot{X} = \left[\dot{Q}Q^{-1}, X\right]$$

Thanks to this Lax-pair formulation, we see that the flow acts by isospectral transformations.

REMARK 3.2.1 (The Toda flow preserves $\mathcal{T}$). *Positivity of the extra-diagonal is preserved because equation (3.2.3) implies:*

$$\dot{b_i} = b_i\,(a_i - a_{i+1}) \tag{3.2.4}$$

3.2.2. Long time behavior: the sorting property. We assume that X_0 is diagonalizable with distinct eigenvalues $(\Lambda_1, \dots, \Lambda_n)$ with

$$\Lambda_1 > \Lambda_2 > \dots > \Lambda_n \ .$$

The diagonalized form of the initial data is given by:

$$X_0 = Q_0 \Lambda Q_0^* \ .$$

It is known in numerical analysis that flow performs continuously the QR algorithm and thus gives the spectrum in long time. Similar results hold for other families of isospectral transformations (see the survey [**Wat84**]).

THEOREM 3.2.2 (Symes [Sym82] in tridiagonal case, [Chu84] in complex case). *The Toda flow diagonalizes matrices in the following sense:*

- *The Toda flow interpolates in continuous time the Arnoldi-Lanczos-QR algorithm used in numerical analysis – See Chapter 6 for more information.*
- *We have the following long time behavior:*

$$\lim_{t\to\infty} X_t = \Lambda$$

$$\lim_{t\to-\infty} X_t = w_0 \Lambda w_0$$

where w_0 is the permutation matrix reversing the order of the canonical basis. In the context of reflection groups, w_0 is the longest element in the symmetric group $\mathfrak{S}_n$ when written as a product of transpositions $(i\ i+1)$.

From this result arises the idea of keeping track of the distribution of the matrix T_β in Eq. (3.1.3), continuously throughout the flow.

REMARK 3.2.3. *A classical remark in eigenvalue problems is that Q has a special structure because it conjugates a diagonal matrix to tridiagonal matrix. In fact, Q_t can be entirely recovered from the first row of the matrix. See Theorem 7.2.1 in Partlett [Par98]. This first row u is important because it plays the role of angle coordinates in the integrable Toda flow.*

3.2.3. Flow on tridiagonal matrices. Let us now explain how the specialisation to tridiagonal matrices gives the original Toda Hamiltonian flow. Let n identical particles, seen as point masses with mass normalized to 1. The configuration space is then $\mathbb{R}^{2n}$. A configuration is a pair $(p, q) \in \mathbb{R}^{2n}$ where p are momenta and q are positions.

The dynamical system defined by Toda [Tod89] is the Hamiltonian system associated to

$$H = \frac{\|p\|^2}{2} + V(q)\ , \tag{3.2.5}$$

with V the Toda potential:

$$V(q) = \sum_{i=1}^{n-1} 2e^{-(q_i - q_{i+1})}\ .$$

Therefore, the equations of motion are given by:

$$\dot{q}_j = \frac{\partial H}{\partial p_j} = p_j\ , \tag{3.2.6}$$

$$\dot{p}_j = -\frac{\partial H}{\partial q_j} = 2e^{-(q_j - q_{j+1})} - 2e^{-(q_{j-1} - q_j)}\ . \tag{3.2.7}$$

The previous equations are valid for all indices $1 \le j \le n$ by by considering $q_0 = -\infty$ and $q_{n+1} = \infty$.

Because the center of mass has a uniform dynamic, we can assume that it is fixed and reduce the configuration space to $\mathbb{R}^{2(n-1)}$. Now, if one introduces the Flaschka variables

$$a_i = p_i$$

$$b_i = 2e^{-\frac{1}{2}(q_i - q_{i+1})}$$

and forms the tridiagonal matrix

$$X = \begin{pmatrix} a_1 & b_1 & 0 & \dots & 0 & 0 \\ b_1 & a_2 & b_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & \dots & b_{n-1} & a_n \end{pmatrix},$$

then the Hamiltonian dynamic (3.2.6) (3.2.7) is exactly equivalent to the Toda flow 3.2.3 acting on real tridiagonal matrices. Note that in the Flaschka variables, the Hamiltonian takes the form

$$H = \frac{1}{2}\sum_{i=1}^{n} a_i^2 + \sum_{i=1}^{n} b_i^2 = \frac{1}{2}\operatorname{Tr} X^2 \ .$$

From the diagonalisation property in Theorem 3.2.2, we easily deduce the crude scattering behavior as $t \to \infty$:

$$p_i(t) = \Lambda_i + o(1) \quad , \quad q_i(t) = \Lambda_i t + o(t) \ , \tag{3.2.8}$$

and as $t \to -\infty$:

$$p_i(t) = \Lambda_{n-i+1} + o(1) \quad , \quad q_i(t) = \Lambda_{n-i+1} t + o(t) \ . \tag{3.2.9}$$

If one is interested in the scattering map for momenta i.e the relation between the behaviors as $t \to \pm\infty$, it is given by reordering eigenvalues in opposite order. This tantamounts to the multiplication by the permutation matrix w_0. The result holds in fact more generally from the works of Goodman and Wallach for Toda lattices in other Lie types [GW84, Subsection 2.3].

3.3. Moser's scattering

For our purposes, we are much more interested in the scattering of positions. In a really beautiful paper [Mos75], Moser refines the $o(t)$ error in Eq. (3.2.8) and Eq. (3.2.9). He finds there exists a $\delta > 0$, depending on eigenvalue gaps, such that $t \to \infty$:

$$q_i(t) = \Lambda_i t + \beta_i^+ + o(e^{-\delta|t|})$$

and as $t \to -\infty$:

$$q_i(t) = \Lambda_{n-i+1} t + \beta_i^- + o(e^{-\delta|t|})$$

and the differences $\beta^+_{n-i+1} - \beta^-_i$ are related to a logarithmic interaction potential. The exact expression is given in [Mos75] eq. (4.3) and (4.4):

$$\beta^+_{n-i+1} - \beta^-_i = 2\sum_{j<i} \log|\Lambda_i - \Lambda_j| - 2\sum_{i<j} \log|\Lambda_i - \Lambda_j|$$

Hence, the second order of the scattering in position variables reveals the logarithmic interaction potential for eigenvalues. That was the starting point of our investigations.

Solution by inverse scattering. In order to completely solve the Toda flow, at the theoretical level, one starts by computing the diagonalization of the initial data X_0. Since Λ is given by the infinite time behavior of the system, it is called the scattering data. Inverse scattering consists in using Λ in order to compute the finite time solution, which is given by the computation of Q.

Nevertheless, due to the special structure of Q discussed in Remark 3.2.3, there is no need to compute the entire matrix. It suffices to keep track of the vector:

$$u := Q^{-1}e$$

where $e = e_1$ is the first vector in $\mathbb{C}^n$.

PROPOSITION 3.3.1. *The vector u follows the dynamic:*

$$u_t = \frac{e^{\Lambda t}u_0}{\|e^{\Lambda t}u_0\|}$$

PROOF. Since

$$\dot{Q} = -\Pi_{\mathfrak{k}}XQ = -\left(X - [X]_0 - [X]_-^* - [X]_+\right)Q = -Q\Lambda + \left([X]_0 + [X^* + X]_+\right)Q \ ,$$

we have that:

$$\frac{d}{dt}\left(Q_t e^{\Lambda t}\right) = \left([X]_0 + [X^* + X]_+\right)Q_t e^{\Lambda t}$$

The previous equation is a right-invariant autonomous equation, with upper triangular increments. As a consequence, there exists an upper triangular matrix T_t with positive diagonal such that:

$$Q_t e^{\Lambda t} = T_t Q_0 \ ,$$

and hence:

$$u_t = Q_t^{-1}e = e^{\Lambda t}Q_0^{-1}T_t^{-1}e = \frac{e^{\Lambda t}u_0}{[T_t]_{11}} \ .$$

The proof is finished as u_t needs to be of norm 1. □

3.3.1. Precise scattering asymptotics. As mentioned before, one sees the appearance of the Vandermonde in the second order scattering asymptotics for positions, in the real tridiagonal case, thanks to Moser's result [Mos75]. His approach relied on real analyticity of the flow and seems difficult to adapt or to generalize. We will rather use an orthogonal polynomial technique that expresses the action variable in a form more amenable to asymptotics. The technique is used on real tridiagonal matrices to express orthogonal polynomials thanks to the coefficients in the three term recurrence (See Chapter II in [Sze75]). Its application to the tridiagonal Toda has been implemented in handwritten lecture notes of Deift, the author managed to get his hands on.

Let us introduce the Gram determinant, using the unit vector $e = e_1$ again:

$$\Delta_k = \det\left(\langle X^{i-1}e, X^{j-1}e\rangle\right)_{i,j=1}^{k} \ .$$

The scattering of positions can be observed from the convergence of the b_i to zero, and the analogue of Moser's scattering result comes from the asymptotic analysis of Δ_k as:

$$\Delta_k = \prod_{i=1}^{k} b_i^{2(k-i)} \ . \tag{3.3.1}$$

Indeed, notice that, for all $k \leq n$:

$$X^k e = \left(\prod_{i=1}^{k} b_i\right) e_{k+1} + h_k$$

where $h_k \in \mathrm{Span}_{\mathbb{C}}(e_1, \ldots, e_k)$. Hence:

$$e \wedge Xe \wedge X^2 e \wedge \cdots \wedge X^{k-1} e = \left(\prod_{l=1}^{k-1} \prod_{i=1}^{l} b_i \right) e_1 \wedge e_2 \wedge e_3 \wedge \cdots \wedge e_k$$

and

$$\Delta_k = \det\left(\langle X^{i-1} e, X^{j-1} e \rangle \right)_{i,j=1}^{k} = \| e \wedge Xe \wedge X^2 e \wedge \cdots \wedge X^{k-1} e \|^2 = \prod_{i=1}^{k} b_i^{2(k-i)}$$

THEOREM 3.3.2 (Precise scattering asymptotics). *As $t \to \infty$:*

$$\Delta_k(t) = e^{2\left(\sum_{l=1}^{k} \Lambda_1 - \Lambda_l\right)t} \left| \Delta(\Lambda_1, \ldots, \Lambda_k) \right|^2 \prod_{l=1}^{k} \left| \frac{u_0(l)}{u_0(1)} \right|^2 (1 + o(1))$$

In particular, the initial angle coordinates can be read from asymptotics:

$$\left| \frac{u_0(l)}{u_0(1)} \right| \sim e^{(\Lambda_1 - \Lambda_k)t} \frac{\left| \Delta(\Lambda_1, \ldots, \Lambda_{k-1}) \right|}{\left| \Delta(\Lambda_1, \ldots, \Lambda_k) \right|} \sqrt{\frac{\Delta_k(t)}{\Delta_{k-1}(t)}}$$

This is nothing but Moser's result with a different proof. Our proof (actually Deift's) is proved by introducing the probability measure on $\mathbb{R}$ - often called the "spectral measure" in the literature devoted to Jacobi operators:

$$\mu_t(d\lambda) = \sum_{k=1}^{n} \left| \langle e_k, u_t \rangle \right|^2 \delta_{\Lambda_k}(d\lambda),$$

thanks to which the Gram determinant has a nice formula:

PROPOSITION 3.3.3.

$$\Delta_k(t) = \int_{\lambda_1 > \lambda_2 > \cdots > \lambda_k} \left| \Delta(\lambda_1, \ldots, \lambda_k) \right|^2 \prod_{l=1}^{k} \mu_t(d\lambda_l)$$

In particular, we have the exact result:

$$\Delta_n(t) = \left| \Delta(\Lambda_1, \ldots, \Lambda_n) \right|^2 \prod_{l=1}^{n} \left| \langle e_l, u_t \rangle \right|^2 .$$

PROOF. We have:

$$\begin{aligned} \langle X_t^{i-1} e, X_t^{j-1} e \rangle &= \langle Q_t \Lambda^{i-1} Q_t^* e, Q_t \Lambda^{j-1} Q_t^* e \rangle \\ &= \langle \Lambda^{i-1} u_t, \Lambda^{j-1} u_t \rangle \\ &= \sum_k \Lambda_k^{i-1+j-1} \left| \langle e_k, u_t \rangle \right|^2 \\ &= \int_{\mathbb{R}} \lambda^{i-1+j-1} \mu_t(d\lambda) \ . \end{aligned}$$

And therefore, by multilinearity of the determinant with respect to columns:

$$\begin{aligned} \Delta_k(t) &= \det\left(\langle X_t^{i-1} e, X_t^{j-1} e \rangle \right)_{i,j=1}^{k} \\ &= \det\left(\int_{\mathbb{R}} \lambda_j^{i-1+j-1} \mu_t(d\lambda_j) \right)_{i,j=1}^{k} \end{aligned}$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^k} \det\left(\lambda_j^{i-1}\lambda_j^{j-1}\right)_{i,j=1}^k \prod_{l=1}^k \mu_t\left(d\lambda_l\right) \\
&= \int_{\mathbb{R}^k} \det\left(\lambda_j^{i-1}\right)_{i,j=1}^k \prod_{j=1}^k \lambda_j^{j-1} \prod_{l=1}^k \mu_t\left(d\lambda_l\right) \\
&= \int_{\mathbb{R}^k} \Delta(\lambda_1,\ldots,\lambda_k) \prod_{l=1}^k \lambda_l^{l-1} \prod_{l=1}^k \mu_t\left(d\lambda_l\right) \ .
\end{aligned}
$$

Using a standard anti-symmetrization trick:

$$
\begin{aligned}
\Delta_k(t) =& \frac{1}{k!} \sum_{\sigma\in W} \int_{\mathbb{R}^k} \varepsilon(\sigma)\Delta\left(\lambda_1,\ldots,\lambda_k\right) \prod_{l=1}^k \lambda_{\sigma(l)}^{l-1} \prod_{l=1}^k \mu_t\left(d\lambda_l\right) \\
=& \frac{1}{k!} \int_{\mathbb{R}^k} \left|\Delta\left(\lambda_1,\ldots,\lambda_k\right)\right|^2 \prod_{l=1}^k \mu_t\left(d\lambda_l\right) \ .
\end{aligned}
$$

And by symmetry, we can order the integration variables:

$$
\Delta_k(t) = \int_{\lambda_1>\lambda_2>\cdots>\lambda_k} \left|\Delta\left(\lambda_1,\ldots,\lambda_k\right)\right|^2 \prod_{l=1}^k \mu_t\left(d\lambda_l\right) \ .
$$

□

Moreover, μ_t has a simple dynamic:

PROPOSITION 3.3.4.

$$
\mu_t\left(d\lambda\right) = \frac{e^{2\lambda t}}{\left\|e^{\Lambda t}u_0\right\|^2}\mu_0(d\lambda)
$$

PROOF. From the the dynamic of u in Proposition 3.3.1, we obtain:

$$
\begin{aligned}
\mu_t\left(d\lambda\right) =& \sum_k \left|\langle e_k, u_t\rangle\right|^2 \delta_{\Lambda_k}(d\lambda) \\
=& \frac{1}{\left\|e^{\Lambda t}u_0\right\|^2} \sum_k e^{2\Lambda_k t}\left|\langle e_k, u_0\rangle\right|^2 \delta_{\Lambda_k}(d\lambda) \\
=& \frac{1}{\left\|e^{\Lambda t}u_0\right\|^2} e^{2\lambda t}\mu_0(d\lambda)
\end{aligned}
$$

□

The Gram determinant has a nice formula in terms of the previous measure:

PROOF OF THEOREM 3.3.2. The combination of the two previous propositions gives:

$$
\Delta_k(t) = \left\|e^{\Lambda t}u_0\right\|^{-2k} \int_{\lambda_1>\lambda_2>\cdots>\lambda_k} e^{2\left(\sum_{l=1}^k \lambda_l\right)t} \left|\Delta\left(\lambda_1,\ldots,\lambda_k\right)\right|^2 \prod_{l=1}^k \mu_0\left(d\lambda_l\right)
$$

As $t\to\infty$, the integral's dominant terms are obtained by picking up only the k largest eigenvalues $\Lambda_1 > \Lambda_2 > \cdots > \Lambda_k$. Hence the asymptotics:

$$
\Delta_k(t) = \left\|e^{\Lambda t}u_0\right\|^{-2k} \int_{\lambda_1>\lambda_2>\cdots>\lambda_k} e^{2\left(\sum_{l=1}^k \lambda_l\right)t} \left|\Delta\left(\lambda_1,\ldots,\lambda_k\right)\right|^2 \prod_{l=1}^k \mu_0\left(d\lambda_l\right)
$$

$$=\|e^{\Lambda t}u_0\|^{-2k}e^{2\left(\sum_{l=1}^k \Lambda_l\right)t}\left|\Delta\left(\Lambda_1,\dots,\Lambda_k\right)\right|^2\prod_{l=1}^k\left|\langle e_l,u_0\rangle\right|^2(1+o(1))$$

Combining that fact with the asymptotics for $\|e^{\Lambda t}u_0\|$ yields the result. □

In particular, Moser's result appears explicitly in:

$$\begin{aligned}&b_k^2(t)\\ =&\frac{\Delta_k\Delta_{k-2}}{\Delta_{k-1}^2}\\ =&e^{2(\Lambda_k-\Lambda_{k-1})t}\frac{\left|\Delta\left(\Lambda_1,\dots,\Lambda_k\right)\right|^2\left|\Delta\left(\Lambda_1,\dots,\Lambda_{k-1}\right)\right|^2}{\left|\Delta\left(\Lambda_1,\dots,\Lambda_{k-2}\right)\right|^4}\frac{\left|u_0(k)\right|^2\left|u_0(k-1)\right|^2}{\left|u_0(k-2)\right|^4}(1+o(1))\end{aligned}$$

3.3.2. Scattering map. Consider the vector field on $\mathcal{T}$:

$$\forall x\in\mathcal{T},\ V_x^{free}:=-\sum_{i=1}^{n-1}E_{i+1,i}b_i\left(a_i-a_{i+1}\right)$$

whose flow gives:

$$e^{tV^{free}}\cdot b_i=b_ie^{-(a_i-a_{i+1})t}$$

In the scattering regime, diagonals are constant and the dynamic of position is linear in time. Therefore, the "free" dynamic corresponding to isolated particles is given by the flow $e^{tV^{free}}$. The question at hand is the behavior of the sequence of diffeomorphisms:

$$e^{-tV}\cdot e^{tV^{free}}$$

as $t\to\infty$. Clearly, the flow generated by V^{free} is not isospectral, but in the regime where the extra-diagonals are small, it almost is.

THEOREM 3.3.5. *Let Δ be the set of matrices with increasing diagonal entries. There exists a map $\mathcal{S}:\mathcal{T}\cap\Delta\to\mathcal{T}$ such that:*

$$\mathcal{S}=\lim_{t\to\infty}e^{-tV}\cdot e^{tV^{free}}$$

We have:

$$\mathcal{S}x=Q[x]_0Q^*$$

where the first row of Q, u_0 satisfies:

$$\langle e_k,u_0\rangle=\frac{\prod_{i=1}^{k-1}b_i(0)}{\prod_{i=1}^{k-1}\left|\Lambda_k-\Lambda_i\right|}\langle e_1,u_0\rangle\ .$$

PROOF. Let x be a matrix in $\mathcal{T}\cap\Delta$. Then:

$$e^{tV^{free}}\cdot x=[x]_0+\sum_i x_{i+1,i}e^{-(x_{i,i}-x_{i+1,i+1})t}\left(E_{i+1,i}+E_{i,i+1}\right)$$

Therefore, $e^{tV^{free}}\cdot x$ converges exponentially fast to a diagonal matrix, for which the sorted eigenvalues lie on the diagonal. The conserved quantities by the flow e^{-tV} are therefore exponentially close to $[x]_0$.

Now consider an asymptotic behavior, obtained after a free evolution:

$$b_k(t)=b_k(0)e^{(\Lambda_{k+1}-\Lambda_k)t}\left(1+o(1)\right)$$

Thanks to Theorem 3.3.2, this corresponds to a scattering state at time t for the Toda evolution which had at time 0 the angle coordinates:

$$\begin{aligned}\frac{|v_0(k)|}{|v_0(1)|} &= e^{(\Lambda_1-\Lambda_k)t}\sqrt{\frac{\Delta_k(t)}{\Delta_{k-1}(t)}}\frac{|\Delta(\Lambda_1,\dots,\Lambda_{k-1})|}{|\Delta(\Lambda_1,\dots,\Lambda_k)|}\left(1+o(1)\right)\\ &= e^{(\Lambda_1-\Lambda_k)t}\left(\prod_{i=1}^{k-1}b_i(t)\right)\frac{|\Delta(\Lambda_1,\dots,\Lambda_{k-1})|}{|\Delta(\Lambda_1,\dots,\Lambda_k)|}(1+o(1))\\ &= \frac{\prod_{i=1}^{k-1}b_i(0)}{\prod_{i=1}^{k-1}|\Lambda_k-\Lambda_i|}(1+o(1))\end{aligned}$$

Therefore $e^{-tV}\cdot e^{tV^{free}}\cdot x$ converges to the element with action variables $[x]_0$ and angle coordinates v_0. □

From the previous proof and from Proposition 3.3.3, we notice that:

$$\Delta_n(\mathcal{S}x) = |\Delta(\Lambda_1,\dots,\Lambda_n)|^2\prod_{k=1}^{n}|v_0(k)|^2 = |v_0(1)|^{2n}\prod_{i=1}^{n-1}b_i^{2(n-i)} \tag{3.3.2}$$

3.3.3. Invariant differential forms. Define the volume form on $\mathcal{T}$:

$$\omega_{\mathcal{T}} := \wedge_{i=1}^{n-1}\frac{db_i}{b_i}\wedge\omega_{\mathfrak{a}} \tag{3.3.3}$$

where

$$\omega_{\mathfrak{a}} = \wedge_{i=1}^{n}da_i\ .$$

As we will use these forms as reference integration measures, it is crucial that are invariant under the flow.

PROPOSITION 3.3.6 (Form invariance). *The Toda flow preserves the volume form $\omega_{\mathcal{T}}$:*

$$\mathcal{L}_V\omega_{\mathcal{T}} = 0$$

In particular:

$$\left(e^{tV}\right)_*\omega_{\mathcal{T}} = 0$$

PROOF. In the positions and momenta coordinates of the classical Toda Hamiltonian, the form $i_*\omega_{\mathcal{T}}$ can be written as:

$$i_*\omega_{\mathcal{T}} = (\wedge_{i=1}^{n}dp_i)\wedge\left(\wedge_{i=1}^{n-1}dq_i - dq_{i+1}\right)$$

which is the Liouville form of the Toda system. In this case, not only $i_*\omega_{\mathcal{T}}$ is preserved but there is an underlying symplectic form that is preserved. □

3.4. Novel proof of Theorem 3.1.1 via scattering

Consider a random matrix $X_0\in\mathcal{T}$ distributed as in (3.1.3). We are interested in the distribution of eigenvalues Λ as well as the distribution of the vector u_0. Clearly, from the properties of the Toda flow, for every bounded continuous function:

$$\mathbb{E}\left(f(\Lambda,u_0)\right) = \lim_{t\to\infty}\mathbb{E}\left(f\left([X_t]_0,u_0\right)\right)$$

Now, from the explicit expression of χ and Gaussian distributions, we have:

$$\mathbb{E}\left(f\left([X_t]_0,u_0\right)\right)$$

$$=C_{n,\beta}\int_{\mathcal{T}} f\left([e^{tV}\cdot X_0]_0, v_0\left(X_0\right)\right)\prod_{i=1}^{n-1} b_i(X_0)^{2\beta(n-i)}e^{-\frac{1}{2}\|X_0\|^2}\omega_{\mathcal{T}}(dX_0)$$

where

$$C_{n,\beta}=\frac{1}{(2\pi)^{n(n+1)}\prod_{j=1}^{n-1}\Gamma(\frac{1}{2}j\beta)}\ .$$

Thanks to Equation (3.3.1):

$$\begin{aligned}&\mathbb{E}\left(f\left([X_t]_0,u_0\right)\right)\\ =&C_{n,\beta}\int_{\mathcal{T}} f\left([e^{tV}\cdot X_0]_0,u_0\left(X_0\right)\right)\Delta_n(X_0)^{\beta}e^{-\frac{1}{2}\|X_0\|^2}\omega_{\mathcal{T}}(dX_0)\end{aligned}$$

Then, we make successively the change of variable by the flow e^{tV} and $e^{tV^{free}}$. Thanks to the invariance of measures:

$$\begin{aligned}&\mathbb{E}\left(f\left([X_t]_0,u_0\right)\right)\\ =&C_{n,\beta}\int_{\mathcal{T}} f\left([e^{tV}\cdot X_0]_0,u_0\left(e^{-tV}\cdot e^{tV}\cdot X_0\right)\right)\\ &\qquad\Delta_n(e^{-tV}\cdot e^{tV}\cdot X_0)^{\beta}e^{-\frac{1}{2}\|X_0\|^2}\omega_{\mathcal{T}}(dX_0)\\ =&C_{n,\beta}\int_{\mathcal{T}} f\left([X_0]_0,u_0\left(e^{-tV}\cdot X_0\right)\right)\\ &\qquad\Delta_n(e^{-tV}\cdot X_0)^{\beta}e^{-\frac{1}{2}\|X_0\|^2}\omega_{\mathcal{T}}(dX_0)\\ =&C_{n,\beta}\int_{\mathcal{T}} f\left([X_0]_0,u_0\left(e^{-tV}\cdot e^{tV^{free}}\cdot X_0\right)\right)\\ &\qquad\Delta_n(e^{-tV}\cdot e^{tV^{free}}\cdot X_0)^{\beta}e^{-\frac{1}{2}\|e^{tV^{free}}\cdot X_0\|^2}\omega_{\mathcal{T}}(dX_0)\ .\end{aligned}$$

Now, as $t\to\infty$, we have the appearance of the scattering map $\mathcal{S}=\lim_{t\to\infty}e^{-tV}\cdot e^{tV^{free}}$. By assuming f supported on $\mathcal{T}\cap\Delta$ in the first variable, there is no need to bother about the initial data X_0 having its eigenvalues ordered. Moreover, the coefficients of the extra-diagonal in $e^{tV^{free}}\cdot X_0$ asymptotically vanish. Hence:

$$\begin{aligned}&\mathbb{E}\left(f(\Lambda,u_0)\right)\\ =&C_{n,\beta}\int_{\mathcal{T}} f\left([X_0]_0,u_0\left(\mathcal{S}X_0\right)\right)\Delta_n(\mathcal{S}X_0)^{\beta}e^{-\frac{1}{2}\|[X_0]_0\|^2}\omega_{\mathcal{T}}(dX_0)\\ =&C_{n,\beta}\int_{\lambda_1>\lambda_2>\cdots>\lambda_n} d\Lambda\ e^{-\frac{1}{2}\sum|\Lambda_i|^2}\\ &\int_{(\mathbb{R}^+)^{n-1}}\wedge_{i=1}^{n-1}\frac{db_i}{b_i}\ f\left(\Lambda,u_0\left(\mathcal{S}X_0\right)\right)\Delta_n(\mathcal{S}X_0)^{\beta}\ .\end{aligned}$$

Finally, we conclude thanks to the identity (3.3.2) that:

$$\begin{aligned}&\mathbb{E}\left(f(\Lambda,u_0)\right)\\ =&C_{n,\beta}\int_{\lambda_1>\lambda_2>\cdots>\lambda_n} d\Lambda e^{-\frac{1}{2}\sum|\Lambda_i|^2}\\ &\int_{(\mathbb{R}^+)^{n-1}}\wedge_{i=1}^{n-1}\frac{db_i}{b_i}\ \left(\prod_{i=1}^{n-1}b_i^{2\beta(n-i)}\right)|u_0(1)|^{2\beta n}f\left(\Lambda,u_0\left(\mathcal{S}X_0\right)\right)\ .\end{aligned}$$

Moreover, by changing the integration variables to c_k:

$$\frac{|u_0(k+1)|}{|u_0(1)|} = \frac{\prod_{i=1}^{k-1} b_i}{\prod_{i=1}^{k-1} |\Lambda_k - \Lambda_i|} = c_k \ ,$$

and knowing that:

$$\prod_{i=1}^{n-1} b_i^{(n-i)} = |\Delta(\Lambda_1, \dots, \Lambda_n)| \prod_{i=1}^{n-1} c_i \ ,$$

we obtain:

$$\begin{aligned}
&\mathbb{E}\left(f(\Lambda, u_0)\right) \\
=&C_{n,\beta} \int_{\lambda_1 > \lambda_2 > \cdots > \lambda_n} d\Lambda \, |\Delta(\Lambda_1, \dots, \Lambda_n)|^{2\beta} \, e^{-\frac{1}{2}\sum |\Lambda_i|^2} \\
&\int_{(\mathbb{R}^+)^{n-1}} \wedge_{i=1}^{n-1} \frac{dc_i}{c_i} \left(\prod_{i=1}^{n-1} c_i^{2\beta} \right) |u_0(1)|^{2\beta n} f\left(\Lambda, |u_0(1)|\,(1, c_1, c_2 \dots)\right) \ .
\end{aligned}$$

We recognize the distribution of the Gaussian β-ensemble for the spectrum Λ. Moreover, the distribution of the vector u_0 is given by the β-Dirichlet distribution.

This concludes the proof of Theorem 3.1.1.

CHAPTER 4

Pitman-type theorems and representation theory

4.1. Quick representation-theoretic survey of Pitman's theorem

Let us discuss the classical Pitman theorem with a focus on its relationship to representation theory. To me, it is always very amusing that Pitman's theorem from probability theory [**Pit75**], while simple in its statement, can take us all the way to the representation theory of Lie groups and quantum groups, with the associated non-commutative geometry.

THEOREM 4.1.1 (Pitman's 2M-X Theorem, Discrete version).
Let $(X_n; n \in \mathbb{N})$ be a simple random walk in $\mathbb{Z}$, i.e. increments are independent and

$$\forall n \in \mathbb{Z}_+,\ \mathbb{P}\left(X_{n+1} - X_n = 1\right) = 1 - \mathbb{P}\left(X_{n+1} - X_n = -1\right) = \frac{1}{2}\ .$$

Then the process $(\Lambda_n^\infty; n \in \mathbb{N})$ defined as

$$\Lambda_n^\infty := X_n - 2 \inf_{0 \le k \le n} X_k$$

is a Markov chain on $\mathbb{N}$ with transition kernel given by Q:

$$Q\left(\lambda, \lambda+1\right) = \frac{\lambda+2}{2(\lambda+1)}\ , \quad Q\left(\lambda, \lambda-1\right) = \frac{\lambda}{2(\lambda+1)}\ . \tag{4.1.1}$$

Moreover, the missing information is stationary and equidistributed in law in the sense that for all $n \in \mathbb{N}$:

$$\mathcal{L}\left(X_n \mid \mathcal{F}_n^{\Lambda^\infty},\ \Lambda_n^\infty = \lambda\right) \ = \ \frac{1}{\lambda+1} \sum_{\substack{-\lambda \le k \le \lambda \\ \lambda - k\ even}} \delta_k\ .$$

Pitman's original proof uses the combinatorics of random walks and is formulated in terms of the running maximum instead of the running infimum. Both are equivalent upon replacing X by $-X$, hence the common name of "Pitman's $2M - X$ Theorem", where the capital letter M stands for "Maximum".

From the discrete version, one obtains a Brownian version thanks to a simple application of Donsker's invariance principle and by computing the diffusive rescaling of the Markov kernel Q.

THEOREM 4.1.2 (Pitman's 2M-X Theorem, Continuous version). *Let $(X_t\ ;\ t \in \mathbb{R}_+)$ be a standard Brownian motion. Then the process $(\Lambda_t^\infty\ ;\ t \in \mathbb{R}_+)$ defined as*

$$\Lambda_t^\infty := X_t - 2 \inf_{0 \le s \le t} X_s$$

is a Bessel 3 process, that is to say it has the same distribution as

$$\Lambda_t^0 := \sqrt{X_t^2 + Y_t^2 + Z_t^2}\ ,$$

where (X, Y, Z) is a Euclidean Brownian motion on $\mathbb{R}^3$.

Moreover, the missing information is stationary and equidistributed in law in the sense that:

$$\mathcal{L}\left(X_t \mid \mathcal{F}_t^{\Lambda^\infty},\ \Lambda_t^\infty = \lambda\right) \ = \ \frac{1}{2\lambda}\mathbb{1}_{[-\lambda,\lambda]}(x)dx \ .$$

REMARK 4.1.3. *The reader trained in probability theory knows that the Markov property is very fragile and can be easily broken, while*

$$\left(-\inf_{0\leq s\leq t} X_s \ ;\ t\in\mathbb{R}_+\right)$$

is the archetype of non-Markovian behavior. As such, Pitman's theorem is rather peculiar. It is also very rigid, as

$$\left(X_t - k\inf_{0\leq s\leq t} X_s \ ;\ t\in\mathbb{R}_+\right)$$

enjoys the Markov property only for $k = 0,\ 1,$ and 2 [MO04] *; the latter case being by far the most interesting in our opinion.*

REMARK 4.1.4. *The reader trained in Lie theory might appreciate the following comment. Recall that for every $\lambda \in \mathbb{N}$, there is a unique irreducible representation $V(\lambda)$ of $SL_2(\mathbb{C})$ with dimension $\dim V(\lambda) = \lambda + 1$. This representation $V(\lambda) \approx \mathrm{Sym}^{(n=\lambda)}(\mathbb{C}^2)$ is the n-th symmetric power of the fundamental representation of $SL_2(\mathbb{C})$. Furthermore, the Clebsch-Gordan rule:*

$$V(\lambda)\otimes\mathbb{C}^2 = V(\lambda+1) + V(\lambda-1)$$

translates to the following identity, upon computing the dimensions:

$$(\lambda+1)2 = (\lambda+2) + \lambda \ .$$

Upon normalizing the previous equality, one finds the transition probabilities of Eq. (4.1.1)*. As such, the process Λ appears to be respecting the structure constants of the tensor product, for representations of $SL_2(\mathbb{C})$. Also, the factor 2 is the same as $\alpha(\alpha^\vee) = 2$, where α is a root and $\alpha^\vee$ is the associated coroot in a root system.*

In fact, direct proofs of Theorem 4.1.2 at the level of continuous-time stochastic processes are available. Jeulin [RY13] has an approach that uses filtration enlargement techniques and [RP81] makes use of intertwinings of Markov kernels. These two proofs led to a flurry of very interesting probabilistic developments.

If other proofs and generalizations abound, we want to focus on two specific approaches where the complex group $SL_2(\mathbb{C})$ plays an important role. The approach by Bougerol and Jeulin [BJ02] is based on a geometric construction while the approach by Biane [Bia06, Bia09] is based on the representation theory of that group. The goal of this chapter is to exhibit a direct relationship between the two, via semi-classical limits.

The Lie algebra of SU_2 is

$$\mathfrak{su}_2 := T_e SU_2 \ , \tag{4.1.2}$$

which coincides with anti-Hermitian matrices. The complexification of $\mathfrak{su}_2$ is the Lie algebra of $SL_2(\mathbb{C})$:

$$\mathfrak{sl}_2 := T_e SL_2(\mathbb{C}) = \mathfrak{su}_2\otimes\mathbb{C} = \mathrm{Span}_{\mathbb{C}}\left(E,F,H\right) \ , \tag{4.1.3}$$

where:

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} ; E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} ; F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} . \tag{4.1.4}$$

Bougerol and Jeulin's approach via curvature deformation. In the paper [BJ02], Bougerol and Jeulin take a parameter $r > 0$ and consider a left-invariant process $g^r = (g^r_t \ ; \ t \geq 0)$ on the symmetric space $\mathbb{H}^3 = SL_2(\mathbb{C})/SU_2$. Because of the Gram-Schmidt decomposition, we make the identification $\mathbb{H}^3 = SL_2(\mathbb{C})/SU_2 \approx NA$, where NA is the subgroup of lower triangular matrices with positive diagonals. More precisely:

$$A := \left\{ \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \mid a \in \mathbb{R}^*_+ \right\} , \quad \text{and} \quad N := \left\{ \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix} \mid b \in \mathbb{C} \right\} .$$

The corresponding Lie algebras are denoted by $\mathfrak{a} := T_eA = \mathbb{R}H$ and $\mathfrak{n} := T_eN = \mathbb{R}F \oplus \mathbb{R}iF$.

In that identification, the process g^r satisfies the left-invariant stochastic differential equation (SDE for short)

$$\forall t \geq 0, \ g^r_t = \begin{pmatrix} \frac{1}{2} r dX_t & 0 \\ r(dY_t + iZ_t) & -\frac{1}{2} r dX_t \end{pmatrix} \circ g^r_t ,$$

where (X, Y, Z) is a standard Euclidean Brownian motion on $\mathbb{R}^3$. Here, the symbol $\circ$ refers to the Stratonovich integration convention. Solving explicitly the SDE yields for all $t \geq 0$:

$$g^r_t = \begin{pmatrix} e^{\frac{1}{2} r X_t} & 0 \\ r e^{\frac{1}{2} r X_t} \int_0^t e^{-r X_s} d(Y_s + iZ_s) & e^{-\frac{1}{2} r X_t} \end{pmatrix} . \tag{4.1.5}$$

The reader unfamiliar with stochastic integration should see the above equation as a definition for the process g^r. More importantly, the parameter $r > 0$ should be seen as a curvature parameter. The full explanation is given in the paper [CC21] where we see that we are considering the hyperbolic space $\mathbb{H}^3$ as the space with constant sectional curvature $-\frac{1}{2}r^2$. As such, there is no harm in loosely referring to r as curvature. At this stage, let us only mention the following. We have, as $r \to 0$:

$$g^r_t = \mathrm{Id} + r \begin{pmatrix} \frac{1}{2} X_t & 0 \\ Y_t + iZ_t & -\frac{1}{2} X_t \end{pmatrix} + o(r) =: \mathrm{Id} + r x^0_t + o(r) ,$$

and thus appears a three dimensional Brownian motion $\left(x^0_t = \frac{\partial g^r_t}{\partial r}|_{r=0} \ ; \ t \geq 0 \right)$ on $\mathfrak{a} \oplus \mathfrak{n} \approx \mathbb{R}^3$, which is a flat space. Because of Brownian motion's time-scaling properties, rescaling r amounts to speeding up the Brownian motion and hence the associated vector fields. As the process g^r moves more erratically as $r > 0$ grows larger, the non-commutativity of the underlying space NA becomes more apparent. One could say that the space increases in curvature, which is a key element in the following result by Bougerol and Jeulin.

Their result holds for all complex semi-simple groups G, but in the context of $G = SL_2(\mathbb{C})$, we have:

THEOREM 4.1.5 (Bougerol-Jeulin, [BJ02]). *Let $\frac{1}{2} r \Lambda^r_t$ be the radial part of g^r_t, i.e. $\exp(r\Lambda^r_t)$ is the largest singular value of g^r_t or equivalently that $\Lambda^r \geq 0$ and*

there exists $(k_1, k_2) : \mathbb{R}_+ \to SU_2 \times SU_2$ *such that*

$$g_t^r = k_1(t) \begin{pmatrix} e^{\frac{1}{2}r\Lambda_t^r} & 0 \\ 0 & e^{-\frac{1}{2}r\Lambda_t^r} \end{pmatrix} k_2(t) \ .$$

Then, Λ^r *is a process whose distribution does not depend on* $r > 0$*. It is explicitly given by:*

$$\Lambda_t^r = \frac{1}{r} \operatorname{Argcosh} \left[\frac{1}{2} r^2 \left| e^{\frac{1}{2}rX_t} \int_0^t e^{-rX_s} (dY_s + i dZ_s) \right|^2 + \cosh(rX_t) \right] \ , \tag{4.1.6}$$

where $\operatorname{Argcosh}(x) = \log\left(x + \sqrt{x^2 - 1}\right)$ *is the inverse of* $\cosh : \mathbb{R}_+ \to [1, \infty)$*. Moreover, for all* $t > 0$*, we have the limits in probability:*

$$\begin{cases} \Lambda_t^{r=0} & := & \mathbb{P} - \lim_{r \to 0} \Lambda_t^r & = & \sqrt{X_t^2 + Y_t^2 + Z_t^2} \ , \\ \Lambda_t^{r=\infty} & := & \mathbb{P} - \lim_{r \to \infty} \Lambda_t^r & = & X_t - 2 \inf_{0 \le s \le t} X_s \ . \end{cases} \tag{4.1.7}$$

In particular, these processes are both Bessel processes of dimension 3*.*

Because Bougerol and Jeulin treat the general case, for a general complex semi-simple Lie group G, extracting the above statement is not a trivial task. As part of our unifying picture, we shall provide a complete proof in the case of $SL_2(\mathbb{C})$, where the key arguments are simplified while giving a few illuminating computations. The only novelty in our treatment of Theorem 4.1.5 is in the proof that the law $\mathcal{L}(\Lambda^r)$ does not depend on $r > 0$. This fact is rather subtle and so is the argument of Bougerol and Jeulin. In [CC21] we give a short argument based on the rigidity of quantum groups and the results developed there. We also reinterpret the argument of Bougerol and Jeulin through the lens of spherical harmonic analysis, thereby showing where the curvature and the rigidity of quantum groups are hidden. By rigidity of quantum groups, we simply mean that key features of the representation theory such as characters and structure constants do not depend on the deformation parameter.

In fact, this same argument is at the heart of our approach to aimed at bridging directed last passage percolation and RMT in Section 4.4.

In any case, the important remark is that the Pitman transform on paths

$$\mathcal{P} : \ X \mapsto \left(X_t - 2 \inf_{0 \le s \le t} X_s \ ; \ t \ge 0 \right)$$

shows up in infinite (negative) curvature, while the norm process in $\mathbb{R}^3$ appears in flat curvature. The interpretation of the parameter r as curvature is mainly absent from the literature except in the very astute remark in the final paragraphs of [BJ02, Section 1].

4.2. Representation theory as the quantization of RMT

The standard framework for classical mechanics is to consider a symplectic manifold (M, ω) which, thanks to the Darboux theorem is locally isomorphic to the standard symplectic space $(\mathbb{R}^{2n}, \omega_{\mathbb{R}^{2n}})$ with n position coordinates and n momenta. More generally, one considers a Poisson manifold $(M, \{\cdot, \cdot\})$. The algebra of classical observables is $\mathcal{C}^\infty(M)$ and M is referred as the "phase space". Given a Hamilonian $H \in \mathcal{C}^\infty(M)$, its generates a flow via Hamilton's equations

$$\dot{f} = \{f, H\} \ , \tag{4.2.1}$$

which implicitly defines the dynamic $(x(t)\ ;\ t \in \mathbb{R})$ with

$$\forall f \in \mathcal{C}^\infty(M),\ \ f(t) = f(x(t))\ .$$

Notice that in this framework, a Hamiltonian H is basically any function in phase space, making this seem void of physical meaning.

Recall quantization is the (non-functorial) process which consists in replacing the classical objects of Hamiltonian mechanics by the quantum objects of quantum mechanics. This correspondence is not one-to-one but it is fair to say that

- Points in phase space M become vectors in a Hilbert space $\mathcal{H}$, called quantum states or wave-functions.
- Classical observables in the commutative algebra $\mathcal{C}^\infty(M)$ become operators in a non-commutative algebra $\mathrm{Op}^\hbar(M)$ acting on $\mathcal{H}$. Elements in $\mathrm{Op}^\hbar(M)$ are called quantum observables or measurement operators.
- Here $\hbar > 0$ is the Planck constant and formally we should have

  $$\text{“}\mathrm{Op}^\hbar(M) \xrightarrow{\hbar\to 0} \mathcal{C}^\infty(M)\text{”}\ .$$

  We will see an instance of this in the form of a map $\bmod \hbar$ and statements about semi-classical limits which formalize this.
- The commutator in the algebra $\mathrm{Op}^\hbar$ retrieves up to first order the Poisson bracket i.e. if $\mathcal{F}_1, \mathcal{F}_2 \in \mathrm{Op}^\hbar$ and $f_i = \mathcal{F}_i \bmod \hbar$ then

  $$[\mathcal{F}_1, \mathcal{F}_2] = \hbar\{f_1, f_2\} + o(\hbar)\ .$$

  We remain elusive on the meaning of the $o(\hbar)$ term, but it is a key ingredient in semi-classical limits. It does not make sense until we have considered representations and states allowing us to evaluate magnitudes of objects.

There are finer subtleties when dealing with geometric quantization theory, and we refer to excellent book [Woo92].

Elements of Kirillov's orbit method. Only in these few paragraphs, we shall give definitions in the general Lie type before very quickly specializing to $GL_n(\mathbb{C})$. Given a Lie group G, its Lie algebra is written $\mathfrak{g}$, and the space of linear forms on $\mathfrak{g}$ is the dual $\mathfrak{g}^*$. The dual Lie algebra $\mathfrak{g}^*$ is naturally endowed with a Poisson structure which foliates the space into orbits [CP95, Definition-Proposition 1.1.2]

$$\mathfrak{g}^* = \bigsqcup_\lambda \mathcal{O}(\lambda)\ , \tag{4.2.2}$$

where $\mathcal{O}(\lambda)$ is the orbit of $\lambda \in \mathfrak{g}^*$ under the co-adjoint action of G. Each orbit is symplectic and endowed with a symplectic form ω_{KKS} called the Kirillov-Kostant-Souriau (KKS) form. The group G naturally acts on $\mathfrak{g}^*$ via the co-adjoint action which is Poisson.

Following [Kir99, Kir04], the representation theory of a group G should be seen as the quantization of the Hamilonian mechanics taking place on $\mathbb{C}[\mathfrak{g}^*]$.

On the one hand, let $\mathcal{C}^\infty(\mathfrak{g}^*)$ be the algebra of smooth functions on $\mathfrak{g}^*$ and we have the inclusion of sub-algebras $\mathbb{C}[\mathfrak{g}^*] \hookrightarrow \mathcal{C}^\infty(\mathfrak{g}^*)$, $\mathbb{C}[\mathfrak{g}^*]$ being polynomial functions. As a semi-classical limit, $\mathcal{C}^\infty(\mathfrak{g}^*)$ becomes a Poisson algebra once endowed with the KKS bracket $\{\cdot,\cdot\}_0 : \mathcal{C}^\infty(\mathfrak{g}^*) \times \mathcal{C}^\infty(\mathfrak{g}^*) \to \mathcal{C}^\infty(\mathfrak{g}^*)$. By definition, a Poisson bracket is a

derivation in both variables. Therefore, because of the Leibniz rule, the Poisson bracket is entirely determined by its values on linear functions:

$$\forall X \in \mathfrak{g} \approx (\mathfrak{g}^*)^*, \;\; f_X(\cdot) := \langle X, \cdot \rangle \; .$$

On linear forms, the KKS bracket is defined as:

$$\{f_X, f_Y\}_0 := f_{[X,Y]} = \langle [X,Y], \cdot \rangle \; . \tag{4.2.3}$$

On the other hand, recall that if $(\mathcal{G}, *_{\mathcal{G}})$ is a group, then the group law $*_{\mathcal{G}}$ can be encoded thanks to a coproduct on algebras of functions. By definition, the coproduct Δ associated to $(\mathcal{G}, *_{\mathcal{G}})$ is the map:

$$\begin{array}{rccc} \Delta : & \mathcal{C}^\infty(\mathcal{G}) & \to & \mathcal{C}^\infty(\mathcal{G} \times \mathcal{G}) \\ & f & \mapsto & ((g_1, g_2) \mapsto f(g_1 *_{\mathcal{G}} g_2)) \end{array} \; . \tag{4.2.4}$$

Since Δ is a morphism of algebras, it needs to be specified on generators only. If $\mathcal{A} \subset \mathcal{C}^\infty(\mathcal{G})$ is a dense sub-algebra such that for all $f \in \mathcal{A}$, $\Delta(f)$ is a separable function, which is written in Sweedler's notation:

$$\Delta(f)(g_1, g_2) = \sum_{(f)} f_1(g_1) f_2(g_2) \; ,$$

then we can actually write $\Delta : \mathcal{A} \to \mathcal{A} \otimes \mathcal{A}$. This is the customary choice in order to work algebraically. Here, consider $(\mathfrak{g}^*, +)$ to be *an Abelian group*, which amounts to the trivial coproduct Δ_0 defined on linear functions $X \in \mathfrak{g} \approx (\mathfrak{g}^*)^*$ via:

$$\begin{array}{rccc} \Delta_0 : & \mathbb{C}[\mathfrak{g}^*] & \to & \mathbb{C}[\mathfrak{g}^*] \otimes \mathbb{C}[\mathfrak{g}^*] \\ & X & \mapsto & X \otimes 1 + 1 \otimes X \end{array} \; . \tag{4.2.5}$$

Here we shall be interested in the linear space $\mathfrak{u}_n^*$, dual to the Lie algebra of U_n which is given by anti-Hermitian matrices. On the other hand, $\mathfrak{gl}_n = M_n(\mathbb{C})$ is the Lie algebra of $GL_n(\mathbb{C})$. In the jargon of Lie theory, $\mathfrak{u}_n$ is a real form associated to $\mathfrak{gl}_n = \mathfrak{u}_n \otimes \mathbb{C}$, and we will work with $\mathfrak{gl}_n$ for the sake of pedagogy. However all the content of the following section generalizes seemlessly to any complex semi-simple Lie algebra $\mathfrak{g}$ and its compact real form $\mathfrak{k}$.

4.2.1. Spherical functions as limit of characters. The following fact is classical and is the simplest illustration of the Kirillov orbit method. On the one hand, we have the characters of the unitary group U_n are the Schur polynomials

$$s_\lambda(x) = \mathrm{Tr}_{V(\lambda)}\left(\mathrm{diag}(x_1, x_2, \ldots, x_n)\right) = \frac{\det x_j^{\lambda_i + N - i}}{\det x_j^{N-i}} \; . \tag{4.2.6}$$

the second equality being the Weyl character formula. On the other hand, we have the spherical integral

$$\mathrm{HCIZ}(\lambda, a) = \int_{U_n} dk \; e^{\langle a, [k \lambda k^*]_0 \rangle} = \frac{\det\left(e^{a_i \lambda_j}\right)}{\Delta(a)\Delta(\lambda)} \; , \tag{4.2.7}$$

where Δ denotes the Vandermonde determinant. This spherical integral is also known as the Harish-Chandra-Itzykson-Zuber function in the RMT community and the Kirillov orbital in the representation theory community. It is my opinion that it should rather be called the Kirillov-Itzykson-Zuber since the Harish-Chandra spherical function actually appears in a multiplicative context for the bi-invariant harmonic analysis on $GL_n(\mathbb{C})$. There is more to say about that which we leave for Subsection 4.3.3.

The following proposition says that spherical integrals are basically characters of very large representations. We use a Planck constant $\hbar$ for the rescaling parameter as that will be shown later in Theorem 4.2.4, proving that this limit is in fact a semi-classical limit.

PROPOSITION 4.2.1. *For $\lambda \in \mathbb{R}^n$ with $\lambda_1 > \lambda_2 > \cdots > \lambda_n$, we have:*

$$\lim_{\hbar \to 0} s_{\lambda/\hbar}(e^{\hbar a}) = \Delta(\lambda)\ \mathrm{HCIZ}(\lambda, a)\ .$$

PROOF. This is a straightforward computation putting together the Weyl character formula (4.2.6) and the determinantal formula for the spherical function (4.2.7).

In spite of this simple proof which relies on existing formulae, this is actually a rather deep statement. Interestingly the determinantal formula has been proven by Itzykson-Zuber using Dyson's Brownian motion [**IZ80**]. On the other hand, this can also be obtained using the Duistermaat-Heckman localisation theorem [**DH82**, **BV85**]. □

4.2.2. Universal enveloping algebra with Planck constant. Consider the universal enveloping algebra with the explicit adjunction of the Planck constant $\hbar > 0$:

$$\mathcal{U}^\hbar(\mathfrak{gl}_n) := T(\mathfrak{gl}_n) \,/\, \{x \otimes y - y \otimes x - \hbar[x,y]\}\ . \tag{4.2.8}$$

As such the algebra $\mathcal{U}^\hbar(\mathfrak{gl}_n)$ can be seen as a non-commutative algebra generated by the $E_{i,j}$. The non-commutative polynomials in the $E_{i,j}$'s have commutation relations that respect the Lie bracket $[\cdot,\cdot]$. Then $\left(\mathcal{U}^\hbar(\mathfrak{gl}_n), \dagger\right)$ is a real form with $\dagger$ being the involutive anti-morphism defined by $E_{i,j}^\dagger = E_{j,i}$. In the spirit of quantum mechanics, where self-adjoint operators quantize real functions on phase space, a self-adjoint element is an $x \in \mathcal{U}^\hbar(\mathfrak{gl}_n)$ such that $x^\dagger = x$.

Going commutative, define the map "$\mathrm{mod}\,\hbar$" : $\mathcal{U}^\hbar(\mathfrak{gl}_n) \to \mathrm{Sym}[\mathfrak{gl}_n] = \mathbb{C}[\mathfrak{gl}_n^*]$ as the operation consisting in seeing non-commutative polynomials as commutative. The $E_{i,j}$'s in this case become coordinate functions on $\mathfrak{gl}_n^*$. The presence of the anti-involution $\dagger$ tells us that self-adjoint elements are of the form

$$E_{j,j}, \quad E_{j,k} + E_{k,j}, \quad iE_{j,k} - iE_{k,j}\ ,$$

for $1 \leq j,k \leq n$. If we make the identification $\mathfrak{gl}_n^* \approx \mathfrak{gl}_n = M_n(\mathbb{C})$, this tells us that we are looking at complex matrices that are actually Hermitian. The foliation given in Eq. (4.2.2) is nothing but the one given by the eigenvalues of the matrix. In other words, the orbits are the Hermitian conjugacy classes by the unitary group.

Let $\mathfrak{h} := \mathrm{Span}_{\mathbb{C}}(E_{i,i}\ ;\ 1 \leq i \leq n)$ be the Lie subalgebra of diagonal matrices in $\mathfrak{gl}_n$. Because the Lie bracket vanishes on $\mathfrak{h}$, we have

$$\mathcal{U}^\hbar(\mathfrak{h}) \approx \mathbb{C}[\mathfrak{h}^*] \hookrightarrow \mathcal{U}^\hbar(\mathfrak{gl}_n)\ .$$

REMARK 4.2.2 (On the use of the $\hbar$-adic topology, and field of scalars). *Many sources make sense of the semi-classical limit $\hbar \to 0$ by endowing $\mathcal{U}^\hbar(\mathfrak{gl}_n)$ with the $\hbar$-adic topology, or the degree. The same goes for quantum groups. Another choice is to consider $\hbar$ as a distinguished variable and work with a field of scalars that is $\mathbb{C}(\hbar)$, i.e. rational functions in $\hbar$.*

I shall refrain from doing that and prefer remaining loose in discussing the limit $\hbar \to 0$, until we take representations. It is my point of view that it does not make

sense to say that an abstract operator in an algebra is small until it has been represented as an operator acting on a Hilbert space $\mathcal{H}$, and with a norm dictated by the natural state – here the trace on $\mathcal{H}$.

In any case, I do not like the abstract non-sense of formal power series.

REMARK 4.2.3 (On bad identifications leading to abstract non-sense, following a discussion with Stephan de Bièvre). *As a side comment, let us describe a situation similar in spirit to the tedious effort of figuring out where the Plank constant $\hbar$ needs to be inserted in quantum groups.*

Consider a physical system which we model by a chain of quantum harmonic oscillators. These oscillators have physical characterics such as mass m and frequency ω. The so-called ground state, also known as the vacuum state, is a product of Gaussians with variance proportional to $\frac{\hbar}{m\omega}$. Yet the algebraic Fock space description always identifies the vacuum state as the scalar 1. This identification yields the wrong impression that all vacuum states are the same – especially that the name "vacuum state" hints to the the idea that nothing is there. Nevertheless, the physical vacuum state is clearly not the same for chains of oscillators with different masses and frequencies.

Representation theory: In this paragraph, we record the unessential changes in the representation theory of $\left(\mathcal{U}^{\hbar}\left(\mathfrak{gl}_n\right),\dagger\right)=\mathcal{U}^{\hbar}\left(\mathfrak{u}_n\right)$ caused by adding an $\hbar>0$ constant.

Recall from the classical references [**FH13**, **Sim96**] that representations of the unitary group U_n are essentially equivalent to the representation of the enveloping algebra $\mathcal{U}^{\hbar=1}\left(\mathfrak{u}_n\right)$, which itself is equivalent to the algebraic representations of $\mathcal{U}^{\hbar=1}\left(\mathfrak{gl}_n\right)$. In order to see how the representation theory changes upon explicitly adding the Planck constant $\hbar>0$, we can introduce the isomorphism of algebras

$$\Phi:\ \mathcal{U}^{\hbar=1}\left(\mathfrak{gl}_n\right)\ \to\ \mathcal{U}^{\hbar}\left(\mathfrak{gl}_n\right)$$

defined on $\mathfrak{gl}_n$ via $\forall X\in\mathfrak{gl}_n, \Phi(X)=X/\hbar$, before being extended via the morphism property.

By definition, the weight lattice is the union of possible spectra for the commuting operators in $\mathcal{U}^{\hbar}\left(\mathfrak{h}\right)$, in all possible representations. Upon rescaling thanks to the isomorphism Φ, the weight lattice is rescaled from $\mathbb{Z}^n$ to $\hbar\mathbb{Z}^n$. Highest weights are therefore of the form $\Lambda^{\hbar}=\hbar\lfloor\Lambda/\hbar\rfloor\in\hbar\mathbb{Z}^n$ for $\Lambda\in\mathbb{R}^n$, and $\Lambda_1\geq\Lambda_2\geq\cdots\geq\Lambda_n$. Here, the floor operation $\lfloor\cdot\rfloor$ is understood as pointwise. We write $V(\Lambda^{\hbar})$ for the module with highest weight $\Lambda^{\hbar}$. The orbit method tells as that the modules $\left(V(\Lambda^{\hbar});\ \Lambda\in\mathbb{R}^n\right)$ should be the quantization of the foliation (4.2.2) into orbits.

The usual quadratic Casimir of $\mathcal{U}^{\hbar}\left(\mathfrak{gl}_n\right)$ belongs to the center and is written:

$$C_2:=\sum_{i,j}E_{i,j}^{\dagger}E_{i,j}=\sum_i E_{i,i}^2+2\sum_{i<j}E_{i,j}^{\dagger}E_{i,j}+\hbar\rho^{\vee}\ ,$$

where $\rho^{\vee}=\sum_{j=1}^n(n-j)E_{j,j}$. It acts on $V(\Lambda^{\hbar})$ as the constant:

$$(C_2)_{|V(\Lambda^{\hbar})}:=\langle\Lambda^{\hbar},\Lambda^{\hbar}+\hbar\rho\rangle=\sum_{i=1}^n\left[\left(\Lambda_i^{\hbar}\right)^2+\hbar(n-i)\Lambda_i^{\hbar}\right]\ . \tag{4.2.9}$$

Co-product: The co-product $\Delta_0:\mathcal{U}^{\hbar}\left(\mathfrak{gl}_n\right)\to\mathcal{U}^{\hbar}\left(\mathfrak{gl}_n\right)\otimes\mathcal{U}^{\hbar}\left(\mathfrak{gl}_n\right)$ is defined on $\mathfrak{gl}_n$ as

$$\forall X\in\mathfrak{gl}_n,\ \Delta_0\left(X\right)=1\otimes X+X\otimes 1\ ,$$

then it is extended as a morphism of algebras. Sweedler's notation consists in writing the image of $\mathcal{F}$ as

$$\Delta_0(\mathcal{F}) = \sum_{(\mathcal{F})} \mathcal{F}_1 \otimes \mathcal{F}_2 \ . \tag{4.2.10}$$

Δ_0 is used to define the tensor product of two modules V, W as the vector space $V \otimes W$ endowed with the action

$$\forall \mathcal{F} \in \mathcal{U}^\hbar\left(\mathfrak{gl}_n\right), \ \mathcal{F} \cdot (v \otimes w) := \sum_{(\mathcal{F})} \mathcal{F}_1 \cdot v \otimes \mathcal{F}_2 \cdot w \ .$$

Notice that Δ_0 has nothing inherently quantum or representation-theoretic. Indeed, it coincides with the coproduct on $\mathbb{C}\left[\mathfrak{gl}_n^*\right]$ given in Eq. (4.2.5). When applied to commutative polynomials, Δ_0 is nothing but a tool to encode additive convolution:

$$\begin{array}{rccc} \Delta_0 : & \mathbb{C}\left[\mathfrak{gl}_n^*\right] & \to & \mathbb{C}\left[\mathfrak{gl}_n^*\right] \otimes \mathbb{C}\left[\mathfrak{gl}_n^*\right] \approx \mathbb{C}\left[\mathfrak{gl}_n^* \times \mathfrak{gl}_n^*\right] \\ & f & \mapsto & ((x,y) \mapsto f(x+y)) \end{array} \tag{4.2.11}$$

Identifications: If one is eager to adopt the point of view of RMT, one makes the identification of the dual Lie algebra to Hermitian matrices, $\mathfrak{u}_n^* \approx H_n$, via the non-degenerate linear form:

$$\forall (u,h) \in \mathfrak{u}_n \times H_n, \langle u, h \rangle = -\operatorname{Tr}(uh) \ .$$

In this identification, the aforementioned coadjoint action on $\mathfrak{u}_n^*$ is nothing but the conjugation action on H_n. Althought simpler to work with, this identification is misleading. Or rather, this identification has misled me over the years. And we will revisit that when discussing quantum groups in Section 4.3 where $\mathfrak{u}_n^*$ will be identified with the Lie algebra of lower triangular matrices $\mathfrak{n} \oplus \mathfrak{a} = T_e NA$.

4.2.3. Geometry of orbits, Hamiltonian orbits. A classical fact is that the orbits of Hamiltonian flows using the KKS symplectic structure are exactly coadjoint orbits. See for example [**CP95**, Example 1.1.3 p.19]. In the above identification $\mathfrak{u}_n^* \approx H_n$, coadjoint orbits are conjugation orbits. By virtue of the Spectral Theorem, orbits are indexed by their spectra i.e. vectors $\Lambda \in \mathbb{R}^n$ such that $\Lambda_1 \geq \Lambda_2 \geq \cdots \geq \Lambda_n$. We denote orbits by

$$\mathcal{O}(\Lambda) := \{ \ k \operatorname{diag}(\Lambda) k^* \mid k \in U_n \} \ . \tag{4.2.12}$$

4.2.4. Semi-classical limits. These results are often considered folklore in literature. And over the years, I must admit I never found good references containing them. Nevertheless I think these statements truly capture the essence of Kirillov's orbit method, that representation theory is truly the quantization of coadjoint orbits.

Here we present proofs which seamlessly generalize to any complex semi-simple Lie algebra $\mathfrak{g}$. Our setup is that of $\mathfrak{gl}_n$ for the same of readability. Given a finite-dimensional vector space V, the trace is denoted Tr_V and the normalized trace is $\operatorname{tr}_V := \frac{1}{\dim V} \operatorname{Tr}_V$.

THEOREM 4.2.4. *Let $\mathcal{F} \in \mathcal{U}^\hbar\left(\mathfrak{gl}_n\right)$ and $\Lambda \in \mathbb{R}^n$ dominant, we have*

$$\begin{aligned} \lim_{\hbar \to 0} \operatorname{tr}_{V(\Lambda^\hbar)} \mathcal{F} &= \int_{\mathcal{O}(\Lambda)} \left(\mathcal{F} \bmod \hbar\right)(x) \frac{\omega_{KKS}^N}{N!}(dx) \\ &= \int_{U_n} \left(\mathcal{F} \bmod \hbar\right)(k \Lambda k^*) \ dk \ . \end{aligned}$$

Here ω_{KKS} is the Kirillov-Kostant-Souriau symplectic form and $\frac{\omega_{KKS}^N}{N!}$ is the induced volume form. N is half the dimension of the orbit.

PROOF. **Step 1: For $\mathcal{F} \in \mathcal{U}^\hbar(\mathfrak{h})$.** In that case $\mathcal{F}$ is a polynomial in the variables $E_{j,j}$. Rather than treating separately each monomial, it is better to treat directly the exponential generating series $\exp\left(\sum_j a_j E_{j,j}\right)$, which belongs to a completion of $\mathcal{U}^\hbar(\mathfrak{h})$. We can write:

$$\mathrm{tr}_{V(\Lambda^\hbar)} e^{\sum_j a_j E_{j,j}} = \mathrm{tr}_{V(\Lambda^\hbar)} \mathrm{diag}\left(e^{a_1}, e^{a_2}, \dots, e^{a_1}\right) = s_{\Lambda^\hbar/\hbar}\left(e^{a_1}, \dots, e^{a_n}\right) \ .$$

This is exactly the quantity appearing in Proposition 4.2.1 and which rescales to the Kirillov-Itzykson-Zuber integral:

$$\lim_{\hbar\to 0} \mathrm{tr}_{V(\Lambda^\hbar)} \exp\left(\sum_j a_j E_{j,j}\right) = \mathrm{HCIZ}(\lambda, a) = \int_{U_n} dk \ e^{\langle \lambda, [kak^*]_0\rangle} \ .$$

By seeing the trace as an integral against weight multiplicities, we can use dominated convergence in order to differentiate using the operator $\prod_j \frac{\partial^{n_j}}{\partial a_j^{n_j}}$. We obtain:

$$\begin{aligned}
\lim_{\hbar\to 0} \mathrm{tr}_{V(\Lambda^\hbar)} \left(\prod_j E_{j,j}^{n_j}\right) &= \lim_{\hbar\to 0} \left(\prod_j \frac{\partial^{n_j}}{\partial a_j^{n_j}} \mathrm{tr}_{V(\Lambda^\hbar)} \exp\left(\sum_j a_j E_{j,j}\right)\right)_{|a=0} \\
&= \left(\prod_j \frac{\partial^{n_j}}{\partial a_j^{n_j}} \int_{U_n} dk \ e^{\langle \lambda, [kak^*]_0\rangle}\right)_{|a=0} \\
&= \int_{U_n} dk \ \prod_j E_{j,j}(k\Lambda k^*)^{n_{i,j}} \ .
\end{aligned}$$

This is the required result for $\mathcal{F} \in \mathcal{U}^\hbar(\mathfrak{h})$.

Step 2: A compactness argument.

Let us prove in this step that for any $\mathcal{F}$, the quantity $\mathrm{Tr}_{V(\Lambda^\hbar)} \mathcal{F}$ is bounded as $\hbar \to 0$. By linearity, it suffices to prove that for $\mathcal{F}$ being a non-commutative mononomial. This is reduced further using non-commutative Hölder inequalities [**PX97**, **PX03**] to $\mathcal{F} = E_{i,j}$.

Then

$$\begin{aligned}
\left|\mathrm{tr}_{V(\Lambda^\hbar)} E_{i,j}\right| &\leq \sqrt{\mathrm{tr}_{V(\Lambda^\hbar)} E_{i,j} E_{i,j}^*} \\
&\leq \sqrt{\mathrm{tr}_{V(\Lambda^\hbar)} \sum_{i,j} E_{i,j} E_{i,j}^*} \ .
\end{aligned}$$

This quantity is (almost) constant as it is equal to the quadratic Casimir constant given Eq. (4.2.9). As such:

$$\left|\mathrm{tr}_{V(\Lambda^\hbar)} E_{i,j}\right| \leq \sqrt{(C_2)_{|V(\Lambda^\hbar)}} \ ,$$

which remains bounded as $\hbar \to 0$.

Step 3: Existence of limits up to extraction.

First notice that the value of any limiting point only depends on $\mathcal{F} \mod \hbar$. Indeed if $\mathcal{F}_1 - \mathcal{F}_2 = \hbar\mathcal{F}$, with $\mathcal{F}$ being polynomial in $\hbar$, then

$$\left|\mathrm{tr}_{V(\Lambda^\hbar)}\,\mathcal{F}_1 - \mathrm{tr}_{V(\Lambda^\hbar)}\,\mathcal{F}_2\right| \longrightarrow 0\ ,$$

because $\mathrm{tr}_{V(\Lambda^\hbar)}\,\mathcal{F}$ remains bounded. Therefore, for any $\mathcal{F}$, $\mathrm{tr}_{V(\Lambda^\hbar)}\,\mathcal{F}$ has a non-empty set of limit points, which depend only on $f = \mathcal{F} \mod \hbar$.

Step 4: Uniquess and identification of the limit. Exactly like in the proof of the classical Prokhorov theorem, we can use the diagonal extraction argument in order to assume that there a limit point for all of

$$\left(\mathrm{tr}_{V(\Lambda^{\hbar_n})}\,\mathcal{F};\ \mathcal{F} \in \mathcal{U}^\hbar\left(\mathfrak{gl}_n\right)\right)\ .$$

As such, we obtain that there is linear form $\mu : \mathbb{C}\left[\mathfrak{gl}_n^*\right] \to \mathbb{C}$ and a subsequence $\hbar_n$ such that

$$\forall \mathcal{F} \in \mathcal{U}^\hbar\left(\mathfrak{gl}_n\right),\ \lim_{n\to\infty} \mathrm{tr}_{V(\Lambda^{\hbar_n})}\,\mathcal{F} = \mu\left(f\right)\ .$$

By the Riesz representation theorem, this linear form is represented by a measure.

All in all, we are trying to identify the measure μ. Notice that for any two elements $\mathcal{F}_1$, $\mathcal{F}_2$ in $\mathcal{U}^\hbar\left(\mathfrak{gl}_n\right)$, we have by the circular property of the trace:

$$0 = \frac{1}{\hbar}\,\mathrm{tr}_{V(\Lambda^\hbar)}\left[\mathcal{F}_1\mathcal{F}_2 - \mathcal{F}_2\mathcal{F}_1\right]$$

Now upon writing $f_i = \mathcal{F}_i \mod \hbar$, and the basic fundamental property:

$$\mathcal{F}_1\mathcal{F}_2 - \mathcal{F}_2\mathcal{F}_1 = \hbar\{f_1, f_2\} + \mathcal{O}(\hbar)\ ,$$

we find $\mu\left(\{f_1, f_2\}\right) = 0$. By launching a Hamiltonian flow as in Eq. (4.2.1) with a fixed Hamiltonian $H \in \mathbb{C}\left[\mathfrak{gl}_n^*\right]$, we have

$$\frac{d}{dt}\left(\mu(f)\right) = \mu(\dot{f}) = \mu(\{f, H\}_0) = 0\ .$$

As such, $t \mapsto \mu(f_t)$ is constant. Because Hamiltonian orbits are exactly conjugation orbits (See Subsection 4.2.3), there is a function $k_H : \mathbb{R} \to U_n$ such that

$$\int_{H_n} \mu(dx) f\left(x\right) = \int_{H_n} \mu(dx) f_t\left(x\right) = \int_{H_n} \mu(dx) f\left(k_H(t) x k_H(t)^*\right)\ .$$

The exact coincidence between the entire orbits yields that for any f and any $k \in U_n$, we have:

$$\mu(f) = \int_{H_n} \mu(dx)\ f\left(kxk^*\right)\ .$$

Upon averaging over the unitary group:

$$\mu(f) = \int_{H_n} \mu(dx) \int_{U_n} dk\ f\left(k\,\mathrm{Spec}(x)k^*\right)\ .$$

In the end, there exists a measure $\nu = \mathrm{Spec}_*\,\mu$ such that:

$$\mu(f) = \int_{\lambda_1 \geq \cdots \geq \lambda_n} \nu(d\lambda) \int_{U_n} dk\ f\left(k\lambda k^*\right)\ .$$

We are done upon proving that $\nu = \delta_\Lambda$, i.e. that the measure μ is concentrated on the orbit $\mathcal{O}(\Lambda)$. This is true by virtue of computing the Casimir invariants or using the central characters.

Another argument that does not yield a self-contained proof is to invoke Step 1 or equivalently Proposition 4.2.1: the limit of characters is a spherical integral on the appropriate orbit. □

Now, for the semi-classical limit of a tensor product, which actually reflects the additive convolution of two unitarily invariant matrices.

THEOREM 4.2.5. *Let $\mathcal{F} \in \mathcal{U}^{\hbar}(\mathfrak{gl}_n)$, we have*

$$\begin{aligned}
&\lim_{\hbar\to 0} \mathrm{tr}_{V\left(\Lambda_1^{\hbar}\right)\otimes V\left(\Lambda_2^{\hbar}\right)} \mathcal{F} \\
&= \int_{\mathcal{O}(\Lambda_1)\times\mathcal{O}(\Lambda_2)} \left(\mathcal{F} \bmod \hbar\right)(x+y) \left(\frac{\omega^N}{N!} \otimes \frac{\omega^N}{N!}\right)(dxdy) \\
&= \int_{U_n\times U_n} \left(\mathcal{F} \bmod \hbar\right)\left(k_1\Lambda_1 k_1^* + k_2\Lambda_2 k_2^*\right) dk_1 dk_2 \ .
\end{aligned}$$

PROOF. This is a shadow of the co-product given on $\mathcal{U}^{\hbar}(\mathfrak{gl}_n)$. In Sweedler's notation, we write:

$$\Delta(\mathcal{F}) = \sum_{(\mathcal{F})} \mathcal{F}_1 \otimes \mathcal{F}_2 \ .$$

Upon considering $(e_k^i)_k$ as the basis of $V(\Lambda_i^{\hbar})$, for $i = 1, 2$, this gives:

$$\begin{aligned}
\mathrm{Tr}_{V(\Lambda_1^{\hbar})\otimes V(\Lambda_2^{\hbar})}(\mathcal{F}) &= \sum_{i,j} \langle \Delta\mathcal{F} \cdot e_i^1 \otimes e_j^2, \left(e_i^1 \otimes e_j^2\right)^* \rangle \\
&= \sum_{(\mathcal{F})} \sum_{i,j} \langle \mathcal{F}_1 e_i^1, \left(e_i^1\right)^* \rangle \langle \mathcal{F}_2 \cdot e_j^2, \left(e_j^2\right)^* \rangle \\
&= \sum_{(\mathcal{F})} \mathrm{Tr}_{V(\Lambda_1^{\hbar})}(\mathcal{F}_1) \ \mathrm{Tr}_{V(\Lambda_2^{\hbar})}(\mathcal{F}_2) \ .
\end{aligned}$$

Now, for the dimension, we have:

$$\dim V(\Lambda_1^{\hbar}) \otimes V(\Lambda_2^{\hbar}) = \dim V(\Lambda_1^{\hbar}) \ \dim V(\Lambda_2^{\hbar}) \ ,$$

and upon dividing by that quantity, we obtain the equality:

$$\begin{aligned}
\mathrm{tr}_{V(\Lambda_1^{\hbar})\otimes V(\Lambda_2^{\hbar})}(\mathcal{F}) &= \frac{\mathrm{Tr}_{V(\Lambda_1^{\hbar})\otimes V(\Lambda_2^{\hbar})}(\mathcal{F})}{\dim V(\Lambda_1^{\hbar}) \otimes V(\Lambda_2^{\hbar})} \\
&= \sum_{(\mathcal{F})} \frac{\mathrm{Tr}_{V(\Lambda_1^{\hbar})}(\mathcal{F}_1)}{\dim V(\Lambda_1^{\hbar})} \frac{\mathrm{Tr}_{V(\Lambda_2^{\hbar})}(\mathcal{F}_2)}{\dim V(\Lambda_2^{\hbar})} \\
&= \sum_{(\mathcal{F})} \mathrm{tr}_{V(\Lambda_1^{\hbar})}(\mathcal{F}_1) \, \mathrm{tr}_{V(\Lambda_2^{\hbar})}(\mathcal{F}_2) \ .
\end{aligned}$$

We are ready to invoke the semi-classical for one orbit given in Theorem 4.2.4. Let $f_j := \mathcal{F}_j \bmod \hbar$ for $j = 1, 2$. Also consider a random variable x uniformly distributed on $\mathcal{O}(\Lambda_1)$ and an independent random variable y uniformly distributed on $\mathcal{O}(\Lambda_2)$. We have, as $\hbar \to 0$:

$$\begin{aligned}
\mathrm{tr}_{V(\Lambda_1^{\hbar})\otimes V(\Lambda_2^{\hbar})}(\mathcal{F}) &= \frac{\mathrm{Tr}_{V(\Lambda_1^{\hbar})\otimes V(\Lambda_2^{\hbar})}(\mathcal{F})}{\mathrm{Tr}_{V(\Lambda_1^{\hbar})\otimes V(\Lambda_2^{\hbar})}(1)} \\
&= o(1) + \sum_{(\mathcal{F})} \mathbb{E}\left(f_1(x)\right) \mathbb{E}\left(f_2(y)\right)
\end{aligned}$$

$$= o(1) + \mathbb{E}\left(\sum_{(\mathcal{F})} f_1(x) f_2(y)\right) \ .$$

Now, recall from the Eq. (4.2.11) that the coproduct Δ on $\mathcal{U}^{\hbar}(\mathfrak{gl}_n)$ has nothing inherently quantum! It is the coproduct encoding the addivitive convolution. And as such, by definition:

$$\sum_{(\mathcal{F})} f_1(x) f_2(y) = f(x+y) \ ,$$

where $f = \mathcal{F} \bmod \hbar$. The result follows. □

Although simple, this statement explains why Littlewood-Richardson coefficients have semi-classical limits being given by the Horn problem.

4.3. Semi-classical limits of quantum groups

We conclude this subsection by stating that Pitman's theorem, in its discrete version, has to do with quantum random walks on $\mathcal{U}_q(\mathfrak{sl}_2)$ and taking q from $q = 1$ to $q = 0$, where crystals do appear. In fact, everything can be conveniently recast in terms of the Littelmann path model [**Lit95a**, **Lit95b**], which is a combinatorial model for crystals. The random walks at hand are readily identified with crystal elements. For an overview, see the introduction of my PhD thesis [**Chh13**].

We are ready to state the problem that is addressed in the Section:

QUESTION 4.3.1. *If the Pitman transform $\mathcal{P}$ is intimately related to crystals, appearing at the level of the representation theory of $\mathcal{U}_q(\mathfrak{sl}_2)$ at $q = 0$, why does it also appear in the geometric context of Bougerol and Jeulin?*

Why would there be crystal-like phenomenons by taking curvature to infinity ($r \to \infty$) in a symmetric space $\mathbb{H}^3 = SL_2(\mathbb{C})/SU_2 \approx NA$?

It is certainly desirable to have single global picture, with an interplay between both the representation theory of $\mathcal{U}_q(\mathfrak{sl}_2)$, as $q > 0$ varies, and the geometry of the symmetric space $\mathbb{H}^3 = SL_2(\mathbb{C})/SU_2$ with varying curvatures $r > 0$. Such a unifying point of view should also extend to dynamics, by relating Biane's quantum random walks and the dynamic of Bougerol-Jeulin on $\mathbb{H}^3$.

At this point, let us summarize the landscape:

- On the one hand, at $q = 1$, there is Biane's construction of quantum random walks [**Bia91**]. The diffusive limit is Brownian motion on the space $\mathfrak{su}_2^*$, which can be seen as a flat space with zero curvature ($r = 0$).
- On the other hand, at $q = 0$, using Kashiwara crystals, for example in the path model form, one recovers Pitman's theorem. The latter is also recovered upon taking a Brownian motion on the symmetric space $\mathbb{H}^3 = SL_2(\mathbb{C})/SU_2$ and taking the curvature to infinity ($r \to \infty$).

Thus, we want to interpolate the two different regimes, and perhaps reinterpret the parameter q in quantum groups as a curvature parameter. The most fruitful idea in trying to answer Question 4.3.1 is to discard the idea that $q = e^h$ in the Drinfeld-Jimbo quantum group $\mathcal{U}_q(\mathfrak{sl}_2)$, with h being a Planck constant. The following conversation will take us back to the genesis of quantum groups, which we feel is necessary in order to really distinguish what is quantum and what is not. We begin

by introducing two important ingredients $\mathcal{U}_q^\hbar(\mathfrak{sl}_2)$ and $\mathbb{C}[(SU_2^*)_r]$. These are tailored so that the formal diagram in Figure 4.3.1 commutes.

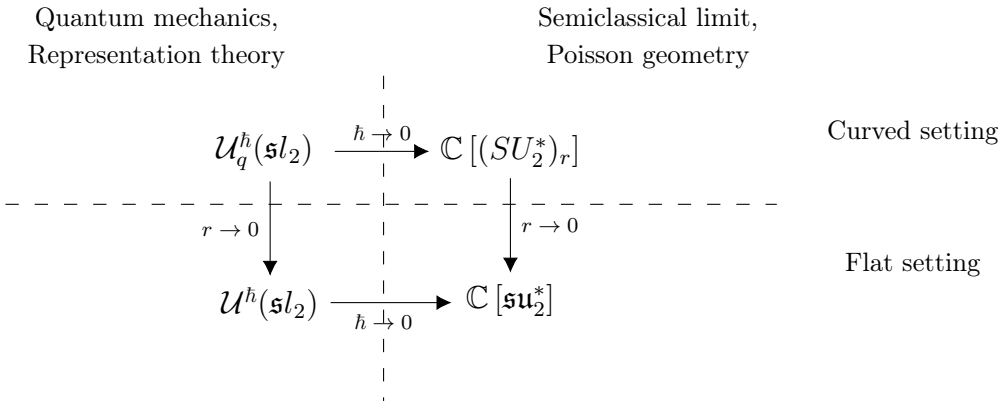

FIGURE 4.3.1. A formal commutative diagram

Quoting Kirillov [Pra05, p.305], who attributes the statement to Drinfeld, the first approximation to quantum groups as classical objects are Poisson-Lie groups. This leads us to the first ingredient, that is a family of Poisson-Lie groups $(SU_2^*)_r$ with varying curvatures $r > 0$. $\mathbb{C}[(SU_2^*)_r]$ will denote the coordinate algebra. In order for such an object to appear as a semi-classical limit, we have to revisit the usual presentation of quantum groups. We require a different presentation $\mathcal{U}_q^\hbar(\mathfrak{sl}_2)$ of the Jimbo-Drinfeld quantum group with two parameters $\hbar > 0$ and $q = e^{-r}$.

4.3.1. Definitions.

4.3.1.1. *Revisiting the Drinfeld-Jimbo quantum group.* We define $\mathcal{U}_q^\hbar(\mathfrak{sl}_2)$ with $q = e^{-r}$ as follows. As explained before, $r > 0$ has to be understood as curvature and $\hbar > 0$ is the actual Planck constant. We set

$$\mathcal{U}_q^\hbar(\mathfrak{sl}_2) := \langle K^{\frac{1}{2}}, K^{-\frac{1}{2}}, E, F\rangle/\mathcal{R} \tag{4.3.1}$$

where this time $K^{\frac{1}{2}} = q^{\frac{1}{2}H} = e^{-\frac{1}{2}rH}$ and $\mathcal{R}$ is the two-sided ideal generated by the relations:

(4.3.2)

$$K^{\frac{1}{2}}EK^{-\frac{1}{2}} = q^\hbar E\,, \quad K^{\frac{1}{2}}FK^{-\frac{1}{2}} = q^{-\hbar}F\,, \quad EF-FE = \hbar\frac{K^{-1}-K}{2r} = \hbar\frac{e^{rH}-e^{-rH}}{2r}\,.$$

Furthermore, $\mathcal{U}_q^\hbar(\mathfrak{sl}_2)$ is a Hopf algebra. The co-product is $\Delta_r\colon \mathcal{U}_q^\hbar(\mathfrak{sl}_2) \to \mathcal{U}_q^\hbar(\mathfrak{sl}_2)\otimes\mathcal{U}_q^\hbar(\mathfrak{sl}_2)$:

$$\left\{\begin{array}{rcl} \Delta_r\left(e^{\frac{1}{2}rH}\right) & = & e^{\frac{1}{2}rH}\otimes e^{\frac{1}{2}rH}\,, \\ \Delta_r(F) & = & F\otimes e^{\frac{1}{2}rH} + e^{-\frac{1}{2}rH}\otimes F\,, \\ \Delta_r(E) & = & E\otimes e^{\frac{1}{2}rH} + e^{-\frac{1}{2}rH}\otimes E\,, \end{array}\right. \tag{4.3.3}$$

while the antipode and counit maps $S_r^\hbar, \varepsilon_r : \mathcal{U}_q^\hbar(\mathfrak{sl}_2) \to \mathcal{U}_q^\hbar(\mathfrak{sl}_2)$ are given by:

$$S_r^\hbar\left(e^{\pm\frac{1}{2}rH}\right) = e^{\mp\frac{1}{2}rH},\ S_r^\hbar(E) = -q^\hbar E,\ S_r^\hbar(F) = -q^{-\hbar}F, \tag{4.3.4}$$

$$\varepsilon_r\left(e^{\pm\frac{1}{2}rH}\right) = 1,\ \varepsilon_r(E) = \varepsilon_r(F) = 0\,. \tag{4.3.5}$$

It is easy to check that, over $\mathbb{C}$, there is a Hopf algebra isomorphism $\Phi\colon \mathcal{U}_{q^\hbar}(\mathfrak{sl}_2) \to \mathcal{U}_q^\hbar(\mathfrak{sl}_2)$, between the classical presentation of Drinfeld-Jimbo that can be found in the literature [CP95] and ours (4.3.1), such that:

$$\Phi(H) = \frac{H}{\hbar}\ , \quad \Phi(E) = E\sqrt{\frac{2r}{\hbar(q^{-\hbar}-q^{\hbar})}}\ , \quad \Phi(F) = F\sqrt{\frac{2r}{\hbar(q^{-\hbar}-q^{\hbar})}}\ . \tag{4.3.6}$$

Strictly speaking, the first equation holds upon continuously extending Φ to a completion so that $K = q^{\hbar H} \in \mathcal{U}_{q^\hbar}(\mathfrak{sl}_2)$ maps to:

$$\Phi(K) = \Phi(e^{-r\hbar H}) = e^{-rH} = K\ .$$

As such, the usual Casimir element:

$$C^q := EF + \frac{q^{-1}K + qK^{-1}}{(q-q^{-1})^2} \in \mathcal{U}_q(\mathfrak{sl}_2) \tag{4.3.7}$$

is changed to

$$C^{q^\hbar} := EF + \frac{q^{-\hbar}K + q^{\hbar}K^{-1}}{(q^{\hbar}-q^{-\hbar})^2} \in \mathcal{U}_{q^\hbar}(\mathfrak{sl}_2)\ ,$$

which maps via the isomorphism Φ and a rescaling to

(4.3.8)

$$\begin{aligned} C^{r,\hbar} :=\ & r\hbar\left(q^{-\hbar}-q^{\hbar}\right)\Phi\left(C^{q^\hbar}\right) = \frac{1}{2}\left(4r^2EF + \left(q^{-\hbar}K + q^{\hbar}K^{-1}\right)\frac{2r\hbar}{(q^{-\hbar}-q^{\hbar})}\right) \\ &= \frac{1}{2}\left(4r^2EF + \left(e^{r\hbar}K + e^{-r\hbar}K^{-1}\right)\frac{2r\hbar}{(e^{r\hbar}-e^{-r\hbar})}\right)\ . \end{aligned}$$

Naturally, $C^{r,\hbar}$ generates the center of $\mathcal{U}_q^\hbar(\mathfrak{sl}_2)$ by [Kas12, Theorem VI.4.8]. We also define the element $\Lambda^{r,\hbar}$ belonging to a completion of $\mathcal{U}_q^\hbar(\mathfrak{sl}_2)$ as

$$\Lambda^{r,\hbar} := \frac{1}{r}\operatorname{Argcosh}\left(\frac{e^{r\hbar}-e^{-r\hbar}}{2r\hbar}C^{r,\hbar}\right) - \hbar\ . \tag{4.3.9}$$

As explained in [CC21], the definition of $\Lambda^{r,\hbar}$ has been tailored so that $\Lambda^{r,\hbar}$ acts as the *appropriate* constant in any fixed irreducible representation of $\mathcal{U}_q^\hbar(\mathfrak{sl}_2)$.

Finally, the analogue of choosing a real form for a Lie algebra in the context of Hopf algebras is exactly the choice of an anti-involution $\dagger$. We recommend the discussion in [KS12, Section 1.2.7] regarding that matter. Here, the compact real form of $\mathcal{U}_q^\hbar(\mathfrak{sl}_2)$ is defined as the pair $\left(\mathcal{U}_q^\hbar(\mathfrak{sl}_2), \dagger\right)$ where $\dagger$ is the algebra anti-involution given by [KS12, p.59]:

$$K^\dagger = K, \qquad E^\dagger = F, \qquad F^\dagger = E. \tag{4.3.10}$$

This real form is compatible with the real form we shall choose for Poisson-Lie groups.

4.3.1.2. *Poisson-Lie groups with varying curvatures $r > 0$.* Consider $B \subset SL_2(\mathbb{C})$ as the Borel subgroup:

$$B := \left\{ \begin{pmatrix} a & 0 \\ b & a^{-1} \end{pmatrix} \mid a \in \mathbb{C}^*,\ b \in \mathbb{C} \right\}\ ,$$

while B^+ is the transpose. If $b \in B \cup B^+$, then $[b]_0$ denotes the projection onto the diagonal. The following complex group will play an important role:

$$SL_2^* := \left\{(b, b^+) \in B \times B^+ \mid [b]_0 = [b^+]_0^{-1}\right\}\ , \tag{4.3.11}$$

which is called the Poisson-Lie group dual to $SL_2(\mathbb{C})$, equipped with the standard structure (see [CP95] or [KS97]). The group law is the pair-wise matrix multiplication. Its Lie algebra

$$\mathfrak{sl}_2^* := T_e SL_2^* = \mathfrak{b} \oplus_{\mathfrak{h}} \mathfrak{b}^+ \ , \tag{4.3.12}$$

is made of the two triangular subalgebras $\mathfrak{b} = T_e B$ and $\mathfrak{b}^+ = T_e B^+$, with the diagonal parts in $\mathfrak{h}$ being opposite. Here $\mathfrak{h} := \mathfrak{a} + i\mathfrak{a} \subset \mathfrak{sl}_2$ is simply the Abelian subalgebra of complex diagonal matrices.

In order to have varying curvatures $r > 0$ and interpolate with the trivial Poisson-Lie group $\mathfrak{sl}_2^*$, we define $(SL_2^*)_r$ as the Lie group with Lie algebra $\left(\mathfrak{sl}_2^*,\ r[\cdot,\cdot]_{\mathfrak{sl}_2^*}\right)$. This is nothing more than SL_2^* as a space but with a different group law. We define $\mathbb{C}\left[(SL_2^*)_r\right]$ as the polynomial algebra generated by the variables $e^{\frac{1}{2}rH}$, $e^{-\frac{1}{2}rH} = \left(e^{\frac{1}{2}rH}\right)^{-1}$, E and F:

$$\mathbb{C}\left[(SL_2^*)_r\right] := \mathbb{C}\left[e^{\frac{1}{2}rH}, e^{-\frac{1}{2}rH}, E, F\right] \ . \tag{4.3.13}$$

In turn, these variables are seen as coordinate functions by writing:

$$\forall g \in (SL_2^*)_r \ , \ g = \left(\begin{pmatrix} e^{\frac{1}{2}rH(g)} & 0 \\ 2r\ F(g) & e^{-\frac{1}{2}rH(g)} \end{pmatrix}, \begin{pmatrix} e^{-\frac{1}{2}rH(g)} & 2r\ E(g) \\ 0 & e^{\frac{1}{2}rH(g)} \end{pmatrix}\right) \ . \tag{4.3.14}$$

When convenient, we will drop the dependence in g for $f(g) = f \in \{H, E, F\}$, as in the definition of the coordinate algebra $\mathbb{C}\left[(SL_2^*)_r\right]$.

Now define:

$$(SU_2^*)_r := \left\{ g \in (SL_2^*)_r \mid H(g) \in \mathbb{R},\ E(g) = \overline{F(g)} \right\} \ . \tag{4.3.15}$$

This is clearly a subgroup of $(SL_2^*)_r$ and we will see in the next section that it is the Poisson-Lie group dual to SU_2, via an involution $\dagger$ which respects the duality at the level of Hopf algebras. Its curvature will also be shown to vary with $r > 0$. In view of the definition of the elements of $(SU_2^*)_r$, all the information is contained in the lower Borel subgroup with positive diagonals, leading to a natural identification

$$(SU_2^*)_r \approx NA \ .$$

The corresponding coordinate algebra is naturally denoted $\mathbb{C}[(SU_2^*)_r]$.

There is also the following, more analytic, presentation of $(SU_2^*)_r$. Notice that the exponential map is a diffeomorphism $\exp : \mathfrak{su}_2^* \approx \mathfrak{n} \oplus \mathfrak{a} \xrightarrow{\sim} NA \approx (SU_2^*)_r$. As such, we can identify $\mathfrak{su}_2^*$ and $(SU_2^*)_r$ as topological spaces. Then we define a group law with a parameter $r > 0$ via:

$$\forall (X, Y) \in \mathfrak{su}_2^* \times \mathfrak{su}_2^*,\ X *_r Y := \frac{1}{r} \log\left(e^{rX} e^{rY}\right) \ . \tag{4.3.16}$$

The new group is denoted by $\left((SU_2^*)_r, *_r\right)$ and its Lie bracket is naturally $r\,[\cdot,\cdot]_{\mathfrak{sl}_2^*}$, i.e. the rescaling of the original bracket by a factor $r > 0$. Clearly, as $r \to 0$, the group $\left((SU_2^*)_r, *_r\right)$ becomes the Abelian group $(\mathfrak{su}_2^*, +)$.

4.3.2. Main result. In order to construct the quantum walk on $\mathcal{U}_q^{\hbar}(\mathfrak{sl}_2)$, this latter algebra is taken as the algebra of observables for one increment. Our algebra of non-commutative random variables is the inductive limit

$$\mathcal{A}^{r,\hbar} := \varinjlim_n \left(\mathcal{U}_q^{\hbar}\left(\mathfrak{sl}_2\right)\right)^{\otimes n} \ . \tag{4.3.17}$$

This is exactly as the algebra of functions for an i.i.d. product of random variables.

Given that $\mathcal{U}_q^\hbar(\mathfrak{sl}_2)$ is isomorphic to the usual Jimbo-Drinfeld quantum group, the representations are essentially the same. The specifics are not needed for now. The state τ is the product state using the standard representation $\mathbb{C}^2$. The pair $\left(\mathcal{A}^{r,\hbar},\tau\right)$ will be our working non-commutative probability space.

In the definition of our measurement operators we have to use the coproduct Δ_r. As such, the morphism of algebras $M_n : \mathcal{U}_q^\hbar(\mathfrak{sl}_2) \to \mathcal{A}^{r,\hbar}$ defined for discrete times $n \in \mathbb{N}$ are as follows. $M_0 = \varepsilon_r$ is given by the counit, and

$$\begin{cases} M_1 &= 1 \ , \\ M_n &= (M_{n-1} \otimes 1) \circ \Delta_r \ , \ \text{for } n \geq 2. \end{cases} \tag{4.3.18}$$

Since we want the random walk to classically start from the identity, and the quantum version consists in expressing everything dually at the level of measurement operators, one sees that M_0 has to be taken as the counit.

We make the convention that $\forall t \geq 0,\ M_t := M_{\lfloor t \rfloor}$. As a random walk on $\mathcal{U}_q^\hbar(\mathfrak{sl}_2)$, we define three non-commutative processes via:

$$\forall t \in \mathbb{R}_+,\ S_t^{r,\hbar} := M_{t/\hbar^2}(S)$$

for each generator $S \in \{H, E, F\}$ of the quantum group $\mathcal{U}_q^\hbar(\mathfrak{sl}_2)$. This three-dimensional non-commutative process is neatly repackaged in matrices of $(SU_2^*)_r$ with non-commutative entries:

$$\forall t \in \mathbb{R}_+,\ g_t^{r,\hbar} \tag{4.3.19}$$

$$:= \left(\begin{pmatrix} e^{\frac{1}{2}rH_t^{r,\hbar}} & 0 \\ 2r\ F_t^{r,\hbar} & e^{-\frac{1}{2}rH_t^{r,\hbar}} \end{pmatrix}, \begin{pmatrix} e^{-\frac{1}{2}rH_t^{r,\hbar}} & 2r\ E_t^{r,\hbar} \\ 0 & e^{\frac{1}{2}rH_t^{r,\hbar}} \end{pmatrix} \right) \in (SU_2^*)_r \otimes \mathcal{A}^{r,\hbar} \ . \tag{4.3.20}$$

By stating that the above quantity is in $(SU_2^*)_r \otimes \mathcal{A}^{r,\hbar}$, we are implicitly saying that the lower and upper triangular parts are †-conjugate. Therefore $g_t^{r,\hbar}$ can be seen as an element in NA with operator-valued entries. The quantum dynamic $\left(\Lambda_t^{r,\hbar}\ ;\ t \geq 0\right)$ is defined from the measurement of the Casimir element (4.3.8) thanks to the explicit expression:

$$\forall t \in \mathbb{R}_+,\ \frac{2r\hbar}{e^{r\hbar} - e^{-r\hbar}} \cosh\left(r\hbar + r\Lambda_t^{r,\hbar}\right) := M_{t/\hbar^2}\left(C^{r,\hbar}\right) \ . \tag{4.3.21}$$

This is equivalent to directly setting $\Lambda_t^{r,\hbar} := M_{t/\hbar^2}\left(\Lambda^{r,\hbar}\right)$ after continuously extending the measurement operators to the completion where $\Lambda^{r,\hbar}$ lives (see Eq. (4.3.9)).

We are ready to state the main result of [CC21], which unifies the results of Biane on the one hand and Bougerol-Jeulin on the other hand. Since the crystal regime $q = e^{-r} \to 0$ is tractable in both quantum and semi-classical settings, it explains why crystal-like phenomena appear upon taking infinite curvature limits. This recovers indeed Pitman's $2M - X$ Theorem in the discrete and continuous versions.

THEOREM 4.3.2 (Main Theorem of [CC21]). *In the sense of (possibly non-commutative) moments, we have the following convergences in law between processes indexed by $t \in \mathbb{R}_+$:*

$$\begin{array}{ccc}
\Lambda_t^{\infty,\hbar} = X_t^\hbar - 2\inf_{0\le s\le t} X_t^\hbar & \xrightarrow{\hbar\to 0} & \Lambda_t^\infty = X_t - 2\inf_{0\le s\le t} X_s \\
\textit{Pitman's Theorem 4.1.1} & & \textit{Pitman's Theorem 4.1.2} \\
\textit{(discrete case)} & & \textit{(continuous case)} \\
\uparrow r\to\infty & & \uparrow r\to\infty \\
g_t^{r,\hbar} \in (SU_2^*)_r \otimes \mathcal{A}^{r,\hbar} & & g_t^r \in (SU_2^*)_r \otimes L^{\infty-}(\Omega) \\
\Lambda_t^{r,\hbar} & \xrightarrow{\hbar\to 0} & \Lambda_t^r = \tfrac{1}{r}\,\mathrm{Argcosh}\circ \mathrm{tr}\left(g_t^r\,(g_t^r)^\dagger\right) \\
\textit{Quantum random walks} & & \textit{Bougerol-Jeulin's convolution dynamic} \\
\textit{on } \mathcal{U}_q^\hbar(\mathfrak{sl}_2) \textit{ as in Eq. (4.3.19)} & & \textit{and its radial part as in Theorem 4.1.5} \\
\downarrow r\to 0 & & \downarrow r\to 0 \\
x_t^{0,\hbar} = \begin{pmatrix} \frac{1}{2}X_t^\hbar & 0 \\ Y_t^\hbar + iZ_t^\hbar & -\frac{1}{2}X_t^\hbar \end{pmatrix} \in \mathfrak{su}_2^* \otimes \mathcal{A}^{0,\hbar} & & x_t^0 = \begin{pmatrix} \frac{1}{2}X_t & 0 \\ Y_t + iZ_t & -\frac{1}{2}X_t \end{pmatrix} \in \mathfrak{su}_2^* \otimes L^{\infty-}(\Omega) \\
\Lambda_t^{0,\hbar} = \sqrt{\tfrac{1}{2}\hbar^2 + (X_t^\hbar)^2 + (Y_t^\hbar)^2 + (Z_t^\hbar)^2} & \xrightarrow{\hbar\to 0} & \Lambda_t^0 = \sqrt{X_t^2 + Y_t^2 + Z_t^2} \\
\textit{Biane's quantum random walks} & & \textit{Flat Brownian Motion on } \mathfrak{su}_2^* \approx \mathbb{R}^3 \\
\textit{on } \mathcal{U}^\hbar(\mathfrak{sl}_2) \textit{ as in [Bia09]} & & \textit{and its radial part}
\end{array}$$

Moreover, on both quantum and semi-classical pictures, i.e. for $\hbar > 0$ and $\hbar = 0$, the dynamic of $\Lambda^{r,\hbar}$ does not depend on r.

The paper [CC21] also describes measures on orbits with a curvature parameter $r > 0$ such that analogous theorems to 4.2.4 and 4.2.5 hold.

In those theorems, the algebra of non-commutative variables is the quantum group $\mathcal{U}_q^\hbar(\mathfrak{sl}_2)$ and the limiting orbits are dressing action orbits in a curved space with curvature parameter $r > 0$.

4.3.3. Elements of generalization beyond quantum $\mathfrak{sl}_2$. An interesting project is to write down the full proofs of the semiclassical limits for general type $\mathcal{U}_q^\hbar(\mathfrak{g})$ or at least for $\mathfrak{g} = \mathfrak{gl}_n$. It would also be desirable to explicit the relationship to the Robinson-Schensted correspondence and crystals. However this would take us too far in the realm of representation theory.

Leaving the quantum world, let us now conclude this Section 4.3 with some elements on the geometry of semiclassical orbits, in the case of $\mathfrak{gl}_n$. This will allow us to describe the potential generalization of Theorem 4.3.2 to $\mathfrak{gl}_n$, when dealing only with the semi-classical aspects which are on the RHS of the commutative diagram.

Dressing orbits. As in the rank one case $SL_2(\mathbb{C})$, consider the subgroup $NA \subset GL_n(\mathbb{C})$ of lower triangular matrices with positive diagonal. Of course $NA \approx GL_n(\mathbb{C})/U_n$ is the symmetric space where Bougerol-Jeulin [BJ02] construct a Brownian motion which they study in the infinite curvature regime. The group U_n acts on the left of NA by the so-called dressing orbit

$$\forall k \in U_n, \forall na \in NA,\ \exists! k\cdot na \in NA,\ kna \in (k\cdot na)U_n\ .$$

This is nothing but the NA part in the Gram-Schmidt decomposition of kna. For future use, define the map

$$\begin{array}{rccc} \mathrm{a}^{(r)}: & GL_n(\mathbb{C}) & \to & \mathbb{R}^n \\ & g & \mapsto & a(g) \end{array} \tag{4.3.22}$$

by saying that $\mathrm{a}^{(r)}(g) \in \mathbb{R}^n$ is the unique vector such that

$$g \in N \exp\left(\frac{1}{2} r \mathrm{a}^{(r)}(g)\right) U_n \ .$$

Equivalently, $\exp\left(r\mathrm{a}^{(r)}(g)\right)$ is the diagonal part in the Cholesky decomposition of gg^*.

As a space define for any $\Lambda \in \mathbb{R}^n$, the curved orbit $\mathcal{O}_r(\Lambda)$ as the dressing orbit passing through $e^{\frac{1}{2} r \ \mathrm{diag}\Lambda}$:

$$\mathcal{O}_r(\Lambda) := \left\{ k \cdot e^{\frac{1}{2} r \mathrm{diag}\Lambda} \mid k \in U_n \right\} \ . \tag{4.3.23}$$

Of course, we have a foliation of the curved space NA analogous to Eq. (4.2.2):

$$NA = \bigsqcup_{\Lambda} \mathcal{O}_r(\Lambda) \ . \tag{4.3.24}$$

Another of seeing $\mathcal{O}_r(\Lambda)$ is as a space of positive definite matrices with given singular values through the map:

$$\begin{array}{ccc} \mathcal{O}_r\left(\Lambda\right) & \to & \{M \in M_n(\mathbb{C}) \mid M = M^*, \ M \text{ positive definite}\} \\ g & \mapsto & gg^* \ . \end{array}$$

Notice that as $r \to 0$, we can identify $\mathcal{O}_{r=0}\left(\Lambda\right)$ with Hermitian matrices and $a^{(r=0)} = \mathrm{diag}$ is the map which selects the diagonal. We will simply prefer the NA picture because it has a natural group law.

Harish-Chandra's spherical integral. From now on, let us drop the notation of the spherical integral as $(z, \Lambda) \mapsto \mathrm{HCIZ}(z, \Lambda)$. For reasons that will soon be clear, we shall now denote it by:

$$\varphi_z^{(r=0)}(\Lambda) := \mathrm{HCIZ}(z, \Lambda).$$

Recall that

$$\varphi_z^{(r=0)}(\Lambda) = \mathbb{E}_{k \sim \mathrm{Haar}}\left(e^{\langle z, \mathrm{diag}(k\Lambda k^*)\rangle}\right) = \frac{\det\left(e^{z_i \Lambda_j}\right)}{\Delta(z)\Delta(\Lambda)} \ .$$

Now, the definition of the Harish-Chandra spherical integral for $GL_n(\mathbb{C})$ is [**Hel22**, Chapter IV, Theorem 4.3]:

$$\phi_z(\Lambda) := \int_{U_n} dk \ \exp\left(\langle z + \rho, \mathrm{a}^{(r=1)}(ke^{\frac{1}{2}\Lambda})\rangle\right) \ .$$

Here $\rho = (n-1, \ldots, 2, 1, 0)$ is again the Weyl vector. And from Harish-Chandra formula for the complex group $GL_n(\mathbb{C})$ [**Hel22**, Chapter IV, Theorem 5.7], which is much more simple than for general real Lie groups, we have

$$\phi_z(\Lambda) = c(z) \frac{\det\left(e^{z_i \Lambda_j}\right)}{\Delta(e^{\Lambda})} \ ,$$

where the c-function is $c(z) = \prod_{1 \leq i < j \leq n} \frac{j-i}{z_j - z_i} = \frac{\Delta(\rho)}{\Delta(z)}$. Notice the following relationship with the Kirillov-Itzykson-Zuber integral:

$$\phi_z(\Lambda) = \frac{\Delta(\Lambda)\Delta(\rho)}{\Delta(e^{\Lambda})} \varphi_z^{(r=0)}(\Lambda) \ .$$

This motivates the following definition of spherical function with curvature parameter r:

$$\varphi_z^{(r)}(\Lambda) := \int_{U_n} dk \ \exp\left(\langle z + r\rho, \mathrm{a}^{(r)}(ke^{\frac{1}{2}r\Lambda})\rangle\right) \ . \tag{4.3.25}$$

We have

$$\begin{aligned}
\varphi_z^{(r)}(\Lambda) &= \int_{U_n} dk \ \exp\left(\langle z + r\rho, \frac{1}{r}\mathrm{a}^{(r=1)}(ke^{\frac{1}{2}r\Lambda})\rangle\right) \\
&= \int_{U_n} dk \ \exp\left(\langle \frac{z}{r} + \rho, \mathrm{a}^{(r=1)}(ke^{\frac{1}{2}r\Lambda})\rangle\right) \\
&= \varphi_{\frac{z}{r}}^{(r=1)}(r\Lambda) \\
&= c(z/r)\frac{\det\left(e^{z_i\Lambda_j}\right)}{\Delta(e^{r\Lambda})} \\
&= \frac{\Delta(r\rho)}{\Delta(z)}\frac{\det\left(e^{z_i\Lambda_j}\right)}{\Delta(e^{r\Lambda})} \\
&= \frac{\Delta(r\rho)\Delta(\Lambda)}{\Delta(e^{r\Lambda})}\varphi_z^{(r=0)}(\Lambda) \ .
\end{aligned}$$

This confirms that we have chosen the correct interpolation between the classical Harish-Chandra spherical integral $\varphi_z^{(r=1)} = \phi_z$ and the Kirillov-Itzykson-Zuber integral $\lim_{r\to 0}\varphi_z^{(r)} = \varphi_z^{(r=0)}$. Our choice uses a curvature parameter r to interpolate between the symmetric space $NA \approx GL_n(\mathbb{C})/U_n$ and the flat space $\mathfrak{n} \oplus \mathfrak{a} \approx H_n$.

REMARK 4.3.3. *Notice that the following quantity does not depend on r:*

$$\varphi_z^{(r)}(\Lambda)/\varphi_0^{(r)}(\Lambda) \ ,$$

which hints to the existence of harmonic analysis that does not depend on the curvature parameter r. More precisely spherical additive convolution and spherical multiplicative convolution have the same harmonic analysis.

Quasi-invariant measures on orbits $\mathcal{O}_r(\Lambda)$. On the orbits, we have the natural left-Haar invariant measure, which is the law of

$$g_{\mathrm{Haar}} = k \cdot e^{\frac{1}{2}r\Lambda} \ ,$$

with k Haar distributed on U_n. Rather, we shall prefer the following quasi-invariant measure, which is a tilting of the left-Haar invariant measure. Our quasi-invariant measure with curvature parameter r is the law of $g \in \mathcal{O}_r(\Lambda) \subset NA$ and

$$\mathbb{E}\left(f(g)\right) = \frac{1}{\varphi_0^{(r)}(\Lambda)}\int_{U_n} dk \ \exp\left(\langle r\rho, \mathrm{a}^{(r)}(k \cdot e^{\frac{1}{2}r\Lambda})\rangle\right) \ f(k \cdot e^{\frac{1}{2}r\Lambda}) \ .$$

REMARK 4.3.4. *By picking $f = \exp\left(\langle z, \mathrm{a}^{(r)}(\cdot)\rangle\right)$, we see that elements g sampled according to this canonical measure satisfy*

$$\mathbb{E}\left(\exp\left(\langle z, \mathrm{a}^{(r)}(g)\rangle\right)\right) = \frac{\varphi_z^{(r)}(\Lambda)}{\varphi_0^{(r)}(\Lambda)} \ .$$

Remark 4.3.3 basically says that the law of $\mathrm{a}^{(r)}(g)$ does not depend on the curvature parameter $r \in \mathbb{R}$.

4.4. Relationship to some integrable models in mathematical physics

4.4.1. Zoo of models: Directed percolation, Polymers, TASEP.

DLPP. The acronym DLPP stands for "Directed Last Passage Percolation". It is a percolation model in random environment. We consider the grid $\mathbb{Z}^2$ and the environment is the collection

$$\omega := (\omega_{i,j} \ ; \ (i,j) \in \mathbb{N} \times \mathbb{N})$$

of i.i.d. random variables with light tails.

Consider the set of directed paths $\Pi_{M,N}$ from $(0,0)$ to (M,N). We restrict the set of possible π to paths with only right and up steps, hence the adjective "directed", see Fig. 4.4.1. The quantity of interest is the last passage time

$$L_{M,N} := \max_{\pi \in \Pi_{M,N}} \mathcal{E}(\pi) \ ,$$

with energy of a path being $\mathcal{E}(\pi) = \sum_{(i,j)\in\pi} \omega_{i,j}$.

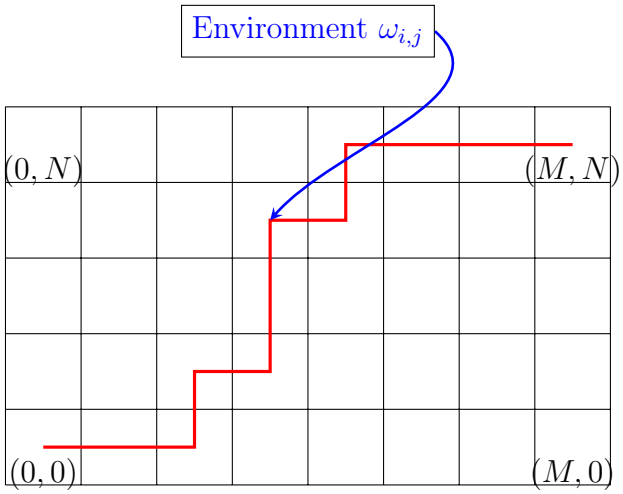

FIGURE 4.4.1. Directed path π in red from $(0,0)$ to (M,N)

Brownian DLPP.

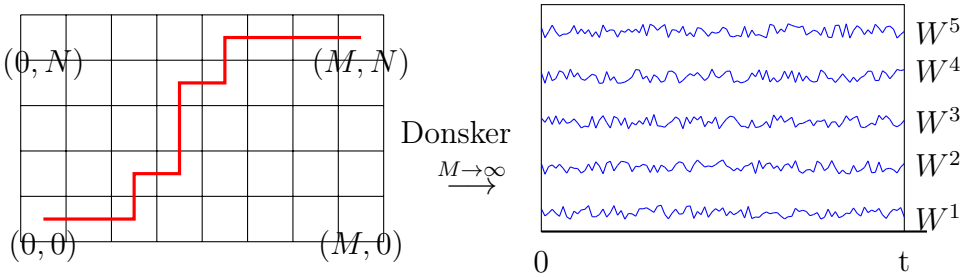

FIGURE 4.4.2. Rescaling the usual DLPP to the Brownian DLPP

By centering and diffusive rescaling, $L_{M,N}$ becomes

$$L_{t,N} = \lim_{M\to\infty} \frac{L_{\lfloor Mt \rfloor, N} - (M+N)\mathbb{E}\omega}{\sqrt{M \operatorname{Var}(\omega)}} = \max_{0=t_0<t_1<\cdots<t_N=t} \sum_{i=1}^{N} W^i_{t_i} - W^i_{t_{i-1}} \ .$$

REMARK 4.4.1. *If $N = 2$, there are independent Brownian motions $X = \frac{W^2 - W^1}{\sqrt{2}}$, $Y = \frac{W^2 + W^1}{\sqrt{2}}$ such that:*

$$L_{t,N=2} = \max_{0 \le s \le t} W_s^1 + W_t^2 - W_s^2 \ = \frac{1}{\sqrt{2}} Y_t + \frac{1}{\sqrt{2}} \left(X_t - 2 \min_{0 \le s \le t} X_s \right) \ .$$

Directed polymers. Let $\beta = \frac{1}{k_B T} > 0$ be a positive inverse temperature parameter. It is natural to consider the Gibbs measure associated to the energy functional of DLPP. This yields a measure on directed paths $\pi \in \Pi_{M,N}$:

$$\mathbb{Q}^{\beta,\omega}(\pi) := \frac{\exp\left(\beta \mathcal{E}(\pi)\right)}{Z^{\beta}_{M,N}} \ .$$

This measure depends on the environment ω as indicated in superscript. The same goes for the normalization constant, which is called the partition function following the jargon of statistical physics and it is the random variable

$$Z^{\beta}_{M,N} = \sum_{\pi \in \Pi_{M,N}} e^{\beta \mathcal{E}(\pi)} \ . \tag{4.4.1}$$

In the zero temperature regime, $\beta \to \infty$, the measure $\mathbb{Q}^{\beta,\omega}$ naturally concentrates on the path with largest energy, hence yielding the optimizer of DLPP functional. The last passage time, on the other hand, is the limit of the free energy:

$$L_{M,N} = \lim_{\beta \to \infty} \frac{1}{\beta} \log Z^{\beta}_{M,N} \ .$$

O'Connell-Yor directed polymer. Exactly as in Fig. 4.4.2, one can rescale the partition function $Z^{\beta}_{M,N}$ to a model which does not depend on the fine features of the environment, in order to obtain the so-called O'Connell-Yor semi-directed polymer. One has to rescale the inverse temperature parameter as

$$\beta_M = \frac{\beta}{\sqrt{M \operatorname{Var} \omega}} \ ,$$

in order to obtain

$$\begin{aligned} & Z^{\beta}_{t,N} \\ &= \lim_{M \to \infty} Z^{\beta_M}_{\lfloor Mt \rfloor, N} \\ &= \int \dots \int_{0 = s_0 < s_1 < \dots < s_N = t} \exp\left(\beta \left(\sum_{i=1}^{N} W^i_{s_i} - W^i_{s_{i-1}} \right) \right) ds_1 ds_2 \dots ds_{N-1} \end{aligned}$$

Again, the zero temperature limit of O'Connell-Yor is nothing but the Brownian percolation.

Connection to particle systems.

When the environment is made of exponential random variables, i.e. $\omega_{i,j} \overset{\mathcal{L}}{=} \mathbf{e}$ with $\mathbb{P}(\mathbf{e} \in dx) = \mathbb{1}_{\mathbb{R}_+}(x) e^{-x} dx$, there are natural correspondences

$$\text{TASEP} \approx \text{Corner growth} \approx \text{DLPP} \ .$$

TASEP stands for "Totally Asymmetric Exclusion Process", and it is a particle system X with state space $\{0,1\}^{\mathbb{Z}}$. The initial condition is $X_0 = (\mathbb{1}_{k \le 0})_{k \in \mathbb{Z}}$. And the dynamic is so that particles jump to the right, at a neighboring site, if the site

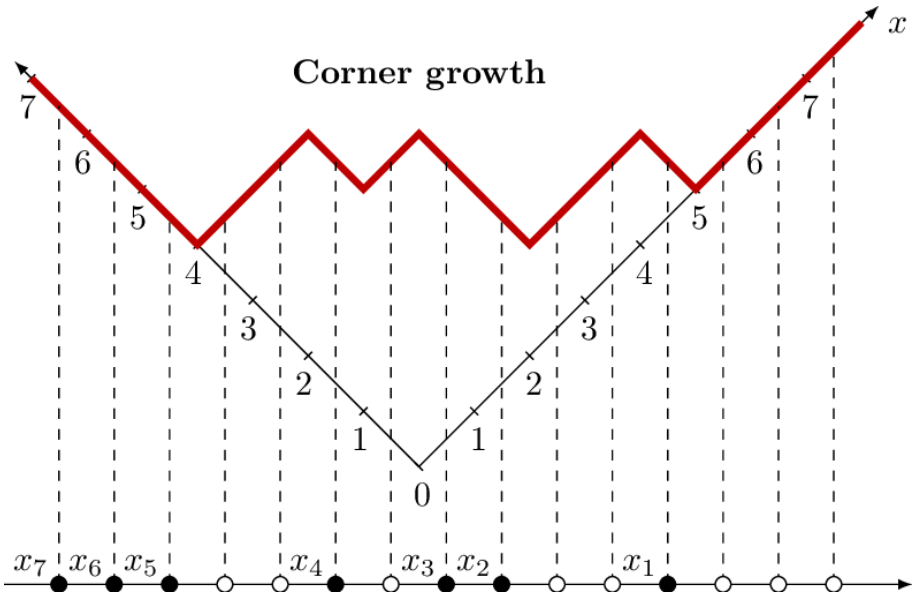

FIGURE 4.4.3. Coupling TASEP & Corner growth

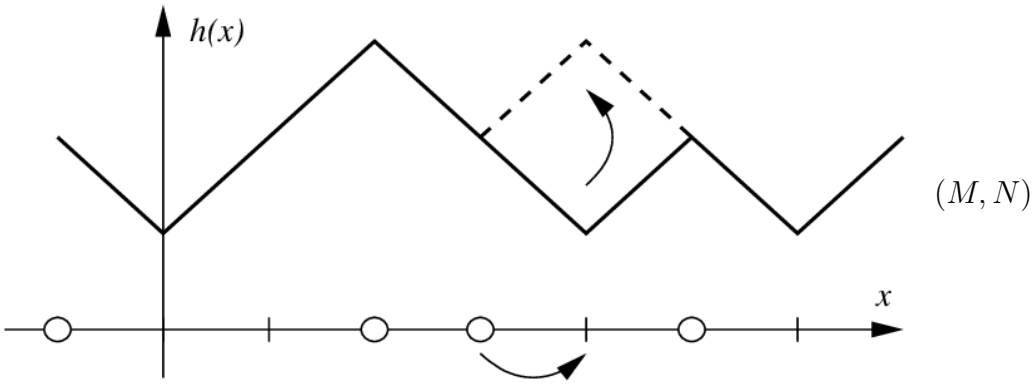

FIGURE 4.4.4. Transition above box (M, N)

unoccupied. In order to have a continuous time Markovian dynamic, each site in $\mathbb{Z}$ has a Poisson clock giving the times of jumps. As such, the waiting times are necessarily exponential random variables in order to preserve the Markov property.

The first correspondence between TASEP and corner growth models is given in Fig. 4.4.3 and is rather self explanatory. On the bottom one sees a particle configuration $\eta \in \{0,1\}^{\mathbb{Z}}$ defined by

$$\forall x \in \mathbb{Z},\ \eta(x) = \sum_k \mathbb{1}_{X_k = x}\ ,$$

where $X_1, X_2, \dots$ are the positions of particles. The correspondence consists in mapping a particle configuration $\eta \in \{0,1\}^{\mathbb{Z}}$ to a piece-wise affine function which is interpreted as a growing interface. Slope is respectively -1 at an occupied site and $+1$ at an unoccupied site. A particle jumping from left to right becomes a corner growth event as in Fig. 4.4.4. Notice that this correspondence works for any particle system on the line with nearest neighbor interaction for example, as in the ASEP where particles are allowed to jump from right to left. A particle jump from right

to left becomes a corner deflation in the corner growth model. In turn, the height function h encodes the current of the particle system η.

The second correspondence only works in the case of TASEP. The last passage time $L_{M,N}$ in DLPP is just the total time necessary for the interface $h(t,\cdot)$ to go above box (M,N). This is a correspondence in law and it requires the memoryless property of the exponential waiting times. As such, the only conclusion that be drawn is the equality

$$\forall t \geq 0, \forall N \in \mathbb{N},\ \mathbb{P}\left(h(t,0) \leq N\right) = \mathbb{P}\left(L_{N,N} \leq t\right)\ ,$$

in the case of TASEP. For ASEP, one can only obtain upper bounds on the current in this fashion.

4.4.2. Pitman-type theorems as a bridge between RMT and DLPP. In the case $N = 2$, Remark 4.4.1 tells us that DLPP contains the Pitman transform. The relationship to RMT appears in the following remark.

REMARK 4.4.2. *Consider a standard Hermitian Brownian motion on 2×2 matrices. This is the process $\mathrm{GUE}_t^{(N=2)}$ as defined in the introduction Section 2.1. It is easy to see that there are independent standard Brownian motions such that:*

$$\begin{aligned}\mathrm{GUE}_t^{(N=2)} &= \frac{1}{\sqrt{2}}\begin{pmatrix}\sqrt{2}B_t^1 & B_t^{12}+iB_t^{21}\\ B_t^{12}-iB_t^{21} & \sqrt{2}B_t^2\end{pmatrix}\\ &= \frac{B_t^1+B_t^2}{2} + \frac{1}{\sqrt{2}}\begin{pmatrix}W_t & B_t^{12}+iB_t^{21}\\ B_t^{12}-iB_t^{21} & -W_t\end{pmatrix}\ .\end{aligned}$$

Hence

$$\begin{aligned}\lambda_1\left(\mathrm{GUE}_t^{(N=2)}\right) &= \frac{B_t^1+B_t^2}{2} + \frac{1}{\sqrt{2}}\sqrt{W_t^2+(B_t^{12})^2+(B_t^{21})^2}\\ &= \frac{Y_t}{\sqrt{2}} + \frac{1}{\sqrt{2}}\mathrm{BES}_t^3\ .\end{aligned}$$

As such the Pitman's Theorem 4.1.2 says exactly that the largest eigenvalue of an 2×2 GUE equals in law the last passage time with two rows, for the Brownian directed percolation.

From GUE to Brownian percolation. In order to move from $N = 2$ to general N, one needs higher rank generalizations of Pitman's theorem.

THEOREM 4.4.3 (Baryshnikov [Bar01], Tracy-Widom [GTW01], Bougerol–Jeulin [BJ02]). *We have equality in law as processes Assuming exponential weights ω for DLPP:*

$$\left(L_{t,N}\ ;\ t \geq 0\right) \overset{\mathcal{L}}{=} \left(\lambda_1(\mathrm{GUE}^{(\mathrm{N})})\ ;\ t \geq 0\right)\ .$$

Notice that this equality is between non-Markovian processes. The full equality in law is between the $\mathrm{Spec}\left(\mathrm{GUE}^{(\mathrm{N})}\right)$ and a generalized Pitman transform for which the last passage time is only the first coordinate. As implicit in this body of work, this generalized Pitman transform encodes a continuous Robinson-Schensted correspondence.

The generalization to any Lie type is done in [BBO05, BBO09] using the machinery of crystals, which generalizes the Robinson-Schented correspondence. This is beyond the scope of this document and does not seem that relevant in relationship to DLPP or polymer models.

Johansson's Proposition 1.4.

THEOREM 4.4.4 (Johansson [Joh00] for fixed M, Warren-Dieker [DW09]). *Assuming exponential weights ω for DLPP:*

$$(L_{M,N}\ ;\ M \geq 1) \overset{\mathcal{L}}{=} (\lambda_1(\mathbb{W}_{M,N})\ ;\ M \geq 1)\ .$$

Let us mention that the process version of was conjectured first by Borodin-Péché [BP08]. And this yields Tracy-Widom fluctuations effortlessly by simply invoking the classical Theorem 2.1.3. In any case, it is my opinion that such correspondences are precious given how difficult it is to prove Tracy-Widom fluctuations.

4.4.3. Novel proof of Johansson's Proposition 1.4. using curvature deformation. Let us now present a proof which goes through harmonic analysis. More precisely, this proof exploits the rigidity of harmonic analysis as curvature varies – à la "Bougerol-Jeulin". As a prerequisite to this proof, we require the setup detailed in Subsection 4.3.3, where we give the definition of (additive) invariant ensembles and introduce our definition of (multiplicative) quasi-invariant ensembles.

The starting point is the following somewhat trivial yet very key lemma.

LEMMA 4.4.5 (Laws of diagonals, laws of spectra).

Zero curvature statement $r = 0$.

Consider an invariant ensemble in the space of Hermitian matrices H_n. Let $a^{(r=0)}$ be its diagonal part, and $\Lambda^{(r=0)}$ be its spectrum. Then the law of the ensemble is equivalently determined by the law of the spectrum $\mathcal{L}\left(\Lambda^{(r=0)}\right)$ or the law of the diagonal $\mathcal{L}\left(a^{(r=0)}\right)$.

Statement with curvature parameter r.

Consider a quasi-invariant ensemble in the space $NA \subset GL_n(\mathbb{C})$. If g is in that ensemble, $g \in \mathcal{O}_r\left(\Lambda^{(r)}\right)$ defines the spectrum $\Lambda^{(r)}$. And $g \in Ne^{\frac{1}{2}ra^{(r)}}U_n$ defines the diagonal $a^{(r)}$. Then the law of a quasi-invariant ensemble invariant is equivalently determined by the law of singular values $\mathcal{L}\left(\Lambda^{(r)}\right)$ or the law of the diagonal $\mathcal{L}\left(a^{(r)}\right)$.

Moreover, the injective map $\mathcal{L}\left(\Lambda^{(r)}\right) \mapsto \mathcal{L}\left(a^{(r)}\right)$ does not depend on the parameter r.

PROOF. We will simply refer to $\Lambda^{(r)}$ as the spectrum and refer to $a^{(r)}$ as the diagonal, irrespective of the curvature parameter r.

Step 0: Fixing spectra fixes everything.

If M_n is Hermitian and invariant, then $M_n = u_n\Lambda_n u_n^*$ with u_n Haar distributed and independent. Therefore specifying the law of Λ_n fixes everything.

Likewise, if $g \in NA$, and we write $g = u_n \cdot e^{\frac{1}{2}r\Lambda_n}$, then conditionally to $\Lambda_n = \lambda$, the law of u_n is specified by virtue of saying that g is quasi-invariant.

Step 1: Fixing the diagonal with $r = 0$.

Now reciprocally, suppose M_n is invariant and that $\operatorname{diag} M_n$ has a fixed given law. That means that we are given a function $h : \mathbb{R}^n \to \mathbb{C}$ such that

$$h(z) = \mathbb{E}\left(e^{i\langle z, \operatorname{diag}(M_n)\rangle}\right)\ .$$

By virtue of M_n being invariant, we can force the appearance of the spherical integral in the following way

$$h(z) = \mathbb{E}\left(e^{i\langle z, \operatorname{diag}(u_n\Lambda_n u_n^*)\rangle}\right)$$

$$= \int \mathbb{P}\left(\Lambda_n \in d\lambda\right) \mathbb{E}\left(e^{i\langle z, \mathrm{diag}(u_n \lambda u_n^*)\rangle}\right)$$
$$= \int \mathbb{P}\left(\Lambda_n \in d\lambda\right) \varphi_{iz}^{(r=0)}(\lambda) \ .$$

This relation can be inverted via the spherical transform, which hinges on the identity:

$$\int_{\mathbb{R}^n} dz \ \varphi_{iz}^{(r=0)}(\lambda)\varphi_{-iz}^{(r=0)}(\mu) \left|\Delta(z)\right|^2 = n! \left|\Delta(\mu)\right|^2 \delta_{\lambda=\mu} \ .$$

This formula makes sense in the sense of distributions and requires λ, μ in the Weyl chamber $C = \{\Lambda \in \mathbb{R}^n \ , \ \Lambda_i > \Lambda_{i+1}\}$. For a quick proof, one can recover this either from the orthogonality properties of Schur polynomials and Proposition 4.2.1, or from the Fourier transform.

Step 2: Fixing the diagonal with r non-vanishing.

Now, if $M_n \in NA$ is quasi-invariant, then $M_n = u_n \cdot e^{\frac{1}{2}r\Lambda_n^{(r)}} \in N_n e^{\frac{1}{2}r a_n^{(r)}} U_n$. Suppose that we are given the law of a_n, so that the characteristic function

$$h(z) = \mathbb{E}\left(e^{i\langle z, a_n^{(r)}\rangle}\right)$$

is known. Let us show that the law of Λ_n can be recovered. By hypothesis, M_n is quasi-invariant so that

$$h(z) = \mathbb{E}\left(e^{i\langle z, \mathrm{a}^{(r)}(M_n)\rangle}\right)$$
$$= \int \mathbb{P}\left(\Lambda_n^{(r)} \in d\lambda\right) \mathbb{E}\left(\exp\left(i\langle z, \mathrm{a}^{(r)}(u_n \cdot e^{\frac{1}{2}r\lambda})\rangle\right)\right)$$
$$= \int \mathbb{P}\left(\Lambda_n^{(r)} \in d\lambda\right) \frac{\varphi_{iz}^{(r)}(\lambda)}{\varphi_0^{(r)}(\lambda)} \ .$$

Because of the Remarks 4.3.3 and 4.3.4, we find exactly the same relationship between h and $\mathbb{P}\left(\Lambda_n \in d\lambda\right)$ as in the flat case:

$$h(z) = \int \mathbb{P}\left(\Lambda_n^{(r)} \in d\lambda\right) \varphi_{iz}^{(r=0)}(\lambda) \ .$$

Along the way, we have proven that the map $\mathcal{L}(\Lambda_n^{(r)}) \mapsto \mathcal{L}(a_n^{(r)})$ is invertible and does not depend on r. □

REMARK 4.4.6 (Important notational change). *So far, N denoted the space of lower unipotent matrices as a subgroup of $NA \subset GL_N(\mathbb{C})$. In order to fit with the notations of our integrable models, N will now denote the number of rows in the environment.*

We apologize to our reader for this egregious offense. We have not found a way to do this at a reasonable cost.

Constructing a curvature-independent dynamic. Recall from Section 2.1 that the Wishart matrices $(\mathbb{W}_{M,N} \ ; \ M \geq 1)$ constitute a random walk with independent and invariant increments. An increment is of the form $\xi\xi^*$ with ξ Gaussian standard vector.

As a corollary of the proof, a great idea is to do the following.

- Let $b_1^{(r)}, b_2^{(r)}, \dots$ as i.i.d. random variables in the space of lower triangular matrices in $GL_N(\mathbb{C})$ with common quasi-invariant distribution $\mathcal{L}\left(b^{(r)}\right)$.
- By writing $b^{(r)} \in \mathcal{O}_r\left(\lambda_N^{(r)}\right)$ and $a_N^{(r)} = \mathrm{a}^{(r)}(g^{(r)})$, we can equivalently choose the law of $a_n^{(r)}$ or the law of $\lambda_N^{(r)}$.
- We choose these laws to correspond to the Wishart case, i.e. we need these to correspond to the law of diagonal and spectrum of $\xi\xi^*$. That is to say that $\lambda_N^{(r)} = \chi_N e_1$ where χ_N is distributed as a χ distribution with N degrees of freedom. Or equivalently $a_N^{(r)}$ vector of i.i.d. exponential standard random variables.

Then form the product:

$$B_{M,N}^{(r)} := b_M^{(r)} \dots b_2^{(r)} b_1^{(r)} \ . \tag{4.4.2}$$

We look at the singular values $\Lambda_{M,N}^{(r)}$. Since $B_{M,N}^{(r)}$ belongs to a quasi-invariant ensemble, and $\mathrm{a}^{(r)}(B_{M,N}^{(r)})$ matches the flat case, we have the equality in law

$$\Lambda_{M,N}^{(r)} \overset{\mathcal{L}}{=} \mathrm{Spec}\left(\mathbb{W}_{M,N}\right) \ .$$

By repeating this reasoning from a step to the next, we upgrade this equality to an equality in law between processes

$$\left(\Lambda_{M,N}^{(r)} \ ; \ M \geq 1\right) \overset{\mathcal{L}}{=} \left(\mathrm{Spec}\left(\mathbb{W}_{M,N}\right) \ ; \ M \geq 1\right) \ . \tag{4.4.3}$$

This is the discrete time analogue of Theorem 4.1.5 by Bougerol-Jeulin, highlighting an equality in law independently of the curvature parameter r. All we need to do now is understand the infinite curvature regime $r \to \infty$.

Semi-explicit form of Eq. (4.4.2) as $r \to \infty$. The content of a generic element $b^{(r)}$ is not very explicit for finite r. However, in the limit $r \to \infty$, the content of $b^{(r)}$ is very explicit.

Let us denote the diagonal part of $b^{(r)}$ as $a^{(r)} = \left(e^{\frac{1}{2}r\omega_1}, e^{\frac{1}{2}r\omega_2}, \dots, e^{\frac{1}{2}r\omega_N}\right)$.

Lemma 4.4.7. *There are uniform i.i.d. random variables $\Theta_1, \Theta_2, \dots$ on the circle such that*

$$b^{(r)} = \sum_{1 \leq i \leq j \leq N} E_{j,i} \prod_{i \leq k \leq j} \Theta_k e^{\frac{1}{2}r\omega_k} \left(1 + o_{\mathbb{P}}(1)\right) \ .$$

Here $o_{\mathbb{P}}(1)$ denotes a quantity which converges to 0 in probability as $r \to \infty$.

Proof. Notice that we have now added phases on the diagonal of $b^{(r)}$. This is of no issue as, technically, $b^{(r)}$ is a lower triangular representative of an equivalence class in $GL_N(\mathbb{C})/U_N$. As such, it is determined up to multiplication by a diagonal unitary matrix.

Moreover, quasi-invariance implies that $b^{(r)} \overset{\mathcal{L}}{=} b^{(r)} t$ for t diagonal and unitary, hence the uniform phases.

Now, we have

$$b^{(r)}\left(b^{(r)}\right)^* = k e^{r\chi_N E_{1,1}} k^* = I_N + \left(e^{r\chi_N} - 1\right) k E_{1,1} k^* \ ,$$

where k is unitary. As such, we have a vector $v \in \mathbb{C}^N$ such that

$$b^{(r)}\left(b^{(r)}\right)^* = I_N + vv^* \ . \tag{4.4.4}$$

As such, we need to compute the Cholesky decomposition of matrices of the form $I_N + vv^*$.

If L_i is the top-left $i \times i$ minor of $b^{(r)}$, we have:

$$e^{\frac{1}{2}r\omega_i} = \left|\frac{\det L_i}{\det L_{i-1}}\right| = \sqrt{\frac{\det L_i L_i^*}{\det L_{i-1}L_{i-1}}} = \sqrt{\frac{1+\|v_{[1:i]}\|^2}{1+\|v_{[1:(i-1)]}\|^2}},$$

where $v_{[1:i]}$ is the vector of the first i coordinates of v.

As such, as $r \to \infty$, we have

$$\|v_{[1:i]}\| = (1+o_{\mathbb{P}}(1))\, e^{\frac{1}{2}r\omega_i}\|v_{[1:(i-1)]}\| \ .$$

Therefore, a similar relation exists between the coordinates of v. We find that there are phases Θ_i's such that

$$v_i = (1+o_{\mathbb{P}}(1))\, e^{\frac{1}{2}r\omega_i}\Theta_i v_{i-1} \ .$$

In the end, we find that Eq. (4.4.4) becomes

$$b^{(r)}\left(b^{(r)}\right)^* = I_N + (1+o_{\mathbb{P}}(1))\left(\prod_{k\le i}\Theta_k e^{\frac{1}{2}r\omega_k}\prod_{k\le j}\overline{\Theta_k}e^{\frac{1}{2}r\omega_k}\right)_{1\le i,j\le N} .$$

This asymptotic form has an easy Cholesky decomposition to compute, which yields the desired result. □

As such, we have the following corollary upon making matrix products.

COROLLARY 4.4.8. *If $b_i^{(r)}$ as $a^{(r)} = \left(e^{\frac{1}{2}r\omega_{i,1}}, e^{\frac{1}{2}r\omega_{i,2}}, \dots, e^{\frac{1}{2}r\omega_{i,N}}\right)$, then the matrix coefficients of $B_{M,N}^{(r)}$ are asymptotically given by:*

$$\left[B_{M,N}^{(r)}\right]_{i,j} = (1+o_{\mathbb{P}}(1))\sum_{\pi\in\Pi_{M,i-j+1}} e^{\frac{1}{2}r\mathcal{E}(\pi)}\Theta_\pi \ ,$$

where

$$\mathcal{E}(\pi) = \sum_{(i,j)\in\pi}\omega_{i,j}, \quad \Theta_\pi = \prod_{(i,j)\in\pi}\Theta_{i,j} \ .$$

Notice that this asymptotic form is exactly the partition function encountered in Subsection 4.4 as Eq. (4.4.1), but with random phases.

Conclusion. We finish the proof by noticing that the infinite curvature regime $r \to \infty$ is exactly the same as the zero temperature regime $\beta = \frac{1}{k_B T} \to \infty$, modulo the details of handling phases. Notice that the variance of $\left[B_{M,N}^{(r)}\right]_{i,j}$ conditionally to the environment is

$$\mathrm{Var}\left(\left[B_{M,N}^{(r)}\right]_{i,j} \mid \omega\right) = \sum_{\pi\in\Pi_{M,i-j+1}} e^{r\mathcal{E}(\pi)} = Z_{M,i-j+1}^{\beta=r} \ .$$

As $r \to \infty$, the largest coefficient is $Z_{M,N}^{\beta=r}$. This corresponds to the coefficient $(i,j) = (N,1)$. Therefore, the largest singular value of $B_{M,N}^{(r)}$ is asymptotically given by

$$\left(\Lambda_{M,N}^{(r)}\right)_1$$

$$
\begin{aligned}
=&\frac{1}{r}\log\lambda_1\left(B^{(r)}_{M,N}\left(B^{(r)}_{M,N}\right)^*\right)\\
=&\frac{1}{r}\log\sup_{\|x\|=1}\left\|B^{(r)}_{M,N}x\right\|^2\\
=&o_{\mathbb{P}}(1)+\frac{1}{r}\log\left|\left[B^{(r)}_{M,N}\right]_{N,1}\right|^2\\
=&o_{\mathbb{P}}(1)+\frac{1}{r}\log Z^{\beta=r}_{M,N}+\frac{1}{r}\log\frac{\left|\left[B^{(r)}_{M,N}\right]_{N,1}\right|^2}{\mathrm{Var}\left(\left[B^{(r)}_{M,N}\right]_{N,1}\mid\omega\right)}\\
=&o_{\mathbb{P}}(1)+L_{M,N}\ .
\end{aligned}
$$

On this last step, we have used the fact that the zero limit temperature of partition function is the last passage time, and the fact that the last term vanishes in probability.

This computation combined with Eq. (4.4.3) yields the desired result. Notice we have proven Johansson's equality in law, without ever actually computing this law!

CHAPTER 5

Strong noise limits in classical and quantum filtering

My work on classical and quantum filtering is part of a long term collaboration with Benoist, Bernardin, Chetrite, Najnudel and Pellegrini [**BCC+23**, **BBC+21**, **BCNP22**]. Quantum trajectories are SDEs giving the state of a quantum system perturbed by a measurement process. In fact, these SDEs result from the filtering of this state: in short, it is the result of a "quantum filtering" theory very similar to the classical filtering equations of Kushner-Stratonovich-Zakai. The more intense a measurement is, the greater the noise of the SDE. And it is in this high noise regime that theoretical physicists Bauer, Bernard and Tilloy have described an interesting metastability phenomenon: in the high noise regime, our (continuous) quantum trajectory looks like a jump process, decorated with very fine spikes. From a probabilistic point of view, it is the proof of convergence and the rigorous description of such processes that interests us. Solving this open question also requires the use of exotic topologies.

The first paper [**BCC+23**] largely solves the question in dimension 1, by invoking the tools of stochastic calculus and the theory of Itô excursions.

In the second paper [**BCNP22**], we take as a starting point the fact that the SDEs of classical filtering are the same as the quantum versions. As such, we can apply our technology to that case and prove a peculiar phase transition: in the study of smoothed filters, there is a sharp transition in the large noise regime, between perfect filtering and the spiking regime.

The third paper [**BBC+21**] tackles the higher dimensional. It only demonstrates the convergence to a jump process, in any dimension, after a smoothing procedure encapsulated in the Meyer-Zheng topology.

We start by Section 5.1 of motivations and setup, which explains classical and quantum filtering. Then Sections 5.2, 5.3 and 5.4 present the results of the aforementioned papers, in order.

5.1. Motivations and setup

5.1.1. The Shiryaev-Wonham filter from classical filtering. Let us start by presenting the Shiryaev-Wonham filter and refer to [**Won64**, **Lip01**, **VH07**] for more extensive material.

The most simple setup, called the "signal plus noise" model, is the one where we observe a process $\mathbf{y}^\gamma = (\mathbf{y}_t^\gamma \ ; \ t \geq 0)$ of the form

$$d\mathbf{y}_t^\gamma = \mathbf{x}_t dt + \frac{1}{\sqrt{\gamma}} \, dB_t \tag{5.1.1}$$

where $B = (B_t \ ; \ t \geq 0)$ is a standard Wiener process, $\gamma > 0$ is the noise level and $\mathbf{x}$ is a hidden process of interest. This continuous analogue of Hidden Markov Chains, is known as the Wonham filter when $\mathbf{x} \in \{0, 1\}$ is a continuous–time Markov chain. And the goal is naturally to estimate $\mathbf{x}$.

General setup with n states. More generally, consider a Markov process $\mathbf{x} = (\mathbf{x}_t \,;\, t \geq 0)$ on a finite state space $E = \{x_1, x_2, \ldots, x_n\}$ and a continuous observation process $\mathbf{y}^\gamma$ of the usual additive form "signal plus noise":

$$d\mathbf{y}_t^\gamma := G(\mathbf{x}_t)\, dt + \frac{1}{\sqrt{\gamma}} dB_t \ . \tag{5.1.2}$$

Here $G : E \to \mathbb{R}$ is a function taking distinct values for identifiability purposes. The filtered state is given by:

$$\rho_t^\gamma(x_i) := \mathbb{P}\left(\mathbf{x}_t = x_i \mid (\mathbf{y}_s^\gamma)_{s \leq t}\right) \ .$$

The generator of $\mathbf{x}$ is denoted by $\mathcal{L}$. The claim of the Shiryaev-Wonham filter is that the filtering equation becomes:

$$\begin{aligned} d\rho_t^\gamma(x_i) = & \sum_j \left(\rho_t^\gamma(x_j)\mathcal{L}(x_j, x_i) - \rho_t^\gamma(x_i)\mathcal{L}(x_i, x_j)\right) dt \\ & + \sqrt{\gamma}\rho_t^\gamma(x_i)\left(G(x_i) - \langle \rho_t^\gamma, G\rangle\right) dW_t \ . \end{aligned} \tag{5.1.3}$$

Here W is the innovation process, and is a $\mathcal{F}^{\mathbf{y}}$-standard Brownian motion. The quantity $\langle \rho_t^\gamma, G\rangle$ denotes the expectation of G with respect to the probability measure ρ_t^γ. Throughout the paper, we only consider $E = \{0, 1\}$, i.e. the two state regime.

Two states. In this case, all the information is contained in

$$\pi_t^\gamma := \rho_t^\gamma(1) = \mathbb{P}\left(\mathbf{x}_t = 1 \mid (\mathbf{y}_s^\gamma)_{s \leq t}\right) \ .$$

Using the notation

$$\mathcal{L} = \begin{pmatrix} -\lambda_{0,1} & \lambda_{0,1} \\ \lambda_{1,0} & -\lambda_{1,0} \end{pmatrix} \ .$$

we have indeed that Eq. (5.1.3) can be rewritten as

$$d\pi_t^\gamma = -\lambda\left(\pi_t^\gamma - p\right) dt + \sqrt{\gamma}\sigma\pi_t^\gamma\left(1 - \pi_t^\gamma\right) dW_t \ ,$$

where

$$\lambda = \lambda_{0,1} + \lambda_{1,0} \ , \quad p = \lambda_{1,0}/\lambda \ , \quad \sigma = G^1 - G^0 \ . \tag{5.1.4}$$

Without loss of generality, we shall assume $\sigma = 1$ in the rest of the paper. Also $(G^0, G^1) = (0, 1)$. In the end, our setup is indeed given by the previously given, which we repeat for convenience and further reference:

$$d\mathbf{y}_t^\gamma = \mathbf{x}_t dt + \frac{1}{\sqrt{\gamma}} dB_t \ , \tag{5.1.5}$$

$$d\pi_t^\gamma = -\lambda\left(\pi_t^\gamma - p\right) dt + \sqrt{\gamma}\pi_t^\gamma\left(1 - \pi_t^\gamma\right) dW_t \ . \tag{5.1.6}$$

REMARK 5.1.1. *The invariant probability measure μ of the Markov process $\mathbf{x}$ solves*

$$\mathcal{L}^*\mu = 0 \Longleftrightarrow \mu = \begin{bmatrix} p \\ 1 - p \end{bmatrix} \ .$$

Without any computation, this is intuitively clear, as setting $\gamma \to 0$ yields an extremely strong observation noise and no noise in the filtering equation:

$$d\pi_t^{\gamma=0} = -\lambda(\pi_t^{\gamma=0} - p)dt$$

whose asymptotic value is p. Informally, this says that, in the absence of information, the best estimation of the law $\mathcal{L}(\mathbf{x}_t)$ in long time is the invariant measure.

This is essentially the content of [Chi06, Theorem 4]*, which holds for a Shiryaev-Wonham filter with any finite number of states.*

5.1.2. Quantum filtering for quantum open systems. Consider a quantum system with a finite dimensional Hilbert space $\mathbb{C}^n$. Its state is described by a density matrix belonging to

$$S^+ := \left\{ \rho \in M_n(\mathbb{C}) \mid \rho = \rho^\dagger, \mathrm{Tr}(\rho) = 1, \ \rho \geq 0 \right\} \ .$$

The matrix $\rho \in S^+$ is also called the density of states for a quantum system with n states.

Measurements correspond to self-adjoint operators A, and an experiment yields a random output with mean value following the Born rule:

$$\langle A \rangle = \mathrm{Tr}\left(\rho A\right) \ . \tag{5.1.7}$$

More precisely, upon diagonalizing the operator in a orthonormal basis as $A = \sum_i E_i |\psi_i\rangle\langle\psi_i|$, we have

$$\mathbb{P}\left(A = E_i\right) = \langle\psi_i|\rho|\psi_i\rangle \ .$$

When dealing with a pure state $\rho = |\psi\rangle\langle\psi|$, i.e. a rank projector in the direction of $|\psi\rangle \in \mathbb{C}^n$ the Born rule (5.1.7) takes the usual form for isolated systems

$$\langle A \rangle = \langle\psi|A|\psi\rangle \ .$$

The description with a density of states is more suitable for quantum open systems.

Haroche's experiment. Let us comment the famous experiment of Haroche and his group [Har13]. The setup is described in Figure 5.1.1. Although there are many aspects that are beyond my understanding, the idea is to nuance the axiom of wave-function collapse.

Haroche's experiment is an engineering "tour de force" that allows to observe the distinction between direct and indirect measurement. Recall that after a direct measurement yielding an energy level E_i, the state ρ collapses to $\rho^{\mathrm{Collapse}} = |\psi_i\rangle\langle\psi_i|$. An indirect measurement occurs when the system is not directly measured, but rather interacts with another system which is measured.

In Haroche's experiment, the quantum system of interest is light (photons) trapped inside of a reflecting cavity at 0.8 Kelvin. The cavity's mirrors are of sufficiently good quality that light can bounce a few zillion times without being absorbed, and thus remain trapped for macroscopic amounts of time. Then one sends Rydberg atoms which interact with the trapped photons. At the exit of the system, the Rydberg atoms are directly measured. In fact one can deduce this way how many photons are in the cavity, and hence measuring light without directly measuring it. More precisely, one can deduce the density of states ρ of the light.

During the experiment, the state of light is described by a stochastic process $(\rho_t \ ; \ t \geq 0)$ which evolves depending on the flux of Rydberg atoms. And in fact, $\lim_{t\to\infty} \rho_t = \rho^{\mathrm{Collapse}}$. In the end, the indirect measurement yields the same result as a direct measurement. But the process $(\rho_t \ ; \ t \geq 0)$ is a stochastic process, and the collapse is not instantaneous.

General stochastic Linblad equations. A quantum trajectory is a stochastic process $(\rho_t \ ; \ t \geq 0)$ in S^+ describing state evolution of the quantum system. If such this quantum system is *simultaneously* subject to continuous measurements and to other interactions, for example free Hamiltonian evolution or contact with

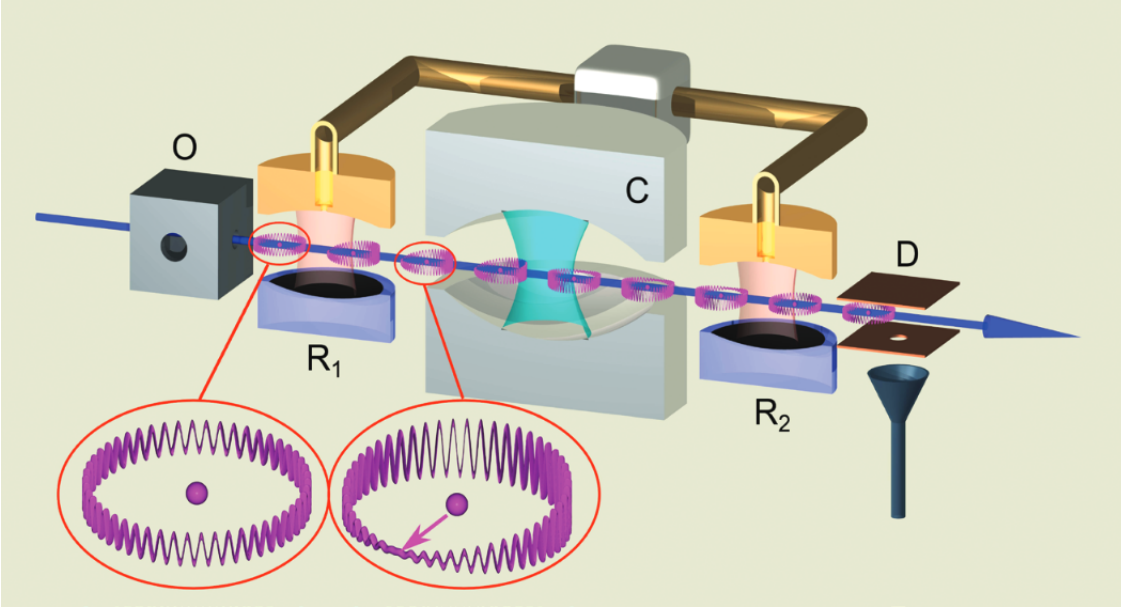

FIGURE 5.1.1. Haroche's experiment from [Har13]. The cavity is (C), where photons are trapped by super-reflecting mirrors. The source (O) sends in Rydberg atoms. The microwave resonators R_1 and R_2 with the Rydberg atoms. R_1 prepares the Rydberg atoms in a very specific superposition of two states. R_2 measures the state of the Rydberg atoms after their have interacted with light.

some other system, then its density matrix $\rho_t \in S^+$ is expected to follow a stochastic Lindblad equation, also called a Belavkin equation:

$$d\rho_t = -i\,[H,\rho_t]\,dt + \sum_{k,l=1}^{n} L[M_{k,l}](\rho_t)dt + \gamma L[N](\rho_t)dt + \sqrt{\gamma}D[N](\rho_t)dW_t\ , \tag{5.1.8}$$

with

$$\begin{cases} L\,[O]\,(\rho) \equiv O\rho O^\dagger - \frac{1}{2}\left(\rho O^\dagger O + O^\dagger O\rho\right)\ , \\ D\,[O]\,(\rho) \equiv O\rho + \rho O^\dagger - \mathrm{Tr}\left[\left(O+O^\dagger\right)\rho\right]\rho \end{cases}$$

and $(W_t; t \geq 0)$ a standard Brownian motion.

Here $O \in M_n(\mathbb{C})$ is called a measurement operator and the application L is called a Lindbladian. It is a super operator i.e an application mapping a matrix $M \in M_n(\mathbb{C})$ to a linear operator $L[M] : M_n(\mathbb{C}) \to M_n(\mathbb{C})$. There are n^2 Lindbladian indexed by k, l but some of them may vanish. The first term involving H in (5.1.8) is due to the free Hamiltonian evolution, the terms $\sum_{k,l=1}^{n} L[M_{k,l}]$ are due to the interaction with some environment (thermal bath for example) while the two last terms are result from the measurement process. In particular, the two last terms depend on some parameter $\gamma > 0$ which represents the intensity of the measurement process. The Belavkin equation (5.1.8) is driven by a single Wiener process but in a more general setting, it makes sense to consider Belavkin equations driven by several Wiener processes.

It is a computational exercice to show that $\mathrm{Tr}\,[\rho_0] = 1$ implies $\mathrm{Tr}\,[\rho_t] = 1$ and that $\rho_0 = \rho_0^\dagger$ implies $\rho_t = \rho_t^\dagger$ at any time $t \geq 0$. The positivity property, i.e. the fact that $(\rho_t; t \geq 0)$ lives in S^+, is obvious.

Guiding example with two energy levels. For $n = 2$, we use the parameterization:

$$\rho_t = \begin{pmatrix} q_t & \overline{p_t} \\ p_t & 1 - q_t \end{pmatrix} . \tag{5.1.9}$$

Under certain hypotheses, we obtain decoherence $p_t = 0$. And all the information is in the scalar process $X_t^\gamma = q_t$.

Following [**BB14**, **BBT15a**, **TBB15**, **BBT16a**] by Bauer, Bernard and Tilloy, we consider a quantum system with two energy levels $\{E_0,\ E_1\}$, a.k.a. a "qubit", in a thermal bath at temperature β^{-1} and subject to continuous indirect measurements [**HR06**, **WM10a**] of the energy with intensity $\gamma > 0$. The indirect nature of these measurements prevents the complete wave-function collapse which occurs in a direct measurement, according to the principles of quantum mechanics [**Hal13**, Section 3.6, Axiom 4]. Let us then denote by X_t^γ the probability of measuring the energy E_0 at time t, upon a *hypothetical* direct measurement. The process $X^\gamma = (X_t^\gamma\ ;\ t \geq 0)$ solves the following SDE obtained from simplifying the Belavkin equation:

$$dX_t^\gamma = -\,\lambda(X_t^\gamma - p)dt + \sqrt{\gamma} X_t^\gamma\,(1 - X_t^\gamma)\,dW_t\ . \tag{5.1.10}$$

In this context, the large γ limit corresponds to the strong measurement regime. Furthermore, $\lambda > 0$ is the coupling strength with the thermal bath and $p = \frac{e^{-\beta E_0}}{e^{-\beta E_0} + e^{-\beta E_1}}$ is the probability of being at the energy level E_0 according to a Gibbs measure. Heuristically, Eq. (5.1.10) expresses a competition between the drift term favoring a convergence towards p and the stochastic term favoring an absorption in $\{0, 1\}$. In physical jargon, one says that there is a competition between thermalization and collapsing.

5.1.3. Phenomenology.

Duality between weak and strong noise. Notice that the observation equation (5.1.5) has a factor $\frac{1}{\sqrt{\gamma}}$, while the filtering equation (5.1.6) has a factor $\sqrt{\gamma}$. This is a well-known duality between the weak noise limit in the observation process and the strong noise limit in the filtered state.

In fact, when analyzing the derivation of the Wonham-Shiryaev filter, this is simply due to writing:

$$d\mathbf{y}_t^\gamma = \frac{1}{\sqrt{\gamma}}\,(dB_t + \sqrt{\gamma}\mathbf{x}_t dt) =: \frac{1}{\sqrt{\gamma}} dW_t^{\mathbb{Q}}\ ,$$

and using the Girsanov transform to construct a new measure $\mathbb{Q}$, for the Kallianpur-Streibel formula, under which $W^{\mathbb{Q}}$ is a Brownian motion – [**VH07**, Chapter 7].

Comparing classical and quantum. Notice that the filtering equations (5.1.6) and (5.1.10) are rigorously identical in their formulation as SDEs. Nevertheless, the physical meaning of their parameters cannot be more different. In the classical filering problem, p and λ parametrize the law of the hidden process $\mathbf{x}$.

Simulations. An effective simulation at large γ of the solution X^γ to the two-state SDE (5.1.10) is given in Figure 5.1.2. From the figure, one observes:

(1) what amounts to a jump process on $\{0, 1\}$. Most of the time, X^γ lives on a thin layer around these points where the noise vanishes.
(2) there is a decoration by spikes. These spikes are very thin since smoothing via convolution blurs them completely.

For the higher dimensional case, we consider n states. By picking initial conditions that insure decohence, the density matrix ρ_t^γ is diagonal. If we call X^γ the diagonal of ρ^γ, then X^γ belongs to the simplex of probability measures $\Delta_n = \{\pi \in \mathbb{R}_+^n \mid \sum_i \pi_i = 1\}$. The extremal states are identified with the vertices of the simplex. A simulation for $n = 3$ is given in Fig. 5.1.3.

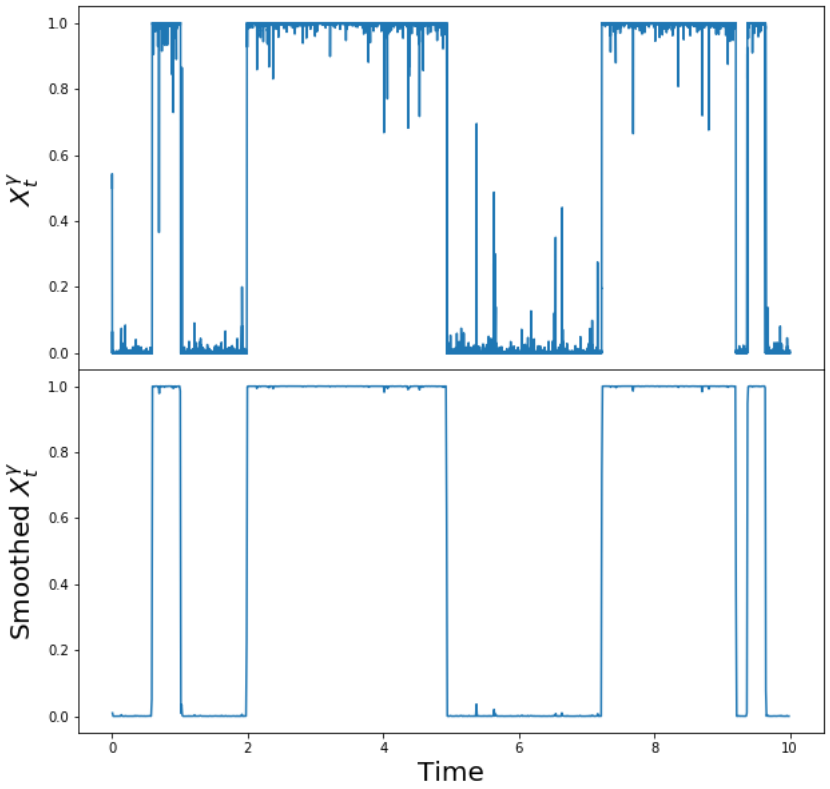

FIGURE 5.1.2. Numerical simulation of the process $(X_t^\gamma; t \geq 0)$ and its smoothing for $\gamma = 10^4$. Parameters are $\lambda = 1.0$ and $p = 0.5$. There are 10^6 time steps. Smoothing is via averaging over 1000 steps. The code is available at the online repository
https://github.com/redachhaibi/quantumCollapse

Further remarks: Regarding the first aspect, the convergence of X^γ to a Markov jump process has been addressed during the last years [BB14, BBT15a, BCF+19] and holds at the level of semi-groups. We refer to this phenomenon as a *local collapse*. From a physical point of view, this is a metastable situation caused by the aforementioned competition between thermalization and collapsing.

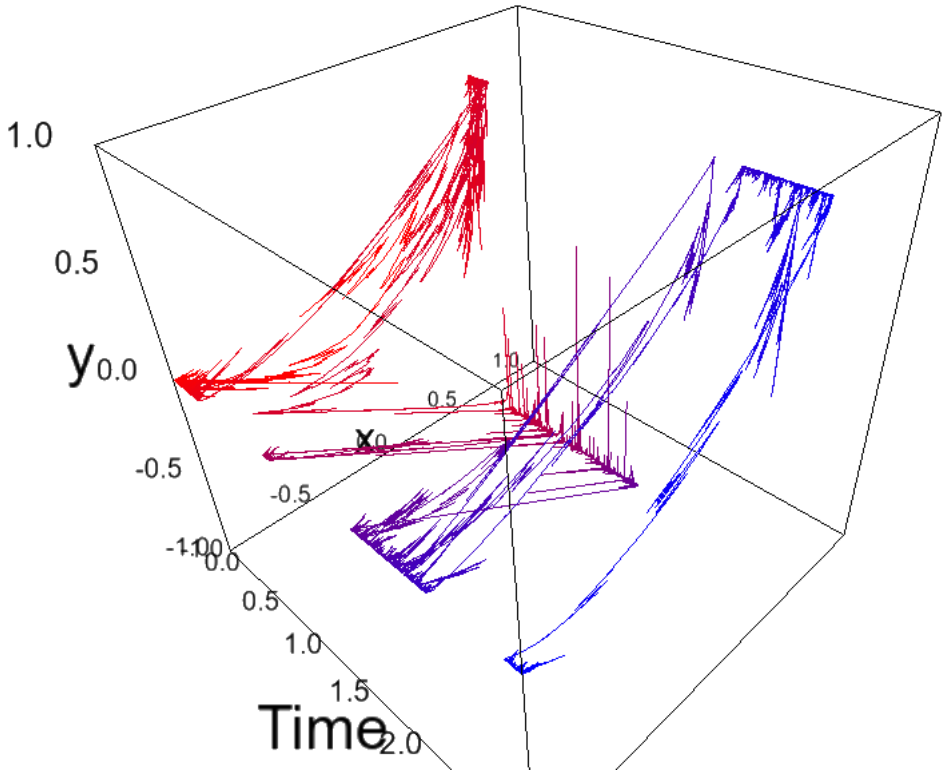

FIGURE 5.1.3. The full process X^γ on the (projected) simplex of probability measures $\Delta_3 = \{\pi \in \mathbb{R}^3_+ \mid \sum_i \pi_i = 1\}$. Same realization as the previous figures. The three corners of the simplex are the often visited points. For better readability, color evolves from red to blue with time.
Like the previous figure, the code is available at the repository
https://github.com/redachhaibi/quantumCollapse

Let us also mention the paper [KL19], where the authors prove for specific SDEs the approximation by a Markov jump process via the study of hitting times and their asymptotics in $\gamma \to \infty$.

The second aspect is more surprising and much less understood. It was first described in [TBB15, BBT16a] and then studied in greater depth in [BB18]. In fact, fluctuations around the local collapse *do* persist in the strong noise limit and take the form of "spikes" decorating the Markov jump process. So far, there is only a limited understanding of the convergence topology and the precise statistics of these spikes.

A general approach developed in [BBT16a, BB18] concerns a change of time (a zooming) which allows to consider the presence of spikes. More precisely, the spikes are explained in terms of excursions of a reflected Brownian motion which appears in the strong noise limit. In particular, in order to obtain their result, the authors of [BB18] prove an effective approximate version of the Skorohod Lemma.

From this body of literature arise the following questions:

QUESTION 5.1.2.

- *How to formalize the limiting phenomena? At this point, the precise mathematical nature of the "spike process" is unclear. Even the name "process" is unwarranted for now.*
- *Is there a limit theorem for the process X^γ as $\gamma \to \infty$? A satisfying answer should be two-fold. On the one hand, we need the convergence to a Markov jump process, which holds only upon smoothing. On the other hand, the convergence to the spike process needs to happen in a non-standard topology, which we need to describe.*

A reasonable answer to these questions should have the following features.

- One would require a precise statement for the convergence of the process X^γ, so that both a Markov jump process $\mathbf{x}$ and a spike process $\mathbb{X}$ are obtained in the limit. We discuss the choice of appropriate topologies in the next Subsection 5.1.4.
- One would require a parsimonious description of the law of the limiting processes.

5.1.4. A key ingredient: the choice of topologies. An important aspect is indeed the choice of topology in order to reflect what we see on the figures 5.1.2 and 5.1.3. For the convergence towards a Markov process, this should only hold up to smoothing. And thus we would require a "Lebesgue-type" topology. In order to capture the spikes, one would need a "uniform-type" topology at the level of geometric shapes, since the limiting phenomenon does not appear to be a function.

Let us start by defining and motivating our choice of "Lebesgue-type", and "uniform-type" topologies.

The Meyer-Zheng topology: This is our weak "Lebesgue-type" topology.

DEFINITION 5.1.3 (Meyer-Zheng topology). *Consider a Euclidean space $(E, \|\cdot\|)$ and denote by $\mathbb{L}^0 := \mathbb{L}^0(\mathbb{R}_+; E)$ the space of E-valued Borel functions on $\mathbb{R}_+$*[1]*. Given a sequence $(w_n)_{n\geq 0}$ of elements of $\mathbb{L}^0$, the following assertions are equivalent and define the convergence in Meyer-Zheng topology of $(w_n)_n$ to $w \in \mathbb{L}^0$:*

- *For all bounded continuous functions $f : \mathbb{R}_+ \times E \to \mathbb{R}$,*

$$\lim_{n\to\infty} \int_0^\infty f\left(t, w_n(t)\right) e^{-t} dt \;\; = \int_0^\infty f\left(t, w(t)\right) e^{-t} dt \;\; .$$

- *For $\lambda(dt) = \; e^{-t}dt$, we have that for all $\varepsilon > 0$,*

$$\lim_{n\to\infty} \lambda\left(\{s \in \mathbb{R}_+ \;\mid\; \|w_n(s) - w(s)\| \geq \varepsilon\}\right) = 0 \;\; ,$$

- *$\lim_{n\to\infty} \mathrm{d}\left(w_n, w\right) = 0$ where d is defined by*

$$\mathrm{d}(w, w') := \int_0^\infty \left\{1 \wedge \|w(t) - w'(t)\|\right\} \, e^{-t} \; dt \;\; .$$

The distance d metrizes the Meyer-Zheng topology on $\mathbb{L}^0$ and $(\mathbb{L}^0, \mathrm{d})$ is a Polish space.

[1] To be more precise $\mathbb{L}^0$ is a quotient space where two functions are considered as equal if they coincide almost everywhere with respect to the Lebesgue measure.

Throughout the text, the reader only needs to have in mind the first characterization of the Meyer-Zheng topology with $E = \mathbb{R}$ or $E = \mathbb{R}^n$. Nevertheless, it is reassuring to know that it is nothing but the convergence in (Lebesgue) measure in an $\mathbb{L}^0$ space.

The Hausdorff topology: In this case, we explicitly take E to be a finite dimensional Euclidean space. Let $H > 0$ be a finite time horizon. The Hausdorff metric induces a complete topology on the collection of closed sets of $[0, H] \times E$ – see [Mun00, Ex. 7, p.280]. It is defined on closed sets $A \subset \mathbb{R} \times E$ and $B \subset E$ via:

$$\mathrm{d}_{\mathbb{H}}(A, B) := \inf \{\varepsilon > 0 \mid A \subset B + \varepsilon \mathbb{B} \ , \ B \subset A + \varepsilon \mathbb{B}\} \ , \tag{5.1.11}$$

where $\mathbb{B} \subset \mathbb{R} \times E$ is the unit ball. In order to motivate this choice, let us consider the restriction of the Hausdorff distance to the graphs of two important classes of functions, $\mathcal{C}$ and $\mathcal{D}$. $\mathcal{C}$ (resp. $\mathcal{D}$) is the space of continuous (resp. càdlàg) maps $f : [0, H] \to \mathbb{R}$.

- Either one can restrict the Hausdorff distance to the graphs of maps in $\mathcal{C}$. In this case, it gives exactly the topology of uniform convergence on $\mathcal{C}$. Indeed, this is a consequence of the following inequality. If f_n is a sequence in $\mathcal{C}$, $\delta_n := d_{\mathbb{H}}(\mathcal{G}(f_n), \mathcal{G}(f))$, $\mathcal{G}(g)$ is the graph of $g \in \mathcal{C}$ and $\omega_f(\delta)$ is the (uniform) modulus continuity of f, then
$$\delta_n \leq \|f_n - f\|_\infty \leq \delta_n + \omega_f(\delta_n) \ .$$
- Or one can restrict to the (completed) graphs of càdlàg maps in $\mathcal{D}$. In this case, Hausdorff convergence of completed graphs is the Skhorohod M_2 topology. See [Whi02, Section 11.5] for a definition and the relationship to the other Skorohod topologies on $\mathcal{D}$.

In the end, the spaces $\mathcal{C}$ and $\mathcal{D}$ endowed with the Hausdorff topology are nothing but the canonical Polish spaces for stochastic processes (see the standard probability textbook [Bil13a]). Therefore, it is natural to analyze the accumulation points of $\mathcal{G}(X^\gamma)$ for the Hausdorff metric on graphs. As an example of observables which are continuous with respect to this topology, we have the hitting times of open sets.

Furthermore, the spike phenomenon is captured by neither the uniform topology nor the Skorohod topology. Otherwise, since these spaces are complete and therefore closed, any accumulation point of the graphs $\mathcal{G}(X^\gamma)$, which is the graph of a function, would be the graph of a continuous or càdlàg function. And as one can guess from Fig. 5.1.2, the limiting spike process $\mathbb{X}$ will not be a bona fide function.

Finally, letting aside all the mathematical abstract non-sense, graphs encode functions as we actually see them. Nothing is more natural than the uniform topology on the shapes we see, hence the choice of the Hausdorff topology.

5.2. Strong noise limits in one-dimensional diffusions

In this Section, we follow [BCC+23] and study one-dimensional diffusions

$$\begin{cases} X_0^\gamma = & x_0 \ , \\ dX_t^\gamma = & b(X_t^\gamma)dt + \sqrt{\gamma}\sigma(X_t^\gamma)dW_t \ . \end{cases} \tag{5.2.1}$$

where W is a standard Wiener process and b, σ are smooth functions. Throughout the paper, we assume the Itô convention. Contrary to the usual weak noise limit ($\gamma \to 0$), developed in the so-called Freidlin-Wentzell theory [FW12], we are interested in the regime where the parameter γ goes to infinity. Our initial motivation

comes from continuous quantum measurements and, as such, let us start by the guiding example which inspired us. Only then we will present the more general setting.

Previous papers [**BB18**, **KL19**] perform a rigorous study only after an uncontrolled perturbative analysis around one of the stable points i.e. points where the noise vanishes. Thus such papers restrict themselves to SDEs living in $[0,\infty)$ with particular coefficients which satisfy

$$\sigma(x) > 0 \quad \text{for all } x \in (0,\infty), \quad \sigma(0) = 0, \quad b(0) > 0.$$

This does not cover the two-boundary case such as Eq. (5.1.10), relevant for quantum mechanics. Moreover, a precise statement describing the spikes has been missing.

In the paper [**BCC+23**], we do not perform any approximation, treat generic coefficients and give a precise description of the spike process. We provide a general technique to study the strong noise limit $\gamma \to \infty$ of one-dimensional SDEs with two possible setups. In the first half of the paper, we have the following working hypotheses:

ASSUMPTION 5.2.1. *We assume that the drift term b and the diffusion coefficient σ are Lipschitz continuous so that the SDE (5.2.1) admits strong solutions. Moreover*

$$\begin{aligned} &\sigma(x) > 0 \quad \text{for all } x \in (0,1), \quad \sigma(0) = \sigma(1) = 0 \ , \\ &b(0) > 0, \quad b(1) < 0 \ . \end{aligned}$$

Recalling that x_0 denotes the initial position, we naturally consider $x_0 \in (0,1)$: the starting point needs to be between points where the noise vanishes.

Our main results in this setup are provided in Theorem 5.2.2. It shows first the convergence of the process $(X_t^\gamma\,;\, t \geq 0)$ to a jump Markov process $(\mathbf{x}_t\,;\, t \geq 0)$ as $\gamma \to \infty$. A reader used to problems of weak convergence of stochastic processes will notice that the previous convergence cannot hold in the usual Skorohod topology since $(X_t^\gamma\,;\, t \geq 0)$ has continuous paths while $(\mathbf{x}_t\,;\, t \geq 0)$ has only càdlàg trajectories. The statement holds only upon smoothing, which is equivalent to the convergence of semi-groups. Hence the precise statement is that for every compactly supported continuous function f of time and space

$$\lim_{\gamma\to\infty} \int_0^\infty f(t, X_t^\gamma)\, dt \;=\; \int_0^\infty f(t, \mathbf{x}_t)\, dt \quad \mathbb{P}-\text{a.s.}$$

Almost sure convergence is due to a particular coupling of X^γ for different γ.

The previous convergence does not detect the spikes that are observed in the numerical simulation given in Fig. 5.1.2. Therefore, in order to mathematically capture them, we need to find a strong "uniform-type" topology. Our solution uses the Hausdorff metric on the graphs of functions. And the second part of our theorem establishes the convergence of $(X_t^\gamma\,;\, t \geq 0)$ to a spike process $\mathbb{X}$ defined thanks to a decoration of the Markov jump process. Within this approach, we obtain the complete picture with the spikes and we make precise the statistics of the involved processes.

5.2.1. Two limiting processes. We start by defining the two processes which shall appear in the main theorem.

On the one hand, we define $(\mathbf{x}_t \ ; \ t \geq 0)$ as the continuous time (càdlàg) Markov process with state space $\{0,1\}$ and jump rates W:

$$W^{0,1} = |b(0)| \, , \qquad W^{1,0} = |b(1)| \ .$$

Here, we wrote $W^{i,j}$ for the jumping rate from state i to state j. The initial position is sampled according to

$$\mathbb{P}\left(\mathbf{x}_0 = 1\right) = 1 - \mathbb{P}\left(\mathbf{x}_0 = 0\right) = x_0 \ .$$

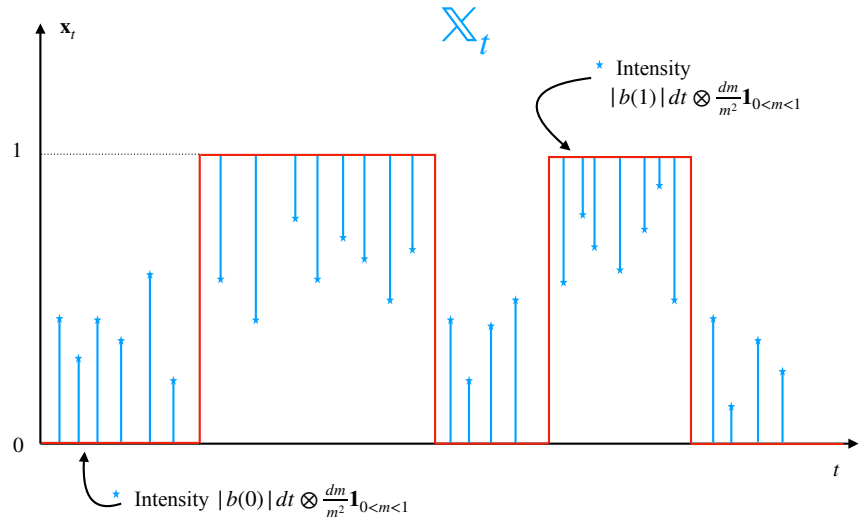

FIGURE 5.2.1. Sketch of the two limiting processes. The Markov pure jump process $\mathbf{x}$ is in red, and the set-valued spike process $\mathbb{X}$ is in blue.

On the other hand, we define the spike process as a *set-valued* random path $\mathbb{X} : \mathbb{R}_+ \to \mathcal{P}\left([0,1]\right)$, where $\mathcal{P}\left([0,1]\right)$ is the power set of the segment $[0,1]$. For a comprehensive sketch, see Figure 5.2.1. It is formally obtained as follows:

- Sample a random initial segment $\mathbb{X}_0$ as

$$\mathbb{X}_0 = \begin{cases} [Y,1] \text{ when } \mathbf{x}_0 = 1, & \mathbb{P}\left(Y \in dy \mid \mathbf{x}_0 = 1\right) = \frac{1-x_0}{x_0} \mathbb{1}_{\{0<y<x_0\}} \frac{dy}{(1-y)^2} \ , \\ [0,Y] \text{ when } \mathbf{x}_0 = 0, & \mathbb{P}\left(Y \in dy \mid \mathbf{x}_0 = 0\right) = \frac{x_0}{1-x_0} \mathbb{1}_{\{x_0<y<1\}} \frac{dy}{y^2} \ . \end{cases}$$

- Sample $\left(t, \widetilde{M}_t\right)$ following a Poisson point process on $\mathbb{R}_+ \times [0,1]$ with intensity

$$\left(dt \otimes \frac{dm}{m^2} \mathbb{1}_{\{0 \leq m < 1\}} \right) \ .$$

Then, by progressively rescaling time for $\left(t, \widetilde{M}_t\right)$ by

$$\begin{cases} |b(0)|^{-1} & \text{when} \quad \mathbf{x}_t = 0 \ , \\ |b(1)|^{-1} & \text{when} \quad \mathbf{x}_t = 1 \ , \end{cases}$$

we obtain a Poisson point process with random intensity which we denote by (t, M_t).

Equivalently, one can sample two independent Poisson point processes on $\mathbb{R}_+ \times [0,1]$ with the above intensity, and then rescale them in time by $|b(0)|^{-1}$ and $|b(1)|^{-1}$ respectively. The process (t, M_t) can also obtained by picking either the first or the second one, depending on the current value of $\mathbf{x}$.

- Finally

$$\mathbb{X}_t = \begin{cases} [0, M_t] & \text{if} \quad \mathbf{x}_t = \mathbf{x}_{t^-} = 0 \ , \\ [1 - M_t, 1] & \text{if} \quad \mathbf{x}_t = \mathbf{x}_{t^-} = 1 \ , \\ [0,1] & \text{if} \quad \mathbf{x}_t \neq \mathbf{x}_{t^-} \ . \end{cases}$$

Notice that by virtue of (t, M_t) being a Poisson point process with finite intensity away from zero, there are no points with the same abscissa and only countably many $t \in \mathbb{R}_+$ with $M_t > 0$. If there is no point with abscissa $t \in \mathbb{R}_+$, then it is natural to set $M_t = 0$ and thus $\mathbb{X}_t = \{\mathbf{x}_t\}$. This convention accounts for the infinite measure at zero.

5.2.2. Main result. We can now state:

THEOREM 5.2.2 (Main Theorem of [BCC+23]). *Under Assumptions 5.2.1, it is possible to couple the processes $(\mathbf{x}, \mathbb{X})$ and X^γ for all values of $\gamma > 0$ on the same probability space, so that the following limits hold almost surely.*

- *Upon smoothing via a continuous function with compact support $f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$, we have the almost sure convergence:*

$$\lim_{\gamma \to \infty} \int_0^\infty f(t, X_t^\gamma)\, dt = \int_0^\infty f(t, \mathbf{x}_t)\, dt \ . \tag{5.2.2}$$

- *In the sense of Hausdorff convergence of closed sets, for all $H > 0$, we have the almost sure convergence of graphs:*

$$\lim_{\gamma \to \infty} (X_t^\gamma ;\ 0 \leq t \leq H) = (\mathbb{X}_t ;\ 0 \leq t \leq H) \ . \tag{5.2.3}$$

REMARK 5.2.3 (Explanations). *The first part of the theorem can be loosely reformulated by saying that the convergence of X^γ to $\mathbf{x}$ holds upon smoothing, which amounts to deleting the spikes. This smoothing is provided by the Meyer-Zheng topology. Nevertheless, one needs an appropriate notion of convergence in order to capture the spikes, which are infinitely thin in the limit. The second part of theorem says that $\mathrm{d}_{\mathbb{H}}(A^\gamma, B) \to 0$ where $A^\gamma = \mathcal{G}(X^\gamma)$ is the graph of X^γ:*

$$A^\gamma := \{(t, X_t^\gamma) \ ,\ 0 \leq t \leq H\} = \bigsqcup_{0 \leq t \leq H} \{t\} \times \{X_t^\gamma\}$$

and B is given by

$$B := \{(t,x) \ ,\ 0 \leq t \leq H,\ x \in \mathbb{X}_t\} = \bigsqcup_{0 \leq t \leq H} \{t\} \times \mathbb{X}_t \ .$$

The set B is seen as the graph of the multi-valued function $\mathbb{X}$. The fact that B is a closed set comes as a by-product of the proof.

IDEAS OF THE PROOF OF THEOREM 5.2.2. The main ideas behind the proof are very specific to dimension 1:

- Every one-dimensional diffusion is a Brownian motion upon changing space and time.

- And when time is parametrized by the inverse of local time, a Brownian trajectory can be broken into excursions thanks to Itô's Excursion Theory.

More precisely, the first point is achieved by sucessively composing with the scale function in order to obtain a martingale and then invoking the Dambis-Dubins-Schwarz (DDS) Theorem [**RY99**, Chapter V, Theorem 1.6] to obtain a Brownian motion β. There, we make explicit the coupling for different $\gamma > 0$: the processes X^γ are coupled thanks to this single DDS Brownian motion β. After computing the asymptotics of these changes of scale, we are able to force the appearance of the limiting processes $\mathbf{x}$ and $\mathbb{X}$. Only the construction of $\mathbb{X}$ will require Itô's Excursion Theory [**RY99**, Chapter XII]. □

Note that the dynamics of the jump process $\mathbf{x}$ and of the spike process $\mathbb{X}$ depend on the characteristics (b, σ^2) of the initial diffusion (5.2.1) only through the absolute value of the drift b at 0 and 1. Indeed, as long as Assumption 5.2.1 is satisfied, these limiting processes are identical no matter the values of the drift b and the diffusivity σ in the bulk $(0, 1)$. In fact, the only impact of σ is in the selection of the space where the jump process $\mathbf{x}$ will live, i.e. $\{0, 1\}$ in this case.

This can be contrasted with the jump processes appearing in the weak noise limits [**FW12**, Chapter 6]. In this case, with diffusions whose drift is the gradient of a potential, the limiting jump processes depend on the full landscape given by this potential. Therefore, it is fair to say that strong noise limits are far more universal than weak noise limits.

5.2.3. Further remarks.

Beyond Assumptions 5.2.1: Such assumptions can be slightly relaxed, as long as the SDE (5.2.1) continues to have strong solutions and that the points $\{0, 1\}$ are natural boundaries in the sense of Gilman-Skorohod - see [**Kle05**, Section 6.9] for a definition and references. As long as that is satisfied, one could adapt the proofs to a drift b and a volatility σ vanishing at $\{0, 1\}$.

Generalizations to multiple zeros of σ and beyond the Rabi setup: What if $\{0, 1\}$ is not a natural boundary? The answer is conceptually simple but difficult to turn into a comprehensive theorem. Consider the case where σ has multiple zeroes. For the sake of simplicity, assume isolated zeroes of order 1, which is the generic behavior. As long as the process is Markovian, the residual drift will always give a unique direction. Indeed, if $\sigma(x_0) = 0$, then this direction is given by the sign of $b(x_0)$. If the process jumps in random direction, then it is not a one-dimensional diffusion. Thus it is natural to think of a process successively crossing the segments between two zeros of σ, if the sign of the drift b allows it.

Let us call a domain a region between two zeroes. There are only the two following cases which we deem interesting. (1) Either the diffusion is trapped in a single domain as treated in the main theorem (unattainable boundaries on both sides) (2) Or the diffusion goes from domain to domain as in the Rabi case (attainable boundary which becomes entrance boundary on the other side).

Regarding the other cases, we believe that a general theorem can be written. However, if one aims for treating all possible cases, one would need to discuss the 36 cases (6 cases for each boundary) in the classification of boundaries – see Table 6.2 in [**KT81**]. Instead, we settled for two extremal cases rather than discussing the general combinatorics of domain change. An important example such as the Rabi

setup, treated in the second half of the paper [BCC+23], is rich enough to illustrate the robustness of our approach. And a complete description of large noise limits for all one-dimensional diffusions might not be very useful.

Limitations: Our approach uses crucially scale functions and the Dambis-Dubins-Schwarz Theorem, which are one-dimensional tools. In particular, extensions to a multi-dimensional setting are absolutely not straightforward. This is a current subject of investigations.

Coupling: On a side note, the coupling of the processes $(\mathbf{x}, \mathbb{X})$ with X^γ for different $\gamma > 0$ is nothing but a convenient device. Exactly like Skorohod's Representation Theorem [Bil13a, Theorem 6.7], it allows to recast weak convergence to an almost sure convergence. Such a coupling is particularly convenient in order to avoid formalizing the weak convergence of random closed sets in the Hausdorff topology, or the weak convergence in the Meyer-Zheng topology. In fact, the paper [BBC+21] uses this latter aspect, invoking a powerful tightness criterion authored by Meyer and Zheng.

5.3. Filtering

Let us go back to the classical setup of filtering, as it naturally complements the previous Section. Filtering Theory adresses the problem of estimating a *hidden process* $\mathbf{x} = (\mathbf{x}_t \ ; \ t \geq 0)$ which can not be directly observed. At hand, one has access to an *observation process* which is naturally correlated to $\mathbf{x}$. The most simple setup, called the "signal plus noise" model, is the one where the observation process $\mathbf{y}^\gamma = (\mathbf{y}^\gamma_t \ ; \ t \geq 0)$ is of the form (5.1.1).

Moreover it is natural to assume that the noise is intrinsic to the observation system, so that the Brownian motion $B = B^\gamma$ has no reason of being the same for different values of γ. See Figure 5.3.1 for an illustration which visually highlights the difficulty of recognizing a drift despite Brownian motion fluctuations. In this paper we shall focus on the case where $(\mathbf{x}_t \ ; \ t \geq 0)$ is a pure jump Markov process on $\{0, 1\}$ with càdlàg trajectories. We denote λp (resp. $\lambda(1-p)$) the jump rate between 0 and 1 (resp. between 1 and 0), with $p \in (0, 1)$ and $\lambda > 0$. This is the historical setting of the celebrated Wonham filter [Won64, Eq. (19)].

In the mean square sense, the best estimator taking value in $\{0, 1\}$ at time t of $\mathbf{x}_t$, given the observation $(\mathbf{y}^\gamma_s)_{s \leq t}$, is equal to

$$\hat{\mathbf{x}}^\gamma_t = \mathbb{1}_{\{\pi^\gamma_t > \frac{1}{2}\}} \tag{5.3.1}$$

where π^γ_t is the conditional probability

$$\pi^\gamma_t := \mathbb{P}\left(\mathbf{x}_t = 1 \mid (\mathbf{y}^\gamma_s)_{s \leq t}\right) \ . \tag{5.3.2}$$

Our interest lies in the situation where the intensity $1/\sqrt{\gamma}$ of the observation noise is small, i.e. γ is large. At first glance, one could argue that weak noise limits for the observation process are not that interesting because we are dealing with extremely reliable systems since they are subject to very little noise. This paper aims at demonstrating that this regime is interesting from both a theoretical and a practical point of view.

A motivating example. Let us describe a simple situation that falls into that scope and motivates our study. Consider for example a single classical bit – say, inside of a DRAM chip. The value of the bit is subject to changes, some of

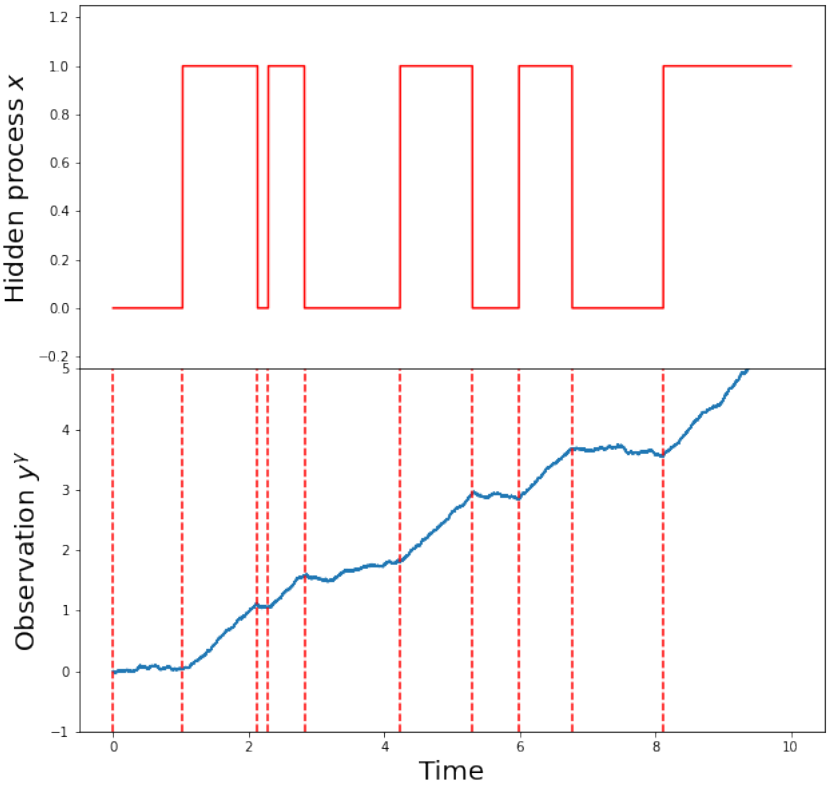

FIGURE 5.3.1. Numerical simulation of the hidden process $\mathbf{x}$ and the observation process $\mathbf{y}^\gamma$ for $\gamma = 10^2$. The challenge is to infer the drift of $\mathbf{y}^\gamma$, in spite of Brownian noise and in a very short window. Parameters are $\lambda = 1.3$ and $p = 0.4$. There are 10^6 time steps to discretize $[0, 10]$. The code is available at the online repository https://github.com/redachhaibi/Spikes-in-Classical-Filtering

which are caused by CPU instructions and computations, some of which are due to errors. The literature points to spontaneous errors due to radiation, heat and various conditions [SPW09]. The value of that process is modeled by the Markov process $\mathbf{x}$ as defined above. Here, the process $\mathbf{y}^\gamma$ is the electric current received by a sensor on the chip, which monitors any changes. Any retroaction, for example code correction in ECC memory [KLK+14, PKHM19], requires the observation during a finite window $\delta > 0$. And the reaction is at best instantaneous. For anything meaningful to happen, everything depends thus on the behavior of:

$$\pi_t^{\delta,\gamma} := \mathbb{P}\left(\mathbf{x}_{t-\delta} = 1 \mid (\mathbf{y}_s^\gamma)_{s\leq t}\right) , \tag{5.3.3}$$

and instead to consider the estimator $\hat{x}_t^\gamma$ given by Eq. (5.3.1), we are left with the estimator

$$\hat{\mathbf{x}}_t^{\delta,\gamma} = \mathbb{1}_{\left\{\pi_t^{\delta,\gamma} > \frac{1}{2}\right\}} .$$

From an engineering point of view, it is the interplay between different time scales which is important in order to design a system with high performance: if the noise is weak, how fast can a feed-back response be? For a given process $z = (z_t\,; t \geq 0)$ with values in $[0,1]$ we denote the hitting time of $(\frac{1}{2}, 1)$ by $T(z) := \inf\left\{t \geq 0\,;\; z_t > \frac{1}{2}\right\}$. Assume for example that initially $\mathbf{x}_0 = 0$. For a given time $t > 0$, a natural problem is to estimate, as $\gamma \to \infty$, the probability to predict a false value of the bit given its value remains equal to 0 during the time interval $[0,t]$, i.e.

$$\mathbb{P}\left(T(\hat{\mathbf{x}}^{\delta,\gamma}) \leq t \mid T(\mathbf{x}) > t\right)\ . \tag{5.3.4}$$

Informal statement of the result. A consequence of the results of this paper is the precise identification of the regimes $\delta := \delta(\gamma)$ for which the probability in (5.3.4) vanishes or not as $\gamma \to \infty$:

- If $\limsup_{\gamma\to\infty} \delta(\gamma)\frac{\gamma}{\log\gamma} < 2$, i.e. δ is too small, the retroaction/control system can be surprised by a so–called spike, causing a misfire in detecting the regime change and the limiting error probability in Eq. (5.3.4) is equal to $1 - \exp(-\lambda p t)$;
- If $\liminf_{\gamma\to\infty} \delta(\gamma)\frac{\gamma}{\log\gamma} > 2$, i.e. δ is sufficiently large, the estimator will be very good at detecting jumps of the Markov process $\mathbf{x}$, the limiting error probability in Eq. (5.3.4) vanishing. However the reaction time will deteriorate.

While the literature usually focuses on L^2 considerations for filtering processes, we focus on this article on pathwise properties of the filtering process under investigation when $\gamma \to \infty$. Indeed, it is clear that the question addressed just above cannot be answered in an L^2 framework only.

Let us now present in some informal way the reasons for which we have this difference of behavior. As it will be recalled later the process $\pi^\gamma = (\pi^\gamma_t\ ;\ t \geq 0)$ satisfies in law

$$d\pi^\gamma_t = -\lambda\,(\pi^\gamma_t - p)\,dt + \sqrt{\gamma}\pi^\gamma_t\,(1 - \pi^\gamma_t)\,dW_t\ , \tag{5.3.5}$$

where $W = (W_t\ ;\ t \geq 0)$ is a Brownian motion with a now strong parameter $\sqrt{\gamma}$ in front of it. This is the so called Shiryaev-Wonham filtering theory which was presented in the beginning of this chapter. As shown in the previous Section, when γ goes to infinity the process π^γ converges in law to an unusual and singular process in a suitable topology (see Figure 5.3.2). Indeed as exhibited in the figure, the limiting process is the Markov jump process $(\mathbf{x}_t\ ;\ t \geq 0)$ but decorated with vertical lines, called spikes, whose extremities are distributed according an inhomogeneous point Poisson process. As we can observe on Figure 5.3.3, if δ is sufficiently large, the spikes in the process $\pi^{\gamma,\delta}$ are suppressed while if δ is sufficiently small they survive. The spikes are responsible of the non vanishing error probability in Eq. (5.3.4) since they are interpreted by the estimator $\hat{\mathbf{x}}^{\delta,\gamma}$ as a jump from 0 to 1 of the process $\mathbf{x}$. The fact that the transition between the two regimes is precisely $2\frac{\log\gamma}{\gamma}$ is more complicated to explain without going into computational details. Building on our earlier results, we examine hence in this paper the effect of smoothing and the relevance of various time scales required for filtering, smoothing and control in the design of a system with feedback.

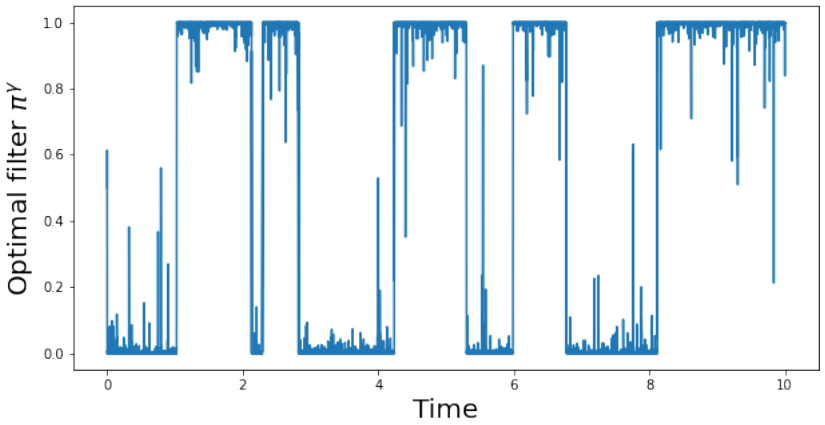

FIGURE 5.3.2. "The whims of the Wonham filter": Informally, on a very short time interval, it is difficult to distinguish between a change in the drift of $\mathbf{y}^\gamma$ and an exceptionnal time of Brownian motion. The figure shows a numerical simulation of the process $(\pi_t^\gamma \ ; \ t \geq 0)$ for the same realization of $\mathbf{x}$ as Fig. 5.3.1. Same time discretization. This time we chose the larger $\gamma = 10^4$ to highlight spikes.

5.3.1. Literature review of filtering theory in the $\gamma \to \infty$ regime. The understanding of the behavior of the classical filter for jump Markov processes with small Brownian observation noise has attracted some attention in the 90's. Most of the work is focused on the long time regime [**Won64**, **KL92**, **KZ96**, **AZ97b**, **AZ97a**, **Ass97**], by studying for example stationary measures, asymptotic stability or transmission rates. In the case where the jump Markov process is replaced by a diffusion process with a signal noise, possibly small, [**Pic86**, **AZ98**] study the efficiency (in the L^2 sense and at fixed time) of some asymptotically optimal filters. In [**PZ05**] are obtained quenched large deviations principles for the distribution of the optimal filter at a fixed time for one dimensional nonlinear filtering in the small observation noise regime – see also [**RBA22**]. In a similar context Atar obtains in [**Ata98**] some non-optimal upper bounds for the asymptotic rate of stability of the filter.

Going through the aforementioned literature one can observe that the term $\log\gamma/\gamma$ already appears in those references. Indeed the quantities of interest include the (average) long time error rate [**Ass97**, Eq. (1.4)]

$$\alpha^* = \lim_{t\to\infty} \frac{1}{t} \int_0^t \min(\pi_s^\gamma, 1 - \pi_s^\gamma) ds$$

or the probability of error in long time ([**Won64**] and [**KZ96**, Theorem 1'])

$$\mathcal{P}_{err}(\gamma) = \lim_{t\to\infty} \inf_{\zeta \in L_\infty(\mathcal{F}_t^{\mathbf{y}})} \mathbb{P}(\zeta \neq \mathbf{x}_t) = \lim_{t\to\infty} \mathbb{P}(\hat{\mathbf{x}}_t \neq \mathbf{x}_t)$$

or the long time mean squared error [**Gol00**]

$$\mathcal{E}_{mse}(\gamma) = \lim_{t\to\infty} \inf_{\zeta \in L_\infty(\mathcal{F}_t^{\mathbf{y}})} \mathbb{E}(\zeta - \mathbf{x}_t)^2 = \lim_{t\to\infty} \mathbb{E}(\pi_t^\gamma - \mathbf{x}_t)^2 \ .$$

Here $\mathcal{F}^{\mathbf{y}}$ denotes the natural filtration of $\mathbf{y} = \mathbf{y}^\gamma$. These quantities are shown to be of order $\frac{\log\gamma}{\gamma}$ up to a constant which is related to the invariant measure of $\mathbf{x}$ and

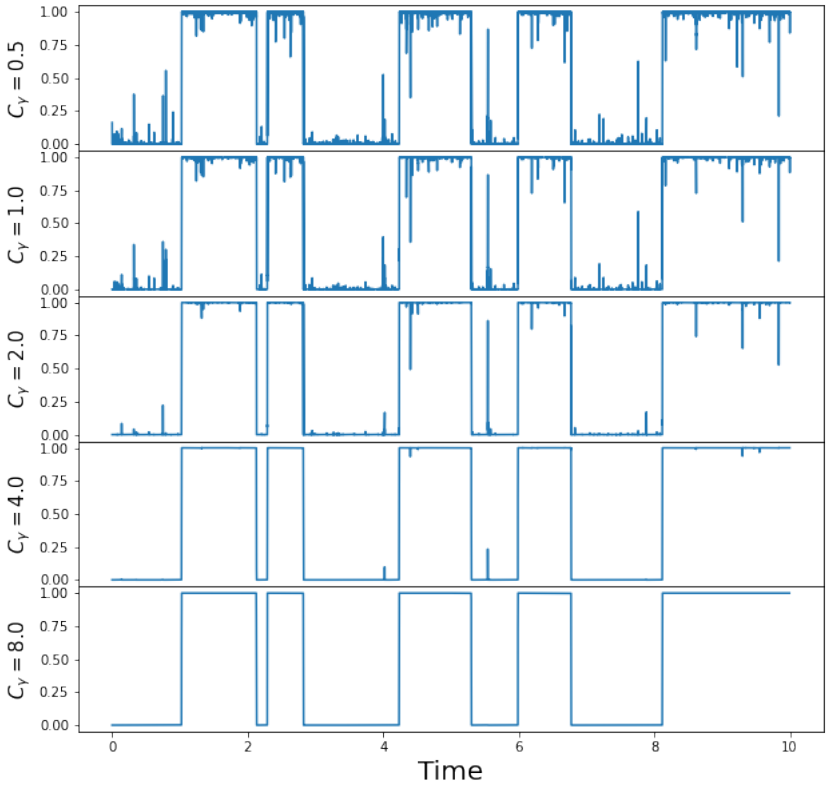

FIGURE 5.3.3. Numerical simulation of the process $\left(\pi_t^{\delta,\gamma}\ ;\ t \geq 0\right)$ for the same realization of $\mathbf{x}$ as Fig. 5.3.1. Same time discretisation. We have $\gamma = 10^4$ and $\delta_\gamma = C\frac{\log\gamma}{\gamma}$, with $C \in \{\frac{1}{2}, 1, 2, 4, 8\}$.

some relative entropy but which is definitively not 2 – see [**Gol00**, Eq. (3)]. Note that all these quantities are of asymptotic nature and their analysis goes through the invariant measure. Beyond the appearance of the quantity $\frac{\log\gamma}{\gamma}$, which is fortuitous, our results are of a completely different nature since we want to obtain a sharp result on a fixed finite time interval. Also, due to the spiking phenomenon and the singularity of the involved processes, there is no chance that the limits can be exchanged.

To the best of the authors' knowledge, this paper is the first of its kind to aim for a trajectorial description of the limit, in the context of classical filtering theory. However, the spiking phenomenon has first been identified in the context of quantum filtering [**Mab09**, Fig. 2] and more specifically, for the control and error correction of qubits. The spiking phenomenon is already seen as a possible source of error where correction can be made while no error has occurred. To quote [**Mab09**, Section 4], when discussing the relevance of the optimal Wonham filter in the strong noise regime, it "is not a good measure of the information content of the system, as it is very sensitive to the whims of the filter".

5.3.2. Statement of the problem and Main Theorem about a phase transition. Recall that the observation equation and the filtering equation are respectively (5.1.5) and (5.1.6). And Eq. (5.3.5) falls in the scope of Theorem 5.2.2 which treats the strong noise limits of a large class of one-dimensional SDEs.

A mathematical statement: The convergences were established thanks to a convenient (but fictitious) coupling of the processes $(\pi^\gamma \ ; \ \gamma > 0)$ for different $\gamma > 0$. In contrast, the filtering problem has a natural coupling for different $\gamma > 0$ which is given by the observation equation (5.1.1). In this context, let us state a small adaptation of an already established result.

THEOREM 5.3.1 (Variant of the Main Theorem 5.2.2). *There is a two-faceted convergence.*

(1) **In the $\mathbb{L}^0$ topology and in probability**, *we have the following convergence :*

$$(\pi_t^\gamma \ ; 0 \leq t \leq H) \overset{\gamma \to \infty}{\longrightarrow} (\mathbf{x}_t \ ; 0 \leq t \leq H) \ .$$

Equivalently, that is to say

$$\forall \varepsilon > 0, \quad \lim_{\gamma \to \infty} \mathbb{P}\left(\mathrm{d}_{\mathbb{L}}(\pi^\gamma, \mathbf{x}) > \varepsilon\right) = 0 \ .$$

Here $\mathbf{x}_0 \in \{0,1\}$ is Bernoulli distributed with parameter π_0^γ the initial condition[2] *of π^γ.*

(2) **In the Hausdorff topology for graphs and in law**, *we have that the graph of $(\pi_t^\gamma \ ; 0 \leq t \leq H)$ converges to a spike process $\mathbb{X} = \bigsqcup_{t \in [0,H]} (\{t\} \times \mathbb{X}_t)$ described by Fig. 5.2.1.*

(3) **In the Hausdorff topology for graphs and in law**, *we have that the graph of $\hat{\mathbf{x}}^\gamma = (\hat{\mathbf{x}}_t^\gamma \ ; 0 \leq t \leq H)$, defined by Eq. (5.3.1), converges to another singular random closed set $\hat{\mathbb{X}} = \bigsqcup_{t \in [0,H]} (\{t\} \times \hat{\mathbb{X}}_t)$ where*

$$\hat{\mathbb{X}}_t = \{0,1\} \mathbb{1}_{\{\mathbb{X}_t \cap [0,\frac{1}{2}) \neq \emptyset, \mathbb{X}_t \cap (\frac{1}{2},1] \neq \emptyset\}} + \{0\} \mathbb{1}_{\{\mathbb{X}_t \subset [0,\frac{1}{2})\}} + \{1\} \mathbb{1}_{\{\mathbb{X}_t \subset (\frac{1}{2},1]\}} \ .$$

Notice that the first convergence is in the weaker Lebesgue-type topology and holds in probability i.e. on the same probability space. The second and third convergences are in the stronger uniform-type topology, however they only hold in law.

POINTERS TO THE PROOF. The second point is indeed a direct corollary of [BCC+23] since almost sure convergence after a coupling implies convergence in law, regardless of the coupling.

The third point is also immediate modulo certain subtleties. Recalling that $\hat{\mathbf{x}}_t^\gamma = \mathbb{1}_{\{\pi_t^\gamma > \frac{1}{2}\}}$ and that the graph of π^γ converges to the random closed set $\mathbb{X}$, it suffices to apply the Mapping Theorem [Bil13b, Theorem 2.7]. Indeed, a spike $\mathbb{X}_t \subset [0,1]$ is mapped to either $\{0\}$, $\{1\}$ or $\{0,1\}$ when examining the range of the indicator $\mathbb{1}_{\{\cdot > \frac{1}{2}\}}$ on $\mathbb{X}_t$. However, when invoking the Mapping Theorem, one needs to check that discontinuity points of the map $\mathbb{1}_{\{\cdot > \frac{1}{2}\}}$ have measure zero for the law of $\mathbb{X}$. This is indeed true since there are no spikes of height $\frac{1}{2}$ almost surely – recall that the spike process $\mathbb{X}$ is described in terms of Poisson processes.

The first point, although simpler and intuitive, does not come Theorem 5.2.2. In the case of filtering, the process $\mathbf{x}$ is intrinsically defined, and we require the use

[2] We assume π_0^γ independent of γ.

of the speficic coupling given by the additive model (5.1.1). Let us show how the result is reduced to a single claim. The result is readily obtained from the Markov inequality and the $L^1(\Omega)$ convergence:

$$\lim_{\gamma\to\infty} \mathbb{E}\ \mathrm{d}_{\mathbb{L}}\left(\pi^\gamma, \mathbf{x}\right) = 0\ .$$

The above convergence itself only requires the definition of $\mathrm{d}_{\mathbb{L}}$ in Definition 5.1.3, Lebesgue's dominated convergence theorem and the claim

$$\forall t > 0, \quad \lim_{\gamma\to\infty} \mathbb{E}\left|\pi_t^\gamma - \mathbf{x}_t\right|^2 = 0\ . \tag{5.3.6}$$

In order to prove Claim (5.3.6), recall that by definition π_t^γ is a conditional expectation:

$$\begin{aligned}\pi_t^\gamma &= \mathbb{P}\left(\mathbf{x}_t = 1 \mid (\mathbf{y}_s^\gamma)_{s\le t}\right)\\ &= \mathrm{argmin}_{c\in\mathcal{F}_t^{\mathbf{y}}}\mathbb{E}(\mathbb{1}_{\mathbf{x}_t=1} - c)^2\\ &= \mathrm{argmin}_{c\in\mathcal{F}_t^{\mathbf{y}}}\mathbb{E}(\mathbf{x}_t - c)^2\ .\end{aligned}$$

At this stage, let $\varepsilon > 0$ and let us introduce the process $z^\varepsilon = (z_t^\varepsilon\ ;\ t \ge \varepsilon)$ defined for all $t \ge \varepsilon$ by

$$z_t^\varepsilon = \frac{1}{\varepsilon}\int_{t-\varepsilon}^t d\mathbf{y}_s^\gamma\ .$$

This process is clearly $(\mathcal{F}_t^{\mathbf{y}})_{t\ge\varepsilon}$ adapted, so for all $t \ge \varepsilon$, by definition of π_t^γ

$$\begin{aligned}\mathbb{E}\left|\pi_t^\gamma - \mathbf{x}_t\right|^2 &\le \mathbb{E}\left|z_t^\varepsilon - \mathbf{x}_t\right|^2\\ &= \mathbb{E}\left|\frac{1}{\varepsilon}\int_{t-\varepsilon}^t d\mathbf{y}_s^\gamma - \mathbf{x}_t\right|^2\\ &= \mathbb{E}\left|\frac{1}{\varepsilon}\int_{t-\varepsilon}^t \mathbf{x}_s ds - \mathbf{x}_t + \frac{1}{\varepsilon\sqrt{\gamma}}\int_{t-\varepsilon}^t dB_s\right|^2\\ &\le 2\mathbb{E}\left|\frac{1}{\varepsilon}\int_{t-\varepsilon}^t \mathbf{x}_s ds - \mathbf{x}_t\right|^2 + 2\mathbb{E}\left|\frac{1}{\varepsilon\sqrt{\gamma}}\int_{t-\varepsilon}^t dB_s\right|^2\\ &= 2\mathbb{E}\left|\frac{1}{\varepsilon}\int_{t-\varepsilon}^t (\mathbf{x}_s - \mathbf{x}_t) ds\right|^2 + \frac{2}{\varepsilon\gamma}\\ &\le 2\mathbb{E}\left|\frac{1}{\varepsilon}\int_{t-\varepsilon}^t \mathbb{1}_{\{\mathbf{x}_s\neq\mathbf{x}_t\}} ds\right|^2 + \frac{2}{\varepsilon\gamma}\\ &\le 2\mathbb{P}\left(\mathbf{x}\text{ jumps at least one time during } [t-\varepsilon, t]\right) + \frac{2}{\varepsilon\gamma}\ .\end{aligned}$$

Note that we have used that for $\varepsilon \le s \le t$,

$$\{\mathbf{x}_s \neq \mathbf{x}_t\} \subset \{\mathbf{x}\text{ jumps at least one time during } [t-\varepsilon, t]\}.$$

Taking $\gamma \to \infty$ then $\varepsilon \to 0$ proves Claim (5.3.6). □

We can now formally state the question of interest:

QUESTION 5.3.2. *For different regimes of $\delta = \delta_\gamma$ and γ, how do the spikes behave in the stochastic process (5.3.3)? Basically, we need an understanding of the tradeoff between spiking and smoothing. The intuition is that there are two regimes:*

- *The slow feedback regime: the smoothing window δ is large enough so that the optimal estimator $\pi^{\delta,\gamma}$ correctly estimates the hidden process $\mathbf{x}$.*
- *The fast feedback regime: the smoothing window δ is too small so that $\pi^{\delta,\gamma}$ does not correctly estimate the hidden process $\mathbf{x}$. One does observe the effect of spikes.*

Our finding is that there is sharp transition between the slow feedback regime and the fast feedback regime:

THEOREM 5.3.3 (Main theorem of [BCNP22]). *As long as $\delta_\gamma \to 0$, we have the convergence in the $\mathbb{L}^0$ topology and in probability, as in the first item of Theorem 5.3.1:*

$$\left(\pi_t^{\delta_\gamma,\gamma}\ ; 0 \le t \le H\right) \overset{\gamma\to\infty}{\longrightarrow} (\mathbf{x}_t\ ; 0 \le t \le H)\ . \tag{5.3.7}$$

However, in the stronger topologies, there exists a sharp transition when writing:

$$\delta_\gamma = C\frac{\log\gamma}{\gamma}\ .$$

There are two constants $0 < C_- < 2 < C_+$ such that following convergences hold in the Hausdorff topology on graphs in $[0,H]\times[0,1]$.

- *(Fast feedback regime) If $C < C_-$, smoothing does not occur and we have convergence in law to the spike process:*

$$\lim_{\gamma\to\infty} \pi^{\delta_\gamma,\gamma} = \mathbb{X}\ .$$

- *(Slow feedback regime) If $C > C_+$, smoothing occurs and we have convergence:*

$$\lim_{\gamma\to\infty} \pi^{\delta_\gamma,\gamma} = \mathbf{x}\ .$$

  This convergence holds equivalently for the usual M_2 Skorohod topology and for the Hausdorf topology on graphs.

REMARK 5.3.4 (On the transition). *Without much change in the proof, one can consider $C = C_\gamma$ depending on γ. In that setting, the fast feed-back regime and the slow feed-back regime correspond respectively to*

$$\limsup_{\gamma\to\infty} C_\gamma < C_- \ \text{ and } \ \liminf_{\gamma\to\infty} C_\gamma > C_+\ .$$

Furthermore, we strongly believe that the transition happens at exactly $C = 2$, but there are technical issues to prove that.

5.4. Homogeneization in multiple dimensions

Let us start by considerations from quantum mechanics which motivate the stochastic differential equations (SDEs) studied in this paper, as well as their strong noise limits.

5.4.1. Physical motivations.

Semigroups associated to open quantum systems: In quantum optics, the evolution of a d-level atom is often described using a Markov approximation. Then the system state, encoded into a $d \times d$ density matrix (*i.e.* a positive semidefinite matrix of trace 1), evolves by the action of a semigroup of completely positive[3] trace preserving linear maps whose generator $\mathcal{L}$ is called a Lindbladian. More precisely, the evolution of the system's density matrix is solution of the linear ordinary differential equation (ODE)

$$d\bar{\rho}_t = \mathcal{L}(\bar{\rho}_t)dt, \quad \bar{\rho}_0 \in \{\rho \in M_d(\mathbb{C}) : \rho \geq 0, \quad \operatorname{tr}\rho = 1\}. \tag{5.4.1}$$

Such equations are known as *quantum master equations*. In a typical quantum optics experiment, one may identify three different contributions to the evolution of the atom. A first contribution is the Hamiltonian dynamic that an experimenter would like to realize. A second one is the unavoidable environment perturbation that often leads the atom to a steady state. The third one is the effect of any instrument that the experimenter may put in contact with the atom to track its state. For more details, we refer the reader to [**BP02**].

In this article we are interested in situations where the dynamics generator, $\mathcal{L} \equiv \mathcal{L}_\gamma$, is associated to three well separated time scales. The separation is done through some parameter $\gamma > 0$:

$$\mathcal{L}_\gamma = \mathcal{L}^{(0)} + \gamma\mathcal{L}^{(1)} + \gamma^2\mathcal{L}^{(2)} \ .$$

To motivate such a setting, let us consider experiments similar to the famous one realized by Haroche's group [**GBD+07**]. In such experiments the aim is to track the unitary dynamic of a d-energy level quantum system when it is well-isolated from its environment. The dynamic induced by the environment is modeled by $\mathcal{L}^{(0)}$, the unitary dynamic by $\mathcal{L}^{(1)}$ and the effect of the instrument by $\mathcal{L}^{(2)}$. Here the large γ limit corresponds to a fast decoherence, at speed γ^2, induced by the instrument compared to the slower steady state relaxation induced by the environment, with speed $\gamma^0 = 1$. To counteract the Zeno effect, the relevant scale of the unitary dynamic is the intermediary speed $\gamma^1 = \gamma$.

This choice of scaling of the Lindbladian is not limited to such experimental situations. For different examples of dynamics verifying our choice of scales, see [**Per98**, Section 4.3].

Stochastic semigroups in the presence of measurements: Equation (5.4.1) only describes the evolution of a quantum system without reading measurement outcomes coming from the instruments. Taking them into account leads to a stochastic process $\rho^\gamma = (\rho_t^\gamma \ ; \ t \geq 0)$ called a quantum trajectory and which takes values in density matrices. This process is solution to an SDE called a *stochastic quantum master equation*. The drift part of this SDE is given by $\mathcal{L}_\gamma(\rho_t^\gamma)$. The noise part results from conditioning upon the measurement outcomes. Such models are often used to describe experiments in quantum optics – see [**WM10b**, **BP02**]. In the present article we limit ourselves to diffusive quantum trajectories. In that case the SDE takes the form, in the Itô convention,

$$d\rho_t^\gamma = \mathcal{L}_\gamma(\rho_t^\gamma) + \sigma_\gamma(\rho_t^\gamma)dW_t, \tag{5.4.2}$$

[3]Completely positive maps $\Phi : M_d(\mathbb{C}) \to M_d(\mathbb{C})$ are linear maps such that for any $n \in \mathbb{N}$, $\Phi \otimes \mathrm{Id}_{M_n(\mathbb{C})} : M_d(\mathbb{C}) \otimes M_n(\mathbb{C}) \to M_d(\mathbb{C}) \otimes M_n(\mathbb{C})$ is positive.

where the volatility σ_γ is a quadratic function of density matrices and W is a standard Brownian motion. The average evolution of ρ^γ solution of (5.4.2) is given by the solution $\overline{\rho}^\gamma$ of (5.4.1).

SDEs such as (5.4.2) where first introduced as effective stochastic models for wave function collapse – see [**Gis84**, **Pea84**, **Dio88**] and references therein. Their Poisson noise version was defined as a numerical tool to compute the average evolution $\overline{\rho}^\gamma$ in [**DCM92**]. Since then, different justifications were given for the fact that they model quantum systems which are subject to continuous indirect measurements. Historically, the first one is based on quantum stochastic calculus and quantum filtering [**Bel89**]. In that setting, the interaction of the measurement apparatus and the environment with the open system is unitary and described by a quantum SDE [**Par92**]. We refer the reader to [**BVHJ07**] for an accessible introduction to quantum filtering.

A second, more phenomenological, approach [**BH95**, **BG09**] starts with a linear SDE extending the deterministic linear equation (5.4.1). By normalizing the resulting process in the set of positive semidefinite operators, and after a Girsanov transform, one obtains the SDE (5.4.4) for density matrices.

Another approach is based on the continuous-time limit of fast quantum repeated measurements. Introducing proper scaling, discrete time quantum trajectories converge weakly, in the continuous time limit, towards processes solution of SDEs such as (5.4.4) – see [**Pel10**, **BBB12**] and references therein.

5.4.2. Our contribution. In [**BBC⁺21**], we generalize to arbitrary finite dimensions the first half of the aforementioned Theorem 5.2.2, that is to say the convergence towards the jump process between pointer states. In order to have an intrinsic proof, it is desirable to invoke the classical machinery of weak convergence of stochastic processes and to avoid using any coupling. Also, since we focus only on the jumps between the states, the spikes need to be discarded. As shown in [**BBT16b**, **BCC⁺23**], only countably many spikes appear in the limit, each being infinitely thin. As such, spikes are of zero Lebesgue measure and disappear upon averaging. Therefore, an ideal candidate for this task is the topology of convergence in (Lebesgue) measure as explained in Subsection 5.1.4.

The study of the convergence in law of stochastic processes in this topology was pioneered by Meyer and Zheng in [**MZ84**] – see [**Kur91**] for further developments and [**Reb87**] for an application to weak noise limits. This topology is also called pseudo-paths topology and is much weaker than the usual Skorokhod topology.

Our main result is stated in Theorem 5.4.3. It shows that in the Meyer-Zheng topology, in the limit of large γ, the quantum trajectory we study converges in law to a Markov process on the pointer states with explicit rates. Not only this provides an extension but also a mathematically complete and rigorous proof of the pioneering works of [**BBT15b**].

We also establish a general homogeneization result for semigroups on finite-dimensional Hilbert spaces that is instrumental to the proof of Theorem 5.4.3. In the usual homogeneization references such as [**Pap78**, **CD99**, **PS08**, **BLP11**], there is a trivial distinction between a slow and a fast variable and it is then assumed that by fixing the slow variable the fast process is ergodic. The novelty of our homogenization result is that it holds for abstract semigroups and moreover the state space is not a

priori the direct product of slow and fast variables. In particular, we show that it applies to semigroups generated by Lindbladians $\mathcal{L}_\gamma$.

5.4.3. Definitions.

5.4.3.1. *Notation.* We denote by $\langle \cdot, \cdot \rangle$ the standard scalar product on $\mathbb{C}^d$ and $\|\cdot\|$ the corresponding norm; $M_d(\mathbb{C})$ the set of $d \times d$ complex matrices, X^* the conjugate transpose of $X \in M_d(\mathbb{C})$. The Hilbert-Schmidt inner product $(X, Y) \in M_d(\mathbb{C})^2 \mapsto \langle X, Y \rangle := \mathrm{tr}(X^*Y)$ transforms $M_d(\mathbb{C})$ into a Hilbert space and the associated norm is also denoted by $\|\cdot\|$. The set $\mathcal{S} = \{\rho \in M_d(\mathbb{C}) \mid \rho \geq 0,\ \mathrm{tr}\,\rho = 1\}$ is a compact convex set whose elements are called density matrices. For any two matrices X, Y of $M_d(\mathbb{C})$, $[X, Y] := XY - YX$ is the commutator while $\{X, Y\} := XY + YX$ is the anti-commutator. An endomorphism $\mathcal{L}$ on $M_d(\mathbb{C})$ is sometimes called a super-operator while a matrix $X \in M_d(\mathbb{C})$ is called an operator. The algebra of super-operators $(\mathrm{End}(M_d(\mathbb{C})), +, \circ)$ is equipped with the operator norm (with respect to the Hilbert-Schmidt norm on $M_d(\mathbb{C})$) and denoted also by $\|\cdot\|$. We usually reserve the notation $\circ$ for super-operators to emphasize the distinction with operators. The adjoint w.r.t the Hilbert-Schmit scalar product of a linear operator $\mathcal{L} : M_d(\mathbb{C}) \to M_d(\mathbb{C})$ is denoted by $\mathcal{L}^*$. For $x \in \mathbb{C}^\ell$ and $A := (A_k)_{k=1}^\ell \in M_d(\mathbb{C})^\ell$, we denote $A \cdot x = \sum_{k=1}^\ell A_k x_k$ and the action of $\mathcal{L} \in \mathrm{End}(M_d(\mathbb{C}))$ on such A is understood component-wise, i.e. $\mathcal{L}(A) = (\mathcal{L}(A_k))_{k=1}^\ell$.

5.4.3.2. *Lindbladians.* By definition a Lindbladian $\mathcal{L} : M_d(\mathbb{C}) \to M_d(\mathbb{C})$ is the generator of a continuous semigroup of completely positive trace-preserving maps, which describes the Markovian evolution of a quantum open system. Following [**GKS76**, **Lin76**], a Lindblad super-operator admits a GKSL[4] decomposition i.e. for all $X \in M_d(\mathbb{C})$:

$$\mathcal{L}(X) = -\imath[H, X] + \sum_{k=1}^{\ell} \left(L_k X L_k^* - \tfrac{1}{2}\{L_k^* L_k, X\} \right) , \tag{5.4.3}$$

where $(H, (L_k)_{k=1}^\ell)$ are matrices of $M_d(\mathbb{C})$ such that $H^* = H$. We call the first matrix, H, the Hamiltonian and the operators $(L_k)_{k=1}^\ell$, Kraus operators.

Diffusive quantum trajectories with three time scales In this paper, for any $\gamma > 0$, we consider diffusive quantum trajectories given by the Itô SDE

$$\begin{aligned} d\rho_t^\gamma &= \mathcal{L}_\gamma(\rho_t^\gamma)dt + \sum_{\alpha=0,1,2} \gamma^{\frac{\alpha}{2}} \sigma^{(\alpha)}(\rho_t^\gamma) \cdot dW_t^\alpha \\ &= \mathcal{L}_\gamma(\rho_t^\gamma)dt + \sigma^{(0)}(\rho_t^\gamma) \cdot dW_t^0 + \gamma^{\frac{1}{2}} \sigma^{(1)}(\rho_t^\gamma) \cdot dW_t^1 + \gamma \sigma^{(2)}(\rho_t^\gamma) \cdot dW_t^2 \end{aligned} \tag{5.4.4}$$

with initial condition $\rho_0^\gamma = \varrho \in \mathcal{S}$. Throughout the paper, the Itô convention for SDEs is in place. The drift $\mathcal{L}_\gamma(\rho_t^\gamma)$ of this equation is given by the Lindblad super-operator $\mathcal{L}_\gamma$ having the form

$$\mathcal{L}_\gamma := \mathcal{L}^{(0)} + \gamma \mathcal{L}^{(1)} + \gamma^2 \mathcal{L}^{(2)} . \tag{5.4.5}$$

We denote by $(H^{(0)}, (L_k^{(0)})_{k=1}^{\ell_0})$, $(H^{(1)}, (L_k^{(1)})_{k=1}^{\ell_1})$ and $(H^{(2)}, (L_k^{(2)})_{k=1}^{\ell_2})$ the GKSL decompositions of the Lindbladians $\mathcal{L}^{(0)}$, $\mathcal{L}^{(1)}$ and $\mathcal{L}^{(2)}$ respectively.

In the noise part $\sum_{\alpha=0,1,2} \gamma^{\frac{\alpha}{2}} \sigma^{(\alpha)}(\rho_t^\gamma) \cdot dW_t^\alpha$, the processes $(W^\alpha, \alpha = 0, 1, 2)$ are independent ℓ_α-dimensional (standard) Wiener processes and the maps $\sigma^{(\alpha)} : \rho \in$

[4]The acronym GKSL stands for Gorini-Kossakowski-Sudarshan-Lindblad.

$\mathcal{S} \mapsto \left(\sigma_k^{(\alpha)}(\rho)\right)_{k=1}^{\ell_\alpha} \in M_d(\mathbb{C})^{\ell_\alpha}$ are the three quadratic maps defined component-wise by

$$\sigma_k^{(\alpha)}(\rho) = \eta_\alpha(k)\left(L_k^{(\alpha)}\rho + \rho L_k^{(\alpha)*} - \mathrm{tr}[(L_k^{(\alpha)*} + L_k^{(\alpha)})\rho]\,\rho\right) , \tag{5.4.6}$$

for $\alpha = 0, 1, 2$ and $k = 1, \ldots, \ell_\alpha$. Here $(\eta_\alpha(k))_{k=1}^{\ell_\alpha} \in [0,1]^{\ell_\alpha}$ are given numbers. We refer the reader to the notations in Section 5.4.3.1 for the meaning of $\sigma^{(\alpha)}(\rho_t^\gamma)\cdot dW_t^\alpha$.

The proof of existence and uniqueness of the strong solution to Eq. (5.4.4) can be found in [**BH95**, **Pel08**, **BG09**, **Pel10**]. In these references, it is also proven that $\rho^\gamma \in C\left(\mathbb{R}_+ ; \mathcal{S}\right)$ almost surely.

Since the SDE (5.4.4) has a linear drift, it follows that the average evolution of ρ^γ is expressed in terms of the semigroup generated by $\mathcal{L}_\gamma$:

$$\forall t \geq 0, \quad \mathbb{E}(\rho_t^\gamma) = e^{t\mathcal{L}_\gamma}\varrho\ . \tag{5.4.7}$$

The asymptotic analysis of this average semigroup in fact plays a crucial role in the proof of the main result.

In terms of interpretation of indirect measurement, the Wiener process W^α results from the output signal of measurements. The numbers $\eta_\alpha(k)$ are introduced in order to encapsulate in a single form the measurement and thermalization aspects. More precisely $\eta_\alpha(k) = 1$ corresponds to perfectly read measurements, $\eta_\alpha(k) \in (0,1)$ to imperfectly read measurements and $\eta_\alpha(k) = 0$ to unread measurements or to model contributions from a thermal bath.

Assumptions. Let us now state and discuss our working assumptions for the main result.

Assumption 5.4.1 (Quantum Non-Demolition (QND) assumption). *The operators $H^{(2)}$, $(L_k^{(2)})_{k=1}^{\ell_2}$ and $(L_k^{(1)})_{k=1}^{\ell_1}$ are all diagonalizable in a common orthonormal basis $(e_i)_{i=1}^d$ of $\mathbb{C}^d$, called the pointer basis.*

Observe that no assumption is made on the Hamiltonian $H^{(1)}$ nor on the Kraus operators and Hamiltonian involved in $\mathcal{L}^{(0)}$. Also, Assumption 5.4.1 is equivalent to requiring that the $*$-algebra generated by the Kraus operators $(L_k^{(2)})_{k=1}^{\ell_2}$, the Hamiltonian $H^{(2)}$ as well as the Kraus operators $(L_k^{(1)})_{k=1}^{\ell_1}$ is commutative.[5]

From a physical perspective, the QND assumption is standard. It is at the cornerstone of the experiment [**GBD+07**] where QND measurements are used to count the number of photons in a cavity without destroying them. It is shown that it reproduces the wave function collapse in long time – see [**BB11**, **BBB13**, **BP14**] and references therein. This condition is tailored to preserve the pointer states during the quantum measurement process. More precisely, under the QND Assumption 5.4.1, in the case $\mathcal{L}^{(0)} = \mathcal{L}^{(1)} = 0$, if the initial state is a pointer state, i.e. $\varrho \in \{E_{i,i} := e_i e_i^*\}_{i=1}^d$, then it is not affected by the indirect measurement in the sense that the state remains unchanged by the stochastic evolution (5.4.4). Note that this behavior is very specific to such models since measurement usually induces a feedback on the quantum system.

A simple yet crucial computation detailed in a separate Lemma in [**BBC+21**] shows then that under the QND Assumption 5.4.1 the super-operator $\mathcal{L}^{(2)} : M_d(\mathbb{C}) \to$

[5] In other words a C^* commutative subalgebra of $M_d(\mathbb{C})$.

$M_d(\mathbb{C})$ is diagonalizable with eigenvectors $\left(E_{i,j} := e_i e_j^*\right)_{i,j=1}^d$ and associated eigenvalues:

$$\tau_{i,j} = -\frac{1}{2}\sum_{k=1}^{\ell}\left|(L_k^{(2)})_{i,i} - (L_k^{(2)})_{j,j}\right|^2 - \imath\left(H_{i,i}^{(2)} - H_{j,j}^{(2)} + \sum_{k=1}^{\ell}\Im\left(\overline{(L_k^{(2)})_{i,i}}(L_k^{(2)})_{j,j}\right)\right) . \tag{5.4.8}$$

Here and in the following, if $X \in M_d(\mathbb{C})$, the notation $X_{i,j}$ always refer to the coordinates of X in the pointer basis $(e_i)_{i=1}^d$. Observe also that the family $(E_{i,j})_{i,j=1}^d$ forms an orthonormal basis of $M_d(\mathbb{C})$.

ASSUMPTION 5.4.2 (Identifiability condition). *For any $i, j \in \{1, \dots, d\}$ such that $i \neq j$, there exists $k \in \{1, \dots, \ell_2\}$ such that $\eta_2(k) > 0$ and*

$$\Re(L_k^{(2)})_{i,i} \neq \Re(L_k^{(2)})_{j,j} .$$

In fact, from Eq. (5.4.8), this assumption together with the QND assumption imply the non-existence of purely imaginary eigenvalues $\tau_{i,j}$ for the super-operator $\mathcal{L}^{(2)}$. We shall see that this will play an important role.

Our motivation to qualify this assumption as identifiability originates again from the theory of non-demolition measurements. Indeed, following [**BBB13**, **BP14**], if the QND Assumption 5.4.1 and the identifiability condition of Assumption 5.4.2 hold, for any $\gamma > 0$, the quantum trajectory obtained when setting $\mathcal{L}^{(0)} = \mathcal{L}^{(1)} = 0$ converges almost surely, as t grows, to a random pointer state, reproducing a non-degenerate projective measurement along the pointer basis. If the identifiability Assumption 5.4.2 does not hold, the limiting random state may exist but will correspond to a degenerate measurement.

Statement:

THEOREM 5.4.3 (Main theorem of [**BBC+21**]). *For any $\gamma > 0$ let ρ^γ be the continuous processes on $\mathcal{S}$ solution of Eq. (5.4.4) starting from ϱ. Under the QND Assumption 5.4.1 and the identifiability condition in Assumption 5.4.2, we have*[6]*:*

$$\lim_{\gamma\to\infty} \rho^\gamma = \mathbf{x}\mathbf{x}^*, \quad \text{weakly in } \left(\mathbb{L}^0(\mathbb{R}_+; M_d(\mathbb{C})),\ \mathrm{d}\right) ,$$

where $\mathbf{x} := (\mathbf{x}_t\ ;\ t \geq 0)$ is a pure jump continuous-time Markov process on the pointer basis $(e_i)_{i=1}^d$ with initial distribution μ_ϱ defined by

$$\mu_\varrho : e_i \mapsto \langle e_i, \varrho e_i\rangle .$$

Furthermore, the generator T of the Markov process $\mathbf{x}$ is explicit. The transition rate from e_i to e_j, $i \neq j$, is given by

$$T_{i,j} = \sum_{k=1}^{\ell_0}|(L_k^{(0)})_{j,i}|^2 + \frac{\left|\left(H^{(1)}\right)_{i,j}\right|^2}{|\tau_{i,j}|^2}\sum_{k=1}^{\ell_2}\left|(L_k^{(2)})_{i,i} - (L_k^{(2)})_{j,j}\right|^2 . \tag{5.4.9}$$

[6]We recall that if χ is a topological space and $(X_n)_n$ is a sequence of χ-valued random variables, we say it converges weakly (or in law) to the χ-valued random variable X if and only if for any bounded continuous function $f : \chi \to \mathbb{R}$, $\lim_{n\to\infty}\mathbb{E}\big[f(X_n)\big] = \mathbb{E}\big[f(X)\big]$.

Here $\tau_{i,j}$ is the eigenvalue of $\mathcal{L}^{(2)}$ corresponding to the eigenvector $E_{i,j} = e_i e_j^*$ given in Eq. (5.4.8).

STRATEGY OF PROOF. The approach is structured as follows.

In a first step, we gave a general homogeneization result for semigroups in the form of Theorem 3.1 in [BBC+21]. The proof follows the philosophy pioneered by Nakajima-Zwanzig. Then we apply it to the case of Lindblad super-operators. Let us mention [BCF+17, Theorem 2.2] as an inspiration for the proof and that our result is consistent with [MGLG16, ABFJ16]. There, we show that in the large γ limit, the dynamic of the semigroup $e^{t\mathcal{L}_\gamma}$ reduces to a dynamic generated by an operator $\mathcal{L}_\infty$ whose expression is explicitly given in terms of $\mathcal{L}^{(0)}$, $\mathcal{L}^{(1)}$ and $\mathcal{L}^{(2)}$. Thanks to Eq. (5.4.7), this leads to the convergence of the mean $\mathbb{E}(\rho_t^\gamma)$. Although this may seem to be very partial information, it is sufficient to identify the generator T.

Only then, we tackled the proof which follows the usual approach for the weak convergence of stochastic processes: we used a tightness criterion in the Meyer-Zheng topology and then identify the limit via its finite-dimensional distributions. Interestingly, the convergence of the mean is bootstrapped to the convergence of finite-dimensional distributions thanks to the Markov property and the collapsing on pointer states $(E_{i,i})_{i=1}^d$. □

5.4.4. Further remarks.

Convention on $\mathcal{L}$ and T: The Markov generator T defined in Eq. (5.4.9) follows the usual probabilistic convention in the sense that $T\mathbb{1} = 0$. On the contrary, to simplify notations, for the various Lindblad generators $\mathcal{L}$, we use the convention that they generate trace-preserving maps, thus their duals with respect to the Hilbert–Schmidt inner product verify $\mathcal{L}^*(\mathrm{id}) = 0$, which is equivalent to $\mathrm{tr} \circ \mathcal{L} = 0$.

Generalizations: Let us mention two possible extensions of the setting of this paper. In principle, our results and methods of proof carry to these cases mutatis mutandis. However, such extensions are not included as this would considerably decrease the readability of the paper.

In Eq. (5.4.5), one could consider a further dependence in γ by replacing $\mathcal{L}^{(0)}$ by $\mathcal{L}_\gamma^{(0)}$ such that a limit holds as γ goes to infinity.

Also, throughout the paper, we limit ourselves to diffusive noises. But the setting can be extended to include Poisson noises in the SDE (5.4.4). In passing, let us mention an interesting result in this direction using a completely different method. In the particular case $\mathcal{L}^{(0)} = 0$ and for Poisson noises only, instead of Wiener ones, an analogous result to the Main Theorem 5.4.3 is possible, building on [BCF+17, Theorem 2.3 item (b)]. Indeed, in that article, it is obtained that there exists $C > 0$ such that for any $t > 0$, $\mathbb{E}(\|\rho_t^\gamma - Y_t^\gamma\|) \leq C\gamma^{-\frac{1}{2}}\sqrt{|\log\gamma|}$ with $(Y^\gamma)_{\gamma>0}$ converging in law to $\mathbf{x}\mathbf{x}^*$, in Skorokhod's topology, as $\gamma \to \infty$. Since $\mathcal{S}$ is compact, integrating both side of the inequality with respect to the probability measure $\lambda(dt) = e^{-t}dt$ on $\mathbb{R}_+$, the result follows from $\mathbb{L}^1(\mathbb{R}_+, \lambda)$ convergence.

The noise vanishes on pointer states: The following intuition dictates that the noise vanishing on pointer states is crucial in order to have emergence of jump processes from strong noise limits as in Theorem 5.4.3. The idea is that, as γ grows larger, the process ρ^γ will spend more time in a thin layer around the points where the noise vanishes. Because of the QND Assumption 5.4.1 and the structure of the

maps σ_α given in Eq. (5.4.6), the noise in the SDE (5.4.4) vanishes exactly on the pointer states $(E_{i,i})_{i=1}^d$, hence the intuition of a limiting process taking values in $(E_{i,i})_{i=1}^d$.

Reversibility properties of the limit generator T: In the context of dynamics where there is a clear distinction between slow variables and fast variables, consider the reduced dynamic in slow variables, obtained by the elimination of fast variables by homogeneization. A common belief in statistical physics is that such a reduced dynamic is in general "more irreversible" than the initial dynamic [**Mac89**, **Leb99**, **GK04**, **Lav04**, **Bal05**] – and regardless of the reversibility of this initial dynamic. The seminal example is provided by the (irreversible) Boltzmann equation which is derived by a kinetic limit from a (reversible) microscopic dynamic ruled by Newton's equations of motion [**Cer88**].

In our context, from both mathematical and physical perspectives, this leads naturally to ask the question of the links between the reversibility properties of the SDE (5.4.4) and the reversibility properties of our effective Markov process $\mathbf{x} := (\mathbf{x}_t \ ; \ t \geq 0)$. This SDE is generically non-reversible but it may happen, in various situations, that the effective Markov process however is, highlighting a possible moderation to the aforementioned popular belief, at least in this particular quantum context. Indeed, it is for example easy to check that if $H^{(1)} = 0$ and there exists a probability $p := (p_i)_{i=1}^d$ such that for any $1 \leq k \leq \ell_0$, $p_i|\big(L_k^{(0)}\big)_{j,i}|^2 = p_j|\big(L_k^{(0)}\big)_{i,j}|^2$, then T is reversible with respect to the probability p, while the SDE is not. Therefore it would be interesting to understand what are the conditions to impose on the Kraus and Hamiltonian operators, and more importantly their physical meaning, in order to obtain a reversible T. In the previously mentioned example, the condition is reminiscent of the one resulting from a weak coupling limit [**Ali76**, **Dav76**, **AL07**] of a quantum system interacting with a heat bath at thermal equilibrium, showing this condition has probably some deeper physical interpretation.

CHAPTER 6

Perspectives

Now that we have presented some of the main results since I have settled down in Toulouse, it is time to turn to the future. In this chapter, I will present some of the perspectives that I have in mind.

6.1. Perspectives about the Toda system

In this section, we draw perspectives that naturally follow after the content of Chapter 3.

6.1.1. Towards hydrodynamics following Deift and Spohn. Let us start by making more precise the first statement of Theorem 3.2.2. When asking a computer to diagonalize a real symmetric matrix, the standard algorithm is Arnoldi-Lanczos-QR.scheme, which performs as follows.

- Input: $M = (m_{i,j})_{1\leq i,j\leq n}$ symmetric real.
- Step 1: Trotter reduction
$$M \rightsquigarrow J = QMQ^*$$
with Q orthogonal, J Jacobi with identical spectrum as M. Recall that a Jacobi matrix is a tridiagonal matrix with non-negative off-diagonal entries.
- Step 2: Recall that the QR decomposition is the Gram-Schmidt orthogonalization of the columns of a matrix, yielding an orthogonal matrix Q and an upper triangular matrix R. The Arnoldi-Lanczos-QR iterative scheme consists in applying the QR decomposition to the matrix J at each step:
$$\rightsquigarrow J_0 = J = Q_0R_0$$
$$\rightsquigarrow J_1 = Q_0^*J_0Q_0 = Q_1R_1$$
$$\rightsquigarrow J_2 = Q_1^*J_1Q_1 = Q_2R_2$$
$$\dots$$
$$\rightsquigarrow J_k = Q_{k-1}^*J_{k-1}Q_{k-1} \ .$$

Now, for a formal statement:

THEOREM 6.1.1.
- *(Arnoldi/Lanczos) If eigenvalues are simple, the algorithm performs diagonalization:*
$$J_k \overset{k\to\infty}{\longrightarrow} Diag(\Lambda_1, \Lambda_2, \dots, \Lambda_n) \ .$$
- *(Symes 1982) The Arnoldi-Lanczos-QR iteration scheme is exactly the Toda flow at integer times:*
$$\forall k \in \mathbb{N} \ , J_k = J^{(n),Toda}(t = k) \ .$$

This latter striking fact has been nicknamed "the stroboscopic effect" by the mathematician Deift: Advancing the Toda flow on a Jacobi matrix by one unit of time is equivalent to one step of the Arnoldi-Lanczos-QR scheme. Therefore, the Toda lattice *is exactly* the Arnoldi-Lanczos-QR scheme which powers some of the most useful algorithms. Hence this natural question, which was asked by Deift.

QUESTION 6.1.2 (Problem 4 in [D+17]). *Starting from a random initial condition for the Toda lattice, is there a description of the (large time) limiting dynamic for large matrices?*

Of course, the same question at integer times amounts to analyzing the Arnoldi-Lanczos-QR. If answered properly, this question would shed light on the behavior of the diagonalization algorithm for large matrices. This is part of a larger program aimed at understanding RMT universality in algorithms [DMOT14].

In fact, virtually the same question was asked by theoretical physicist Spohn [Spo20], while motivated by the question of hydrodynamical limits. More generaly, he is recently interested in the hydrodynamical limits of Hamiltonian systems with infinitely many conserved quantities. But what is a hydrodynamical limit? That is a macroscopic description, usually a PDE, derived from the microscopic behavior of particles. For example, rather than describing the individual positions of a large number of particles, one derives an evolution equation for the particle density. For a recent panorama of the state of the art, we recommend [QY] for an account focused on Varadhan's work, and [SR09] for a point of view focused on the analysis of PDEs and on the legacy of Boltzmann.

Although Spohn derived the equations for the densities of conserved quantities in the Toda lattice, in terms of the current's local average, that was only at the physical level of rigor. A crucial ingredient which was assumed is the persistence of the local equilibrium property, which is a non-trivial mathematical question – see the more extensive review [Spo21] for details.

QUESTION 6.1.3. *Prove the local equilibrium property for the Toda lattice and thus complete Spohn's hydrodynamical description.*

When comparing the two questions 6.1.2 and 6.1.3, we see that Spohn's formulation is formulated in the language of statistical physics. Moreover, the technical issue is clearly identified. Nevertheless, it is essentially the same question. And I think this question is an excellent candidate for a first step in hydrodynamical limits. Indeed, the main objects are close to my historical research interests: RMT and stochastic integrability.

6.1.2. General Lie type tridiagonal models and MMO intergrals. The definition in general Lie type requires a few definitions: let $\mathfrak{a}$ be the Cartan Lie algebra of a complex Lie group with rank r, the Weyl group W acts on $\mathfrak{a}$, the norm $\|\cdot\|$ is given by the Killing form and d_j are the degrees of the generators of the algebra of invariants $\mathbb{C}[\mathfrak{a}]^W$. With these notations, the Macdonald-Mehta-Opdam (MMO) integral states that:

$$\int_{\mathfrak{a}} dx \ |\Delta(x)|^\beta \, e^{-\frac{1}{2}\|x\|^2} = (2\pi)^{\frac{1}{2}r} \prod_{j=1}^{r} \frac{\Gamma\left(1+\frac{\beta}{2}d_j\right)}{\Gamma\left(1+\frac{\beta}{2}\right)} \ , \tag{6.1.1}$$

where the Vandermonde is replaced by the Weyl denominator $\Delta(x) = \prod_{\beta\in\Phi^+}\langle\beta, x\rangle$. Here Φ^+ are the positive roots of the root system Φ associated to W.

REMARK 6.1.4. *The MMO formula is valid more generally for W a finite Coxeter group, i.e., also for the non-crystallographic Coxeter groups H_3, H_4 and $I_2(n)$ for $n \neq 3, 4, 6$. These are the symmetry groups of the dodecahedron, the 120-cell, and regular polygons respectively.*

Howeever, our method does not directly apply since the Toda lattice is a Lie-theoretic construction which makes use of the associated group and there are no Lie groups associated to non-crystallographic root systems. See [**FK06**] *for some constructions related to Toda lattices in the non-crystallographic case, using the reduction method.*

Entirely analogously to Eq. (3.2.5), one defines the generalized Toda lattice for abitrary root systems:

$$H_G = \frac{\|p\|_{\mathfrak{a}}^2}{2} + \sum_{\alpha \in \Delta} \frac{\|\alpha\|^2}{2} e^{-\alpha(q)} \ , \tag{6.1.2}$$

where $\| \cdot \|_{\mathfrak{a}}$ is the norm induced by the Killing form and $\Delta \subset \Phi$ is a choice of simple roots, such that $\Phi \subset \mathrm{Span}_{\mathbb{Z}} \Delta$. In this setup, the generalized Toda lattice plays an important role in representation theory as shown for example in Kostant's work [**Kos79**] and the author's [**Chh13**]. The previous sentence deserves more explanations. The claim from Kostant's work is that integrating the classical Toda system i.e. knowing exactly all possible evolutions is equivalent to knowing the representation theory of the underlying group. Afterall, solving the evolution needs matrix coefficients in all representations. A similar claim is made in [**Chh13**], where the harmonic analysis of the quantum Toda lattice encodes geometric crystals. And geometric crystals are algebro-geometric objects reflecting finer information on the representation theory.

We can now state the question:

QUESTION 6.1.5. *Is it possible to build matrix models like Dumitriu and Edelman's for general root systems? There, the analogue Λ of the spectrum should have a distribution:*

$$\mathbb{P}\left(\Lambda \in dx\right) = \frac{1}{Z^{\beta}(\mathfrak{a})} \left|\Delta(x)\right|^{\beta} e^{-\frac{\beta}{2}\|x\|_{\mathfrak{a}}^2} \, dx \ .$$

It is similar to the Gaussian β-ensemble of Eq. 3.1.1 except that, once again, now the Vandermonde means $\Delta(x) := \prod_{\alpha \in \Phi^+} \langle \alpha, x \rangle$, where Φ^+ are the positive roots.

An interesting goal would be to describe a "tridiagonal" model in the spirit of Theorem 3.1.1. In fact, I believe this would probably not interest the RMT community very much. However, the hope is that it would give a new proof of the MMO integrals (6.1.1).

This takes us to one of the very influencial Macdonald conjectures. The identity (6.1.1) is in fact given as [**Mac82**, Conjecture 6.1]. It is now settled by Opdam [**O$^+$89**, § 6] using his technology of hypergeometric shift operators, which until now seemed unrelated to the Toda system. In the same fashion as Dumitriu-Edelman, answering Question 6.1.5 would provide a geometric proof of the Macdonald-Mehta-Opdam-Selberg integrals for general root systems and a strong relationship to the ubiquitous Toda system.

It is our belief that the correct proof of the MMO integrals should be a simple book-keeping exercise by keeping track of a normalization constant for the tridiagonal model of Question 6.1.5. And the change of variables which correspond to

diagonalization in general Lie type is better encoded in the scattering of the Toda flow.

In passing let us mention that many of the ingredients are available.

- The scattering of positions as in Theorem 3.2.2 is available from the works of Goodman and Wallach [**GW84**, Subsection 2.3].
- More importantly, in the general Lie type, Jacobi matrices are replaced by slices of coadjoint orbits, and Moser's scattering asymptotics for generalized Toda chain are due to Kostant [**Kos79**].

In my opinion, the main difficulty is to understand the geometry of Kostant's coadjoint slices and to choose good coordinates to describe the scattering.

6.2. Perspectives on quantum groups at roots of unity

In this section, we draw perspectives that naturally follow after the content of Chapter 4. While the generalization to higher rank of the paper [**CC21**] is a natural continuation, we do not detail that aspect here. Indeed, I feel that Section 4.4 shows very clearly where that leads – at least on the semi-classical level. Instead, we now focus on another aspect: the quantum group at roots of unity and the positive curvature.

Following [**CP95**, Chapter 11], the representation theory of $\mathcal{U}_q(\mathfrak{sl}_2)$ changes drastically when q is a root of unity. In particular, the category of finite dimensional representations is no longer semisimple. This is a very interesting phenomenon, which I not fully understand. In any case, I do understand that representations at roots of unity are at the basis of the construction of the famous Jones polynomial, which is a knot invariant. Also, its computation requires the class restricted representations of [**CP95**, §11.2].

In the future, I would like to understand what is the proper semi-classical limit, in the spirit of Kirillov's orbit method. Our construction of a semi-classical limit for the quantum group uses the space $\mathbb{H}^3 = SL_2(\mathbb{C})/SU_2 = NA$ and normalizes the curvature to $\kappa = -\frac{1}{2}r^2$. In the quantum picture $q = e^{-r}$ is the deformation parameter in our presentation of the quantum group $\mathcal{U}_q^{\hbar}(\mathfrak{sl}_2)$. Loosely speaking, the exotic representation theory at roots of unity reflects what should happen upon taking $q = e^{ir}$.

One is tempted to perform the "Wick rotation" of the curvature parameter $r \rightsquigarrow ir$. And it strangely makes sense on multiple levels.

- It forces the curvature to becomes positive via $\kappa = \frac{1}{2}r^2 \geq 0$.
- The model space of constant positive curvature, in dimension 3, is now the sphere $S^3 \approx SU_2$.
- The radial part of Brownian motion has generator which changes from

$$\mathcal{L}_{\mathbb{H}^3} = \frac{1}{2}\partial_x^2 + (\partial_x \log \sinh(rx))\,\partial_x$$

in negative curvature to

$$\mathcal{L}_{S^3} = \frac{1}{2}\partial_x^2 + (\partial_x \log \sin(rx))\,\partial_x$$

in positive curvature. It is indeed common to see trigonometric functions in positive curvature, replaced by their hyperbolic counterparts in negative curvature.

Nevertheless, so far, I am failing to see how generators of the quantum group at roots of unity are related to S^3.

6.3. Optimal transport (OT)

Beyond the fact that optimal transport is a popular topic in mathematics, my interest lies specifically in spectral and numerical aspects of the Sinkhorn formulation. As we will see, this shares many similarities with the content of Subsection 4.4.

This is the PhD topic of Anirban BOSE, who is supervised jointly with Serge GRATTON.

6.3.1. Elements of OT. Optimal transport is the general problem of moving one distribution of mass to another as efficiently as possible. See Fig. 6.3.1 for an illustrative drawing and Fig. 6.3.2 for an effective solution on an example.

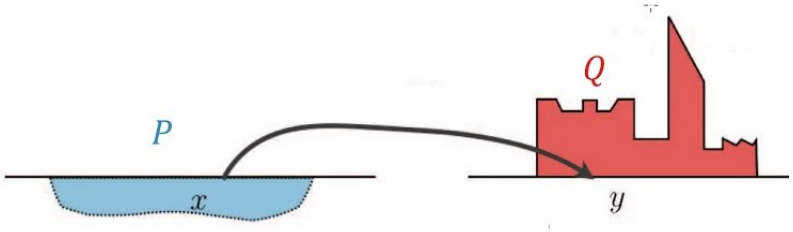

FIGURE 6.3.1. Transport of a mass from one place to another.

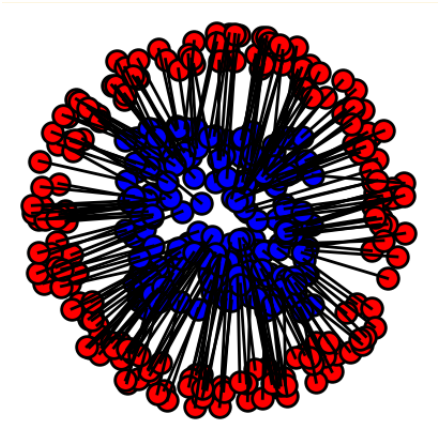

FIGURE 6.3.2. Transport between empirical measures, uniform on an annulus (red) and uniform on a square (blue).

Kantorovich primal and dual formulation. Let $\mathcal{X}$ and $\mathcal{Y}$ be two reference Polish spaces with measures $\alpha \in \mathcal{M}_1(\mathcal{X}), \beta \in \mathcal{M}_1(\mathcal{Y})$ respectively. Let $\mathcal{C} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ be the cost function.

The primal formulation is

$$W^C(\alpha, \beta) := \min_{\pi \in \Pi(\alpha,\beta)} \int_{\mathcal{X}\times\mathcal{Y}} C(x, y)\pi(dxdy), \tag{6.3.1}$$

where $\Pi(\alpha,\beta) := \{\pi \in \mathcal{M}_1(\mathcal{X}\times\mathcal{Y}) : P_{\mathcal{X}\#}\pi = \alpha, P_{\mathcal{Y}\#}\pi = \beta\}$, $P_{\mathcal{X}\#}$ and $P_{\mathcal{Y}\#}$ are the push forwards w.r.t first and second marginals.

The dual formulation is

$$W^C(\alpha,\beta) = \sup_{(f,g)\in R(C)} \int_{\mathcal{X}} f(x)\alpha(dx) + \int_{\mathcal{Y}} g(y)\beta(dy) \tag{6.3.2}$$

where $R(C) := \{(f,g) \in \mathcal{X}\times\mathcal{Y} : f(x)+g(y) \le C(x,y)\}$. For an account of the equivalence between primal and dual formulation, under mild assumptions on the cost function $\mathcal{C}$, we refer to [Vil21]. The pair of functions (f,g) are commonly referred to as Kantorovich potentials.

Sinkhorn formulation or the Entropic regularization of OT.

The OT problems given in Eq. (6.3.1) and Eq. (6.3.2) are linear problems under convex contrains. For better properties, the idea is to add a convex penalization term weighted by a penalization parameter $\varepsilon > 0$, for example an entropy term via the Kullback-Leibler divergence. In this case, the primal problem becomes:

$$W_\varepsilon^C(\alpha,\beta) := \min_{\pi\in\Pi(\alpha,\beta)} \int_{\mathcal{X}\times\mathcal{Y}} C(x,y)\pi(dxdy) + \varepsilon KL\left(\pi\|\alpha\otimes\beta\right), \tag{6.3.3}$$

where $KL(\pi\|\alpha\otimes\beta) := \int_{\mathcal{X}\times\mathcal{Y}} \log\left(\frac{d\pi}{d\alpha\otimes\beta}(x,y)\right)\pi(dxdy)$.

The dual formulation is:

$$\begin{aligned} W_\varepsilon^C(\alpha,\beta) = \sup_{f\in\mathcal{L}_1(\alpha), g\in\mathcal{L}_1(\beta)} & \int_{\mathcal{X}} f(x)\alpha(dx) + \int_{\mathcal{Y}} g(y)\beta(dy) \\ & - \varepsilon \int_{\mathcal{X}\times\mathcal{Y}} \left(e^{\frac{f(x)+g(y)-C(x,y)}{\varepsilon}}\alpha(dx)\beta(dy) - 1\right) . \end{aligned} \tag{6.3.4}$$

For fixed (small) $\varepsilon > 0$, these are respectively a strongly convex minimization problem and a strongly concave maximization problem.

6.3.2. Spectral and numerical aspects. All the illustrations in this Subsection will use the dataset of Fig. 6.3.2: 400 points sampled uniformly on the square and 500 points sampled on an annulus.

In order to numerically solve the OT problems (6.3.3) and (6.3.4), one needs to consider discrete spaces $\mathcal{X} = \{x_1, x_2, \dots, x_m\}$, $\mathcal{Y} = \{y_1, y_2, \dots, y_n\}$. In this setting, the cost function comes a cost matrix

$$\mathcal{C} = (\mathcal{C}(x_i,y_j) = C_{i,j})_{\substack{1\le i\le m\\ 1\le j\le n}} .$$

Also, the two problems become finite dimensional optimization problems. Notice that the convex set of couplings $\Pi(\alpha,\beta)$ is identified to a subspace of the matrices $\mathbb{R}^{n\times m}$, while the Kantorovich potentials f and g are respectively identified to vectors in $\mathbb{R}^n$ and $\mathbb{R}^m$. In this form, it is clear that one should prefer the dual formulation to the primal one for dimensionality reason ($n+m \le nm$ for all positive integers m,n).

The objective function for the dual formulation in this discrete setup is

$$\begin{aligned} Q(f,g) = \langle f,\alpha\rangle + \langle g,\beta\rangle - \varepsilon\left(\sum_{i,j} e^{\frac{f_i+g_j-C_{ij}}{\varepsilon}} - 1\right), \\ f\in\mathbb{R}^m,\ g\in\mathbb{R}^n, \alpha\in\mathcal{M}_+(\mathcal{X}) \cong \mathbb{R}_+^m,\ \beta\in\mathcal{M}_+(\mathcal{Y}) \cong \mathbb{R}_+^n . \end{aligned}$$

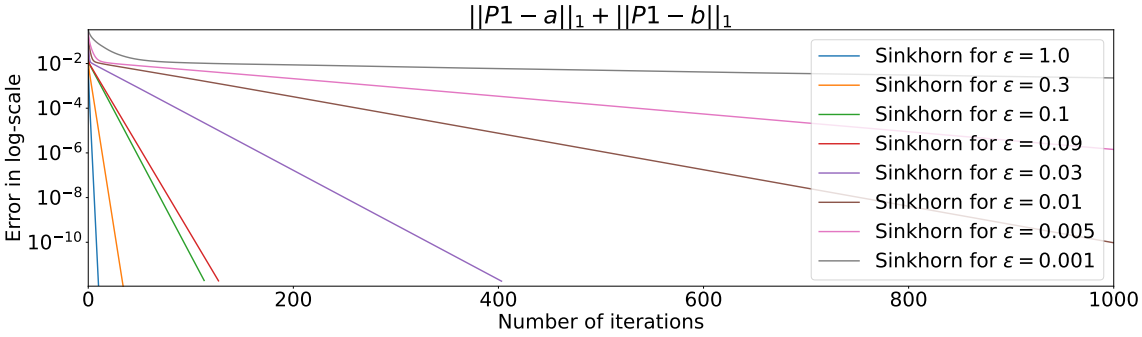

FIGURE 6.3.3. Convergence quality of the Sinkhorn algorithm for varying $\varepsilon > 0$.

Then the Hessian at the optimum potentials (f^*, g^*) is given by:

$$\nabla^2 Q(f^*, g^*) = -\frac{1}{\varepsilon} \begin{pmatrix} \Delta(\alpha) & \pi_\varepsilon \\ \pi_\varepsilon^T & \Delta(\beta) \end{pmatrix},$$

where $(\pi_\varepsilon)_{ij} = e^{\frac{f_i^* + g_j^* - C_{ij}}{\varepsilon}}$, $\Delta = \mathrm{diag} : \mathbb{R}^n \to \mathcal{M}_n(\mathbb{R})$ and $\varepsilon > 0$.

6.3.2.1. *The Problem.* The Sinkhorn algorithm, aka the proportional fitting algorithm, is the go to method for computing the solution, in the dual formulation 6.3.4. We recommend Cuturi [**Cut13**] for an illustration motivated by machine learning, and the book by Cuturi-Peyré [**PC⁺19**] for an extensive review.

It is a fixed-point algorithm with a contraction coefficient $q^* = q^*(\varepsilon)$. Following Vialard [**Via19**, Proposition 19], there is a constant $\kappa > 0$, depending only on the measures and the cost, such that for all $\varepsilon > 0$,

$$q^*(\varepsilon) \leq 1 - \exp\left(-\frac{\kappa}{\varepsilon}\right) ,$$

and actually this bound seems sharp from numerical experiments – see Fig. 6.3.3. One can also try implementing other algorithms, e.g. simple gradient ascent without line search, gradient ascent with line search and Armijo condition and L-BGFS, see Fig. 6.3.6.

In the end, all these algorithms display poor convergence when ε becomes small. And that is the problem we aim at adressing.

6.3.2.2. *A partial solution in the form of damped Newton.* In [**MT21**, **KMT19**], Kitagawa, Merigot and Thibert advocate for the use of damped Newton in a semi-discrete context, meaning that one of the measures is continuous. We propose to follow their lead, while focusing on the numerical stability and performance of the algorithm.

Damped Newton. Let us recall how the algorithm works. Denote Q as the objective function in the dual formulation of the regularized OT problem. The algorithm aims at finding the unique $x^* = (f^*, g^*)$ such that $\nabla Q(x^*) = 0$.

- At iteration k, let the current point be $x_k = (f_k, g_k) \in \mathbb{R}^{n+m}$.
- Find the Newton ascent direction $p_k \in \mathbb{R}^{n+m}$: If the current point is x_k, p_k is obtained by solving the linear system

$$\nabla^2 Q(x_k) p_k = \nabla Q(x_k) \ . \tag{6.3.5}$$

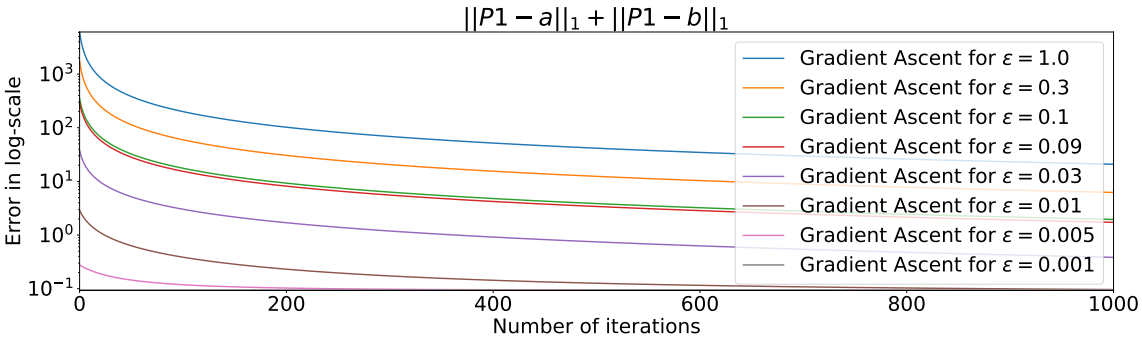

FIGURE 6.3.4. Gradient ascent without line search.

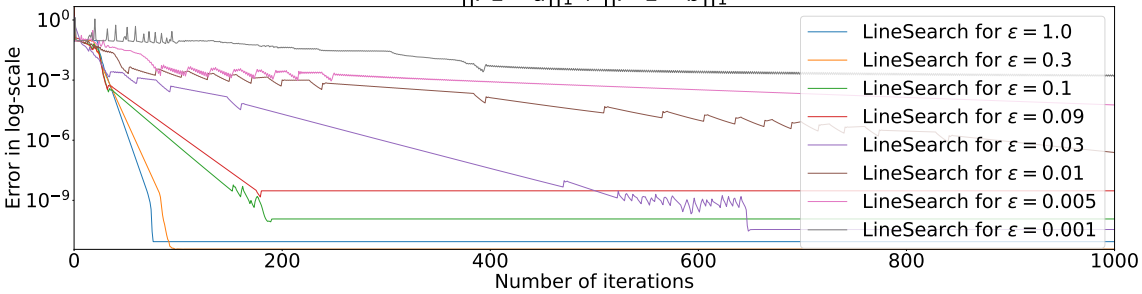

FIGURE 6.3.5. Gradient ascent with line search with learning rate= 0.001.

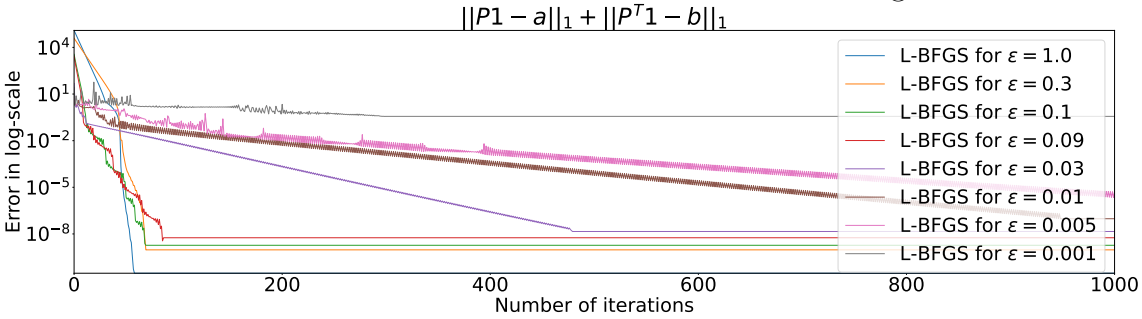

FIGURE 6.3.6. L-BFGS.

- Line search: For parameters $\rho \in (0,1)$, $c \in (0,1)$, we set an initial step size $\alpha_k = 1$. If p_k does not satisfy the Armijo condition:

$$Q(x_k + \alpha_k p_k) \geq Q(x_k) + c\alpha_k \langle p_k, \nabla Q(x_k) \rangle \ , \tag{6.3.6}$$

  we update the step size $\alpha_k \leftarrow \rho \alpha_k$ until condition is met.
- When the condition is met, we update the current point:

$$x_{k+1} \leftarrow x_k + \alpha_k p_k \ .$$

Why going for the damped Newton algorithm? The following pros are illustrated in Fig. (6.3.7).

- Damped Newton is one of the preferred algorithm for solving convex optimization problems due to less iterations for convergence.
- Damped Newton is rather insensitive to ε.

And there is one major inconvenience: The inversion of the Hessian in Eq. (6.3.5) is computationally very expensive. As such each step can be costly, especially if the

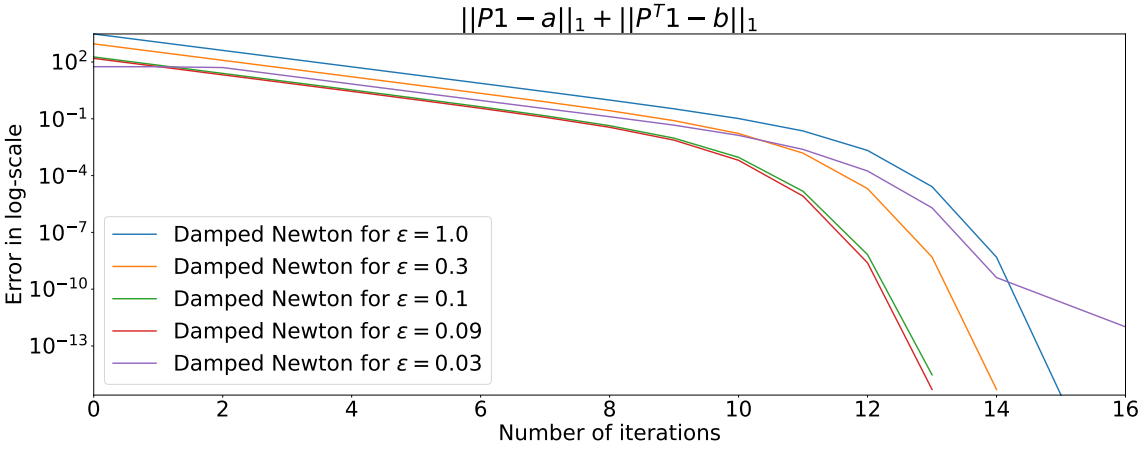

FIGURE 6.3.7. Convergence speed of damped Newton. This illustration shows that the algorithm converges with much fewer steps, in a way that is insensitive to $\varepsilon > 0$.

Hessian is inverted exactly. As such, it is natural to invoke approximate methods rather than exact solvers of linear systems. The inversion of the Hessian can be done faster using iterative solvers, for example CG (Conjugate Gradient) and GMRES (Generalized Minimal Residuals) – see [NW99]. To do so, we will need a better understanding of the structure of the Hessian $\nabla^2 Q(x_k)$.

6.3.2.3. *Theoretical and empirical analysis of the Hessian.* Let us have a look at the Hessian of the objective function. The following proposition gives us structural information about it.

PROPOSITION 6.3.1 (Work in progress, as part of the PhD thesis of Anirban BOSE). *Let us consider the setting of general measures $\alpha \in \mathcal{M}_1(\mathcal{X})$ and $\beta \in \mathcal{M}_1(\mathcal{Y})$. Define the Banach spaces of continuous bounded functions $E := \mathcal{C}_b(\mathcal{X})$ and $F := \mathcal{C}_b(\mathcal{Y})$. At the optimum potentials $x^* = (f^*, g^*)$, the Hessian can be written as*

$$\nabla^2 Q(f^*, g^*) = -\frac{1}{\varepsilon}(I + K) \tag{6.3.7}$$

where $I \in \mathcal{L}(E, E)$ is the identity operator and K is a compact operator with symmetric spectrum in $[-1, 1]$.

Here K is a block form operator given by

$$K = \begin{pmatrix} 0 & \pi_\sigma \\ \pi_\sigma^T & 0 \end{pmatrix},$$

where for $(f_1, g_1) \in E \oplus F$, $\pi_\sigma \in \mathcal{L}(F, E)$ and $\pi_\sigma^T \in \mathcal{L}(E, F)$ are given by

$$(\pi_\sigma g_1)(x) = \int_{\mathbb{R}^d} e^{\frac{f^*(x)+g^*(y)-C(x,y)}{2\sigma^2}} g_1(y)\beta(dy)$$

$$(\pi_\sigma^T f_1)(y) = \int_{\mathbb{R}^d} e^{\frac{f^*(x)+g^*(y)-C(x,y)}{2\sigma^2}} f_1(x)\alpha(dx)$$

and $\langle \pi_\sigma^T f_1, g_1 \rangle_{\mathcal{L}^2(\beta)} = \langle f_1, \pi_\sigma g_1 \rangle_{\mathcal{L}^2(\alpha)}$.

Interestingly, the previous Proposition 6.3.1 tells us that the normalized Hessian $-\varepsilon \nabla Q(x^*) = I + K$ has spectrum in $[0, 2]$. Given that the stability of inversion

of linear systems such as Eq. (6.3.5) is given by the condition number, we now understand that the spectral gap of $I + K$, at zero, is at the heart of the problem.

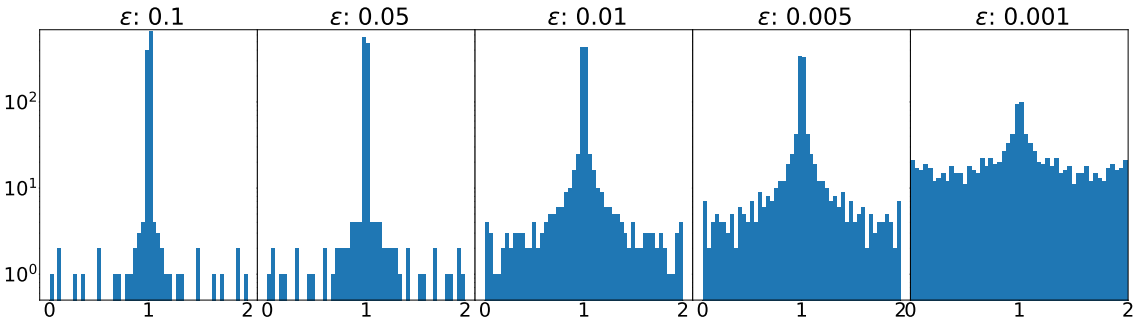

FIGURE 6.3.8. Dispersion of the spectrum of the normalized Hessian $-\varepsilon\nabla Q(x^*) = (I + K)$ as $\varepsilon \to 0$. This causes numerical instability in performing the inversion of the Hessian.

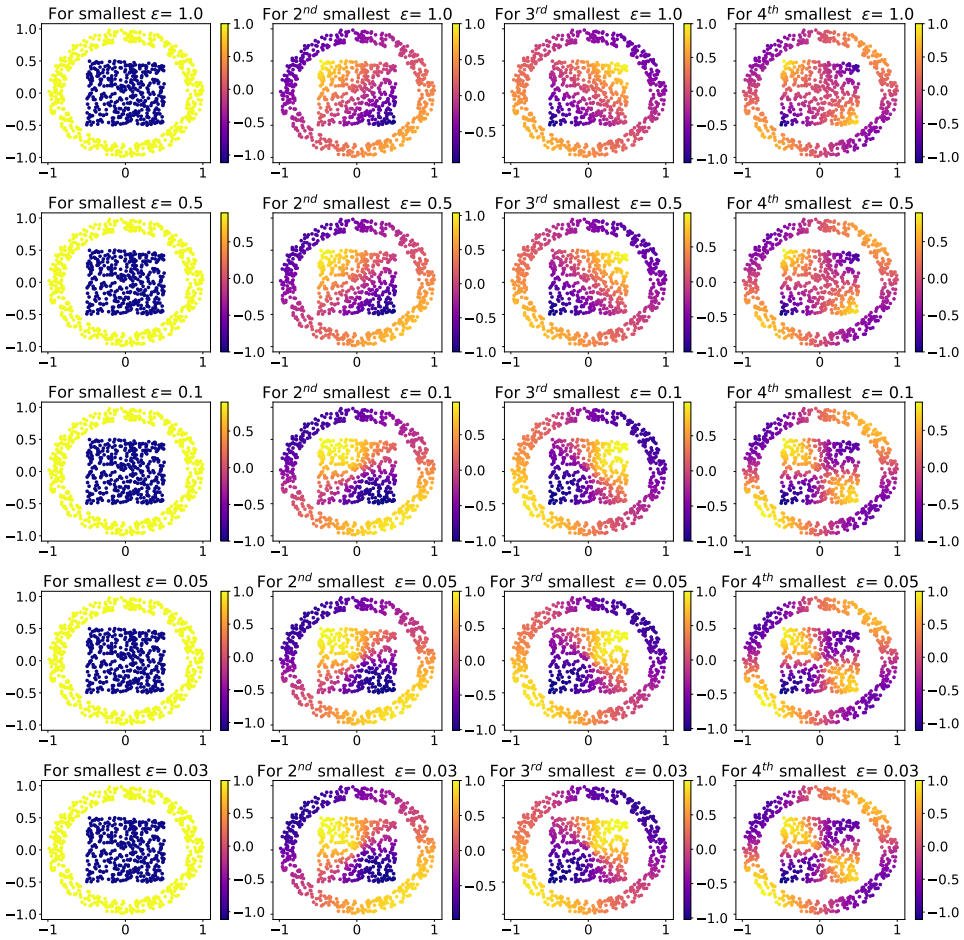

FIGURE 6.3.9. Heatmap of eigenvectors

This is corroborated numerically in Fig. 6.3.8, where one sees the spectral gap at zero closing up as $\varepsilon \to 0$.

6.3.2.4. *The idea: Leveraging the stability of eigenvectors.* As not much can be gained from the eigenvalues, the next idea which proves to be fruitful is to examine eigenvectors.

CONJECTURE 6.3.2 (Verified empirically). *As $\varepsilon \to 0$:*

- *The spectrum is unstable as shown in Fig. 6.3.8.*
- *The eigenvectors on the other hand are stable as shown in Fig. 6.3.9.*

This phenomenon is very reminiscent of the behavior of eigenvalues of Schrödinger operators at low temperature [Mic95]. And it can be leveraged as follows:

- Use eigenvectors of the Hessian obtained from execution of the algorithm for larger values ε to form a preconditioning matrix.
- Use preconditioning to move eigenvalues that are dangerously close to the boundaries of [0,2] to the center and consequently reduce the instability and cost of iterative inversion.

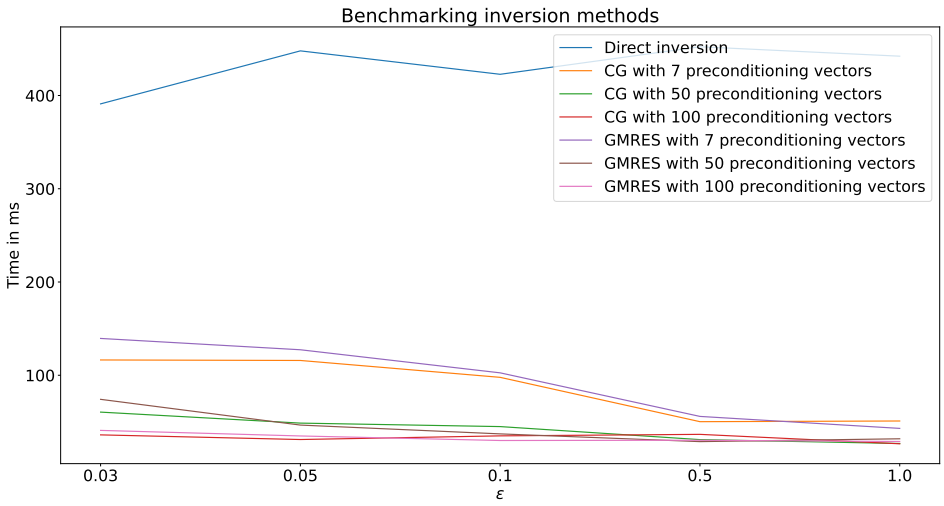

FIGURE 6.3.10. Benchmarking inversions methods after preconditioning. Illustration of time in milliseconds to invert matrix using different inversion methods with a data with square of size 1000 and annulus of size 1100.

But what is preconditioning? Preconditioning is the idea of solving an equivalent for more numerically stable linear system. Let $A \in \mathcal{M}_{m,n}(\mathbb{R})$ and $b \in \mathbb{R}^m$. Consider the linear system in $x \in \mathbb{R}^n$

$$Ax = b \ .$$

Let $Q \in GL(m, \mathbb{R})$ and $P \in GL(n, \mathbb{R})$ be matrices such that we can write

$$Ax = b \iff QAP^{-1}(Px) = Qb \iff A^{'}x^{'} = b^{'} \ .$$

Hence one can change the linear system of equations so that the iterative methods for inversion converge faster. Fig. 6.3.10 illustrates that effect.

6.3.2.5. *Conjecture 6.3.2 is true in the Gaussian case.* As a first step towards understanding the phenomenon described in Conjecture 6.3.2, it is natural to attempt a proof in a particular case. Here, we consider α and β as Gaussian measures on $\mathbb{R}^d$. Also the cost is taken as the square of the Euclidean norm. We write

$$\alpha = \mathcal{N}(a, A), \qquad \beta = \mathcal{N}(b, B) \ ,$$

so that a and b are the means, while A and B are positive definite matrices. Let us give the following theorem which an explicit diagonalization of the operator K using this data. This diagonalization is in term of the normalized Hermite (multivatiate) polynomials

$$\forall x \in \mathbb{R}^d, \forall \mathbf{n} \in \mathbb{N}^d, \ he_{\mathbf{n}}(x) := \prod_i he_{n_i}(x_i) \ ,$$

which are themselves products of the univariate normalized Hermite polynomials he_n.

THEOREM 6.3.3 (Conjecture 6.3.2 is true in the Gaussian case). *Consider the kernel defining the operator K given by*

$$k(x, y) = \exp\left(\frac{f^*(x) + g^*(y) - \|x - y\|^2}{\varepsilon}\right) \ .$$

This kernel has a series expansion:

$$k(x, y) = \sum_{\mathbf{n} \in \mathbb{N}^{\mathbf{d}}} \lambda_{\mathbf{n}}(\varepsilon) P_{\mathbf{n}}(x) Q_{\mathbf{n}}(y) \ ,$$

where $A^{\frac{1}{2}} B^{\frac{1}{2}} = UDV^T$ the SVD decomposition of $A^{\frac{1}{2}} B^{\frac{1}{2}}$, $\varepsilon = 2\sigma^2$ and

$$\lambda_{\mathbf{n}}(\varepsilon) = \prod_{i=1}^{d} \rho_i^{\mathbf{n_i}} \ , \qquad \rho_i = \left(\left(1 + \frac{\sigma^4}{4D_{ii}^2}\right)^{\frac{1}{2}} - \frac{\sigma^2}{2D_{ii}}\right) \ ,$$
$$P_{\mathbf{n}}(x) = he_{\mathbf{n}}(U^T A^{-\frac{1}{2}}(x - a)) \ , \qquad Q_{\mathbf{n}}(y) = he_{\mathbf{n}}(V^T B^{-\frac{1}{2}}(y - a)) \ .$$

As a consequence the operator K has eigenvalues $\left\{\pm\lambda_{\mathbf{n}}(\varepsilon)^2 \ ; \mathbf{n} \in \mathbb{N}^d\right\}$ and eigenvectors that do not depend on $\varepsilon > 0$.

In particular, the spectrum of K is unstable converging to ± 1, while the eigenvectors of K are stable – in fact constant in this case.

Let us say a few words on the proof. The entropic regularized optimal transport problem (6.3.4) as in fact closed-form solutions detailed in a paper by Janati et al. [JMPC20]. The Kantorovich potentials (f^*, g^*) happen to be quadratic forms, so that k is a Gaussian kernel. The series is then performed by adapting the Mehler kernel expansion formula.

6.3.2.6. *Openings.* So far, we presented a theoretical and empirical analysis of the Hessian of Q. We propose to leverage the stability of eigenvectors to accelerate the damped Newton algorithm.

Among the work that remains, we need to

- Incorporate log-domain computations ([PC+19, § 4.4]) that are aimed at stabilizing the algorithm. Indeed, such numerical instabilities are inevitable

when computing sums of the form

$$\sum_i \exp\left(\frac{a_i}{\varepsilon}\right)$$

for small ε.

- Properly benchmark against Sinkhorn and other algorithms for very small ε, which also need stabilization.

6.4. Statistics and Learning Theory

6.4.1. Deep learning with Free Probability. The stability of a neural network is driven by the spectrum of singular values of the Jacobian. By the Jacobian, we really mean the matrix of differentials with respect to all the neural networks' weights. In the line of Pennington et al.'s work [**PSG18**, **BKP+20**], the natural tool is Free Probability Theory (FPT) in order to describe these spectra at the initialization of the training. Indeed FPT is the machinery to theoretically compute the law of large numbers for the spectra of large random matrices.

The issue is that computational methods were lacking for an abstract topic such as free probability. In fact, it was even considered that numerically computing free convolutions is very difficult. Thus in [**CDK22**], we propose a computational solution to the inversion of Voiculescu's S-transforms. Our method is based on the chaining of basins of attraction for the Newton-Raphson algorithm. Not only is the result guaranteed to be correct – unlike the solution previously proposed by Pennington et al., but the method is also very fast.

The contribution of [**CDK22**] this paper are:

- theoretical: as we extend the framework of Free Probability to the necessary rectangular operators.
- numerical: we provide a computational solution to the inversion of Voiculescu's S-transforms. Our method is based on the chaining of basins of attraction for the Newton-Raphson algorithm. Not only is the result guaranteed to be correct – unlike the solution previously proposed by Pennington et al., but the method is also very fast.
- empirical: we prove that FPT indicators correlate to the accuracy of real life neural networks, after training

In the future, we want to generalize this work to neural networks with more general architectures. In order to incorporate skip-connections, it seems that we will need to come up with a similar solution for operator-valued FPT.

6.4.2. High dimensional statistics with free deconvolution. Currently, in collaboration with Gamboa, Kammoun and Velasco, we are revisiting the ideas of computational FPT from a statistical perspective. A preliminary draft of this work is available in [**CGKV23**]. More precisely, we are interested in the estimation of spectra of covariance matrices in high dimension. Again, the natural framework is free probability thanks to free deconvolution. The importance of that problem stems from the fact that *all* basic statistical procedures such as Principal Component Analysis (PCA) start with the estimation of the spectrum, which is blurred by a high dimensional phenomenon well-known and well-studied in Random Matrix Theory.

6.4.3. Deep learning with Riemannian Geometry. A common technique in machine learning is dimension reduction. If PCA is the archetype, it is common nowadays to train a neural network to create a parsimonious representation of a dataset. The network architecture of autoencoders is basically a non-linear method for dimension reduction.

Latent spaces and autoencoders : Consider an autoencoder architecture as in Fig. 6.4.1 that is to say an approximate factorization of the identity through a space M of small dimension:

$$\Phi = \Phi_{\theta_e} : \mathbb{R}^D \to M, \quad \Psi = \Psi_{\theta_d} : M \to \mathbb{R}^D$$

such that for all the data samples $X_1, X_2, \dots$ we have

$$\forall i,\ \hat{X}_i = \Psi \circ \Phi(X_i) \approx X_i \ .$$

The map Φ is called an encoder, while the map Ψ is called a decoder. The underlying parameters of the neural network are denoted θ_e and θ_d. These are nothing but very large real tensors. As such, training is formulated as the usual minimizing problem

$$\mathrm{Argmin}_{(\theta_e,\theta_d)} \mathcal{L}(\theta_e, \theta_d)$$

during training, we minimize the loss using a distance d on $\mathbb{R}^D$:

$$\mathcal{L} = \mathcal{L}(\theta_e, \theta_d) = \sum_i d(X_i, \hat{X}_i) \ .$$

In typical vision applications, the input space is images and as such D can be $32 \times 32 = 1024$. In genomics, the input is a space of gene expressions with $D \approx 2000$. The latent space is of low dimension as small as 2 but almost never larger than a 100. The middle manifold M is called a latent space, since one the decoder can be used in inference to generate images.

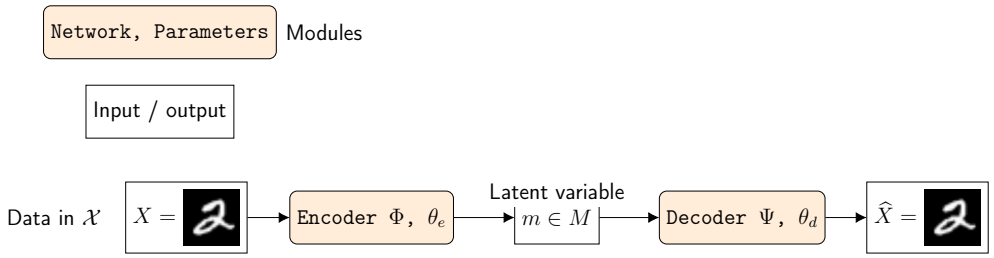

FIGURE 6.4.1. Network architecture of an autoencoder.

In Fig. 6.4.2, one can see how the dataset is encoded into the latent space. And Fig. 6.4.3 gives the so-called "manifold plot" which amounts to choosing a grid in the latent space and generating the corresponding images thanks to the decoder. The reader can indeed check that both images echo each other.

Natural measure and natural metric: The data space $\mathbb{R}^D$ is endowed with a natural metric, which can be taken as the Euclidean $\|\cdot\|^2$ when thinking about images. By pullback, the latent space is endowed with the metric

$$g := \Phi^* \|\cdot\|^2 \ .$$

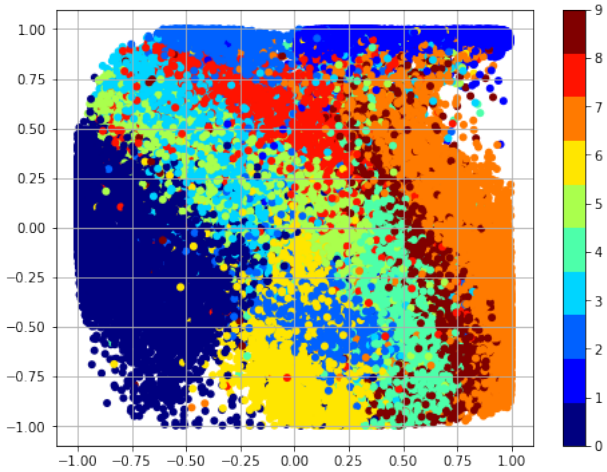

FIGURE 6.4.2. Latent space for an AE on the classical MNIST dataset.

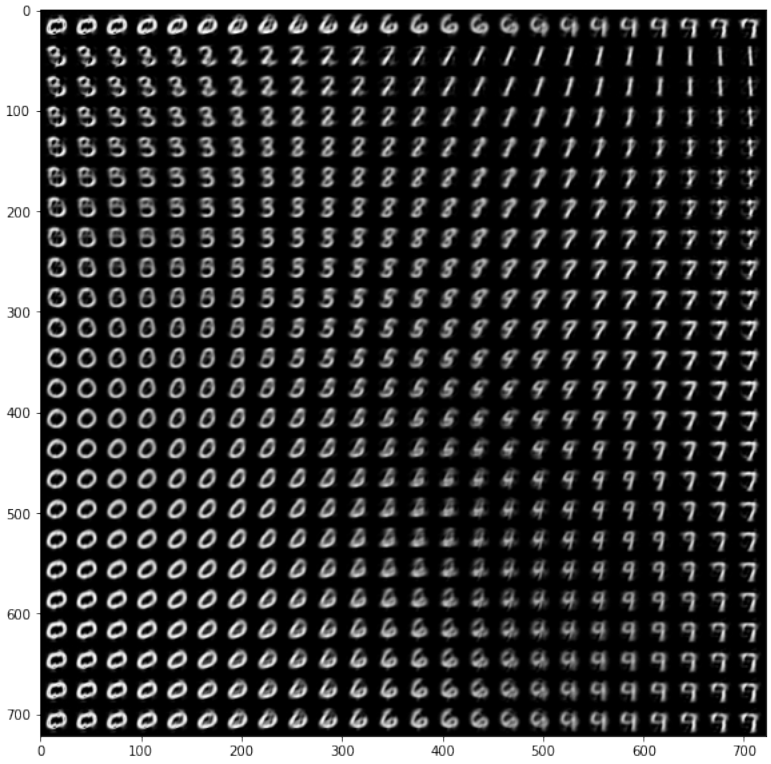

FIGURE 6.4.3. Manifold plot for AE.

Moreover, the natural measure induced by the dataset is clearly

$$\mu := \frac{1}{n}\sum_{i=1}^{n} \delta_{\Phi(X_i)} \ .$$

A first contribution: In [DCTV20], we study the manifold (M, g) and formalize a machine learning problem using Riemannian geometry. In this work, we formalize the creation of correspondences as a computation of geodesics between fibers of a fibered manifold $M = B \times F$. Here B is a base space where embeddings

for classes are. F is a fiber space encoding the variability of MNIST images within each class. For more details see [DCTV20].

Then, we show that the problem is generically well-posed because of the geometric properties of the cut-locus. Finally, we give a modern computational technique that takes advantage of frameworks like PyTorch or Tensorflow.

However, we did notice that the Riemannian manifold (M, g) can be quite disorganised. For example, the geodesics can be very wild. This is illustrated in Fig. 6.4.5 and 6.4.6. There the fibers F_i corresponds to the fiber associated to the class $i \in \{0, 1, \dots, 9\}$, which correspond to the digits of MNIST.

Metric regularization via Ricci-type flows: As continuation of that work, there is now the PhD thesis of Alexey LAZAREV, jointly supervised with Francesco COSTANTINO.

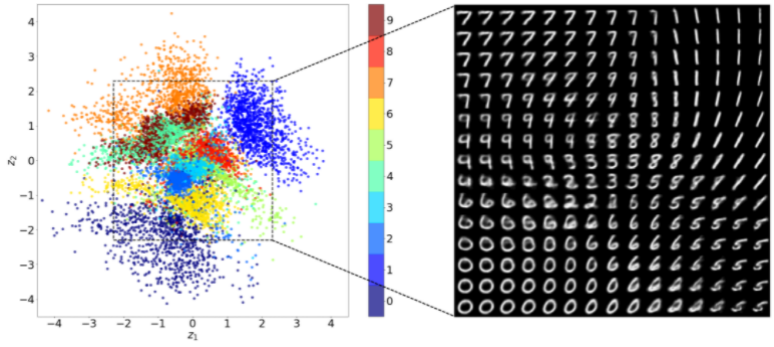

FIGURE 6.4.4. Manifold plot for VAE. The reference measure on the latent space is forced to be standard Gaussian in two dimensions.

Second, how to regularize the metric? Is there a natural entropy to minimize? As the question has been formulated, it is natural to think about the only functional which plays the role of entropy in the realm of Riemannian geometry: Perelman's entropy whose gradient flow is the Ricci flow. Its stationary points are usually of constant curvature and thus very simple model spaces.

Therefore the idea is to launch the Ricci flow on the metric g. One can surmise that the finite difference PDE evolution will be very poorly behaved. Machine learning is powered by automatic differentiation and will prefer the gradient flow formulation. The fact that the Ricci flow can be formulated as the gradient flow on a functional is the main statement of Perelman's paper [Per02]. In fact, there are multiple functionals and we are working on simplified versions.

6.4.4. Kernel learning with RMT. Consider the classical regression problem where we want to estimate a function $f : \mathbb{R}^d \to \mathbb{R}$ from observed data. The dataset is composed of n observables $(x_1, y_1), \dots, (x_n, y_n)$ where $y_j = f(x_j)$.

Kernel regression. A classical approach is to look for an estimator $\hat{f}_n \in H$ with H a Reproducing Kernel Hilbert Space (RKHS) of functions from $\mathbb{R}^d$ to $\mathbb{R}$. Our setup will be through the regularized Kernel regression problem

$$\hat{f}_n := \operatorname{argmin}_{h \in H} \sum_{i=1}^{n} \ell\left(y_i, f(x_i)\right) + \lambda \Omega(\|f\|_H) \ . \tag{6.4.1}$$

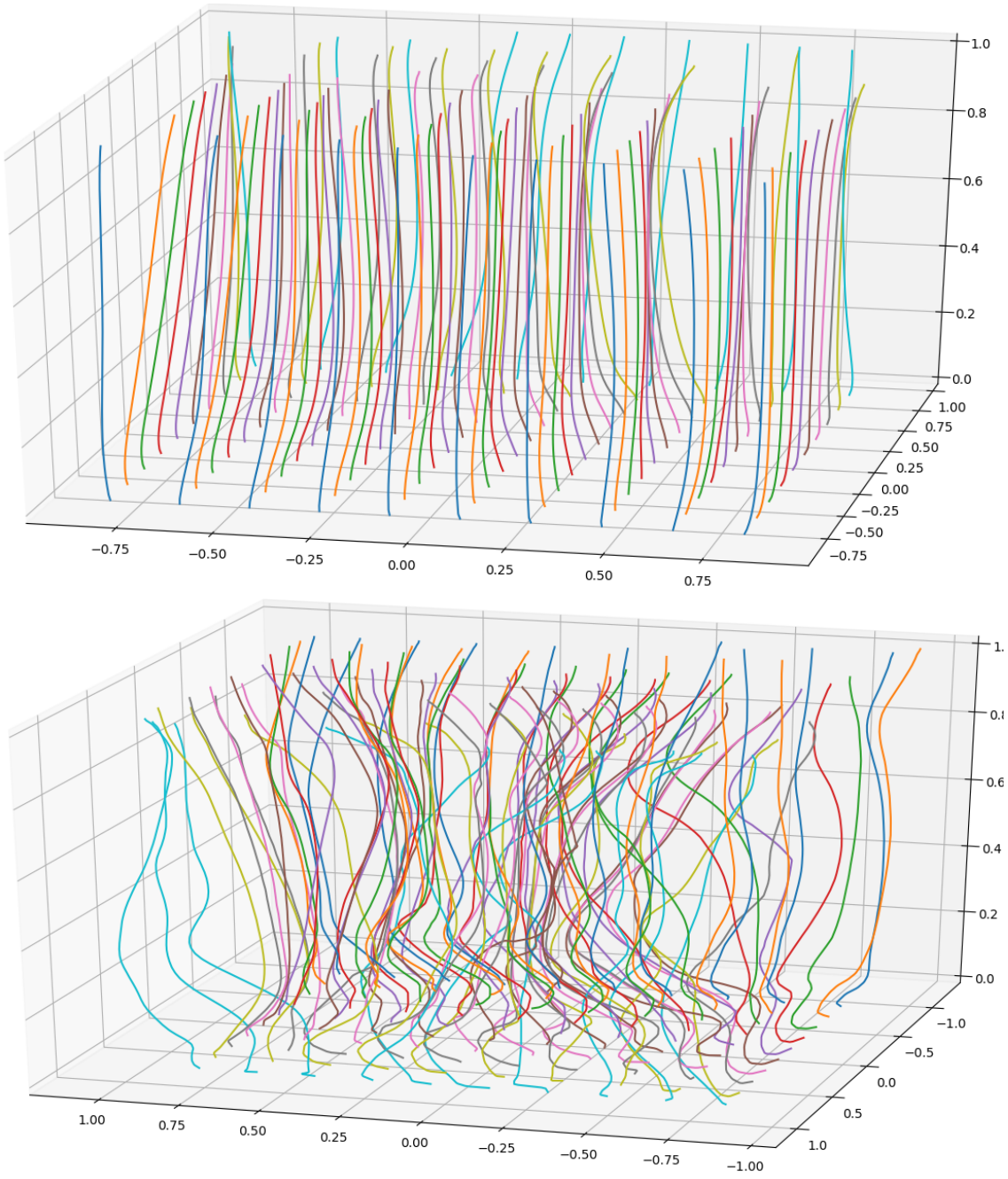

FIGURE 6.4.5. Visualization of geodesic curves between two fibers isomorphic to $[-1,1]^2$. Top: From F_4 to F_9. Bottom: From F_1 to F_0. With z being the height coordinate, the $z = 0$ plane represents the starting fiber, while the $z = 1$ plane represents the destination fiber.

Here ℓ is a loss function and Ω is a regularization function.

In order to obtain a tractable formulation, we restrict to Mean Square Error (MSE) for loss and Ridge regularization, meaning:

$$\hat{f}_n := \operatorname{argmin}_{h \in H} \|\mathbb{Y} - f(\mathbb{X})\|^2_{\mathbb{R}^n} + \lambda \|f\|^2_H \ . \tag{6.4.2}$$

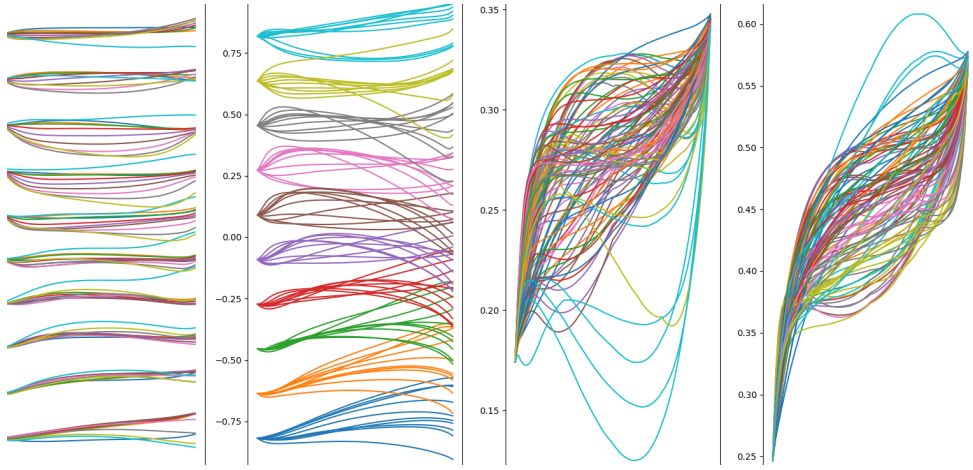

FIGURE 6.4.6. Geodesics from F_4 to F_9. The x axis displays time $t \in [0,1]$. The two left panels display coordinates in fiber space $F = [-1,1]^2$ and the two right panels display coordinates on $B = \mathbb{R}^2$.

Here $\mathbb{Y} \in \mathbb{R}^n$ is the vector of outputs. $\mathbb{X} \in \mathbb{R}^{d \times n}$ is the matrix of inputs. The map f is understood as acting pointwise on columns.

Recall that the RKHS representation theorem states that for all functions $\varphi \in H$:

$$\forall x \in \mathbb{R}^d,\ \varphi(x) = \langle \varphi, k(x,\cdot)\rangle_H\ ,$$

where k is the kernel. In particular, we have $k(x_i,x_j) = \langle k(x_i,\cdot), k(x_j,\cdot)\rangle_H$ and we write $\mathbb{K} := (k(x_i,x_j))_{1\leq i,j\leq n}$. By looking for f as a linear combination of kernel functions, we assume:

$$f = \sum_{j=1}^n \alpha_j k(x_j,\cdot)\ .$$

As such, the regression problem boils down to estimating a finite dimensional parameter α via:

$$\hat{\alpha}_n := \operatorname{argmin}_{\alpha\in\mathbb{R}^n} \|\mathbb{Y} - \mathbb{K}\alpha\|^2_{\mathbb{R}^n} + \lambda\langle \alpha, \mathbb{K}\alpha\rangle_{\mathbb{R}^n} \tag{6.4.3}$$

$$= (\lambda I_n + \mathbb{K})^{-1}\,\mathbb{Y}. \tag{6.4.4}$$

Features learning. Postulating the existence of a matrix factorization $\mathbb{K} = \Phi(\mathbb{X})^T\Phi(\mathbb{X})$ for a certain $\Phi(\mathbb{X}) \in \mathbb{R}^{p\times n}$ is equivalent to the existence of a feature map $\Phi : \mathbb{R}^d \to \mathbb{R}^p$. The data $\Phi(x_i) \in \mathbb{R}^p$ is called the feature associated to x_i , p is the dimension of features and $\Phi(\mathbb{X}) = [\Phi(x_1),\dots,\Phi(x_n)]$.

ASSUMPTION 6.4.1. *The natural setup of interest is:*

$$d \ll p \ll n\ .$$

That is to say that the number of samples n is much larger than the dimension of features p, which is itself much larger than the dimension of inputs d.

By setting $w = \Phi(\mathbb{X})\alpha$, the optimization problem of interest takes the form:

$$\hat{w}_n := \operatorname{argmin}_{w\in\mathbb{R}^p} \|\mathbb{Y} - \Phi(\mathbb{X})^T w\|^2_{\mathbb{R}^n} + \lambda\|w\|^2_{\mathbb{R}^p} \tag{6.4.5}$$

$$= \operatorname{argmin}_{w\in\mathbb{R}^p} \langle w, \left(\lambda I_p + \Phi(\mathbb{X})\Phi(\mathbb{X})^T\right) w\rangle_{\mathbb{R}^p} - 2\langle \Phi(\mathbb{X})\mathbb{Y}, w\rangle_{\mathbb{R}^p} \tag{6.4.6}$$

$$= \left(\lambda I_p + \Phi(\mathbb{X})\Phi(\mathbb{X})^T\right)^{-1}\Phi(\mathbb{X})\mathbb{Y}. \tag{6.4.7}$$

Notice that the equality $\hat{w}_n = \Phi(\mathbb{X})\hat{\alpha}_n$ is equivalent to

$$\left(\lambda I_p + \Phi(\mathbb{X})\Phi(\mathbb{X})^T\right)^{-1}\Phi(\mathbb{X})\mathbb{Y} = \Phi(\mathbb{X})\left(\lambda I_n + \Phi(\mathbb{X})^T\Phi(\mathbb{X})\right)^{-1}\mathbb{Y}\ ,$$

which is an instance of the Woodbury-Sherman-Morrison identity. Equating the solutions of the two problems is basically a variational proof of that fact.

EXAMPLE 6.4.2 (Random Fourier Features).

$$\Phi_{RFF}(x) := \begin{pmatrix} \cos\left(\langle W_1, x\rangle_{\mathbb{R}^d}\right) \\ \cdots \\ \cos\left(\langle W_p, x\rangle_{\mathbb{R}^d}\right) \\ \sin\left(\langle W_1, x\rangle_{\mathbb{R}^d}\right) \\ \cdots \\ \sin\left(\langle W_p, x\rangle_{\mathbb{R}^d}\right) \end{pmatrix}$$

with $W \in \mathbb{R}^{d\times p}$ is a random matrix and the W_j's are the columns of W. In the RFF model, entries of W are all i.i.d. Gaussian $\mathcal{N}(0,\sigma^2)$. Average kernel and variances:

$$\mathbb{E}_W\left[\langle\Phi_{RFF}(x),\Phi_{RFF}(y)\rangle\right] = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right)\ ,$$

$$\mathrm{Var}_W\left[\langle\Phi_{RFF}(x),\Phi_{RFF}(y)\rangle\right] = \frac{1}{p}\left(1-\exp\left(-\frac{\|x-y\|^2}{\sigma^2}\right)\right)^2\ .$$

Here the computation complexity for computing $\hat{w}_n$ is decomposed as follows:

- Creating $\Phi(\mathbb{X})$: $\mathcal{O}(npd)$.
- Forming $\Phi(\mathbb{X})\Phi(\mathbb{X})^T$: $\mathcal{O}(np^2)$, which is however not needed if iterative inversion.
- Inversion $\mathcal{O}(p^3)$.

Fast features. In the context of random Fourier features, and if evaluation algorithms only make use of matrix-vector multiplication, an excellent idea is to use matrices W which are structured.

The full model is for example the Gaussian RFF, where one takes

$$W = (W_{i,j})_{\substack{1\le i\le d\\ 1\le j\le p}} \text{ with } W_{i,j} \overset{\mathcal{L}}{=} \mathcal{N}(0,\sigma^2)\ .$$

We can use the SVD to decompose this matrix as:

$$W = O_1\sqrt{\begin{pmatrix} D & 0_{d\times(p-d)}\end{pmatrix}}O_2\ ,$$

with $O_1 \in O_d(\mathbb{R})$, $O_2 \in O_p(\mathbb{R})$ are orthogonal and $D \in \mathbb{R}^{d\times d}$ is a diagonal matrix.

Recall that classically in random matrix theoy (RMT), we know that O_1 and O_2 are Haar distributed, while D has entries distributed according to the real Wishart (or Laguerre) Ensemble. This gives multiple ways to approximate this law via simpler matrices.

We can now formulate the question at hand, which is the topic of a collaboration with Hachem KADRI and Nizar DEMNI.

QUESTION 6.4.3. *By decomposing the matrix W as above, can we approximate the Gaussian RFF model by simpler models? A natural idea is to approximate:*

- *the Haar distributed O_1 and O_2 via only a few products in the virtual isometry decomposition of Chapter 2.*

- *the diagonal matrix D by deterministic equivalents of the Wishart distribution.*

These approximations should be to be faster to compute than the original model, while keeping the same statistical properties. In particular, this should not affect the kernel learning once the features are computed.

Bibliography

[ABB16] Louis Pierre Arguin, David Belius, and Paul Bourgade. Maximum of the characteristic polynomial of random unitary matrices. *Communications in Mathematical Physics*, pages 1–49, 9 2016.

[ABFJ16] Victor V Albert, Barry Bradlyn, Martin Fraas, and Liang Jiang. Geometry and response of lindbladians. *Physical Review X*, 6(4):041031, 2016.

[AGZ10] Greg W Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*. Number 118. Cambridge university press, 2010.

[Aïd13] Elie Aïdékon. Convergence in law of the minimum of a branching random walk. *Ann. Probab.*, 41(3A):1362–1426, 2013.

[AL07] Robert Alicki and Karl Lendi. *Quantum dynamical semigroups and applications*, volume 717 of *Lecture Notes in Physics*. Springer, Berlin, second edition, 2007.

[Ali76] Robert Alicki. On the detailed balance condition for non-hamiltonian systems. *Reports on Mathematical Physics*, 10(2):249–258, October 1976.

[APS18] Juhan Aru, Ellen Powell, and Avelio Sepúlveda. Critical Liouville measure as a limit of subcritical measures. *ArXiv preprint arXiv:1802.08433*, 2018.

[AS10] E. Aïdékon and Z. Shi. Weak convergence for the minimal position in a branching random walk: a simple proof. *Period. Math. Hungar.*, 61(1-2):43–54, 2010.

[Ass97] David Assaf. Estimating the state of a noisy continuous time markov chain when dynamic sampling is feasible. *The Annals of Applied Probability*, 7(3):822–836, 1997.

[Ata98] Rami Atar. Exponential stability for nonlinear filtering of diffusion processes in a non-compact domain. *Ann. Probab.*, 26(4):1552–1574, 1998.

[AZ97a] Rami Atar and Ofer Zeitouni. Exponential stability for nonlinear filtering. *Ann. Inst. H. Poincaré Probab. Statist.*, 33(6):697–725, 1997.

[AZ97b] Rami Atar and Ofer Zeitouni. Lyapunov exponents for finite state nonlinear filtering. *SIAM J. Control Optim.*, 35(1):36–55, 1997.

[AZ98] Rami Atar and Ofer Zeitouni. A note on the memory length of optimal nonlinear filters. *Systems Control Lett.*, 35(2):131–135, 1998.

[AZ14] Louis-Pierre Arguin and Olivier Zindy. Poisson-Dirichlet statistics for the extremes of a log-correlated Gaussian field. *Ann. Appl. Probab.*, 24(4):1446–1481, 2014.

[B$^+$17] Nathanaël Berestycki et al. An elementary approach to Gaussian multiplicative chaos. *Electronic Communications in Probability*, 22, 2017.

[Bal05] Roger Balian. Information in statistical physics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 36(2):323–353, 2005.

[Bar01] Yu Baryshnikov. Gues and queues. *Probability Theory and Related Fields*, 119:256–274, 2001.

[BB11] Michel Bauer and Denis Bernard. Convergence of repeated quantum nondemolition measurements and wave-function collapse. *Physical Review A*, 84(4):044103, 2011.

[BB14] Michel Bauer and Denis Bernard. Real time imaging of quantum and thermal fluctuations: the case of a two-level system. *Lett. Math. Phys.*, 104(6):707–729, 2014.

[BB18] Michel Bauer and Denis Bernard. Stochastic spikes and strong noise limits of stochastic differential equations. In *Annales Henri Poincaré*, volume 19, pages 653–693. Springer, 2018.

[BBB12] Michel Bauer, Denis Bernard, and Tristan Benoist. Iterated stochastic measurements. *J. Phys. A: Math. Theor.*, 45(49):494020, December 2012.

[BBB13] Michel Bauer, Tristan Benoist, and Denis Bernard. Repeated quantum non-demolition measurements: Convergence and continuous time limit. *Ann. H. Poincaré*, 14(4):639–679, May 2013.

[BBC+21] Tristan Benoist, Cédric Bernardin, Raphaël Chetrite, Reda Chhaibi, Joseph Najnudel, and Clément Pellegrini. Emergence of jumps in quantum trajectories via homogenization. *Communications in Mathematical Physics*, 387(3):1821–1867, 2021.

[BBO05] Philippe Biane, Philippe Bougerol, and Neil O'Connell. Littelmann paths and Brownian paths. *Duke Math. J.*, 130(1):127–167, 2005.

[BBO09] Philippe Biane, Philippe Bougerol, and Neil O'Connell. Continuous crystal and Duistermaat-Heckman measure for Coxeter groups. *Adv. Math.*, 221(5):1522–1583, 2009.

[BBT15a] Michel Bauer, Denis Bernard, and Antoine Tilloy. Computing the rates of measurement-induced quantum jumps. *Journal of Physics A: Mathematical and Theoretical*, 48(25):25FT02, 2015.

[BBT15b] Michel Bauer, Denis Bernard, and Antoine Tilloy. Computing the rates of measurement-induced quantum jumps. *J. Phys. A: Math. Theor.*, 48(25):25FT02, 2015.

[BBT16a] Michel Bauer, Denis Bernard, and Antoine Tilloy. Zooming in on quantum trajectories. *Journal of Physics A: Mathematical and Theoretical*, 49(10):10LT01, 2016.

[BBT16b] Michel Bauer, Denis Bernard, and Antoine Tilloy. Zooming in on quantum trajectories. *Journal of Physics A: Mathematical and Theoretical*, 49(10):10LT01, 2016.

[BCC+23] Cédric Bernardin, Raphaël Chetrite, Reda Chhaibi, Joseph Najnudel, and Clément Pellegrini. Spiking and collapsing in large noise limits of sdes. *The Annals of Applied Probability*, 33(1):417–446, 2023.

[BCF+17] Miguel Ballesteros, Nick Crawford, Martin Fraas, Jürg Fröhlich, and Baptiste Schubnel. Perturbation theory for weak measurements in quantum mechanics, systems with finite-dimensional state space. In *Ann. Henri Poincaré*, pages 1–37. Springer, 2017.

[BCF+19] Miguel Ballesteros, Nick Crawford, Martin Fraas, Jürg Fröhlich, and Baptiste Schubnel. Perturbation theory for weak measurements in quantum mechanics, systems with finite-dimensional state space. In *Annales Henri Poincaré*, volume 20, pages 299–335. Springer, 2019.

[BCNP22] Cédric Bernardin, Reda Chhaibi, Joseph Najnudel, and Clément Pellegrini. To spike or not to spike: the whims of the wonham filter in the strong noise regime. *arXiv preprint arXiv:2211.02032*, 2022.

[BDZ16] Maury Bramson, Jian Ding, and Ofer Zeitouni. Convergence in law of the maximum of the two-dimensional discrete Gaussian free field. *Communications on Pure and Applied Mathematics*, 69(1):62–123, 2016.

[Bel89] V. P. Belavkin. A new wave equation for a continuous nondemolition measurement. *Phys. Lett. A*, 140:355–358, October 1989.

[BG06] Daniel Bump and Alex Gamburd. On the averages of characteristic polynomials from classical groups. *Communications in mathematical physics*, 265(1):227–274, 2006.

[BG09] Alberto Barchielli and Matteo Gregoratti. *Quantum Trajectories and Measurements in Continuous Time: The Diffusive Case*. Springer Science, July 2009.

[BH95] Alberto Barchielli and Alexander S. Holevo. Constructing quantum measurement processes via classical stochastic calculus. *Stoch. Process. Appl.*, 58(2):293–317, August 1995.

[Bia91] Philippe Biane. Quantum random walk on the dual of SU(n). *Probab. Theory Related Fields*, 89(1):117–129, 1991.

[Bia06] Philippe Biane. Le théorème de Pitman, le groupe quantique $\mathrm{SU}_q(2)$, et une question de P. A. Meyer. In *In memoriam Paul-André Meyer: Séminaire de Probabilités XXXIX*, volume 1874 of *Lecture Notes in Math.*, pages 61–75. Springer, Berlin, 2006.

[Bia09] Philippe Biane. From Pitman's theorem to crystals. In *Noncommutativity and singularities*, volume 55 of *Adv. Stud. Pure Math.*, pages 1–13. Math. Soc. Japan, Tokyo, 2009.

[Bil13a] Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons, 2013.

[Bil13b] Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons, 2013.

[BJ02] Philippe Bougerol and Thierry Jeulin. Paths in Weyl chambers and random matrices. *Probability Theory and Related Fields*, 124(4):517–543, 2002.

[BJRV13] Julien Barral, Xiong Jin, Rémi Rhodes, and Vincent Vargas. Gaussian multiplicative chaos and KPZ duality. *Communications in Mathematical Physics*, 323(2):451–485, 2013.

[BK14] David Belius and Nicola Kistler. The subleading order of two dimensional cover times. *Probability Theory and Related Fields*, pages 1–92, 2014.

[BKP+20] Yasaman Bahri, Jonathan Kadmon, Jeffrey Pennington, Sam S Schoenholz, Jascha Sohl-Dickstein, and Surya Ganguli. Statistical mechanics of deep learning. *Annual Review of Condensed Matter Physics*, 11:501–528, 2020.

[BLP11] A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structures*. AMS Chelsea Publishing, Providence, RI, 2011. Corrected reprint of the 1978 original [MR0503330].

[BNN13] Paul Bourgade, Joseph Najnudel, and Ashkan Nikeghbali. A unitary extension of virtual permutations. *International Mathematics Research Notices*, 2013(18):4101–4134, 2013.

[BP02] H. P. Breuer and F. Petruccione. *The theory of open quantum systems*. Oxford University Press, Great Clarendon Street, 2002.

[BP08] Alexei Borodin and Sandrine Péché. Airy kernel with two sets of parameters in directed percolation and random matrix theory. *Journal of Statistical Physics*, 132:275–290, 2008.

[BP14] Tristan Benoist and Clement Pellegrini. Large time behaviour and convergence rate for non demolition quantum trajectories. *Comm. Math. Phys.*, 331(2):703–723, October 2014.

[BV85] Nicole Berline and Michèle Vergne. The equivariant index and kirillov's character formula. *American Journal of Mathematics*, 107(5):1159–1190, 1985.

[BVHJ07] Luc Bouten, Ramon Van Handel, and Matthew R. James. An introduction to quantum filtering. *SIAM J. Control Optim.*, 46(6):2199–2241, January 2007.

[BZ12] Maury Bramson and Ofer Zeitouni. Tightness of the recentered maximum of the two-dimensional discrete Gaussian free field. *Communications on Pure and Applied Mathematics*, 65(1):1–20, 2012.

[CC21] François Chapon and Reda Chhaibi. Quantum sl _2 sl 2, infinite curvature and pitman's 2m-x theorem. *Probability Theory and Related Fields*, 179:835–888, 2021.

[CD99] Doina Cioranescu and Patrizia Donato. *An introduction to homogenization*, volume 17 of *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press, Oxford University Press, New York, 1999.

[CDK22] Reda Chhaibi, Tariq Daouda, and Ezechiel Kahn. Free probability for predicting the performance of feed-forward fully connected neural networks. *Advances in Neural Information Processing Systems*, 35:2439–2450, 2022.

[Cer88] Carlo Cercignani. *The Boltzmann equation and its applications*, volume 67 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1988.

[CGKV23] Reda Chhaibi, Fabrice Gamboa, Slim Kammoun, and Mauricio Velasco. Estimation of large covariance matrices via free deconvolution: computational and statistical aspects, 2023.

[Chh13] Reda Chhaibi. Littelmann path model for geometric crystals, Whittaker functions on Lie groups and Brownian motion. *PhD thesis. Université Paris VI. arXiv preprint arXiv:1302.0902*, 2013.

[Chi06] Pavel Chigansky. On filtering of markov chains in strong noise. *IEEE transactions on information theory*, 52(9):4267–4272, 2006.

[CHN+19] Reda Chhaibi, Emma Hovhannisyan, Joseph Najnudel, Ashkan Nikeghbali, and Brad Rodgers. The limiting characteristic polynomial of classical random matrix ensembles. In *Annales Henri Poincaré*, volume 20, pages 1093–1119. Springer, 2019.

[Chu84] Moody T. Chu. The generalized Toda flow, the QR algorithm and the center manifold theory. *SIAM J. Algebraic Discrete Methods*, 5(2):187–201, 1984.

[CMN18] Reda Chhaibi, Thomas Madaule, and Joseph Najnudel. On the maximum of the c β e field. 2018.

[CN19] Reda Chhaibi and Joseph Najnudel. On the circle, $GMC^\gamma = \varprojlim C\beta E_n$ for $\gamma = \sqrt{\frac{2}{\beta}}$, $(\gamma \leq 1)$. *arXiv preprint arXiv:1904.00578*, 2019.

[CNN17] Reda Chhaibi, Joseph Najnudel, and Ashkan Nikeghbali. The circular unitary ensemble and the riemann zeta function: the microscopic landscape and a new approach to ratios. *Inventiones mathematicae*, 207:23–113, 2017.

[CP95] Vyjayanthi Chari and Andrew N Pressley. *A guide to quantum groups*. Cambridge university press, 1995.

[Cut13] Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. *Advances in neural information processing systems*, 26, 2013.

[D+17] Percy Deift et al. Some open problems in random matrix theory and the theory of integrable systems. ii. *SIGMA. Symmetry, Integrability and Geometry: Methods and Applications*, 13:016, 2017.

[Dav76] E. B. Davies. *Quantum theory of open systems*. Academic Press [Harcourt Brace Jovanovich, Publishers], London-New York, 1976.

[DCM92] Jean Dalibard, Yvan Castin, and Klaus Mölmer. Wave-function approach to dissipative processes in quantum optics. *Phys. Rev. Lett.*, 68(5):580, February 1992.

[DCTV20] Tariq Daouda, Reda Chhaibi, Prudencio Tossou, and Alexandra-Chloé Villani. Geodesics in fibered latent spaces: A geometric approach to learning correspondences between conditions. *arXiv preprint arXiv:2005.07852*, 2020.

[DE02] Ioana Dumitriu and Alan Edelman. Matrix models for beta ensembles. *J. Math. Phys.*, 43(11):5830–5847, 2002.

[DG10] Charles Dunkl and Stephen Griffeth. Generalized Jack polynomials and the representation theory of rational Cherednik algebras. *Selecta Mathematica*, 16(4):791–818, 2010.

[DH82] Johannes J Duistermaat and Gerrit J Heckman. On the variation in the cohomology of the symplectic form of the reduced phase space. *Inventiones mathematicae*, 69(2):259–268, 1982.

[Dio88] L. Diosi. Quantum stochastic processes as models for state vector reduction. *J. Phys. A: Math. Gen.*, 21:2885–2898, July 1988.

[DMOT14] Percy A Deift, Govind Menon, Sheehan Olver, and Thomas Trogdon. Universality in numerical computations with random data. *Proceedings of the National Academy of Sciences*, 111(42):14973–14978, 2014.

[DRS+14] Bertrand Duplantier, Rémi Rhodes, Scott Sheffield, Vincent Vargas, et al. Critical Gaussian multiplicative chaos: convergence of the derivative martingale. *The Annals of Probability*, 42(5):1769–1808, 2014.

[DRZ15] J. Ding, R. Roy, and O. Zeitouni. Convergence of the centered maximum of log-correlated Gaussian fields. *ArXiv e-prints*, March 2015.

[DS94] Persi Diaconis and Mehrdad Shahshahani. On the eigenvalues of random matrices. *J. Appl. Probab.*, 31A:49–62, 1994. Studies in applied probability.

[DW09] AB Dieker and J Warren. On the largest-eigenvalue process for generalized wishart random matrices. *Alea*, 6:369–376, 2009.

[FH13] William Fulton and Joe Harris. *Representation theory: a first course*, volume 129. Springer Science & Business Media, 2013.

[FHK12] Y. V. Fyodorov, G. A. Hiary, and J. P. Keating. Freezing Transition, Characteristic Polynomials of Random Matrices, and the Riemann Zeta Function. *Physical Review Letters*, 108(17):170601, April 2012.

[FK06] Andreas Fring and Christian Korff. Non-crystallographic reduction of generalized calogero–moser models. *Journal of Physics A: Mathematical and General*, 39(5):1115, 2006.

[FK14] Y.-V Fyodorov and J.-P. Keating. Freezing transitions and extreme values: random matrix theory and disordered landscapes. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 372(2007), 20120503, 2014.

[FW12] Mark I. Freidlin and Alexander D. Wentzell. *Random Perturbations of Dynamical Systems*. Springer Berlin Heidelberg, 2012.

[GBD+07] Christine Guerlin, Julien Bernu, Samuel Deleglise, Clement Sayrin, Sebastien Gleyzes, Stefan Kuhr, Michel Brune, Jean-Michel Raimond, and Serge Haroche. Progressive field-state collapse and quantum non-demolition photon counting. *Nature*, 448(7156):889–893, 2007.

[Gis84] N. Gisin. Quantum measurements and stochastic processes. *Phys. Rev. Lett.*, 52(19):1657, May 1984.

[GK04] Alexander N. Gorban and Iliya V. Karlin. Uniqueness of thermodynamic projector and kinetic basis of molecular individualism. *Physica A: Statistical Mechanics and its Applications*, 336(3):391–432, 2004.

[GKS76] Vittorio Gorini, Andrzej Kossakowski, and E. C. G. Sudarshan. Completely positive dynamical semigroups of N-level systems. *J. Math. Phys.*, 17(5):821–825, May 1976.

[Gol00] Georgii Ksenofontovich Golubev. On filtering for a hidden markov chain under square performance criterion. *Problemy Peredachi Informatsii*, 36(3):22–28, 2000.

[GTW01] Janko Gravner, Craig A Tracy, and Harold Widom. Limit theorems for height fluctuations in a class of discrete space and time growth models. *Journal of Statistical Physics*, 102:1085–1132, 2001.

[GW84] Roe Goodman and Nolan R. Wallach. Classical and quantum mechanical systems of Toda-lattice type. II. Solutions of the classical flows. *Comm. Math. Phys.*, 94(2):177–217, 1984.

[Hal13] Brian C. Hall. *Quantum theory for mathematicians*, volume 267. Springer, 2013.

[Har13] Serge Haroche. Nobel lecture: Controlling photons in a box and exploring the quantum to classical boundary. *Reviews of Modern Physics*, 85(3):1083, 2013.

[Hel22] Sigurdur Helgason. *Groups and geometric analysis: integral geometry, invariant differential operators, and spherical functions*, volume 83. American Mathematical Society, 2022.

[HR06] Serge Haroche and Jean-Michel Raimond. *Exploring the quantum*. Oxford Graduate Texts. Oxford University Press, Oxford, 2006. Atoms, cavities and photons.

[HS09] Yueyun Hu and Zhan Shi. Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *Ann. Probab.*, 37(2):742–789, 03 2009.

[IZ80] Claude Itzykson and J-B Zuber. The planar approximation. ii. *Journal of Mathematical Physics*, 21(3):411–421, 1980.

[JMPC20] Hicham Janati, Boris Muzellec, Gabriel Peyré, and Marco Cuturi. Entropic optimal transport between unbalanced gaussian measures has a closed form. *Advances in neural information processing systems*, 33:10468–10479, 2020.

[Joh00] Kurt Johansson. Shape fluctuations and random matrices. *Communications in mathematical physics*, 209:437–476, 2000.

[JS+17] Janne Junnila, Eero Saksman, et al. Uniqueness of critical Gaussian chaos. *Electronic Journal of Probability*, 22, 2017.

[Kas12] Christian Kassel. *Quantum groups*, volume 155. Springer Science & Business Media, 2012.

[Kir99] A. A. Kirillov. Merits and demerits of the orbit method. *Bull. Amer. Math. Soc. (N.S.)*, 36(4):433–488, 1999.

[Kir04] A. A. Kirillov. *Lectures on the orbit method*, volume 64 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2004.

[KL92] Rafail Z. Khasminskii and Betty V. Lazareva. On some filtration procedure for jump Markov process observed in white Gaussian noise. *Ann. Statist.*, 20(4):2153–2160, 1992.

[KL19] Martin Kolb and Matthias Liesenfeld. Stochastic spikes and poisson approximation of one-dimensional stochastic differential equations with applications to continuously measured quantum systems. In *Annales Henri Poincaré*, volume 20, pages 1753–1783. Springer, 2019.

[Kle05] Fima C. Klebaner. *Introduction to stochastic calculus with applications*. World Scientific Publishing Company, 2005.

[KLK+14] Samira Khan, Donghyuk Lee, Yoongu Kim, Alaa R Alameldeen, Chris Wilkerson, and Onur Mutlu. The efficacy of error mitigation techniques for dram retention failures: A

comparative experimental study. *ACM SIGMETRICS Performance Evaluation Review*, 42(1):519–532, 2014.

[KMT19] Jun Kitagawa, Quentin Mérigot, and Boris Thibert. Convergence of a newton algorithm for semi-discrete optimal transport. *Journal of the European Mathematical Society*, 21(9):2603–2651, 2019.

[KN04a] Rowan Killip and Irina Nenciu. Matrix models for circular ensembles. *Int. Math. Res. Not.*, (50):2665–2701, 2004.

[KN04b] Rowan Killip and Irina Nenciu. Matrix models for circular ensembles. *Int. Math. Res. Not.*, (50):2665–2701, 2004.

[Kos79] Bertram Kostant. The solution to a generalized toda lattice and representation theory. *Advances in Mathematics*, 34(3):195–338, 1979.

[KS97] Yvette Kosmann-Schwarzbach. Lie bialgebras, Poisson Lie groups and dressing transformations. In *Integrability of nonlinear systems*, pages 104–170. Springer, 1997.

[KS09] Rowan Killip and Mihai Stoiciu. Eigenvalue statistics for CMV matrices: from Poisson to clock via random matrix ensembles. *Duke Math. J.*, 146(3):361–399, 2009.

[KS12] Anatoli Klimyk and Konrad Schmüdgen. *Quantum groups and their representations*. Springer Science & Business Media, 2012.

[KT81] Samuel Karlin and Howard E Taylor. *A second course in stochastic processes*. Elsevier, 1981.

[Kur91] Thomas G Kurtz. Random time changes and convergence in distribution under the meyer-zheng conditions. *Ann. Probab.*, pages 1010–1034, 1991.

[KZ96] Rafail Khasminskii and Ofer Zeitouni. Asymptotic filtering for finite state Markov chains. *Stochastic Process. Appl.*, 63(1):1–10, 1996.

[Lav04] D. A. Lavis. The spin-echo system reconsidered. *Foundations of Physics*, 34(4):669–688, 2004.

[Leb99] Joel L. Lebowitz. Microscopic origins of irreversible macroscopic behavior. *Physica A: Statistical Mechanics and its Applications*, 263(1):516–527, 1999. Proceedings of the 20th IUPAP International Conference on Statistical Physics.

[Lin76] G. Lindblad. On the generators of quantum dynamical semigroups. *Comm. Math. Phys.*, 48(2):119–130, June 1976.

[Lin02] Torgny Lindvall. *Lectures on the coupling method*. Courier Corporation, 2002.

[Lip01] Liptser, Robert S and Shiryaev, Albert N. *Statistics of random processes: I. General theory*, volume 1. Springer Science & Business Media, 2001.

[Lit95a] Peter Littelmann. The path model for representations of symmetrizable kac-moody algebras. In *Proceedings of the International Congress of Mathematicians*, pages 298–308. Springer, 1995.

[Lit95b] Peter Littelmann. Paths and root operators in representation theory. *Annals of Mathematics*, pages 499–525, 1995.

[Mab09] Hideo Mabuchi. Continuous quantum error correction as classical hybrid control. *New Journal of Physics*, 11(10):105044, oct 2009.

[Mac82] Ian G Macdonald. Some conjectures for root systems. *SIAM Journal on Mathematical Analysis*, 13(6):988–1007, 1982.

[Mac89] Michael C. Mackey. The dynamic origin of increasing entropy. *Rev. Mod. Phys.*, 61:981–1015, Oct 1989.

[Mac98] Ian Grant Macdonald. *Symmetric functions and Hall polynomials*. Oxford university press, 1998.

[Mad15] T. Madaule. Maximum of a log-correlated Gaussian field. *Ann. Inst. H. Poincaré Probab. Statist.*, 51(4):1369–1431, 11 2015.

[MGLG16] Katarzyna Macieszczak, Mădălin Guţă, Igor Lesanovsky, and Juan P Garrahan. Towards a theory of metastability in open quantum dynamics. *Physical review letters*, 116(24):240404, 2016.

[Mic95] Laurent Miclo. Comportement de spectres d'operateurs de schrodinger a basse temperature. *Bulletin des sciences mathématiques*, 119(6):529–554, 1995.

[MNN20] Kenneth Maples, Joseph Najnudel, and Ashkan Nikeghbali. Limit operators for circular ensembles. In *Frontiers in Analysis and Probability: In the Spirit of the Strasbourg-Zürich Meetings*, pages 327–369. Springer, 2020.

[MO04] Hiroyuki Matsumoto and Yukio Ogura. Markov or non-Markov property of $cM - X$ processes. *J. Math. Soc. Japan*, 56(2):519–540, 04 2004.

[Mos75] Jürgen Moser. Finitely many mass points on the line under the influence of an exponential potential–an integrable system. In *Dynamical systems, theory and applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974)*, pages 467–497. Lecture Notes in Phys., Vol. 38. Springer, Berlin, 1975.

[MRV+16] Thomas Madaule, Rémi Rhodes, Vincent Vargas, et al. Glassy phase and freezing of log-correlated Gaussian potentials. *The Annals of Applied Probability*, 26(2):643–690, 2016.

[MT21] Quentin Merigot and Boris Thibert. Optimal transport: discretization and algorithms. In *Handbook of numerical analysis*, volume 22, pages 133–212. Elsevier, 2021.

[Mun00] James R Munkres. *Topology*. Prentice Hall, 2000.

[MZ84] PA Meyer and WA Zheng. Tightness criteria for laws of semimartingales. *Ann. I. H. Poincaré - Pr.*, 20(4):353–372, 1984.

[NW99] Jorge Nocedal and Stephen J Wright. *Numerical optimization*. Springer, 1999.

[O+89] Eric M Opdam et al. Some applications of hypergeometric shift operators. *Inventiones mathematicae*, 98(1):1–18, 1989.

[OV96] Grigori Olshanski and Anatoli Vershik. Ergodic unitarily invariant measures on the space of infinite hermitian matrices. *arXiv preprint math/9601215*, 1996.

[Pap78] G. C. Papanicolaou. Asymptotic analysis of stochastic equations. In *Studies in probability theory*, volume 18 of *MAA Stud. Math.*, pages 111–179. Math. Assoc. America, Washington, D.C., 1978.

[Par92] K. R. Parthasarathy. *An introduction to quantum stochastic calculus*. Birkhäuser, 1992.

[Par98] Beresford N. Parlett. *The symmetric eigenvalue problem*, volume 20 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998. Corrected reprint of the 1980 original.

[PC+19] Gabriel Peyré, Marco Cuturi, et al. Computational optimal transport: With applications to data science. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019.

[Pea84] Philip Pearle. Comment on "quantum measurements and stochastic processes". *Phys. Rev. Lett.*, 53:1775–1775, Oct 1984.

[Pel08] Clément Pellegrini. Existence, uniqueness and approximation of a stochastic Schrödinger equation: the diffusive case. *Ann. Probab.*, 36(6):2332–2353, 2008.

[Pel10] Clément Pellegrini. Markov chains approximation of jump–diffusion stochastic master equations. *Ann. Inst. H. Poincaré: Prob. Stat.*, 46(4):924–948, November 2010.

[Per98] Ian Percival. *Quantum state diffusion*. Cambridge University Press, 1998.

[Per02] Grisha Perelman. The entropy formula for the ricci flow and its geometric applications. *arXiv preprint math/0211159*, 2002.

[Pic86] Jean Picard. Nonlinear filtering of one-dimensional diffusions in the case of a high signal-to-noise ratio. *SIAM J. Appl. Math.*, 46(6):1098–1125, 1986.

[Pit75] James W Pitman. One-dimensional brownian motion and the three-dimensional Bessel process. *Advances in Applied Probability*, 7(3):511–526, 1975.

[Pit06] Jim Pitman. *Combinatorial stochastic processes: Ecole d'eté de probabilités de saint-flour xxxii-2002*. Springer, 2006.

[PKHM19] Minesh Patel, Jeremie S Kim, Hasan Hassan, and Onur Mutlu. Understanding and modeling on-die error correction in modern dram: An experimental study using real devices. In *2019 49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pages 13–25. IEEE, 2019.

[Pow18] Ellen Powell. Critical Gaussian chaos: convergence and uniqueness in the derivative normalisation. *Electronic Journal of Probability*, 23, 2018.

[Pra05] Victor Prasolov. *Surveys in modern mathematics*, volume 321. Cambridge University Press, 2005.

[PS08] Grigoris Pavliotis and Andrew Stuart. *Multiscale methods: averaging and homogenization*. Springer Science & Business Media, 2008.

[PSG18] Jeffrey Pennington, Samuel Schoenholz, and Surya Ganguli. The emergence of spectral universality in deep networks. In *International Conference on Artificial Intelligence and Statistics*, pages 1924–1932. PMLR, 2018.

[PX97] Gilles Pisier and Quanhua Xu. Non-commutative martingale inequalities. *Communications in mathematical physics*, 189(3):667–698, 1997.

[PX03] Gilles Pisier and Quanhua Xu. Non-commutative L^p-spaces. *Handbook of the geometry of Banach spaces*, 2:1459–1517, 2003.

[PZ05] Étienne Pardoux and Ofer Zeitouni. Quenched large deviations for one dimensional nonlinear filtering. *SIAM J. Control Optim.*, 43(4):1272–1297, 2004/05.

[PZ16] E. Paquette and O. Zeitouni. The maximum of the CUE field. *ArXiv e-prints*, February 2016.

[PZ22] Elliot Paquette and Ofer Zeitouni. The extremal landscape for the c β e ensemble. *arXiv preprint arXiv:2209.06743*, 2022.

[QY] Jeremy Quastel and Horng Tzer Yau. Celebratio mathematica.

[RBA22] Anugu Sumith Reddy, Amarjit Budhiraja, and Amit Apte. Some large deviation asymptotics in small noise filtering problems. *SIAM J. Control Optim.*, 60(1):385–409, 2022.

[Reb87] Rolando Rebolledo. Topologie faible et méta-stabilité. In *Séminaire de Probabilités XXI*, pages 544–562. Springer, 1987.

[RP81] L. C. G. Rogers and J. W. Pitman. Markov functions. *Ann. Probab.*, 9(4):573–582, 1981.

[RV13] R. Rhodes and V. Vargas. Gaussian multiplicative chaos and applications: a review. *ArXiv e-prints*, May 2013.

[RY99] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[RY13] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293. Springer Science & Business Media, 2013.

[Sim96] Barry Simon. *Representations of finite and compact groups*. Number 10. American Mathematical Soc., 1996.

[Sim05a] Barry Simon. *Orthogonal polynomials on the unit circle. Part 1*, volume 54 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2005. Classical theory.

[Sim05b] Barry Simon. *Orthogonal polynomials on the unit circle. Part 1*, volume 54 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2005. Classical theory.

[Spo20] Herbert Spohn. Generalized gibbs ensembles of the classical Toda chain. *Journal of Statistical Physics*, 180(1):4–22, 2020.

[Spo21] Herbert Spohn. Hydrodynamic equations for the Toda lattice. *arXiv preprint arXiv:2101.06528*, 2021.

[SPW09] Bianca Schroeder, Eduardo Pinheiro, and Wolf-Dietrich Weber. Dram errors in the wild: a large-scale field study. *ACM SIGMETRICS Performance Evaluation Review*, 37(1):193–204, 2009.

[SR09] Laure Saint-Raymond. *Hydrodynamic limits of the Boltzmann equation*. Number 1971. Springer Science & Business Media, 2009.

[Sym82] W. W. Symes. The QR algorithm and scattering for the finite nonperiodic Toda lattice. *Phys. D*, 4(2):275–280, 1981/82.

[Sze75] Gábor Szegö. *Orthogonal polynomials*. American Mathematical Society, Providence, R.I., fourth edition, 1975. American Mathematical Society, Colloquium Publications, Vol. XXIII.

[TBB15] Antoine Tilloy, Michel Bauer, and Denis Bernard. Spikes in quantum trajectories. *Physical Review A*, 92(5):052111, 2015.

[Tod89] Morikazu Toda. *Theory of nonlinear lattices*, volume 20 of *Springer Series in Solid-State Sciences*. Springer-Verlag, Berlin, second edition, 1989.

[TW94] Craig A Tracy and Harold Widom. Level-spacing distributions and the airy kernel. *Communications in Mathematical Physics*, 159:151–174, 1994.

[VH07] Ramon Van Handel. Stochastic calculus, filtering, and stochastic control. *Course notes., URL http://www. princeton. edu/rvan/acm217/ACM217. pdf*, 14, 2007.

[Via19] François-Xavier Vialard. An elementary introduction to entropic regularization and proximal methods for numerical optimal transport. 2019.

[Vil21] Cédric Villani. *Topics in optimal transportation*, volume 58. American Mathematical Soc., 2021.

[VV22] Benedek Valkó and Bálint Virág. The many faces of the stochastic zeta function. *Geometric and Functional Analysis*, 32(5):1160–1231, 2022.

[Wat84] David S. Watkins. Isospectral flows. *SIAM Rev.*, 26(3):379–391, 1984.

[Whi02] Ward Whitt. *Stochastic-process limits: an introduction to stochastic-process limits and their application to queues*. Springer Science & Business Media, 2002.

[Wig58] Eugene P Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, 67(2):325–327, 1958.

[WM10a] Howard M. Wiseman and Gerard J. Milburn. *Quantum measurement and control*. Cambridge University Press, Cambridge, 2010.

[WM10b] Howard M. Wiseman and Gerard J. Milburn. *Quantum Measurement and Control*. Cambridge University Press, January 2010.

[Won64] W Murray Wonham. Some applications of stochastic differential equations to optimal nonlinear filtering. *Journal of the Society for Industrial and Applied Mathematics, Series A: Control*, 2(3):347–369, 1964.

[Woo92] Nicholas Michael John Woodhouse. *Geometric quantization*. Oxford university press, 1992.