Sur certains objets universels liés à des changements de probabilité et des modèles de matrices aléatoires

Joseph Najnudel

UNIVERSITE PARIS VI - PIERRE ET MARIE CURIE

Mémoire d'habilitation à diriger des recherches

Spécialité : Mathématiques

présenté par : **Joseph NAJNUDEL**

Titre:

Sur certains objets universels liés à des changements de probabilité et des modèles de matrices aléatoires

Soutenue le 7 décembre 2011 devant le jury composé de:

M. Philippe BIANE, *rapporteur*
M. Philippe BOUGEROL, *examinateur*
M. Michel EMERY, *examinateur*
M. Ashkan NIKEGHBALI, *examinateur*
M. Alain ROUAULT, *examinateur*
M. Marc YOR, *examinateur*

Remerciements

Je tiens tout d'abord à remercier Philippe Biane, Bálint Virág et Ofer Zeitouni pour le temps qu'ils ont consacré à la lecture de ce mémoire, et pour les rapports qu'ils ont accepté de rédiger. La lecture de ces rapports a été une aide précieuse pour ma propre appréciation du travail de recherche que j'ai effectué depuis ma thèse.

Je suis également très honoré que Philippe Biane, Philippe Bougerol, Michel Emery, Ashkan Nikeghbali, Alain Rouault et Marc Yor aient accepté de faire partie de mon jury d'habilitation à diriger des recherches, témoignant ainsi de l'intérêt qu'ils portent à mes travaux.

Ceux-ci n'auraient pas pu voir le jour sans mes coauteurs, qui sont aussi des amis, et avec qui j'ai toujours travaillé avec enthousiasme: Paul Bourgade, Chris Hughes, Ashkan Nikeghbali, Bernard Roynette, Felix Rubin, Dan Stroock, Marc Yor, Dirk Zeindler. Je tiens particulièrement à exprimer ma gratitude à Marc Yor, qui par son travail d'encadrement de ma thèse et sa disponibilité, m'a permis de commencer mon travail de recherche mathématique, et à remercier Ashkan Nikeghbali, pour l'amitié dont il toujours fait preuve pour moi, et pour le soutien très fort qu'il m'a constamment apporté, aussi bien personnellement que professionnellement. La quantité de travail que nous avons en commun ces trois dernières années en sont un témoignage.

Je tenais aussi à remercier Alain-Sol Sznitman et Erwin Bolthausen pour leur travail d'encadrement post-doctoral, à l'ETH puis à l'Université de Zürich, ainsi que les personnes que je n'ai pas encore citées et qui ont accepté d'écrire des lettres de recommandation pour moi: Jean Bertoin, Emmanuel Kowalski, Jean-François Le Gall, Lorenzo Zambotti.

Je remercie également les nombreux chercheurs et étudiants avec qui j'ai pu avoir des dicussions, mathématiques ou autres, en particulier à l'Université de Zürich, à l'ETH et à l'Université Paris VI: je ne peux malheureusement pas citer tous les noms...

J'adresse également un remerciement aux membres du personnel administratif et technique de ces institutions, qui par leur disponibilité et leur efficacité, m'ont permis d'avoir les conditions matérielles nécessaires à la bonne réalisation de mon travail.

Enfin, je n'oublie pas de remercier tous les amis dont je n'ai pas parlé jusqu'ici, ainsi que tous les membres de ma famille. J'ai une pensée particulière pour ma mère, et pour mon frère Jean-Samuel, qui m'ont toujours soutenu, particulièrement dans les moments difficiles, et qui m'ont tant aidé depuis si longtemps.

Table des matières

Résumé

Ce mémoire d'habilitation comporte deux parties.

- Dans la première partie, nous étudions les propriétés générales d'une famille de mesures de probabilité obtenues à partir de la loi du mouvement brownien via une certaine procédure appelée pénalisation. Nous construisons une mesure σ-finie donnant une explication globale à une partie des phénomènes observés, et nous généralisons cette mesure à d'autres contextes. Nous faisons également le lien avec certains processus aléatoires modélisant la trajectoire des polymères.

- Dans la deuxième partie, nous étudions différents modèles de matrices aléatoires, construits à partir de matrices de permutations et de matrices unitaires. Nous construisons également des objets infini-dimensionnels associés à ces modèles, afin d'améliorer notre compréhension de leurs propriétés universelles.

Dans une serie d'articles écrits par Roynette, Vallois et Yor (voir [59], [60], [61], [62] et [64] par exemple), de nombreuses mesures de probabilité sont construites sur l'espace $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ des fonctions continues de $\mathbb{R}_+$ vers $\mathbb{R}^d$ (généralement pour $d = 1$). Le procédé de construction, appelé *pénalisation*, est similaire dans la plupart des cas:

- On définit $\mathbb{P}_d$ comme la mesure de Wiener sur $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$.

- On prend une famille $(\Gamma_t)_{t\geq 0}$ de variables aléatoires positives, telles que $0 < \mathbb{E}_{\mathbb{P}_d}[\Gamma_t] < \infty$ pour tout $t \geq 0$.

- On considère la famille de probabilités $(\mathbb{Q}_t)_{t\geq 0}$ définie par

$$\mathbb{Q}_t = \frac{\Gamma_t}{\mathbb{E}_{\mathbb{P}_d}[\Gamma_t]} \cdot \mathbb{P}_d.$$

Un phénomène intéressant a été observé par Roynette, Vallois et Yor: pour chacun des très nombreux exemples de poids $(\Gamma_t)_{t\geq 0}$ qu'ils ont étudiés, il existe une mesure $\mathbb{Q}$ satisfaisant la propriété suivante: pour tout $s \geq 0$ et pour tout ensemble Λ_s dans la tribu engendrée par le processus canonique jusqu'au temps s, $\mathbb{Q}_t(\Lambda_s)$ tend vers $\mathbb{Q}(\Lambda_s)$ quand t tend vers l'infini. Ce résultat est d'autant plus remarquable que le comportement du processus canonique sous cette mesure limite $\mathbb{Q}$ peut être très différent selon les exemples considérés.

Il n'existe pour l'instant aucune explication globale de ce phénomène qui soit applicable à tous les cas. Cependant, dans de nombreux exemples en dimension $d = 1$, la mesure $\mathbb{Q}$ est absolument continue par rapport à une certaine mesure σ-finie $\mathbb{W}$ sur $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$, construite et étudiée en détail dans notre monographie écrite avec Roynette et Yor [48]. La mesure $\mathbb{W}$ peut être définie comme suit:

- Le supremum g des valeurs d'annulation du processus canonique est fini $\mathbb{W}$-presque partout.

- Pour tout $t \geq 0$, et pour toute fonctionnelle F_t, mesurable par rapport à la tribu engendrée par le processus canonique X pris jusqu'au temps t,

$$\mathbb{W}[F_t \mathbb{1}_{g\leq t}] = \mathbb{E}_{\mathbb{P}_1}[F_t |X_t|].$$

La trajectoire du processus canonique sous $\mathbb{W}$ peut également être décrite comme la concaténation d'un mouvement brownien arrêté à un inverse de temps local, et d'un processus de Bessel de dimension 3. La mesure $\mathbb{W}$ satisfait de nombreuses propriétes remarquables et dans [48], nous construisons un analogue de cette mesure adapté aux contextes suivants: le mouvement brownien plan, les diffusions linéaires et les chaînes de Markov. Une autre généralisation est obtenue dans une série de papiers en collaboration avec Nikeghbali (voir [40], [41], [43], [45]), dans le cadre des sous-martingales de classe (Σ), i.e. dont le processus croissant est porté par les zéros. Une partie des résultats obtenus sont liés à certains problèmes de mathématiques financières.

La technique de construction de mesures de probabilités par pénalisation est également liée au *modèle d'Edwards* (voir [17]), qui correspond à une famille de mesures utilisées pour décrire la forme des molécules de certains polymères. Plus précisément, les trajectoires correspondantes sont aléatoires, mais présentent un phénomène d'auto-répulsion: deux parties distinctes du polymère ont tendance à éviter d'occuper la même partie de l'espace. Informellement, pour une trajectoire de longueur t, on considère la mesure $\mathbb{Q}_t$ donnée par le poids de pénalisation

$$\Gamma_t := \exp\left(-\beta \int_0^t \int_0^t \delta(X_s - X_u) ds\, du\right), \tag{1}$$

où X est le processus canonique, δ la mesure de Dirac en zéro et $\beta > 0$ un paramètre fixé à l'avance. Le symbole δ ne représente pas une véritable fonction, aussi l'équation (1) ne donne-t-elle pas une définition rigoureuse de $\mathbb{Q}_t$. La difficulté de la construction effective de cette mesure dépend de la dimension d:

- Pour $d = 1$, Γ_t peut être défini rigoureusement à l'aide des temps locaux du processus canonique. De plus, la mesure $\mathbb{Q}_t$ converge vers une limite $\mathbb{Q}$ (au sens des pénalisations décrit plus haut) quand t tend vers l'infini, ce qui définit une version du modèle d'Edwards unidimensionnel pour des trajectoires infinies. La preuve de cette convergence est le principal résultat de notre article [38].

- Pour $d = 2$, toute tentative raisonnable de donner un sens à Γ_t donne une valeur nulle, mais en divisant formellement ce poids par une constante (nulle elle aussi), on peut se ramener à une situation similaire à la dimension 1, et donc définir une mesure $\mathbb{Q}_t$, absolument continue par rapport à la mesure de Wiener. Le procédé employé est appelé renormalisation de Varadhan (voir [70]).

- Pour $d = 3$, la renormalisation de Varadhan ne suffit pas mais il est encore possible de construire $\mathbb{Q}_t$, comme cela a été fait par Westwater (voir [73] et [74]) puis par Bolthausen (voir [9]).

- Pour $d \geq 4$, le mouvement brownien n'a pas d'auto-intersection et donc la construction du modèle d'Edwards devient triviale.

Le cas le plus intéressant et le plus difficile est la dimension 3, pour laquelle $\mathbb{Q}_t$ est singulière par rapport à la mesure de Wiener $\mathbb{P}_3$. Dans [36], Chap. 5, nous construisons, en dimensions 1 et 2, une modification du modèle d'Edwards qui est singulière par rapport à la mesure de Wiener, contrairement au modèle d'Edwards classique.

La deuxième partie du mémoire est consacrée à des modèles de matrices aléatoires et à certaines de leurs généralisations infini-dimensionnelles. Une introduction à la théorie des matrices aléatoires est par exemple donnée par Mehta [33], ou plus récemment, par Anderson, Guionnet et Zeitouni [2]. Parmi les modèles les plus classiques, on peut noter les deux suivants:

- L'ensemble gaussien unitaire (*Gaussian Unitary Ensemble*), correspondant à des matrices hermitiennes dont les coefficients sont construits à partir de variables gaussiennes indépendantes.

- L'ensemble circulaire unitaire (*Circular Unitary Ensemble*), correspondant à des matrices dont la loi est la mesure de Haar sur un groupe unitaire.

Pour ces deux exemples, le comportement microscopique des valeurs propres en grande dimension fait apparaître un processus ponctuel limite, appelé processus déterminantal de noyau sinus (*determinantal sine-kernel process*). Ce processus peut informellement être défini de la manière suivante: on considère un ensemble de points aléatoire, tel que pour tout $p \geq 1$, et pour tous $t_1, \ldots, t_p \in \mathbb{R}$, la probabilité d'avoir un point au voisinage de t_j pour $j \in \{1, ..., p\}$ est proportionnelle à

$$\rho(t_1, \ldots, t_p) := \det(K(t_i, t_j))_{1 \leq i,j \leq k},$$

où K est une fonction de $\mathbb{R}^2$ vers $\mathbb{R}$, donnée par

$$K(x, y) = \frac{\sin(\pi(x - y))}{\pi(x - y)}.$$

Ce processus déterminantal apparaît en fait dans le comportement limite de nombreux ensembles de matrices aléatoires hermitiennes et unitaires: le fait d'obtenir le même processus limite pour de nombreuses situations différentes est un phénomène appelé *universalité*, et dans le cas décrit ici, il est encore compris de manière incomplète.

Une situation analogue (mais beaucoup plus simple) se produit dans le cadre du théorème central limite: la loi gaussienne intervient de manière universelle, indépendamment de la loi des variables indépendantes qui sont considérées au départ. Dans ce cadre, la compréhension que l'on a du rôle particulier joué par la loi gaussienne est considérablement améliorée par la construction d'un objet infini modélisant le comportement à grande échelle des sommes de variables aléatoires indépendantes: le mouvement brownien. Afin de comprendre de manière analogue le comportement microscopique du spectre des matrices aléatoires, et en particulier leurs propriétés universelles, nous nous intéressons à la construction d'objets limites infini-dimensionnels pouvant être naturellement associés à différents ensembles de matrices. Notre travail se focalise autour de deux types d'ensembles:

- Des ensembles construits à partir de matrices de permutations.

- L'ensemble circulaire unitaire et certaines de ses généralisations.

L'avantage des permutations est qu'on peut directement calculer le spectre de leurs matrices à partir de leur décomposition en cycles. L'ensemble de matrices de permutations le plus simple que l'on puisse considérer est obtenu en choisissant aléatoirement une permutation, suivant la loi uniforme sur le groupe symétrique $\mathfrak{S}_N$ pour $N \geq 1$. Une généralisation de cet ensemble est obtenue en choisissant un paramètre $\theta > 0$, et en donnant à chaque permutation une probabilité proportionnelle à θ^n, n désignant le nombre de cycles de la permutation ($\theta = 1$ correspond à la loi uniforme). Sous la mesure de probabilité ainsi construite, appelée *mesure d'Ewens de paramètre* θ, le comportement asymptotique de la taille des grand cycles est connu: si $(\ell_k)_{k \geq 1}$ est la suite décroissante (complétée par des zéros) des longueurs de cycles associées à une permutation suivant la mesure d'Ewens de paramètre θ sur $\mathfrak{S}_N$, alors $(\ell_k/N)_{k\geq 1}$ converge en loi vers une distribution de Poisson-Dirichlet de paramètre θ. Ce résultat implique que le processus ponctuel des angles propres des matrices de permutation correspondantes, renormalisés à l'aide d'une multiplication par $N/2\pi$ (de manière à avoir un espacement moyen de 1 entre les points), converge en loi (en un sens à préciser) vers un processus ponctuel limite, obtenu comme réunion des suites $(2\pi m/\lambda_k)_{m\in\mathbb{Z}}$ pour $k \geq 1$, $(\lambda_k)_{k\geq 1}$ étant un processus de Poisson-Dirichlet de paramètre θ.

Cette convergence peut être expliquée dans le cadre d'un modèle infini induisant des permutations aléatoires dans tous les groupes symétriques $(\mathfrak{S}_N)_{N\geq 1}$ en même temps. L'objet central de cette construction est la notion de *permutation virtuelle* introduite par Kerov, Olshanski et Vershik (voir [24]) et étudiée en détail par Tsilevich (voir [66] et [67]). Une permutation virtuelle est une suite de permutations $(\sigma_N)_{N\geq 1}$, $\sigma_N \in \mathfrak{S}_N$ telle que pour tout $N \geq 1$, la structure en cycles de σ_N se déduit de celle de σ_{N+1} en enlevant simplement l'élément $N+1$ de son cycle. En considérant une permutation virtuelle aléatoire, on définit ainsi, sur un espace de probabilité commun, une matrice de permutation aléatoire d'ordre N pour tout $N \geq 1$. De plus, pour tout $\theta > 0$, il est possible, grâce à une certaine relation de compatibilité satisfaite par les mesures d'Ewens, de faire en sorte que pour tout $N \geq 1$, la permutation d'ordre N suive la mesure d'Ewens de paramètre θ sur $\mathfrak{S}_N$. Dans ce cas, la convergence en loi du processus ponctuel des angles propres renormalisés décrite ci-dessus se renforce en une convergence presque sûre. Il y a alors un lien déterministe entre le processus ponctuel limite et la permutation virtuelle aléatoire considérée. De plus, dans un papier écrit avec Nikeghbali [44], nous prouvons qu'il est possible d'exprimer ce processus ponctuel limite comme le spectre d'un opérateur aléatoire, représentant le comportement infinitésimal de la permutation virtuelle. En d'autres termes, la convergence en loi des angles propres renormalisés correspondant à une permutation suivant une mesure d'Ewens peut être expliquée par une sorte de convergence presque sûre d'opérateurs. Notons que dans [44], les mesures considérées sont en fait plus générales que celles d'Ewens.

Comme indiqué plus haut, les modèles de matrices de permutations sont pratiques à étudier, grâce aux calculs explicites de valeurs propres qu'ils permettent. Cependant, ces valeurs propres vérifient également des propriétés très particulières: par exemple, ce sont toutes des racines complexes de l'unité d'ordre fini. Pour cette raison, les ensembles de matrices de permutations se comportent de manière très différente des ensembles plus classiques de matrices unitaires et hermitiennes. Un type d'ensemble ayant un comportement intermédiaire est obtenu en partant de matrices de permutations et en remplaçant les coefficients égaux à 1 par des variables indépendantes, identiquement distribuées et non nulles. De tels ensembles, ainsi que les modèles infinis associés, sont étudiés en détail dans [42].

Le cadre de l'ensemble circulaire unitaire est, de notre point de vue, plus difficile à étudier que celui des matrices de permutations. Cependant, dans notre article avec Bourgade et Nikeghbali [11], nous construisons les *isométries virtuelles*, qui sont des suites $(u_N)_{N\geq 1}$ de matrices unitaires, $u_N \in U(N)$, vérifiant une relation de compatibilité généralisant celle qui définit les permutations virtuelles. La notion d'isométrie virtuelle englobe donc celle de permutation virtuelle, et également une notion similaire introduite précédemment par Neretin (voir [49]) et étudiée par Borodin, Olshanski et Vershik (voir [10] et [51]). Une famille de mesures de probabilité remarquables, similaires aux mesures d'Ewens et appelées mesures d'Hua-Pickrell (voir [20], [53] et [54]) peut alors être définie sur l'espace des rotations virtuelles. Cette famille est indexée par un paramètre complexe, et permet de définir, sur un même espace de probabilité, un ensemble de matrices unitaires de dimension

N pour tout $N \geq 1$. Pour une des valeurs du paramètre, la loi obtenue est, pour tout N, la mesure de Haar, i.e. la mesure qui définit l'ensemble circulaire unitaire. Dans ce cas, un résultat classique implique que le processus ponctuel des angles propres renormalisés converge en loi vers un processus déterminantal de noyau sinus. Le cadre des isométries virtuelles permet, comme expliqué dans [11], de remplacer la convergence en loi par une convergence presque sûre, comme dans le cadre des ensembles de matrices de permutation. Cependant, contrairement au cas du groupe symétrique, nous ne connaissons aucune manière naturelle d'interpréter le processus ponctuel limite comme le spectre d'un opérateur aléatoire.

Dans nos perspectives de recherche, nous prévoyons d'étudier la possibilité de construire un tel opérateur, qui modéliserait le comportement local des grandes matrices aléatoires hermitiennes et unitaires. L'existence de cet opérateur permettrait d'améliorer de manière significative notre compréhension des phénomènes d'universalité observés pour de nombreux ensembles de matrices aléatoires. Cela pourrait peut-être également donner des informations sur les liens entre matrices aléatoires et théorie des nombres (fonction zêta de Riemann par exemple), qui ne sont pour le moment pas très bien compris. Nous prévoyons d'explorer ces liens, et plus généralement, d'étudier les connections que l'on peut établir entre les matrices aléatoires et les autres domaines des mathématiques.

Liste de publications

[11] P. Bourgade, J. Najnudel and A. Nikeghbali, *A unitary extension of virtual permutations*, http://arxiv.org/pdf/1102.2633v1, 2011.

[35] J. Najnudel, *Pénalisations de l'araignée brownienne (Penalizations of Walsh Brownian motion)*, Annales de l'Institut Fourier **57** (2007), no. 4, 1063-1093.

[36] J. Najnudel, *Temps locaux et pénalisations browniennes*, Ph.D. thesis, Université Paris VI, 2007.

[37] J. Najnudel, *Penalisations of the Brownian motion with a functional of its local times*, Stochastic Process. Appl. **118** (2008), 1407-1433.

[38] J. Najnudel, *Construction of an Edwards' probability measure on* $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$, Ann. Probab. **38** (2010), no. 6, 2295-2321.

[39] J. Najnudel and A. Nikeghbali, *On penalisation results related with a remarkable class of submartingales*, http://arxiv.org/pdf/0911.4365, 2009.

[40] J. Najnudel and A. Nikeghbali, *On some properties of a universal sigma-finite measure associated with a remarkable class of submartingales*, http://arxiv.org/pdf/0911.2571, to appear in Pub. of the Res. Instit. of Math. Sci. (Kyoto University), 2009.

[41] J. Najnudel and A. Nikeghbali, *On some universal σ-finite measures and some extensions of Doob's optional stopping theorem*, http://arxiv.org/pdf/0906.1782, 2009.

[42] J. Najnudel and A. Nikeghbali, *The distribution of eigenvalues of randomized permutation matrices*, http://arxiv.org/pdf/1005.0402, 2010.

[43] J. Najnudel and A. Nikeghbali, *A new construction of the σ-finite measures associated with submartingales of class* (Σ), Comptes Rend. Acad. Sci. Paris **348** (2010), 311-316.

[44] J. Najnudel and A. Nikeghbali, *On a flow of operators associated to virtual permutations*, http://arxiv.org/pdf/1008.4972, 2010.

[45] J. Najnudel and A. Nikeghbali, *A remarkable σ-finite measure associated with last passage times and penalisation problems*, Contemporary Quantitative Finance, Essays in Honour of Eckhard Platen, Springer, 2010, pp. 77-98.

[46] J. Najnudel and A. Nikeghbali, *A new kind of augmentation of filtrations*, ESAIM: Probability and Statistics **15** (2011), S39-S57.

[47] J. Najnudel, A. Nikeghbali and F. Rubin, *Scaled limit and rate of convergence for the largest eigenvalue from the generalized Cauchy random matrix ensemble*, Jour. of Stat. Phys. **137** (2009), no. 2, 373-406.

[48] J. Najnudel, B. Roynette and M. Yor, *A Global View of Brownian Penalisations*, MSJ Memoirs, vol. 19, Mathematical Society of Japan, Tokyo, 2009.

Part I

Some universal objects related to changes of probability measures and penalization

The typical behavior of a stochastic process can be characterized by the almost sure properties of its trajectories. For example, a Brownian path is almost surely unbounded, it tends to infinity at infinity if the dimension is larger than of equal to 3, and it contains self-intersections if the dimension is smaller than or equal to 3.

A question which naturally arises is the following: how these almost sure properties are modified when one changes the probability law of the process? An general introduction to change of time and change of measure is given by Barndorff-Nielsen and Shyriaev [6]. More specifically, one of the simplest ways to do a change of measure is to replace the initial probability measure by another one, absolutely continuous, whose density is explicitly given. This kind of change is often made in statistical physics (e.g. for the Ising model), where on a given measurable space, one first considers a probability measure $\mathbb{P}$, and then a measure $\mathbb{Q}$ such that

$$\mathbb{Q} := \frac{\exp(-\beta H)}{\mathbb{E}_{\mathbb{P}}[\exp(-\beta H)]} \cdot \mathbb{P},$$

where the parameter $\beta > 0$ represents the inverse of the temperature, and the hamiltonian H represents the energy: informally, the states of high energy are penalized by the exponential factor $e^{-\beta H}$. Another situation involving a similar change of measure is the framework of the Girsanov's theorem, which describes a way to change the law of a semi-martingale, in order to make it a martingale. This result is very classical in mathematical finance.

In any case, if a probability measure $\mathbb{P}$ is changed to a measure $\mathbb{Q}$ given by

$$\mathbb{Q} := \frac{\Gamma}{\mathbb{E}_{\mathbb{P}}[\Gamma]} \cdot \mathbb{P},$$

for $\Gamma \geq 0$ and $0 < \mathbb{E}[\Gamma] < \infty$, then any negligible set with respect to $\mathbb{P}$ remains negligible with respect to $\mathbb{Q}$, i.e. any property satisfied almost surely under $\mathbb{P}$ is also almost sure under $\mathbb{Q}$. Hence, if we want to modify the typical behavior of the model, we need to change the measure $\mathbb{P}$ in a deeper way. As we will study in detail in several different situations, such a deeper change can sometimes be obtained by passing to the limit. More precisely, if $(\Gamma_t)_{t\geq 0}$ is a family of random variables satisfying the same properties as Γ, and if $(\mathbb{Q}_t)_{t\geq 0}$ is the family of probability measures given by

$$\mathbb{Q}_t = \frac{\Gamma_t}{\mathbb{E}_{\mathbb{P}}[\Gamma_t]} \cdot \mathbb{P},$$

then one can ask if $\mathbb{Q}_t$ converges, in a sense which has to be made precise, to a limit probability measure $\mathbb{Q}$, when the parameter t goes to infinity. Note that one can also consider discrete sequences $(\Gamma_n)_{n\geq 1}$. Moreover, the naive definition of the convergence by the fact that

$$\mathbb{Q}_t(A) \underset{t\to\infty}{\longrightarrow} \mathbb{Q}(A)$$

for all measurable sets A is too strong in general: one needs to adapt, in a more subtle way, the notion of weak convergence to the present setting.

The particular case of the Brownian motion has been extensively studied by Roynette, Vallois and Yor in a number of papers (for example, see [59], [60], [61], [62] and [64]), where the changes of measure described above are referred as *penalizations*. In these articles, $\mathbb{P}$ is the Wiener measure on the space $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}^d$ (generally for $d = 1$), and the notion of convergence considered is the following: a sequence $(\mathbb{Q}_t)_{t \geq 0}$ of probability measures on $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ converges to a limit measure $\mathbb{Q}$ if and only if for all $s \geq 0$, and for any event Λ_s in the σ-algebra generated by the canonical process up to time s,

$$\mathbb{Q}_t(\Lambda_s) \underset{t \to \infty}{\longrightarrow} \mathbb{Q}(\Lambda_s).$$

Roynette, Vallois and Yor have considered many examples of families $(\Gamma_t)_{t \geq 1}$ of weights and have proven that in each case, the limit measure $\mathbb{Q}$ exists. Moreover, in the most interesting cases, this measure is singular with respect to the Wiener measure, which implies that the behavior of the canonical process under $\mathbb{Q}$ is radically different from the behavior of the Brownian motion. Here are some examples they have studied:

- $\Gamma_t = \phi(S_t)$, where ϕ is an integrable function from $\mathbb{R}_+$ to $\mathbb{R}_+^*$, and S_t is the supremum of the canonical trajectory up to time t.
- $\Gamma_t = \exp\left(-\int_0^t V(X_u)\, du\right)$, where $(X_u)_{u \geq 0}$ is the canonical process, and V is a function from $\mathbb{R}$ to $\mathbb{R}_+$, such that $0 < \int_{-\infty}^{\infty} (1 + |y|) V(y) dy < \infty$.
- $\Gamma_t = e^{\lambda L_t}$, where λ is a real parameter and L_t is the local time at level zero of the canonical process up to time t.
- More generally, $\Gamma_t = e^{\lambda L_t + \mu X_t}$, where λ and μ are real parameters.

We have also studied other cases, which partially generalize the previous examples:

- The penalization by:
$$\Gamma_t = F\left((L_t^y)_{y \in \mathbb{R}}\right),$$
where $(L_t^y)_{y \in \mathbb{R}}$ is the continuous family of the local times of the canonical process, and where F is a functional from $\mathcal{C}(\mathbb{R}, \mathbb{R}_+)$ to $\mathbb{R}_+$, satisfying some technical conditions detailed in [37].
- The penalization by $\Gamma_t := \exp(-\beta H_t)$, where $\beta > 0$ is a parameter, and where
$$H_t := \int_{-\infty}^{\infty} (L_t^y)^2 dy.$$
This example, studied in [38], also corresponds to a functional of the local times, but it does not satisfy the technical assumptions involved in [37]. The quantity H_t represents the *self-intersection local time* of the canonical trajectory, i.e. the time it spends to intersect itself. The interest of this example is its connection with the so-called *Edwards model* (see [17]), which was introduced to describe the trajectory of some long polymer chains.
- A generalization of the penalization by $\Gamma_t = e^{\lambda L_t + \mu X_t}$, studied in [35] and obtained by replacing the Wiener measure $\mathbb{P}$ by the law of the so-called *Brownian spider* (or *Walsh Brownian motion*). This process, introduced by Walsh [71], and studied in more detail by Barlow, Pitman and Yor [5], is informally a Brownian motion on a finite set of half-lines starting at a common point (for two half-lines, we recover the standard Brownian motion).

In general, we note that despite the fact that all these examples of penalization look very different from each other, the existence of the limit measure is proven for each of them (and for many others). However, the behavior of the canonical process under this limit measure changes a lot with the different cases.

In other words, there seems to exist a property of universality, implying that for any family of weights satisfying very general assumptions, the limit measure exists. In order to understand this phenomenon in a better way, we have continued the work of Roynette, Vallois and Yor, and we have obtained several results which are presented below.

Part I of this thesis is divided into three sections. In Section 1, we detail the example of penalization studied in [38], and more generally, the link between the penalizations and some polymer models. In Section 2, we construct a remarkable σ-finite measure which partially explains the universal properties of the Brownian penalizations, and then we define an analog of this measure in the following frameworks: the two-dimensional Brownian motion, the linear diffusions, and the discrete Markov chains. In Section 3, we develop another generalization of this σ-finite measure, which partially covers the results of Section 2, and which involves a particular class of submartingales, called *the class* (Σ). Some of the results we obtain in Section 3 are related to problems in mathematical finance.

1 Penalizations and polymer models

1.1 Penalization with the self-intersection local time: the infinite, one dimensional Edwards model

In [38], we study the penalization of the one-dimensional Wiener measure $\mathbb{P}$ by the following weight:

$$\Gamma_t := \exp\left(-\beta H_t\right),$$

where $\beta > 0$ is a parameter, and where

$$H_t := \int_{-\infty}^{\infty} (L_t^y)^2 dy.$$

Here, $(L_t^y)_{y \in \mathbb{R}, t \geq 0}$ is the continuous family of local times of the canonical process $(X_t)_{t \geq 0}$: informally, one can write

$$H_t = \int_0^t \int_0^t \delta(X_s - X_u)\, ds\, du,$$

where δ denotes the Dirac measure at zero. In other words, the so-called *self-intersection local time* H_t measures the amount of time spent by the canonical trajectory to intersect itself: the more a trajectory intersects itself on the time interval $[0, t]$, the more this trajectory is penalized. The restriction of the canonical trajectory to the interval $[0, t]$, under the probability measure

$$\mathbb{Q}_t := \frac{\Gamma_t}{\mathbb{E}_{\mathbb{Q}_t}[\Gamma_t]} \cdot \mathbb{P},$$

is a stochastic process with finite time horizon of time, which is called the *one-dimensional Edwards model* (see [17]). Initially, such a model has been introduced in physics in order to represent polymers, the penalization of the self-intersections corresponding to the physical repulsion which occurs between two different parts of a long molecular chain. Of course, in dimension one, the Edwards model is less realistic than in dimension two or three (the polymer should be confined on a line), but it is still interesting, and it has already been studied by Westwater in [75], and by van der Hofstad, den Hollander and König in [68] and [69]. The behavior of the canonical trajectory under $\mathbb{Q}_t$ is ballistic, i.e. linear in t: more precisely, for t going to infinity, the law of X_t/t under $\mathbb{Q}_t$ tends to $\frac{1}{2}(\delta_{b^*\beta^{1/3}} + \delta_{-b^*\beta^{1/3}})$, where δ_x is Dirac measure at x and b^* is a universal constant (approximately equal to 1.1). Note that the splitting of the limit law into two Dirac masses comes from the symmetry of the model with respect to the origin: the typical trajectory goes either up or down, in each case with asymptotic speed $b^*\beta^{1/3}$. There is also a central limit theorem: under $\mathbb{Q}_t$, the law of $\frac{|X_t|-b^*\beta^{1/3}t}{\sqrt{t}}$ tends to the distribution of a centered gaussian random variable, with universal variance (approximately equal to 0.4, in particular smaller than one).

In all these results, one only considers Edwards model with a finite time horizon, even if one looks at the behavior of the model for t going to infinity. It is natural to ask if there exists a similar model with infinite time horizon. One way to do that is to seek a limit for the family of measures $(\mathbb{Q}_t)_{t\geq 0}$, when t goes to infinity. The main result of [38] is the following theorem, which shows that the situation is similar to the other penalization examples stated before:

Theorem 1.1 *There exists a unique probability measure $\mathbb{Q}$ such that for all $s \geq 0$, and for all $\Lambda_s \in \mathcal{F}_s := \sigma\{X_u, 0 \leq u \leq s\}$,*

$$\mathbb{Q}_t(\Lambda_s) \underset{t\to\infty}{\longrightarrow} \mathbb{Q}(\Lambda_s).$$

Moreover, there exists a $\mathbb{P}$-martingale $(M_s)_{s\geq 0}$ such that for all $s \geq 0$, $\Lambda_s \in \mathcal{F}_s$,

$$\mathbb{Q}(\Lambda_s) = \mathbb{E}_{\mathbb{P}}[\mathbb{1}_{\Lambda_s} M_s].$$

The martingale $(M_s)_{s\geq 0}$ is explicitly described in [38]. Its expression is quite complicated: in particular, it involves the eigenfunctions of a differential operator related to the semi-group of the two-dimensional Bessel process.

When one knows that the measure $\mathbb{Q}$ exists, it is natural to study the behavior of the canonical process under $\mathbb{Q}$. However, for the moment, we are not able to obtain a result in this direction. In [38], we conjecture that the behavior is ballistic:

Conjecture 1.2 *Under $\mathbb{Q}$, the canonical process is transient, and*

$$\mathbb{Q}(X_t \underset{t\to\infty}{\longrightarrow} +\infty) = \mathbb{Q}(X_t \underset{t\to\infty}{\longrightarrow} -\infty) = 1/2.$$

Moreover, there exist universal strictly positive constants a and σ such that

$$\frac{|X_t|}{t} \underset{t\to\infty}{\longrightarrow} a\,\beta^{1/3}$$

almost surely, and such that the random variable

$$\frac{|X_t| - a\beta^{1/3}t}{\sqrt{t}}$$

converges in law to a centered Gaussian variable of variance σ^2 (the factor $\beta^{1/3}$ comes from the Brownian scaling).

1.2 The Edwards model and its modifications in dimensions 1, 2 and 3

As written above, the Edwards model is more realistic in dimensions two or three than in dimension one. In any case, let us fix the dimension $d \in \{1, 2, 3\}$, the duration $t \geq 0$ of the finite trajectories, a parameter $\beta > 0$, and let us denote by $\mathbb{P}$ the Wiener measure on $\mathcal{C}([0,t], \mathbb{R}^d)$. The d-dimensional Edwards model should be represented by a measure $\mathbb{Q}$ on the space $\mathcal{C}([0,t], \mathbb{R}^d)$, informally defined by

$$\mathbb{Q} := \frac{\exp(-\beta H)}{\mathbb{E}_{\mathbb{P}}[\exp(-\beta H)]} \cdot \mathbb{P},$$

where

$$H = \int_0^t \int_0^t \delta(X_s - X_u)\, ds\, du,$$

for δ equal to the d-dimensional Dirac measure at zero.

In dimension one, H can be written as a function of the local times of X: the construction of the Edwards model is straightforward, and the measure is absolutely continuous with respect to the Wiener measure $\mathbb{P}$ (recall that we only consider finite trajectories). The almost sure properties of the canonical process are the same under $\mathbb{P}$ and under $\mathbb{Q}$: in particular, the quantity of self-intersection of the trajectories is essentially not changed by the penalization.

In dimension two, the situation is more difficult to deal with, since the local times do not exist anymore. Moreover, an informal computation suggests that the quantity H should be infinite almost surely, due to the contribution of the couples (s, u) such that $X_s = X_u$ and $s - u$ is small: the planar Brownian motion has many self-intersections at short-time scale. However, a solution has been found by Varadhan [70], and informally, it consists to subtract an "infinite constant" from H in order to get a quantity which is almost surely finite. This technique has also been developed by Rosen (see [57] and [58]), and by Hu and Yor (see [19]). A possible construction is the following: one proves that there exists almost surely a random function Γ from $\mathbb{R}^2 \backslash \{(0,0)\}$ to $\mathbb{R}_+$, such that for all bounded, measurable functions f from $\mathbb{R}^2$ to $\mathbb{R}$, equal to zero in the neighborhood of zero,

$$\int_0^t \int_0^t f(X_s - X_u)\, ds\, du = \int_{\mathbb{R}^2} f(x)\, \Gamma(x)\, dx.$$

Informally, $\Gamma(x)$ represents the quantity of couples (s, u) such that $X_s - X_u = x$, i.e. the self-intersections "shifted by x". The fact that H is infinite is confirmed by the almost sure infinite limit of $\Gamma(x)$ when x goes to zero. However, there exists a constant $K > 0$ such that $\gamma : x \mapsto \Gamma(x) - K \log(1/||x||)$ can almost surely be extended to a continuous function from $\mathbb{R}^2$ to $\mathbb{R}$: the *self-intersection local time* is then given by $H' := \gamma(0)$. Moreover, since $\mathbb{E}_{\mathbb{P}}[e^{-\beta H'}] < \infty$, the two-dimensional Edwards model can be defined by

$$\mathbb{Q} := \frac{\exp(-\beta H')}{\mathbb{E}_{\mathbb{P}}[\exp(-\beta H')]} \cdot \mathbb{P}.$$

Again, this measure is absolutely continuous with respect to the Wiener measure.

In dimension three, the quantity H is again infinite: moreover, the Varadhan renormalization is not sufficient to give a convergence. The solution has been first obtained by Westwater in [73] and [74], in a long proof involving in an essential way the dyadic splitting of the interval $[0, t]$. A simpler solution has been found by Bolthausen in [9] and later, Albeverio and Zhou [1] have proven that the two constructions give the same version of the three-dimensional Edwards model, i.e. the same probability measure $\mathbb{Q}$. A remarkable phenomenon which occurs here is that this measure $\mathbb{Q}$ is singular with respect to the Wiener measure: the penalization of the self-intersections changes the trajectories in a deeper way in dimension three than in dimension one or two. However, the change is not deep enough to prevent the paths to intersect themselves: more precisely, the fractal dimension of the double points in the same as for the standard three-dimensional Brownian motion, as proven by Zhou in [80].

A question which can now be asked is the following: does there exist a way to modify the Edwards model in dimension one or two, in order to obtain a probability measure which is singular with respect to the Wiener measure? A solution can be found in [36], Chap. 5, where we prove that for $d \in \{1, 2\}$, for fixed $t \geq 0$ and for $\beta > 0$ small enough, one can construct a measure $\widetilde{\mathbb{Q}}$ informally given by

$$\widetilde{\mathbb{Q}} := \frac{\exp\left(-\beta \widetilde{H}\right)}{\mathbb{E}_{\mathbb{P}}\left[\exp\left(-\beta \widetilde{H}\right)\right]},$$

where

$$\widetilde{H} := \int_0^t \int_0^t \frac{\delta(X_s - X_u)}{|s-u|^{(3-d)/2}} \, ds \, du.$$

In other words, the self-intersections are more penalized if they occur at a short-time scale, with a correcting factor $1/|s-u|$ in dimension one and $1/\sqrt{|s-u|}$ in dimension two. The construction we have made involves a general result proven in [73] and used by Westwater to construct the three-dimensional Edwards model. Any reasonable definition of $\widetilde{H}$ should give an infinite value, but it is possible to define, for all $n \geq 0$, a finite random variable $\widetilde{H}_n$ interpreted as follows:

$$\widetilde{H}_n := \int_{R(n)} \frac{\delta(X_s - X_u)}{|s-u|^{(3-d)/2}} \, ds \, du,$$

where $R(n)$ is the set of $(s, u) \in [0, t] \times [0, t]$ such that the $n + 1$ first binary digits of s/t and u/t do not all coincide. For n increasing to infinity, $R(n)$ increases to a set containing almost every element of $[0, t] \times [0, t]$. Moreover, one can show that $\widetilde{H}_n$ has negative exponential moments of any order: one can define a sequence of measures $\widetilde{\mathbb{Q}}_n$ by

$$\widetilde{\mathbb{Q}}_n := \frac{\exp\left(-\beta \widetilde{H}_n\right)}{\mathbb{E}_{\mathbb{P}}\left[\exp\left(-\beta \widetilde{H}_n\right)\right]} \cdot \mathbb{P}.$$

Since $\widetilde{H}_n$ looks like an approximation of $\widetilde{H}$, it is natural to expect that $\widetilde{\mathbb{Q}}_n$ converges, in a certain sense, to a limit measure which can be taken for $\widetilde{\mathbb{Q}}$ and which should define the modified Edwards model in dimension one or two. The main result of [36], Chap. 5 is the proof that such a convergence takes place:

Theorem 1.3 *For $m \geq 0$, let $\mathcal{F}_m$ be the σ-algebra generated by the random variables $X_{kt/2^m}$, for $k \in \{1, 2, \ldots, 2^m\}$. For fixed t, if $\beta > 0$ is small enough, there exists a unique measure $\widetilde{\mathbb{Q}}$ on $\mathcal{C}([0, t], \mathbb{R}^d)$ such that for all $m \geq 0$, and $\Lambda_m \in \mathcal{F}_m$,*

$$\widetilde{\mathbb{Q}}_n(\Lambda_m) \underset{n \to \infty}{\longrightarrow} \widetilde{\mathbb{Q}}(\Lambda_m).$$

Moreover, the measures $\widetilde{\mathbb{Q}}$ obtained for different values of β are pairwise singular, and they are singular with respect to the Wiener measure.

The type of convergence involved in this result is similar to the convergence considered by Roynette, Vallois and Yor, except that the canonical filtration $(\mathcal{F}_t)_{t\geq 0}$ of $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ is replaced by the "dyadic filtration" $(\mathcal{F}_m)_{m\geq 0}$ of $\mathbb{C}([0,t], \mathbb{R}^d)$.

We do not know whether the measure $\widetilde{\mathbb{Q}}$ exists for any $\beta > 0$ but conjecture that the answer is positive, similarly to the case of the three-dimensional Edwards model. It would be interesting to obtain some information on the canonical trajectory under $\widetilde{\mathbb{Q}}$, in particular on its self-intersections. It is known (see Zhou [80]) that for the three-dimensional Edwards model, the fractal dimension of the double points is the same as under the Wiener measure, and we expect that the situation is similar for our model in dimension one or two. In particular, even if we obtain singular measures with respect to the Wiener measure, we do not expect to avoid the self-intersections in a really significant way.

Another perspective of research would be the study of a penalization informally given by the hamiltonian

$$H_\alpha := \int_0^t \int_0^t \frac{\delta(X_s - X_u)}{|s-u|^\alpha} \, ds \, du,$$

for more general values of α. We do not know what happens in this general case, but we suppose that there exist critical values for α: for example, Varadhan renormalization should be sufficient to construct the model if and only if $\alpha < (3-d)/2$. It may also be possible to construct, as in dimension one, a version of the two- or three-dimensional Edwards model defined on the space $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ of infinite trajectories, by proving that the measures associated to the finite-time versions of the Edwards model converge, in the sense of Roynette, Vallois and Yor, when the length of the trajectories goes to infinity. Such a measure may give some information about the long-term behavior of the Edwards model, which is unknown for the moment. It is conjectured that under the Edwards model on $\mathcal{C}([0,t], \mathbb{R}^d)$, the expectation of $||X_t||$ is equivalent to a constant times $t^{\gamma(d)}$, where $\gamma(2) = 3/4$ and $\gamma(3)$ is approximately equal to 0.588.

2 A universal measure related to Brownian penalization, and its generalizations

As written previously, it has been observed that many different examples of penalization give a limit measure when the time parameter goes to infinity. For the moment, we do not know whether it is possible to obtain a really universal result which would explain this phenomenon for all the examples studied up to now. However, it is possible to unify in a single setting a large class of examples where the Wiener measure is penalized by a functional of the local times. The main topic of the current section is to present a still more general result, gathering all the penalization weights which satisfy some conditions of domination, stated below. In each of the examples covered by this framework, the limit measure is absolutely continuous with respect to a universal σ-finite measure $\mathbb{W}$ on the space $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$, deeply related to the Bessel process of dimension three. After describing this measure, we will show that one can construct similar measures in relation with the following processes: the Brownian motion in dimension two, a certain class of linear diffusions, and a certain class of discrete Markov processes. The main results of this section come from our monograph with Roynette and Yor [48].

2.1 The one-dimensional Brownian setting

A possible way to construct the measure $\mathbb{W}$ is to start from a particular example of Brownian penalizations, originally studied by Roynette, Vallois and Yor in [61]. This example is related to the Feynman-Kac formula, recently studied in a quite similar point of view by Lörinczi, Hiroshima and Betz [31].

One obtains a remarkable phenomenon: all the limit measures obtained in the corresponding framework are absolutely continuous with respect to each other. More precisely, the following result holds:

Theorem 2.1 *Let $\mathbb{P}$ be the Wiener measure on $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$, and let V be a measurable function from $\mathbb{R}$ to $\mathbb{R}_+$ such that $0 < \int_{-\infty}^{\infty} (1 + |y|) V(y) dy < \infty$. Then, for $t \geq 0$ going to infinity, the measure*

$$\mathbb{Q}_t^{(V)} := \frac{\exp\left(-\int_0^t V(X_s)\, ds\right)}{\mathbb{E}_{\mathbb{P}}\left[\exp\left(-\int_0^t V(X_s)\, ds\right)\right]} \cdot \mathbb{P}$$

converges, in the usual sense of penalization, to a limit measure $\mathbb{Q}^{(V)}$. Moreover, the equivalence class of $\mathbb{Q}^{(V)}$ does not depend on the choice of the function V: there exists a σ-finite measure $\mathbb{W}$ on $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$, independent of V, singular with respect to $\mathbb{P}$, with infinite total mass, and such that

$$\mathbb{W}\left(\int_0^\infty V(X_s)\, ds = \infty\right) = 0,$$

$$0 < \mathbb{W}\left[\exp\left(-\int_0^\infty V(X_s)\, ds\right)\right] < \infty$$

and

$$\mathbb{Q}^{(V)} = \frac{\exp\left(-\int_0^\infty V(X_s)\, ds\right)}{\mathbb{W}\left[\exp\left(-\int_0^\infty V(X_s)\, ds\right)\right]} \cdot \mathbb{W}.$$

The measure $\mathbb{W}$ defined by this theorem is infinite: in particular, it is not a probability measure. Because of this fact, $\mathbb{W}$ enjoys some properties which may first look surprising: for example, for all $s \geq 0$ and $\Lambda_s \in \mathcal{F}_s := \sigma\{X_u, u \leq s\}$, the quantity $\mathbb{W}(\Lambda_s)$ is infinite if $\mathbb{P}(\Lambda_s) > 0$, and it is equal to zero if $\mathbb{P}(\Lambda_s) = 0$. Consequently, any set with nontrivial measure should involve the whole canonical trajectory $(X_t)_{t \geq 0}$.

The definition of $\mathbb{W}$ given above is not the only one which can be stated: several properties of $\mathbb{W}$ can be used to characterize the measure, and some of them may show more clearly its universal properties.

Theorem 2.2 *The measure $\mathbb{W}$ is the unique σ-finite measure on $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ enjoying the following properties:*

- *The supremum g of the values of $t \geq 0$ such that $X_t = 0$ is finite, $\mathbb{W}$-almost everywhere.*
- *For all $t \geq 0$, and for all bounded, $\mathcal{F}_t$-measurable random variables F_t,*

$$\mathbb{W}[F_t \mathbb{1}_{g \leq t}] = \mathbb{E}_{\mathbb{P}}[F_t |X_t|]$$

Moreover, $\mathbb{W}$ can be disintegrated with respect to L_∞, the total local time at zero (which is finite $\mathbb{W}$-almost everywhere):

$$\mathbb{W} = \int_0^\infty \mathbb{W}_\ell d\ell,$$

where $\mathbb{W}_\ell$ is a probability measure under which the local time at zero is almost surely equal to ℓ. This probability measure can be defined as follows:

- *One denotes by $(Y_t)_{t\geq 0}$ the standard Brownian motion, and τ_ℓ its inverse local time at ℓ, for the level zero.*
- *One denotes by $(Z_t)_{t\geq 0}$ an independent Bessel process of dimension three.*
- *One denotes by ϵ an independent symmetric Bernoulli random variable.*
- *Then, $\mathbb{W}_\ell$ is the law of a process $(U_t)_{t\geq 0}$ obtained by concatenation of $(Y_t)_{t\leq\tau_\ell}$ and $(\epsilon Z_t)_{t\geq 0}$: $U_t = Y_t$ if $t \leq \tau_\ell$ and $U_t = \epsilon Z_{t-\tau_\ell}$ if $t \geq \tau_\ell$.*

There exists also a disintegration of $\mathbb{W}$ with respect to g, the last hitting time of zero:

$$\mathbb{W} = \int_0^\infty \mathbb{W}^{(v)} \frac{dv}{\sqrt{2\pi v}},$$

where $\mathbb{W}^{(v)}$ is a probability measure under which $g = v$ almost surely. The measure $\mathbb{W}^{(v)}$ can be defined as the law of the concatenation of a Brownian bridge on the interval $[0, v]$, and an independent process with the same law as $(\epsilon Z_t)_{t\geq 0}$.

The two disintegrations of $\mathbb{W}$ given in this theorem have, in particular, the following consequences:

Corollary 2.3 *Almost everywhere under $\mathbb{W}$, the canonical process tends either to $+\infty$ or to $-\infty$. The image measure of $\mathbb{W}$ by the total local time of $(X_t)_{t\geq 0}$ at zero is equal to the Lebesgue measure on $\mathbb{R}_+$, and the image of $\mathbb{W}$ by the last hitting time of zero has, at $v \in \mathbb{R}_+$, a density $1/\sqrt{2\pi v}$ with respect to the Lebesgue measure on $\mathbb{R}_+$.*

Note that the range of the trajectory $(X_t)_{t\geq 0}$ is $\mathbb{W}$-almost everywhere an interval of the form $[a, +\infty)$ for $a \leq 0$ or $(-\infty, a]$ for $a \geq 0$. There also exists a disintegration of $\mathbb{W}$ with respect to the bound a, which can be viewed as a dual form of the disintegration with respect to the local time. This duality can be related to Levy's equivalence theorem which links the supremum and the local time of a Brownian motion.

The following result shows another deep link between the measure $\mathbb{W}$ and the Wiener measure $\mathbb{P}$. The intuitive meaning is that the behaviors of $\mathbb{P}$ and $\mathbb{W}$ cannot be distinguished when we only look at a finite part of the trajectories.

Theorem 2.4 *Let $t \geq 0$, let $\mathbb{P}_t$ be the Wiener measure on $\mathcal{C}([0, t], \mathbb{R}_+)$, and let Φ be the application from $\mathcal{C}([0, t], \mathbb{R}) \times \mathcal{C}(\mathbb{R}_+, \mathbb{R})$ to $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ such that $\Phi((X_s)_{s\leq t}, (Y_u)_{u\geq 0}) = (Z_u)_{u\geq 0}$ where $Z_u = X_u$ for $u \leq t$ and $Z_u = X_t + (Y_{u-t} - Y_0)$ for $u \geq t$. Then, the image of $\mathbb{P}_t \times \mathbb{W}$ by Φ is equal to $\mathbb{W}$.*

Informally, one obtains a process following the σ-finite measure $\mathbb{W}$ by concatening a Brownian motion stopped at a given time, and an "independent" process following $\mathbb{W}$ (of course, to give a meaning to this description, one has to take care of the fact that $\mathbb{W}$ is not a finite measure, and in particular, not a probability measure). Note that this result is also true if we replace the measure $\mathbb{W}$ by any linear combination of the three following measures:

- The Wiener measure $\mathbb{P}$.
- The restriction of $\mathbb{W}$ to the trajectories tending to $+\infty$ at infinity.
- The restriction of $\mathbb{W}$ to the trajectories tending to $-\infty$ at infinity.

We do not know if there are other σ-finite measures enjoying this property.

As stated in Theorem 2.1, a large class of limit measures obtained by penalization are absolutely continuous with respect to $\mathbb{W}$. This result can be deeply extended, as follows. Let $(\Gamma_t)_{t\geq 0}$ be a family of functionals from $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ to $\mathbb{R}_+$. We say that $(\Gamma_t)_{t\geq 0}$ is in the class $\mathcal{C}$ when the following holds:

- For all $t \geq 0$, Γ_t is $\mathcal{F}_t$-measurable.
- Γ_0 is equal to a deterministic constant $C > 0$, and $(\Gamma_t)_{t\geq 0}$ is a decreasing process: in particular, $0 \leq \Gamma_t \leq C$.
- There exists $a \geq 0$ such that for every $t \geq \sigma_a$, with

$$\sigma_a := \sup\{t \geq 0, |X_t| \leq a\},$$

we have $\Gamma_t = \Gamma_{\sigma_a} = \Gamma_\infty$.

- $0 < \mathbb{W}(\Gamma_\infty) = \mathbb{W}(\Gamma_{\sigma_a}) < \infty$.

Note that the class $\mathcal{C}$ is very large. For example, let $L_t^{a_1}, L_t^{a_2}, \dots, L_t^{a_r}$ denote the local times at time t and at fixed levels $a_1, \dots, a_r$, let $V_1, \dots, V_p$ be bounded, measurable functions from $\mathbb{R}$ to $\mathbb{R}$, with compact support, and let S_t be the supremum of $(X_s)_{s\leq t}$. If ϕ is a Borel function from $\mathbb{R}^{r+p+1}$ to $\mathbb{R}_+$, decreasing in each of its arguments, equal to zero if the last argument is large enough, and strictly positive at $(\epsilon, \epsilon, \dots, \epsilon)$ for $\epsilon > 0$ small enough, then the family $(\Gamma_t)_{t\geq 0}$ of functionals defined by

$$\Gamma_t := \phi\left(L_t^{a_1}, L_t^{a_2}, \dots, L_t^{a_r}, \exp\left(-\int_0^t V_1(X_s)\, ds\right), \dots, \exp\left(-\int_0^t V_p(X_s)\, ds\right), S_t\right)$$

is in the class $\mathcal{C}$. Therefore, the class $\mathcal{C}$ includes several examples of penalization discussed before, which look very different from each other: penalization by a functional of the local times, penalization by the function of the unilateral supremum, Feynman-Kac penalization studied by Roynette, Vallois and Yor in [61], etc. However, the technical conditions involved are generally more restrictive in the framework of the class $\mathcal{C}$ than in the separate study of each particular case (for example, contrarily to the setting of [61], $V_1, \dots, V_p$ need to have a compact support in order to give a functional in the class $\mathcal{C}$). Moreover, there are also penalization weights which do not enter at all in the present setting, and for which the limit measure is not absolutely continuous with respect to $\mathbb{W}$. For example, it is the case when we penalize with $\Gamma_t = e^{\gamma L_t}$ for $\gamma > 0$, and even if we have not yet found a rigorous proof, we expect that it is also the case for the one-dimensional Edwards model:

$$\Gamma_t := \exp\left(-\beta \int_{-\infty}^{\infty} (L_t^y)^2 dy\right).$$

The main penalization result of [48] is the following:

Theorem 2.5 *Let* $(\Gamma_t)_{t\geq 0}$ *be a family of functionals in the class* $\mathcal{C}$*. Then, the family* $(\mathbb{Q}_t)_{t\geq 0}$ *of probability measures given by*

$$\mathbb{Q}_t := \frac{\Gamma_t}{\mathbb{E}_{\mathbb{P}}[\Gamma_t]} \cdot \mathbb{P}$$

converges to a limit measure $\mathbb{Q}$ *in the sense of penalization. Moreover,* $\mathbb{Q}$ *is absolutely continuous with respect to* $\mathbb{W}$*: more precisely, one has*

$$\mathbb{Q} = \frac{\Gamma_\infty}{\mathbb{W}[\Gamma_\infty]} \cdot \mathbb{W}.$$

The fact that such a large class of penalization results involves the same measure $\mathbb{W}$ explains some similarities which were not clearly explained before: for example, the intervention of the Bessel process of dimension 3 in the description of several limiting processes previously obtained by Roynette, Vallois and Yor.

Because of our general penalization result, it is interesting to obtain some information about the finite measures (in particular the probability measures) which are absolutely continuous with respect to $\mathbb{W}$. The following theorem gives some of their properties:

Theorem 2.6 *Let* Γ *be a functional from* $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ *to* $\mathbb{R}_+$*, integrable with respect to* $\mathbb{W}$*. Then,* $\mathbb{W}^{(\Gamma)} := \Gamma.\mathbb{W}$ *is a finite measure on* $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$*, and it is a probability measure if and only if* $\mathbb{W}[\Gamma] = 1$*. Moreover, there exists a nononegative* $(\mathbb{P}, (\mathcal{F}_t)_{t\geq 0}))$*-martingale* $(M_t(\Gamma))_{t\geq 0}$ *such that* $M_t(\Gamma)$ *is the density, with respect to* $\mathbb{P}$*, of the restriction of* $\mathbb{W}^{(\Gamma)}$ *to the* σ*-algebra* $\mathcal{F}_t$*. In other words, for all* $t \geq 0$*, and for all bounded,* $\mathcal{F}_t$*-measurable functionals* F_t*, one has:*

$$\mathbb{W}^{(\Gamma)}[F_t] = \mathbb{W}[\Gamma \cdot F_t] = \mathbb{E}_{\mathbb{P}}[M_t(\Gamma) \cdot F_t].$$

The martingale $(M_t(\Gamma))_{t\geq 0}$ *tends* $\mathbb{P}$*-almost surely to zero when* t *goes to infinity: in particular, it is not uniformly integrable.*

Note that there exist some functionals Γ for which the martingale $(M_t(\Gamma))_{t\geq 0}$ can be explicitly computed. For example, in relation with the penalization by a function of the supremum, studied in [60], one can prove that

$$M_t(\Gamma) = \psi(S_t)(S_t - X_t) + \int_{S_t}^{\infty} \psi(y)\, dy,$$

if $\Gamma = \psi(S_\infty)\mathbb{1}_{S_\infty < \infty}$, for an integrable function ψ from $\mathbb{R}_+$ to $\mathbb{R}_+$. One obtains similar results with other examples of penalization.

2.2 The two-dimensional Brownian setting

In [63], Roynette, Vallois and Yor study some examples of Brownian penalizations in dimension larger than or equal to 2. One of the important examples is the analog in dimension 2 of the Feynman-Kac penalizations studied in [61]. Similarly to the one-dimensional case, the following result holds:

Theorem 2.7 *Let* $\mathbb{P}^{(2)}$ *be the Wiener measure on* $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^2)$*, and let* V *be a measurable function from* $\mathbb{R}^2$ *to* $\mathbb{R}_+$*, bounded, with compact support and such that* $\int_{\mathbb{R}^2} V(y)dy > 0$*. Then, for* $t \geq 0$

going to infinity, the measure

$$\mathbb{Q}_t^{(V)} := \frac{\exp\left(-\int_0^t V(X_s)\,ds\right)}{\mathbb{E}_{\mathbb{P}^{(2)}}\left[\exp\left(-\int_0^t V(X_s)\,ds\right)\right]} \cdot \mathbb{P}^{(2)}$$

converges, in the usual sense of penalization, to a limit measure $\mathbb{Q}^{(V)}$*. Moreover, the equivalence class of* $\mathbb{Q}^{(V)}$ *does not depend on the choice of the function* V*: there exists a* σ*-finite measure* $\mathbb{W}^{(2)}$ *on* $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^2)$*, independent of* V*, singular with respect to* $\mathbb{P}^{(2)}$*, with infinite total mass, and such that*

$$\mathbb{W}^{(2)}\left(\int_0^\infty V(X_s)\,ds = \infty\right) = 0,$$

$$0 < \mathbb{W}^{(2)}\left[\exp\left(-\int_0^\infty V(X_s)\,ds\right)\right] < \infty$$

and

$$\mathbb{Q}^{(V)} = \frac{\exp\left(-\int_0^\infty V(X_s)\,ds\right)}{\mathbb{W}^{(2)}\left[\exp\left(-\int_0^\infty V(X_s)\,ds\right)\right]} \cdot \mathbb{W}^{(2)}.$$

The definition of $\mathbb{W}^{(2)}$ deduced from this theorem is, as in dimension one, strongly depending on a particular family of penalization weights. It is then interesting to find a more intrinsic definition. The characterization given at the beginning of Theorem 2.2 can be generalized in dimension 2, as follows:

Theorem 2.8 *The measure* $\mathbb{W}^{(2)}$ *is the unique* σ*-finite measure on* $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^2)$ *enjoying the following properties:*

- *The supremum* g *of the values of* $t \geq 0$ *such that* $||X_t|| \leq 1$ *is finite,* $\mathbb{W}^{(2)}$*-almost everywhere.*
- *For all* $t \geq 0$*, and for all bounded,* $\mathcal{F}_t$*-measurable random variables* F_t*,*

$$\mathbb{W}[F_t \mathbb{1}_{g \leq t}] = \frac{1}{\pi}\mathbb{E}_{\mathbb{P}}[F_t \log_+(||X_t||)],$$

  where $\log_+$ *denotes the positive part of the logarithm.*

Note that the circle of center zero and radius one plays a particular role in Theorem 2.8. This partially comes from the arbitrary choice of g.

There also exists a path description of the measure $\mathbb{W}^{(2)}$, obtained by splitting the trajectory into two pieces: before and after time g. However, this description is more difficult to state than its equivalent in dimension one, given in Theorem 2.2. We need to define the following objects: the local time at the unit circle, its right-continuous inverse, and a two-dimensional random process $(V_t)_{t\geq 0}$, playing a similar role as the signed three-dimensional Bessel process $(\epsilon Z_t)_{t\geq 0}$ in dimension one (see Theorem 2.2).

For $t \in \mathbb{R}_+$ or $t = +\infty$, the local time $(L_t^{(C)})_{t\geq 0}$ at the unit circle is defined as follows:

$$L_t^{(C)} = \lim_{\epsilon \to 0} \frac{1}{2\pi\epsilon} \int_0^t \mathbb{1}_{1-\epsilon \leq ||X_s|| \leq 1+\epsilon} ds,$$

if the corresponding limit exists ($\mathbb{P}^{(2)}$-almost surely and $\mathbb{W}^{(2)}$-almost everywhere, it is the case for all $t \in \mathbb{R}_+$). The right-continuous inverse of $L_t^{(C)}$ is given by $(\tau_\ell^{(C)})_{\ell \geq 0}$, where

$$\tau_\ell^{(C)} := \inf\{t \geq 0, L_t^{(C)} > \ell\}.$$

The process $(V_t)_{t\geq 0}$ is constructed as follows:

- One considers a Brownian motion $(B_u)_{u\geq 0}$ and an independent three-dimensional Bessel process $(Z_u)_{u\geq 0}$.
- One defines the clock process $(H_t)_{t\geq 0}$ by

$$H_t := \inf\{u \geq 0, \int_0^u e^{2Z_s} ds = t\}.$$

- Then, by identifying $\mathbb{R}^2$ with $\mathbb{C}$, one can define:

$$V_t = \exp(Z_{H_t} + iB_{H_t}).$$

Now, it is possible to give a full analog of Theorem 2.2:

Theorem 2.9 *The measure $\mathbb{W}^{(2)}$ can be disintegrated with respect to $L_\infty^{(C)}$, the total local time at the unit circle, which is, $\mathbb{W}^{(2)}$-almost everywhere, finite and well-defined:*

$$\mathbb{W}^{(2)} = \int_0^\infty \mathbb{W}_\ell^{(2)} d\ell,$$

where $\mathbb{W}_\ell^{(2)}$ is a probability measure under which the local time is almost surely equal to ℓ. This probability measure can be defined as follows:

- *One considers the two-dimensional process $(V_t)_{t\geq 0}$, whose law is described just above.*
- *One denotes by $(Y_t)_{t\geq 0}$ an independent standard, two-dimensional Brownian motion, and $\tau_\ell^{(C)}$ its inverse local time at ℓ, on the unit circle.*
- *Then, $\mathbb{W}_\ell^{(2)}$ is the law of a process $(U_t)_{t\geq 0}$ obtained from a concatenation of $(Y_t)_{t\leq \tau_\ell^{(C)}}$ and $(V_t)_{t\geq 0}$: $U_t = Y_t$ if $t \leq \tau_\ell^{(C)}$ and by identifying $\mathbb{R}^2$ and $\mathbb{C}$, $U_t = Y_{\tau_\ell^{(C)}} V_{t-\tau_\ell^{(C)}}$ if $t \geq \tau_\ell$.*

There exists also a disintegration of $\mathbb{W}^{(2)}$ with respect to g, the last hitting time of the unit circle:

$$\mathbb{W}^{(2)} = \int_0^\infty \mathbb{W}^{(2,v)} \frac{e^{-1/v}}{2\pi v} dv,$$

where $\mathbb{W}^{(2,v)}$ is a probability measure under which $g = v$ almost surely. The measure $\mathbb{W}^{(2,v)}$ can be defined as follows:

- *One considers the process $(V_t)_{t\geq 0}$ above.*
- *One takes an independent uniform variable u on the unit circle.*

- *One defines an independent two-dimensional Brownian bridge $(P_s)_{s\leq v}$ of duration v, starting at 0 and ending at 1.*
- *Then, $\mathbb{W}^{(2,v)}$ is the law of a process $(U_t)_{t\geq 0}$ such that $U_t = uP_t$ for $t \leq v$, and $U_t = uV_{t-v}$ for $t \geq v$.*

Similarly to the one-dimensional setting, there are strong links between the measure $\mathbb{W}^{(2)}$, the Wiener measure $\mathbb{P}^{(2)}$ and a particular class of martingales. By doing some straightforward changes, one can state analogs of Theorem 2.4 and Theorem 2.6. It is also natural to expect that there exists a general penalization result similar to Theorem 2.5, even if this problem has not been studied in detail.

2.3 The setting of linear diffusions

In the one-dimensional Brownian setting, the measure $\mathbb{W}$ can be constructed, by concatenation, from a Brownian motion and a Bessel process of dimension 3. This Bessel process can be informally interpreted as a Brownian motion conditioned not to reach zero: of course, such a conditioning is not rigorous since the corresponding event has probability zero. One may now ask if such a construction can be adapted to a more general class of one-dimensional processes, which can be informally conditioned not to vanish, and for which the local time at zero is well-defined.

The setting studied in [48] was first stated by Salminen, Vallois and Yor in [65]. It can be described as follows:

- One considers, for all $x \in \mathbb{R}_+$, a measure $\mathbb{P}_x$ on the space $\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+)$, under which the canonical process $(X_t)_{t\geq 0}$ is a linear diffusion starting from x (note that here, all the trajectories take nonnegative values).
- One supposes that zero is an instantaneously reflecting barrier for this diffusion, whose infinitesimal generator $\mathcal{G}$ is given, for $y > 0$, by the formula
$$\mathcal{G}f(y) = \frac{d}{dm}\frac{d}{dS} f(y),$$
where the scale function S is a continuous, strictly increasing function from $\mathbb{R}_+$ to $\mathbb{R}_+$, such that $S(0) = 0$ and $S(y) \to \infty$ when $y \to \infty$, and where the speed measure m is carried by $\mathbb{R}_+^*$.
- The semi-group of the diffusion admits a density with respect to the measure m: for $t, x, y > 0$,
$$\mathbb{P}_x(X_t \in dy) = p(t,x,y)\, m(dy),$$
where p is continuous with respect to the three variables, and symmetric in the two last variables ($p(t,x,y) = p(t,y,x)$). One also defines the semi-group $\widehat{p}$ of the diffusion killed at its first hitting time of zero:
$$\mathbb{P}_x(X_t \in dy, \forall s \leq t, X_s > 0) = \widehat{p}(t,x,y)\, m(dy).$$
- Almost surely under $\mathbb{P}_x$, one can define a jointly continuous family $(L_t^y)_{t\geq 0, y\geq 0}$ of local times of $(X_t)_{t\geq 0}$, satisfying the density occupation formula: for all Borel functions h from $\mathbb{R}_+$ to $\mathbb{R}_+$, and for all $t \geq 0$,
$$\int_0^t h(X_s)ds = \int_0^\infty h(y) L_t^y m(dy).$$

- Focusing on the level zero and setting $L_t := L_t^0$, one can show that $(S(X_t) - L_t)_{t\geq 0}$ is a martingale under $\mathbb{P}_x$, and one can define the right-continuous inverse local time $(\tau_\ell)_{\ell\geq 0}$ at level zero by

$$\tau_\ell := \inf\{t \geq 0, L_t > \ell\}.$$

From the probability measures $(\mathbb{P}_x)_{x\in\mathbb{R}_+}$ given here, it is then possible to construct another family of measures $(\mathbb{P}_x^+)_{x\in\mathbb{R}_+}$, informally by conditioning with the event that the trajectories do not vanish. The rigorous construction is obtained as follows:

- For $x > 0$, $\mathbb{P}_x^+$ is obtained from $\mathbb{P}_x$ by killing the trajectories when they reach zero, and then by making a Doob h-transform with the scale function S: for all $t \geq 0$ and for any bounded, $\mathcal{F}_t$-measurable functional F_t,

$$\mathbb{P}_x^+[F_t] = \frac{1}{S(x)} \mathbb{E}_{\mathbb{P}_x}\left[F_t S(X_t) \mathbb{1}_{\forall s\leq t, X_s>0}\right].$$

- For $t, x > 0$, the law of X_t under $\mathbb{P}_x^+$ has density $y \mapsto p^+(t,x,y)$ with respect to the speed measure m, where

$$p^+(t,x,y) = \frac{S(y)}{S(x)} \widehat{p}(t,x,y)$$

- The function p^+, defined for $t, x > 0$ and $y \geq 0$, can be extended to the case $x = 0$ by continuity, which defines the semi-group of a diffusion, whose law is equal to $\mathbb{P}_x^+$ for $x > 0$.

- The measure $\mathbb{P}_0^+$ is then defined by starting this diffusion at level zero.

Similarly to the description of the measure $\mathbb{W}$ given in Theorem 2.2, one can construct a σ-finite measure on $\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+)$ from the measures $\mathbb{P}_0$ and $\mathbb{P}_0^+$, by concatenation of the trajectories. One obtains the following result:

Theorem 2.10 *There exists a unique σ-finite measure $\mathbb{W}_0$ on $\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+)$ which satisfies the following properties:*

- *The supremum g of the values of $t \geq 0$ such that $X_t = 0$ is finite, $\mathbb{W}_0$-almost everywhere.*

- *For all $t \geq 0$, and for all bounded, $\mathcal{F}_t$-measurable random variables F_t,*

$$\mathbb{W}_0[F_t \mathbb{1}_{g\leq t}] = \mathbb{E}_{\mathbb{P}_0}[F_t S(X_t)].$$

Moreover, $\mathbb{W}_0$ can be disintegrated with respect to L_∞, the total local time at zero (which is finite $\mathbb{W}_0$-almost everywhere):

$$\mathbb{W}_0 = \int_0^\infty \mathbb{W}_{0,\ell} d\ell,$$

where $\mathbb{W}_{0,\ell}$ is a probability measure under which the local time at zero is almost surely equal to ℓ. This probability measure is the law of a process obtained by concatenation of a diffusion with law $\mathbb{P}_0$, stopped at the inverse local time τ_ℓ, and an independent diffusion with law $\mathbb{P}_0^+$.

There exists also a disintegration of $\mathbb{W}_0$ with respect to g, the last hitting time of zero:

$$\mathbb{W}_0 = \int_0^\infty \mathbb{W}_0^{(v)} \, p(v,0,0) \, dv,$$

where $\mathbb{W}_0^{(v)}$ is a probability measure under which $g = v$ almost surely. If $\mathbb{P}_0^{(v)}$ is the law of the bridge of $\mathbb{P}_0$ on the interval $[0, v]$ (informally, a process with law $\mathbb{P}_0$, stopped at v and conditioned to vanish at this time), then the measure $\mathbb{W}^{(v)}$ can be defined as the law of the concatenation of two independent processes with respective distributions $\mathbb{P}_0^{(v)}$ and $\mathbb{P}_0^+$.

The relation between $\mathbb{W}_0$ and $\mathbb{P}_0$ is similar to the relation between $\mathbb{W}$ and the Wiener measure on $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$. Note that by translating the canonical trajectory, one obtains, from $\mathbb{W}$, a family of σ-finite measures, each of them starting from a different point of $\mathbb{R}$. In the present situation, one can also obtain an analog of $\mathbb{W}_0$ which starts at any point $x \in \mathbb{R}_+$, but the construction is slightly more difficult than for the Brownian setting, since the semi-group of the diffusion is not invariant by translation.

Theorem 2.11 *For all $x \in \mathbb{R}$, there exists a unique σ-finite measure $\mathbb{W}_x$ on $\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+)$ which satisfies the following properties:*

- *The supremum g of the values of $t \geq 0$ such that $X_t = 0$ is finite, $\mathbb{W}_x$-almost everywhere.*
- *For all $t \geq 0$, and for all bounded, $\mathcal{F}_t$-measurable random variables F_t,*
$$\mathbb{W}_x[F_t \mathbb{1}_{g \leq t}] = \mathbb{E}_{\mathbb{P}_x}[F_t S(X_t)]$$

One has the disintegration:
$$\mathbb{W}_x = S(x)\, \mathbb{P}_x^+ + \int_0^\infty \mathbb{W}_{x,\ell}\, d\ell,$$
where $\mathbb{W}_{x,\ell}$ is the law of the concatenation of a process with law $\mathbb{P}_x$, stopped at its first hitting time of zero, and an independent process with law $\mathbb{W}_{0,\ell}$.

The fact that one can construct a measure $\mathbb{W}_x$ for all $x \geq 0$ is important, since it allows to state an analog of Theorem 2.4, proving that $\mathbb{W}_x$ looks like $\mathbb{P}_x$ if we only consider the beginning of the trajectories. The following result can be deduced from the characterization of $\mathbb{W}_x$ given in Theorem 2.11:

Theorem 2.12 *Let $x \geq 0$, $t > 0$, and for $y > 0$, let $\mathbb{P}_{x,y}^{(t)}$ be the law of the bridge of $\mathbb{P}_x$ with duration t and terminal value $X_t = y$. Then, one can disintegrate $\mathbb{W}_x$ with respect to the value of X_t:*
$$\mathbb{W}_x = \int_0^\infty \mathbb{W}_{x,y}^{(t)}\, p(t, x, y)\, dy,$$
where $\mathbb{W}_{x,y}^{(t)}$ is obtained from $\mathbb{P}_{x,y}^{(t)}$ and $\mathbb{W}_y$ by concatenation of the trajectories.

Intuitively, the meaning of Theorem 2.12 is the following: one obtains $\mathbb{W}_x$ by concatenation of a diffusion of law $\mathbb{P}_x$ stopped at time t, and a trajectory "following the measure $\mathbb{W}_{X_t}$" (of course, this last expression is informal).

One can also relate the measures $(\mathbb{W}_x)_{x \geq 0}$ to a certain class of martingales, as in Theorem 2.6, and to penalization results. However, we have not proven a really general theorem of penalization which can be applied to the present setting.

On the other hand, there are some examples of diffusions which are studied in detail in [48]. The first one corresponds to the case where $\mathbb{P}_x$ is the law of a Bessel process of dimension d, for $0 < d < 2$. Under these assumptions, one obtains, for $\mathbb{P}_x^+$, a Bessel process of dimension $4 - d$. If $d = 1$, one recovers the Brownian setting, after taking absolute values everywhere (recall that one only considers trajectories in $\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+)$). It is also possible to do some explicit computation in the case where $\mathbb{P}_x$ is the law of a stable process.

2.4 The Markov chains setting

In this last setting, we focus our study on discrete processes. This restriction allows to start from a very large class of recurrent Markov chains, whereas the previous constructions of universal σ-finite measures were only concerning very particular stochastic processes (Brownian motion and linear diffusions). The intuitive meaning of the σ-finite measures we will construct is the following: in a sense which has to be made precise, we "condition recurrent Markov chains to become transient". This point of view was also previously available: for example, the measure $\mathbb{W}$ can be considered as the "law" of a Brownian motion "conditioned to tend to $+\infty$ or $-\infty$ at infinity" (for a discussion on this point of view, and more generally, on the main properties of the linear diffusions, see Knight [28]).

In order to be more precise, let us define rigorously the objects studied here, as in [48]. We denote by E a countable set, $(X_n)_{n\geq 0}$ the canonical process on $E^{\mathbb{N}}$, $(\mathcal{F}_n)_{n\geq 0}$ its natural filtration, and $\mathcal{F}_\infty$ the σ-algebra generated by $(X_n)_{n\geq 0}$. We define the family $(\mathbb{P}_x)_{x\in E}$ of probability measures on $E^{\mathbb{N}}$ corresponding to a Markov chain whose transition probabilities are denoted $p_{y,z}$ for all $y, z \in E$. The only assumptions made on these probability measures are the following:

- For all $x \in E$, the set of $y \in E$ such that $p_{x,y} > 0$ is finite, i.e. the graph associated to the Markov chain is locally finite.
- For all $x, y \in E$, there exists $n \geq 0$ such that $\mathbb{P}_x(X_n = y) > 0$, i.e. the graph of the Markov chain is connected.
- For all $x \in E$, the canonical process is recurrent under $\mathbb{P}_x$.

In order to "condition $\mathbb{P}_x$ so that $(X_n)_{n\geq 0}$ becomes transient", we need to consider some kind of h-transforms of the Markov chain, as in the examples studied above: the Bessel process of dimension 3 can be deduced from a Brownian motion by a h-transform, and there exists a similar link between $\mathbb{P}_x$ and $\mathbb{P}_x^+$ in the setting of linear diffusions. The first step of our construction is to fix a function which is nonnegative, and harmonic everywhere except at one particular point. More precisely, we consider a point x_0 and a function ϕ from E to $\mathbb{R}$ satisfying the following conditions:

- $\phi(x) \geq 0$ for all $x \in E$, $\phi(x_0) = 0$, and ϕ is not identically zero.
- For all $x \neq x_0$, $\sum_{y\in E} p_{x,y}\, \phi(y) = \mathbb{E}_{\mathbb{P}_x}[\phi(X_1)] = \phi(x)$.

The equivalent of the function ϕ in the one-dimensional Brownian setting is the absolute value, and in the setting of linear diffusions, it is the scale function S. Note that there is no reason for ϕ to be unique, even up to a multiplicative constant.

For a given choice of ϕ, we can now construct a family $(\mathbb{Q}_x)_{x\in E}$ of σ-finite measures, which will enjoy similar properties as the measures $(\mathbb{W}_x)_{x\in E}$ constructed above in the setting of linear diffusions. The measure $\mathbb{Q}_x$ is carried by the trajectories starting at x. The construction can be made in the following way:

- For $r \in (0, 1)$, one defines a function ψ_r from E to $\mathbb{R}_+$ by

$$\psi_r(x) := \frac{r}{1-r}\, \mathbb{E}_{\mathbb{P}_{x_0}}[\phi(X_1)] + \phi(x).$$

- For $y \in E$ and $k \geq -1$, one defines the local time L_k^y at point y and time k by

$$L_k^y := \sum_{0 \leq m \leq k} \mathbb{1}_{X_m = y},$$

in particular $L_{-1}^y = 0$ and $L_0^y = \mathbb{1}_{X_0 = y}$.

- For every $x \in E$, the process $\left(\psi_r(X_n)\, r^{L_{n-1}^{x_0}}\right)_{n \geq 0}$ is a martingale under $\mathbb{P}_x$, and one can construct a finite measure $\mu_x^{(r)}$ such that the density, with respect to $\mathbb{P}_x$, of its restriction to $\mathcal{F}_n$ is equal to $\psi_r(X_n)\, r^{L_{n-1}^{x_0}}$.

- The total local time $L_\infty^{x_0}$ of the canonical process is finite, $\mu_x^{(r)}$-almost everywhere, and then one can define a σ-finite measure $\mathbb{Q}_x^{(r)}$ by

$$\mathbb{Q}_x^{(r)} := r^{-L_\infty^{x_0}} \cdot \mu_x^{(r)}.$$

- The measure $\mathbb{Q}_x^{(r)}$ is independent of the choice of $r \in (0, 1)$ made above, and then for all $x \in E$, one can state $\mathbb{Q}_x := \mathbb{Q}_x^{(r)}$.

With the construction presented here, it is not clear that $\mathbb{Q}_x$ enjoys the same properties as the σ-finite measures constructed previously. The following result proves the similarity of the situations:

Theorem 2.13 *For $n \geq 0$, let F_n be a bounded, $\mathcal{F}_n$-measurable functional, and let g_{x_0} be the last hitting time of x_0 by the canonical process. Then, for all $x \in E$, the measure $\mathbb{Q}_x$ satisfies the following equality:*

$$\mathbb{Q}_x \left[F_n \mathbb{1}_{g_{x_0} < n}\right] = \mathbb{E}_{\mathbb{P}_x} \left[F_n \,\phi(X_n)\right].$$

Moreover, the random time g_{x_0} is finite, $\mathbb{Q}_x$-almost everywhere: more generally, the canonical process is transient under $\mathbb{Q}_x$.

It is also possible to describe the canonical process under $\mathbb{Q}_x$ by splitting the trajectory at time g_{x_0}, and by disintegrating the measure with respect to the total local time at x_0. The decomposition which is obtained involves the three following measures:

- The restriction $\mathbb{Q}_x^{[x_0]}$ of $\mathbb{Q}_x$ to the trajectories which do not hit x_0.
- The restriction $\widetilde{\mathbb{Q}}_{x_0}$ of $\mathbb{Q}_{x_0}$ to the trajectories which do not return to x_0.
- For $k \geq 1$, the law $\mathbb{P}_x^{\tau_k^{(x_0)}}$ of a Markov chain which follows $\mathbb{P}_x$, and which is stopped at the k-th hitting time of x_0.

With this notation, we obtain the following result:

Theorem 2.14 *One can write the measure $\mathbb{Q}_x$ as a sum of finite measures:*

$$\mathbb{Q}_x = \sum_{k \geq 0} \mathbb{Q}_{x,k},$$

where the total local time at x_0 is, for all $k \geq 0$, equal to k, $\mathbb{Q}_{x,k}$-almost everywhere. Moreover, $\mathbb{Q}_{x,0} = \mathbb{Q}_x^{[x_0]}$ and for all $k \geq 1$, $\mathbb{Q}_{x,k}$ is the image of $\mathbb{P}_x^{\tau_k^{(x_0)}} \times \widetilde{\mathbb{Q}}_{x_0}$ by the concatenation of the canonical trajectories.

Similarly to the previous cases (see Theorems 2.4 and 2.6), there are some direct links between $(\mathbb{Q}_x)_{x\in E}$, $(\mathbb{P}_x)_{x\in E}$ and a particular class of martingales. In the present setting, the main result is the following:

Theorem 2.15 *Let Γ be a nonnegative, $\mathcal{F}_\infty$-measurable functional, integrable with respect to $\mathbb{Q}_x$ for all $x \in E$. For $y_0, \dots, y_n \in E$, let us define the quantity*

$$M(\Gamma, y_0, \dots, y_n) := \mathbb{Q}_{y_n}[\Gamma(y_0, y_1, \dots, y_n = X_0, X_1, X_2, \dots)].$$

Then, for all $x \in E$, $n \geq 0$, and for every bounded, $\mathcal{F}_n$-measurable functional F_n, one has

$$\mathbb{Q}_x[\Gamma \cdot F_n] = \mathbb{E}_{\mathbb{P}_x}[M(\Gamma, X_0, \dots, X_n) \cdot F_n],$$

where $(M(\Gamma, X_0, \dots, X_n))_{n\geq 0}$ is a $\mathbb{P}_x$-martingale.

As we have seen before, the family $(\mathbb{Q}_x)_{x\in E}$ of σ-finite measures depends on the choice of a point $x_0 \in E$ and a function ϕ. In fact, the choice of x_0 is not so important as it may seem at first sight. Indeed, if we consider any point $y_0 \in E$, and if we define the function $\phi^{[y_0]}$ from E to $\mathbb{R}_+$ by $\phi^{[y_0]}(x) := \mathbb{Q}_x(L_\infty^{y_0} = 0)$, then our construction can be made by taking the point y_0 and the function $\phi^{[y_0]}$, instead of the point x_0 and the function ϕ, and the σ-finite measures which are obtained in this way are not changed. Note also that if we take $y_0 = x_0$, the function $\phi^{[x_0]}$ is precisely equal to ϕ: in other words, $\phi(x) = \mathbb{Q}_x(L_\infty^{x_0} = 0)$.

In [48], we detail several examples for which the construction described above can be applied:

- The standard random walk, which gives a discrete version of the measure $\mathbb{W}$: in fact, we can take two linearly independent choices for the function ϕ, giving two families of σ-finite measures for which the canonical trajectory tends respectively to $+\infty$ and $-\infty$.

- A random walk on the set of nonnegative integers, which is attracted by zero as soon as it becomes strictly positive: in this case, the attraction becomes a repulsion when $\mathbb{P}_x$ is changed to $\mathbb{Q}_x$.

- A random walk on an infinite binary tree, for which one can obtain many different families of σ-finite measures $(\mathbb{Q}_x)_{x\in E}$, depending on the choice of the function ϕ. The linearly independent families we have found are indexed by the leaves of the tree, which form an uncountable set.

We have also been able to construct $(\mathbb{Q}_x)_{x\in E}$ when $(\mathbb{P}_x)_{x\in E}$ corresponds to the standard random walk on $\mathbb{Z}^2$.

3 A universal measure for general submartingales of class (Σ)

In the previous section, we have observed that, as studied in detail in [48], the same kind of σ-finite measures can be constructed in several contexts, some of them looking very different from each other. However, we still do not have a global result which explains simultaneously all the constructions described above. Moreover, there are other situations for which we expect that similar constructions are possible, even if we do not know how to deal with all the technical details: for example, if one starts with a general Markov process in continuous time.

In all the situations for which we are able to construct "universal σ-finite measures", we have observed that these measures are strongly related to a certain class of martingales. Because of this observation, it is natural to expect that a generalization of our previous results is possible if we focus on a class of stochastic processes which has some suitable links with the martingale property.

The processes which are mainly studied in this section are the so-called *submartingales of class* (Σ). Their definition has been first stated by Yor in [79], and most of its main properties, which play an essential role in our work, are studied by Nikeghbali in [50].

Definition 3.1 *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space. A nonnegative submartingale $(X_t)_{t\geq 0}$ is of class (Σ) if and only if it can be decomposed as $X_t = N_t + A_t$, where $(N_t)_{t\geq 0}$ and $(A_t)_{t\geq 0}$ are $(\mathcal{F}_t)_{t\geq 0}$-adapted processes satisfying the following assumptions:*

- *$(N_t)_{t\geq 0}$ is a càdlàg martingale.*
- *$(A_t)_{t\geq 0}$ is a continuous increasing process, with $A_0 = 0$.*
- *The measure (dA_t) is carried by the set $\{t \geq 0, X_t = 0\}$.*

In the sequel of this section, we will see that if some technical assumptions are satisfied, then one can associate to any submartingale of class (Σ) a σ-finite measure $\mathcal{Q}$ which is related to the initial probability measure $\mathbb{P}$ in the same way as the measure $\mathbb{W}$ above is related to the Wiener measure.

In order to define precisely the assumptions which are involved in our main result, we first present a not very usual way to complete the filtrations, which solves some technical problems related to the extension of probability measures and often neglected in the literature. Then, we construct the σ-finite measure $\mathcal{Q}$ and we study some of its remarkable properties. At the end of this section, we prove some penalization results, available in the present setting.

3.1 A new kind of augmentation of filtrations

If a stochastic process is defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ which does not satisfy any technical assumption, then it is in general not possible to define a "nice" version of this process, satisfying some properties of regularity which turn to be useful for the applications. For example, in general, the martingales do not admit a regular càdlàg version, and consequently, if a Brownian motion is defined on the space $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ endowed with its uncompleted canonical filtration, then it is not possible to define a version $(L_t)_{t\geq 0}$ of its local time at level zero which is both adapted and càdlàg everywhere.

For this reason, one generally considers spaces $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ which satisfy the so-called *usual assumptions*, i.e. such that the following conditions are satisfied:

- For any $\mathbb{P}$-negligible set $A \in \mathcal{F}$, i.e. such that $\mathbb{P}(A) = 0$, one has $A \in \mathcal{F}_0$, and then $A \in \mathcal{F}_t$ for all $t \geq 0$.
- The filtration is right-continuous, i.e. for all $t \geq 0$, $\mathcal{F}_t = \mathcal{F}_{t+}$, where $\mathcal{F}_{t+} = \bigcap_{s>t} \mathcal{F}_s$.

Moreover, if we start with a filtered probability space, it is always possible to transform it to a space satisfying the usual assumptions, by performing the so-called *usual augmentation*, i.e. by replacing, for all $t \geq 0$, $\mathcal{F}_t$ by the filtration generated by $\mathcal{F}_{t+}$ and the negligible sets of $\mathcal{F}$.

However, the usual augmentation turns to cause some problems when we want to extend to the whole σ-algebra $\mathcal{F}$ a consistent family of probability measures, defined on each of the σ-algebras $\mathcal{F}_t$ for $t \geq 0$. For example, let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be the usual augmentation of the space $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ endowed with its canonical filtration and the Wiener measure. For $t \geq 0$, let $\mathbb{Q}_t$ be the probability measure on $\mathcal{F}_t$, whose density with respect to the restriction of $\mathbb{P}$ to $\mathcal{F}_t$ is equal to $e^{X_t - t/2}$, where X is the canonical process. The measures $(\mathbb{Q}_t)_{t\geq 0}$ are compatible with each other, and under $\mathbb{Q}_t$, $(X_s)_{s\leq t}$ is a Brownian motion with drift one. Now, let us assume that this family of probability measures can be extended to a measure $\mathbb{Q}$ defined on the whole σ-algebra $\mathcal{F}$. If A is the event that the canonical trajectory always stays above level -1, then A is $\mathbb{P}$-negligible, hence in $\mathcal{F}_0$, and therefore, $\mathbb{Q}$-negligible since the restrictions of $\mathbb{P}$ and $\mathbb{Q}$ to $\mathcal{F}_0$ are equivalent. This contradicts the fact that under $\mathbb{Q}$, $(X_t)_{t\geq 0}$ should be a Brownian motion with drift one.

With this example, we see that the usual augmentation is an inappropriate tool as soon as one performs the simplest Girsanov transformation. The point is the following: when we complete the filtration with all the $\mathbb{P}$-negligible sets, we loose all information about any probability measure which is singular with respect to $\mathbb{P}$, even if its restriction to $\mathcal{F}_t$ is absolutely continuous. That is why we need to consider a weaker augmentation, obtained by adding a smaller class of negligible sets in a way which avoids simultaneously the two issues described above: it should be possible to define regular versions of martingales, and also to extend compatible families of measures under reasonably general conditions.

The augmentation which is used for this purpose was first introduced by Bichteler in [8], and then studied in more detail in a joint work with Nikeghbali [46]. We say that a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ satisfies the *natural conditions* (or natural assumptions) if and only if the following properties hold:

- For all $t \geq 0$, and any $\mathbb{P}$-negligible set $A \in \mathcal{F}_t$, one has also $A \in \mathcal{F}_0$.
- The filtration $(\mathcal{F}_t)_{t\geq 0}$ is right-continuous.

Similarly to the case of the usual assumptions, one can construct the *natural augmentation* of any filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, by first replacing $\mathcal{F}_t$ by $\mathcal{F}_{t+}$ for all $t \geq 0$, and then by adding all the $\mathbb{P}$-negligible sets of $\mathcal{F}_s$ to $\mathcal{F}_t$, for all $s, t \geq 0$.

The difference between the usual and the natural assumptions is that in the second case, the only $\mathbb{P}$-negligible events which are added in the σ-algebras depend only on the information available in finite time, even if this time is a priori not bounded. In this way, for the example of Girsanov transformation stated above, we can avoid to put in $\mathcal{F}_0$ an event like A, which directly depends on the whole canonical trajectory, and then we have the possibility to define the measure $\mathbb{Q}$ without the contradiction explained before. More precisely, the results of [46] can be summarized as follows:

- The most classical results in stochastic calculus which are usually proven under usual conditions: existence of càdlàg versions of martingales, Doob-Meyer decomposition, début theorem, etc., are in fact true if one only assumes natural conditions.
- Let us suppose that $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is a filtered probability space and that any consistent family $(\mathbb{Q}_t)_{t\geq 0}$ of probability measures, $\mathbb{Q}_t$ defined on $\mathcal{F}_t$, can be extended to a probability measure $\mathbb{Q}$ defined on $\mathcal{F}$. Then, the same situation holds if $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is replaced by its natural augmentation, except that one needs to suppose that for all $t \geq 0$, $\mathbb{Q}_t$ is absolutely continuous with respect to the restriction of $\mathbb{P}$ to $\mathcal{F}_t$.

Because of the second point, it is important to construct some filtered spaces $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0})$ for which the extension of a consistent family of probability measures is always possible. Such conditions, referred in [46] as *property (P)*, have been already studied, in a slightly different form, by Parthasarathy in [52]. Some particular examples of spaces satisfying property (P) are the following:

- The space of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}^d$ for $d \geq 1$, endowed with its canonical filtration.
- The space of càdlàg functions from $\mathbb{R}_+$ to $\mathbb{R}^d$ for $d \geq 1$, endowed with its canonical filtration.

Now, we have all the technical tools needed to construct rigorously the σ-finite measure $\mathcal{Q}$ announced before.

3.2 The measure $\mathcal{Q}$ and its main properties

The σ-finite measure associated to a submartingale in the class (Σ) has been constructed and studied in a series of papers with Nikeghbali (see [40], [41], [43], [45]). The techniques used in this setting are very different from those involved in the results of Section 2: in particular, no Markov or scaling property is needed for the stochastic processes which are considered.

Our main result can be stated as follows:

Theorem 3.2 *Let $(X_t)_{t\geq 0}$ be a submartingale of the class (Σ) defined on the natural augmentation of a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ which satisfies the property (P). Then, there exists a unique σ-finite measure $\mathcal{Q}$, defined on $(\Omega, \mathcal{F}, \mathbb{P})$, such that for $g := \sup\{t \geq 0, X_t = 0\}$:*

- *$\mathcal{Q}[g = \infty] = 0$;*
- *For all $t \geq 0$, and for all $\mathcal{F}_t$-measurable, bounded random variables F_t,*

$$\mathcal{Q}\left[F_t \, \mathbb{1}_{g \leq t}\right] = \mathbb{E}_{\mathbb{P}}\left[F_t X_t\right]. \tag{2}$$

The link between the measure $\mathcal{Q}$ and the measure $\mathbb{W}$ constructed before is clear, especially if we compare Theorem 3.2 to the first part of Theorem 2.2. The reason of this similarity is the fact that under some technical assumptions, the absolute value of a Brownian motion is a class (Σ) submartingale, and the local time at level zero is its increasing process: the measure $\mathbb{W}$ is, up to technicalities, the measure $\mathcal{Q}$ associated to a reflected Brownian motion. Several particular cases of Theorem 3.2 (which are not disjoint from each other) are studied in detail in [41]:

- The case where $(X_t)_{t\geq 0}$ is the absolute value, or the positive part, of a continuous martingale.
- The case where $(X_t)_{t\geq 0}$ is the *draw-down* of a martingale $(M_t)_{t\geq 0}$ starting from zero and without positive jumps, which means that for all $t \geq 0$, $X_t = \sup\{M_s, s \leq t\} - M_t$.
- The case where $(X_t)_{t\geq 0}$ is uniformly integrable. Under this assumption, X_t tends $\mathbb{P}$-almost surely to a limit X_∞ when t goes to infinity, and $\mathcal{Q} = X_\infty \cdot \mathbb{P}$. This result has essentially been proven by Azéma, Meyer and Yor in [3].

If we do not take into account the problems related to the natural augmentation, the first particular case essentially includes the measure $\mathbb{W}$ studied in Chap. 1 of [48], by taking, for $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, the natural augmentation of $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ endowed with its canonical filtration and the Wiener measure, and for $(X_t)_{t\geq 0}$, the absolute value of the canonical process. Moreover, one can also consider, on the same space, the draw-down of the canonical Brownian motion: in this case, one obtains another σ-finite measure, which turns out to be equal to the restriction of $\mathbb{W}$ to the trajectories tending to

$-\infty$ at infinity. It is also possible to construct, with the help of Theorem 3.2, the σ-finite measures introduced in Chap. 3 of [48], and associated to linear diffusions.

Let us notice that the measure $\mathcal{Q}$ has links with some problems in mathematical finance. For example, the results obtained by Cheridito, Nikeghbali and Platen in [13] are directly related to the construction of $\mathcal{Q}$ for the draw-down of the stock price, which turns out to be a martingale under the risk-neutral probability measure. Another motivation in the study of the measure $\mathcal{Q}$ is its relation with some results stated by Madan, Roynette and Yor in [32], and by Bentata and Yor in [7], where the authors study the representation of the price of a European put option in terms of the probability distribution of some last passage time. More precisely, they prove that if $(M_t)_{t\geq 0}$ is a continuous, nonnegative local martingale defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ satisfying the usual assumptions, and such that M_t tends to zero when t tends to infinity, then $(K - M_t)_+ = K\,\mathbb{P}(g_K \leq t \,|\, \mathcal{F}_t)$, where $K \geq 0$ is a constant and $g_K = \sup\{t \geq 0,\, M_t = K\}$. This formula, which informally says that it is enough to know the terminal value of the submartingale $((K - M_t)_+)_{t\geq 0}$ and its last zero g_K to reconstruct it, is in fact equivalent to (2), for $X_t = (K - M_t)_+$, and $\mathcal{Q} = K \cdot \mathbb{P}$. For this reason, the problem of the existence of $\mathcal{Q}$ under general assumptions was already posed in [32] and [7]. Note that this motivation is not a priori related to the penalization problems at the origin of the construction of $\mathbb{W}$ in [48].

Once the measure $\mathcal{Q}$ is constructed, it is natural to see which properties such a σ-finite measure should enjoy. This is the main topic of [40], and its turns out that a part of the properties of the σ-finite measures studied in [48] are also satisfied in the present setting. For example, there is an analog of Theorem 2.6, which can be stated as follows:

Theorem 3.3 *Let as suppose that the assumptions of Theorem 3.2 are satisfied, and let us take the same notation. Let Γ be a $\mathcal{Q}$-integrable, nonnegative functional defined on $(\Omega, \mathcal{F})$. Then, there exists a càdlàg $\mathbb{P}$-martingale $(M_t(\Gamma))_{t\geq 0}$ such that the measure $M^{\Gamma} := \Gamma \cdot \mathcal{Q}$ is the unique finite measure satisfying, for all $t \geq 0$, and for all bounded, $\mathcal{F}_t$-measurable functionals F_t:*

$$M^{\Gamma}[F_t] = \mathbb{P}[F_t \cdot M_t(\Gamma)].$$

There are some cases where $M_t(\Gamma)$ can be explicitly computed: for example, if $\Gamma = f(A_\infty)$ where f is an integrable function from $\mathbb{R}_+$ to $\mathbb{R}_+$, or if $(X_t)_{t\geq 0}$ is a strictly positive martingale. An interesting problem is to find the behavior at infinity of $(M_t(\Gamma))_{t\geq 0}$. We are able to answer to this question if $A_\infty = \infty$, $\mathbb{P}$-almost surely, and this condition is satisfied in the particular case where $(X_t)_{t\geq 0}$ is the absolute value of a Brownian motion and $(A_t)_{t\geq 0}$ its local time at level zero. The behavior of the martingale is in fact different under $\mathbb{P}$ and under $\mathcal{Q}$:

Theorem 3.4 *Let us assume the conditions of Theorem 3.2, let us suppose that $A_\infty = \infty$, $\mathbb{P}$-almost surely, and let Γ be a $\mathcal{Q}$-integrable, nonnegative functional defined on $(\Omega, \mathcal{F})$. Then for t tending to infinity:*

- *Almost surely under $\mathbb{P}$, $M_t(\Gamma)$ tends to zero.*
- *$\mathcal{Q}$-almost everywhere, X_t tends to infinity and $M_t(\Gamma)/X_t$ tends to Γ: in particular, $M_t(\Gamma)$ tends to infinity as soon as $\Gamma > 0$.*

Remark 3.5 *If $\Gamma > 0$, $\mathcal{Q}$-almost everywhere (which, for example, occurs for $\Gamma = e^{-A_\infty}$), then the behaviors of $(X_t)_{t\geq 0}$ under $\mathbb{P}$ and under $\mathcal{Q}$ are incompatible, which confirms the fact that the measures $\mathbb{P}$ and $\mathcal{Q}$ are singular. Note that this fact is a direct consequence of the assumptions we* *made: by definition of $\mathcal{Q}$, $A_\infty < \infty$, $\mathcal{Q}$-almost everywhere, and by assumption, $A_\infty = \infty$, $\mathbb{P}$-almost surely.*

Note that the martingales $(M_t(\Gamma))_{t\geq 0}$ are involved in the following general theorem of decomposition of the nonnegative supermartingales:

Theorem 3.6 *Let us assume the conditions of Theorem 3.2, and let us suppose that $A_\infty = \infty$, $\mathbb{P}$-almost surely. Let Z be a nonnegative, càdlàg $\mathbb{P}$-supermartingale. We denote by Z_∞ the $\mathbb{P}$-almost sure limit of Z_t when t goes to infinity. Then, $\mathcal{Q}$-almost everywhere, the quotient Z_t/X_t is well-defined for t large enough and converges, when t goes to infinity, to a limit z_∞, integrable with respect to $\mathcal{Q}$, and $(Z_t)_{t\geq 0}$ decomposes as*

$$(Z_t = M_t(z_\infty) + \mathbb{P}[Z_\infty|\mathcal{F}_t] + \xi_t)_{t\geq 0}\,,$$

where $(\mathbb{P}[Z_\infty|\mathcal{F}_t])_{t\geq 0}$ denotes a càdlàg version of the conditional expectation of Z_∞ with respect to $\mathcal{F}_t$, and $(\xi_t)_{t\geq 0}$ is a nonnegative, càdlàg $\mathbb{P}$-supermartingale, such that:

- *$Z_\infty \in L^1(\mathcal{F},\mathbb{P})$, hence $\mathbb{P}[Z_\infty|\mathcal{F}_t]$ converges $\mathbb{P}$-almost surely and in $L^1(\mathcal{F},\mathbb{P})$ towards Z_∞.*
- *$\mathcal{Q}$-almost everywhere, $\mathbb{P}[Z_\infty|\mathcal{F}_t]/X_t$ and ξ_t/X_t tend to zero when t goes to infinity.*
- *$\mathbb{P}$-almost surely, $M_t(z_\infty)$ and ξ_t tend to zero when t goes to infinity.*

This decomposition is strongly related to the decomposition of the probability measures on $(\Omega, \mathcal{F})$ into three parts:

- A part which is absolutely continuous with respect to $\mathbb{P}$.
- A part which is absolutely continuous with respect to $\mathcal{Q}$.
- A part which is singular with respect to $\mathbb{P}$ and $\mathcal{Q}$.

3.3 Some penalization results related to the measure $\mathcal{Q}$

Since the measure $\mathcal{Q}$ enjoys similar properties as the σ-finite measures described in the previous section and studied in [48], it is natural to look whether $\mathcal{Q}$ also has some link with penalization problems. This question is studied in [39] and [40]. The general result obtained in [40] comes from an estimate of the expectation of a suitable class of functionals, which can be stated as follows:

Theorem 3.7 *Let us assume the conditions of Theorem 3.2, and let us suppose that $A_\infty = \infty$, $\mathbb{P}$-almost surely. Let $(\Gamma_t)_{t\geq 0}$ be a càdlàg, adapted, nonnegative process such that its limit Γ_∞ exists $\mathcal{Q}$-almost everywhere. We assume that there exists a $\mathcal{Q}$-integrable, nonnegative functional H, such that for all $t \geq 0$, one has $\Gamma_t X_t \leq M_t(H)$, $\mathbb{P}$-almost surely, where $(M_t(H))_{t\geq 0}$ is the martingale introduced in Theorem 3.3. Then, Γ_∞ is $\mathcal{Q}$-integrable and*

$$\mathbb{E}_{\mathbb{P}}[\Gamma_t X_t] \underset{t\to\infty}{\longrightarrow} \mathcal{Q}[\Gamma_\infty].$$

From this estimate, we can easily deduce a general penalization result:

Theorem 3.8 *Let us assume the conditions of Theorem 3.2, and let us suppose that* $A_\infty = \infty$, $\mathbb{P}$*-almost surely. Let* $(\Gamma_t)_{t\geq 0}$ *be a càdlàg, adapted, nonnegative process such that its limit* Γ_∞ *exists* $\mathcal{Q}$*-almost everywhere and is not* $\mathcal{Q}$*-almost everywhere equal to zero. We assume that there exists a* $\mathcal{Q}$*-integrable, nonnegative functional* H*, such that for all* $t \geq 0$*, one has* $\Gamma_t X_t \leq M_t(H)$*,* $\mathbb{P}$*-almost surely, where* $(M_t(H))_{t\geq 0}$ *is the martingale introduced in Theorem 3.3. Then, for* t *sufficiently large,* $0 < \mathbb{E}_{\mathbb{P}}[\Gamma_t X_t] < \infty$ *and one can define a measure* $\mathbb{Q}_t$ *by*

$$\mathbb{Q}_t := \frac{\Gamma_t X_t}{\mathbb{E}_{\mathbb{P}}[\Gamma_t X_t]} \cdot \mathbb{P}.$$

Moreover, there exists a probability measure $\mathbb{Q}_\infty$ *such that for all* $s \geq 0$*, and for all events* $\Lambda_s \in \mathcal{F}_s$*,*

$$\mathbb{Q}_t(\Lambda_s) \underset{t\to\infty}{\longrightarrow} \mathbb{Q}_\infty(\Lambda_s).$$

The measure $\mathbb{Q}_\infty$ *is absolutely continuous with respect to* $\mathcal{Q}$*:*

$$\mathbb{Q}_\infty = \frac{\Gamma_\infty}{\mathcal{Q}[\Gamma_\infty]} \cdot \mathcal{Q}$$

where $0 < \mathcal{Q}[\Gamma_\infty] < \infty$*.*

It is quite satisfactory to get such a general result, but unfortunately, in the case where $(X_t)_{t\geq 0}$ is the absolute value of a Brownian motion, this theorem does not cover any of the classical examples of penalizations studied by Roynette, Vallois and Yor. In order to deal with more classical penalization weights, we need to restrict in an important way the class of processes $(X_t)_{t\geq 0}$ which are considered. This restriction is made in [39], where we only consider a certain class of linear diffusions, as in [48], Chap. 3. The precise setting is the following:

- We define $\Omega := \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+)$, we denote by $(\mathcal{F}_s^0)_{s\geq 0}$ the canonical filtration of Ω and by $\mathcal{F}^0$ the σ-algebra generated by $(\mathcal{F}_s^0)_{s\geq 0}$.
- We consider a probability measure $\mathbb{P}^0$ on $(\Omega, \mathcal{F}^0)$, under which the canonical process is a recurrent diffusion in linear scale, starting from a fixed point $x_0 \geq 0$ and with zero as an instantaneously reflected barrier.
- We suppose that the speed measure of this diffusion is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}_+$, with a continuous density $m : \mathbb{R}_+^* \to \mathbb{R}_+^*$.
- We assume that $m(x)$ is equivalent to cx^β when x goes to infinity, for some $c > 0$ and $\beta > -1$, and we suppose that there exists $C > 0$ such that for all $x > 0$, $m(x) \leq Cx^\beta$ if $\beta \leq 0$, and $m(x) \leq C(1 + x^\beta)$ if $\beta > 0$.
- The filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is defined as the natural augmentation of $(\Omega, \mathcal{F}^0, (\mathcal{F}_t^0)_{t\geq 0}, \mathbb{P}^0)$.

Under $\mathbb{P}$, the canonical process $(X_t)_{t\geq 0}$ is a diffusion in natural scale, and as in [48], Chap. 3, one can show that $(X_t)_{t\geq 0}$ is a class (Σ) submartingale, and that its increasing process is the local time $(L_t)_{t\geq 0}$ at level zero. Therefore, one can construct the σ-finite measure $\mathcal{Q}$ defined in Theorem 3.2. The main penalization result proven in [39] is then the following:

Theorem 3.9 *We suppose that the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_s)_{s\geq 0}, \mathbb{P})$, the diffusion process X and the σ-finite measure $\mathcal{Q}$ are defined as above. Let $(\Gamma_t)_{t\geq 0}$ be a process satisfying the following properties:*

- *$(\Gamma_t)_{t\geq 0}$ is nonnegative, uniformly bounded, nonincreasing, càdlàg and adapted with respect to $(\mathcal{F}_t)_{t\geq 0}$.*
- *There exists $a > 0$ such that for all $t \geq 0$, $\Gamma_t = \Gamma_{g^{[a]}}$ on the set $\{g^{[a]} \leq t\}$, where $g^{[a]}$ denotes the supremum of $s \geq 0$ such that $X_s \leq a$.*
- *One has $0 < \mathcal{Q}[\Gamma_\infty] < \infty$, where Γ_∞ denotes the nonincreasing limit of Γ_t when t goes to infinity.*

Then, for all $t \geq 0$, $0 < \mathbb{E}_{\mathbb{P}}[\Gamma_t] < \infty$ and one can define a probability measure $\mathbb{Q}_t$ on $(\Omega, \mathcal{F})$ by

$$\mathbb{Q}_t := \frac{\Gamma_t}{\mathbb{E}_{\mathbb{P}}[\Gamma_t]} \cdot \mathbb{P}.$$

Moreover, the probability measure

$$\mathbb{Q}_\infty := \frac{\Gamma_\infty}{\mathcal{Q}[\Gamma_\infty]} \cdot \mathcal{Q}$$

is the weak limit of $(\mathbb{Q}_t)_{t\geq 0}$ in the sense of penalization, i.e. for all $s \geq 0$, and for any event $\Lambda_s \in \mathcal{F}_s$,

$$\mathbb{Q}_t[\Lambda_s] \underset{t\to\infty}{\longrightarrow} \mathbb{Q}_\infty[\Lambda_s].$$

This result is essentially a generalization of Theorem 2.5, where the assumptions made on $(\Gamma_t)_{t\geq 0}$ correspond to the class $\mathcal{C}$. In particular, we recover some of the penalization results studied by Roynette, Vallois and Yor.

Part II

Some models of random matrices and their infinite-dimensional extensions

Random matrix theory is a very large and rich mathematical subject, which has much developed in the last decades, and which interconnects various parts of mathematics (e.g. complex analysis, theory of representations, number theory) and theoretical physics (statistical physics, quantum mechanics). The first results on the topic were obtained in 1928 by Wishart [77] in relation with problems in statistics, and an intensive research began with the work of Wigner in the 1950s (see [76]), on the energy levels of heavy nuclei: this research has taken many different directions, several of them are discussed in a book by Mehta [33], and more recently, by Anderson, Guionnet and Zeitouni [2].

Two of the most classical ensembles of random matrices are the following:

- The *Gaussian Unitary Ensemble*, corresponding to a random hermitian matrix, for which the entries, on the diagonal and above, are independent and gaussian.
- The *Circular Unitary Ensemble*, corresponding to a random matrix following the Haar measure on a unitary group.

For these two models, some remarkable results have been proven on the local bahavior of the corresponding eigenvalues, when the dimension goes to infinity. These results involve a limiting point process, called the *determinantal sine-kernel point process*, and informally described as follows: for all integers $p \geq 1$, and for all $t_1, \dots, t_p \in \mathbb{R}$, the probability to have a point in the neighborhood of t_j for $j \in \{1, ..., p\}$ is proportional to

$$\rho(t_1, \dots, t_p) := \det(K(t_i, t_j))_{1 \leq i,j \leq k},$$

where K is the function from $\mathbb{R}^2$ to $\mathbb{R}$ given by

$$K(x, y) = \frac{\sin(\pi(x - y))}{\pi(x - y)}.$$

The determinantal sine-kernel process enjoys a property of *universality*: it appears as a limiting point process for a large class of random matrix models. This situation looks quite similar to the central limit theorem, which universally involves the gaussian distribution, independently of the common law of the random variables considered.

Behind the central limit theorem, there exists a universal limiting object which models the long term behavior of the sum of i.i.d., square-integrable random variables, namely the Brownian motion. The existence of such a process gives a great improvement for our understanding of the universal properties of the gaussian distribution. Similarly, it is natural to expect that the universal properties of the sine-kernel process can be explained by an infinite-dimensional random object modeling the limiting behavior of large random matrices. That is why we focus a part of our research on the infinite-dimensional objects which can be related to different random matrix ensembles.

In Section 4, we study several random objects constructed from permutation matrices. The important advantage of these models is that essentially, all the corresponding eigenvalues are directly related to the cycle lengths of the permutations, thus allowing very explicit computations, which are not always possible for the most classical ensembles. In Section 5, we present some models of unitary and hermitian matrices, and we prove that a part of the infinite objects constructed in Section 4 can be generalized to this setting. In particular, we obtain a deterministic link between the Circular Unitary Ensemble and the determinantal sine-kernel process. In Section 6, we present some other promising directions of research we plan to explore in the near future.

4 The framework of the symmetric group

4.1 The spaces of virtual permutations

The notion of permutations is still meaningful for infinite sets, even if the symmetric groups are generally studied when their orders are finite. However, relatively few results can be obtained for the infinite symmetric groups, since these groups have, in a certain sense, too many elements, which do not enjoy the same properties as the permutations of finite order. For example, a permutation of infinite order cannot, in general, be written in terms of a cycle structure. That is one of the reasons why some other infinite objects have been constructed to generalize the notion of permutation of finite order.

One of these objects can be constructed by taking the inductive limit of all the finite symmetric groups: in this way, one obtains the space containing the permutations of an infinite set which fix all but finitely many elements. This space is a group, which can be written as an increasing union of the finite symmetric groups, and all its elements enjoy essentially the same properties as the finite permutations (for example, decomposition as a product of cycles). These properties are very convenient to use, but on the other hand, the situation is quite disappointing, since one often does not observe really new phenomena, compared to the finite symmetric groups.

Another way to construct infinite analogs of finite permutations is to take a projective limit. This construction has been first introduced by Kerov, Olshanski and Vershik in [24], and then studied in more detail by Tsilevich in [66] and [67]: the object obtained in this way is the space of the so-called *virtual pemutations*. By definition, a virtual permutation is a sequence $(\sigma_N)_{N\geq 1}$ of permutations, σ_N in the N-th symmetric group $\mathfrak{S}_N$, such that for all $N \geq 1$, the cycle structure of σ_N is deduced from the structure of σ_{N+1} by removing the element $N+1$ from its cycle. The virtual permutations are in bijection with the sequences $(k_N)_{N\geq 2}$ such that $k_N \in \{1, 2, \dots, N\}$ for all $N \geq 2$. The construction of $(\sigma_N)_{N\geq 1}$ from $(k_N)_{N\geq 2}$ can be made inductively, as follows:

- σ_1 is the unique permutation of $\mathfrak{S}_1$.
- If $N \geq 1$ and $k_{N+1} = N+1$, then the cycle structure of σ_{N+1} is obtained from σ_N by adding $N+1$ as a fixed point.
- If $N \geq 1$ and $k_{N+1} \leq N$, then the cycle structure of σ_{N+1} is obtained from σ_N by inserting $N+1$ inside the cycle of k_{N+1}, just before this element: $\sigma_{N+1}(N+1) = k_{N+1}$.

This algorithm, sometimes called the *chinese restaurant process* (see Pitman [55], Chap. 3), is directly related to the splitting of a permutation as a product of tranpositions: here, for all $N \geq 2$, one has $\sigma_N = \tau_{N,k_N} \circ \tau_{N-1,k_{N-1}} \circ \cdots \circ \tau_{2,k_2}$, where $\tau_{i,j}$ is the transposition of i and j for $i \neq j$, and the identity for $i = j$.

In the definition of the virtual permutations given above, the natural order on the set of positive integers plays an explicit role: if $(\sigma_N)_{N\geq 1}$ is a virtual permutation, then σ_N permutes the N smallest positive integers. However, in most of the limiting results on virtual permutations, all the integers play in fact the same role: that is why in a joint paper with Nikeghbali [44], we have generalized the notion of virtual permutations to arbitrary sets. The precise definition is the following:

Definition 4.1 *A virtual permutation of a given set E is a family of permutations $(\sigma_I)_{I\in\mathcal{F}(E)}$, indexed by the set $\mathcal{F}(E)$ of the finite subsets of E, such that for all $I \in \mathcal{F}(E)$, σ_I is a permutation of the set I, and for $I, J \in \mathcal{F}(E)$, $I \subset J$, the permutation σ_I is obtained by removing the elements of $J\backslash I$ from the cycle structure of σ_J.*

One can check the two following facts:

- If E is finite, then $(\sigma_I)_{I\in\mathcal{F}(E)}$ can be identified with σ_E, and then a virtual permutation is essentially a permutation.
- If E is the set of positive integers, then a virtual permutation, in the classical sense, is equivalent to a virtual permutation on E in the sense of Definition 4.1. Indeed $(\sigma_N)_{N\geq 1}$ can be identified with $(\sigma_I)_{I\in\mathcal{F}(E)}$, where for $I \in \mathcal{F}(E)$ and for N equal to the largest element of I, σ_I is deduced from σ_N by restricting its cycle structure to I.

Notice that the set $\mathfrak{S}_E$ of the virtual permutations of E is not a group if E is infinite. However, the set $\mathfrak{S}_E^{(0)}$ of the permutations of E which fix all but finitely many elements of E is a group. This group acts on $\mathfrak{S}_E$ by conjugation, as follows: for $\sigma = (\sigma_I)_{I\in\mathcal{F}(E)}$, and for $g \in \mathfrak{S}_E^{(0)}$ fixing each of the elements of $E\backslash J$ for $J \in \mathcal{F}(E)$, one defines $g\sigma g^{-1}$ as the unique virtual permutation $(\tau_I)_{I\in\mathcal{F}(E)}$ such that for $I \in \mathcal{F}(E)$ containing J, $\tau_I = g_I \sigma_I g_I^{-1}$, where g_I denotes the restriction of g to I.

Another interesting fact is that a virtual permutation $(\sigma_I)_{I\in\mathcal{F}(E)}$ on a set E induces a partition of E. Indeed, for all $x, y \in E$, one of the two following possibilities holds:

- For all $I \in \mathcal{F}(E)$ containing x and y, these two elements belong to the same cycle of the permutation σ_I.
- For all $I \in \mathcal{F}(E)$ containing x and y, these two elements belong to different cycles of σ_I.

One can then check that the first possibility induces an equivalence relation on E, and then a partition of this set.

4.2 The central probability measures on $\mathfrak{S}_E$

For the moment, we have not introduced any randomness in the discussion. In order to change this situation, we need to define a σ-algebra $\mathcal{S}_E$ on $\mathfrak{S}_E$, and then a probability measure on $(\mathfrak{S}_E, \mathcal{S}_E)$. From now on, we suppose that E is countable, and that $\mathcal{S}_E$ is the σ-algebra generated by the coordinate maps $(\sigma_I)_{I\in\mathcal{F}(E)} \mapsto \sigma_J$, for all $J \in \mathcal{F}(E)$. Under this assumption, the following result shows that one can construct probability measures on $\mathfrak{S}_E$ from probability measures on $\mathfrak{S}_I$, $I \in \mathcal{F}(E)$, if a certain compatibility property is satisfied:

Proposition 4.2 *For $I \in \mathcal{F}(E)$, let $\mathbb{P}_I$ be a probability measure on $\mathfrak{S}_I$. We assume that for $I, J \in \mathcal{F}(E)$, $I \subset J$, $\mathbb{P}_I$ is the image of $\mathbb{P}_J$ by the map from $\mathfrak{S}_J$ to $\mathfrak{S}_I$ which removes $J\backslash I$ from the cycle structure of the permutations. Then, there exists a unique probability measure $\mathbb{P}$ on $(\mathfrak{S}_E, \mathcal{S}_E)$ such that for all $J \in \mathcal{F}(E)$, the image of $\mathbb{P}$ by the map $(\sigma_I)_{I\in\mathcal{F}(E)} \mapsto \sigma_J$ is equal to $\mathbb{P}_J$.*

Example 4.3 *Let $\theta \in \mathbb{R}_+^*$, and for all $I \in \mathcal{F}(E)$, let $\mathbb{P}_I$ be the Ewens measure of parameter θ on $\mathfrak{S}_I$, i.e. the unique measure such that for all $\sigma \in \mathfrak{S}_I$,*

$$\mathbb{P}_I(\sigma) = \frac{\theta^{n-1}}{(\theta+1)(\theta+2)\dots(\theta+N-1)},$$

where n is the number of cycles of σ and N the cardinality of I. Then, by the classical properties of the Ewens measure, the family of probability measures $(\mathbb{P}_I)_{I \in \mathcal{F}(E)}$ satisfies the compatibility conditions given in Proposition 4.2, which then induces a probability measure $\mathbb{P}$ on $(\mathfrak{S}_E, \mathcal{S}_E)$. The measure $\mathbb{P}$ is also called the Ewens measure of parameter θ. For $\theta = 1$, it is also called the uniform measure on $(\mathfrak{S}_E, \mathcal{S}_E)$.

Note that in this example, the measure $\mathbb{P}_I$ is invariant by conjugation for all $I \in \mathcal{F}(E)$. One deduces that $\mathbb{P}$ is also invariant by conjugation, i.e. $\mathbb{P}$ is a *central measure* on the space $(\mathfrak{S}_E, \mathcal{S}_E)$. One can ask if there exist some other measures satisfying this property. The answer is positive and in [44], we obtain a complete characterization of these measures, which can be summarized as follows:

Theorem 4.4 *Let Λ be the set containing all the nonincreasing sequences $(\lambda_1, \lambda_2, \dots, \lambda_n, \dots)$ of elements of $\mathbb{R}_+$, such that $\sum_{k \geq 1} \lambda_k \leq 1$. For $\lambda := (\lambda_1, \lambda_2, \dots, \lambda_n, \dots) \in \Lambda$, let C be a the disjoint union of the following sets:*

- *A segment L of length $\lambda_0 := 1 - \sum_{k \geq 1} \lambda_k$.*
- *For each $k \geq 1$, a circle C_k of perimeter λ_k (the empty set if $\lambda_k = 0$).*

Let $(X_x)_{x \in E}$ be a family of i.i.d., uniform random points on C, i.e. for any subset A of C which is either an arc of one of the circles C_k for $k \geq 1$, or a segment included in L, the probability that $X_x \in A$ is equal to the length of A. Let $(\sigma_I)_{I \in \mathcal{F}(E)}$ be a random family of permutations, defined as follows:

- *If $I \in \mathcal{F}(E)$, $x \in I$ and $X_x \in C_k$ for some $k \geq 1$, then $\sigma_I(x) = y$, where X_y is the first point of $C_k \cap \{X_z, z \in I\}$ encountered by turning counterclockwise on the circle C_k, starting just after X_x. Note that if X_x is the unique point of $C_k \cap \{X_z, z \in I\}$, then $\sigma_I(x) = x$.*
- *If $I \in \mathcal{F}(E)$, $x \in I$ and $X_x \in L$, then $\sigma_I(x) = x$.*

Then, $(\sigma_I)_{I \in \mathcal{F}(E)}$ is a random virtual permutation, and its distribution $\mathbb{P}_\lambda$ is a central measure. Moreover, for each central measure $\mathbb{P}$ on $(\mathfrak{S}_E, \mathcal{S}_E)$, there exists a unique probability measure ν on Λ (endowed with the σ-algebra generated by the coordinate maps) such that for all $A \in \mathcal{S}_E$,

$$\mathbb{P}(A) = \int_\Lambda \mathbb{P}_\lambda(A) \, d\nu(A), \tag{3}$$

where the integral is well-defined, since $\lambda \mapsto \mathbb{P}_\lambda(A)$ is measurable. Conversely, each probability measure ν on Λ defines a central measure $\mathbb{P}_\nu$ on $(\mathfrak{S}_E, \mathcal{S}_E)$ given by (3).

Example 4.5 *For $\theta \in \mathbb{R}_+^*$, the Ewens measure on $(\mathfrak{S}_E, \mathcal{S}_E)$ is equal to $\mathbb{P}_\nu$, where ν is the law of a Poisson-Dirichlet process of parameter θ.*

The classification of the central measures given in Theorem 4.4 is strongly related to a similar result proven by Tsilevich in [67], using the properties of the exchangeable partitions stated by Kingman (see [25], [26], [27]). The cycle structure of a virtual permutation following a central measure can be characterized in a quite precise way. Since any central measure can be written as a convex combination of the measures $\mathbb{P}_\lambda$ for $\lambda \in \Lambda$, one may focus the discussion on these particular measures.

Theorem 4.6 *Let $\lambda = (\lambda_1, \lambda_2, \dots) \in \Lambda$, and let $\sigma = (\sigma_I)_{I \in \mathcal{F}(E)}$ be a random virtual permutation following the measure $\mathbb{P}_\lambda$. The cycle structure of $(\sigma_I)_{I \in E}$ induces a random exchangeable partition Π of E. If for $x \in E$, $\mathcal{C}_x$ denotes the set of Π containing x, then the random variable $|\mathcal{C}_x \cap I|/|I|$ tends in L^1 to a limit random variable $\lambda(x)$ when the cardinality $|I|$ of $I \in \mathcal{F}(E)$ goes to infinity. Almost surely, the variable $\lambda(x)$ depends only on the set of Π containing x, and $\lambda(x) = 0$ if and only if $\mathcal{C}_x = \{x\}$, i.e. x is a fixed point of σ_I for all $I \in \mathcal{F}(E)$ containing x. Moreover, if $(x_k)_{k \geq 1}$ is a random sequence of points in E containing exactly one element in each infinite set of Π, and if $\lambda(x_k)$ decreases with k, then almost surely, $\lambda(x_k) = \lambda_k$ for all $k \geq 1$.*

More informally, this result says that for a random virtual permutation following the measure $\mathbb{P}_\lambda$, the renormalized cycle lengths are almost surely well-defined, and form the sequence λ when they are taken in decreasing order. By the law of large numbers, this fact is consistent with the description of $\mathbb{P}_\lambda$ given in Theorem 4.4.

Theorem 4.6 gives some information about the cycle lengths of a random virtual permutation. The following result, also consistent with Theorem 4.4, implies that the relative position of the elements of E in their cycle has also a limiting behavior:

Theorem 4.7 *Let $\lambda = (\lambda_1, \lambda_2, \dots) \in \Lambda$, and let $\sigma = (\sigma_I)_{I \in \mathcal{F}(E)}$ be a random virtual permutation following the measure $\mathbb{P}_\lambda$. If x and y are two points in E, then on the event $\{\sigma_{\{x,y\}}(x) = y\}$:*

- *For all $I \in \mathcal{F}(E)$ containing x and y, these two points are in the same cycle of σ_I. If K denotes the length of this cycle, then there exists a unique integer $k_I(x, y) \in \{0, 1, 2, \dots, K-1\}$ such that $\sigma_I^{k_I(x,y)}(x) = y$.*
- *The random variable $k_I(x, y)/|I|$ converges in L^1 to a limit $\Delta(x, y)$ when the cardinality $|I|$ of $I \in \mathcal{F}(E)$ goes to infinity.*

Now, let $\delta(x, y)$ be the congruence class of $\Delta(x, y)$, modulo $\lambda(x)$. Then for $x, y, z \in E$, the following relation holds almost surely:

$$\delta(x, y) + \delta(y, z) = \delta(x, z)$$

modulo $\lambda(x)$, in the case where x, y, z belong to the same cycle of $\sigma_{\{x,y,z\}}$.

Intuitively, $\delta(x, y)$ represents the asymptotic behavior of the number of iterations of σ needed to go from x to y. This random variable induces a random distance d on E, defined as follows:

- If $\sigma_{\{x,y\}}(x) \neq y$, then $d(x, y) = 1$.
- If $\sigma_{\{x,y\}}(x) = y$, then $d(x, y)$ is the minimum of $|\delta|$, for $\delta \equiv \delta(x, y)$ modulo $\lambda(x)$.

The set E can then be completed for the distance d. This gives a random metric space $\widehat{E}$, which can be explicitly described up to an isometry:

Proposition 4.8 *The space $\widehat{E}$ is isometric to the metric space (C, D), where C is the set introduced in Theorem 4.4, and where the distance D is described as follows:*

- *For a, b in a common circle C_k for $k \geq 1$, $D(a, b)$ is the length of the smallest circle arc of C_k which joins a and b.*
- *For $a, b \in C$ not in a common circle C_k for any $k \geq 1$, $D(a, b) = 1$.*

4.3 A flow of operators associated to random virtual permutations

Let $\sigma := (\sigma_I)_{I \in \mathcal{F}(E)}$ be a random virtual permutation following the central measure $\mathbb{P}_\lambda$ for $\lambda \in \Lambda$. Then, by using the complete random metric space $\widehat{E}$ defined just above, one can construct a flow of isometries from $\widehat{E}$ to $\widehat{E}$, which describes the asymptotic behavior of σ when the permutations σ_I are iterated a number of times approximately proportional to the cardinality of I. This construction is given by the following result:

Theorem 4.9 *For $x \in \widehat{E}$ and $\alpha \in \mathbb{R}$, there exists a random element $S^\alpha(x)$ of $\widehat{E}$ such that all the following properties hold:*

- *Almost surely, for all $x \in \widehat{E}$ and for all $\alpha, \beta \in \mathbb{R}$, $S^0(x) = x$ and $S^{\alpha+\beta}(x) = S^\alpha(S^\beta(x))$.*
- *Almost surely, the map $(x, \alpha) \mapsto S^\alpha(x)$ is continuous.*
- *For all $x \in E$, $\alpha \in \mathbb{R}$, the distance between $S^\alpha(x)$ and $\sigma_I^{[\alpha|I|]}(x)$ tends to zero in L^1 when the cardinality $|I|$ of $I \in \mathcal{F}(E)$ goes to infinity ($[\alpha|I|]$ denotes the integer part of $\alpha|I|$).*

If the space $\widehat{E}$ is identified with the set C described in Theorem 4.4, then the transformation S^α fixes the segment L, and turns each circle C_k counterclockwise, by an angle corresponding to an arc of length α.

Moreover, the random flow of transformations $(S^\alpha)_{\alpha \in \mathbb{R}}$ can be viewed as a flow of operators $(T^\alpha)_{\alpha \in \mathbb{R}}$ acting on the random space of continuous functions from $\widehat{E}$ to $\mathbb{C}$. For $\alpha \in \mathbb{R}$, the operator T^α is defined as follows: for any continuous function f from $\widehat{E}$ to $\mathbb{C}$, and for all $y \in \widehat{E}$, one has $(T^\alpha(f))(y) = f(S^\alpha(y))$. As we have seen above, the operator T^α models the asymptotic behavior of $\sigma_I^{[\alpha|I|]}$ for a set $I \in \mathcal{F}(E)$ containing a large number of elements. If we want to get some information on the permutation σ_I itself, without iterating it, we need to consider the operator T^α for α very small, i.e. with order of magnitude $1/|I|$. That is why it is natural to study the infinitesimal generator of $(T^\alpha)_{\alpha \in \mathbb{R}}$.

This operator is defined on the space of *continuously differentiable functions* from $\widehat{E}$ to $\mathbb{C}$, i.e. the space of continuous functions f for which there exists g, continuous, such that for all $y \in \widehat{E}$,

$$\frac{(T^\alpha(f))(y) - f(y)}{\alpha} \underset{\alpha \to 0}{\longrightarrow} g(y).$$

The infinitesimal generator U of $(T^\alpha)_{\alpha \in \mathbb{R}}$ is then the operator from the space of continuously differentiable functions to the space of continuous functions from $\widehat{E}$ to $\mathbb{C}$, which maps f to g.

The spectrum of the random operator U can be explicitly described in terms of the sequence $\lambda = (\lambda_1, \lambda_2, \dots) \in \Lambda$, for which the random virtual permutation $(\sigma_I)_{I \in \mathcal{F}(E)}$ used to construct U follows the law $\mathbb{P}_\lambda$.

Theorem 4.10 *The eigenvalues of U can be obtained by taking the union, with multiplicity, of the sequences $(2im\pi/\lambda_k)_{m\in\mathbb{Z}}$ for all $k \geq 1$ such that $\lambda_k > 0$. Moreover, if $\sum_{k\geq 1} \lambda_k < 1$, 0 is an eigenvalue with infinite multiplicity.*

The spectrum of U can be compared with the asymptotic behavior of the spectrum of the permutation matrix of σ_I for $I \in \mathcal{F}(E)$ with large cardinality. Indeed, under $\mathbb{P}_\lambda$, one can prove, by studying the asymptotic behavior of the cycle lengths of σ_I, that the eigenangles of the corresponding matrix, multiplied by the cardinality of I, converge (in a sense which can be made precise) to the eigenangles of iU. The precise statement is the following:

Theorem 4.11 *Let $(\sigma_I)_{I\in\mathcal{F}(E)}$ be a virtual permutation following a central measure on $\mathfrak{S}_E$. Let X be the spectrum of the random operator iU, and for $I \in \mathcal{F}(E)$, let X_I be the set of the eigenangles of the matrix of σ_I, multiplied by the cardinality of I. For $\gamma \in X$ (resp. $\gamma \in X_I$), let $m(\gamma)$ (resp. $m_I(\gamma)$) be the multiplicity of the corresponding eigenvalue (resp. rescaled eigenangle). Then, X and X_I, $I \in \mathcal{F}(E)$ are included in $\mathbb{R}$, and for all continuous functions f from $\mathbb{R}$ to $\mathbb{R}_+$, with compact support, the following convergence in probability holds:*

$$\sum_{\gamma\in X_I} m_I(\gamma) f(\gamma) \underset{|I|\to\infty}{\longrightarrow} \sum_{\gamma\in X} m(\gamma) f(\gamma).$$

Remark 4.12 *A priori, the random operator U has only been defined under the probability measures $\mathbb{P}_\lambda$ for $\lambda \in \Lambda$. However, U is almost surely well-defined under any central measure $\mathbb{P}$ on $\mathfrak{S}_E$, since $\mathbb{P}$ can always be written as a convex combination of the measures $\mathbb{P}_\lambda$ for $\lambda \in \Lambda$.*

The intuitive meaning of Theorem 4.11 is that the convergence in distribution of the rescaled eigenangles of a random permutation matrix to a limiting point process X, when the dimension goes to infinity, can be explained by the existence of a limiting random operator whose spectrum is given by X. Indeed, for $N \geq 1$ and $\theta \in \mathbb{R}_+^*$, let σ_N be a random permutation of order N, following Ewens measure of parameter θ (for $\theta = 1$, σ_N is uniform on $\mathfrak{S}_N$). The point process of the eigenangles of the matrix of σ_N, multiplied by N, tends in law to a point process X, obtained as the union of the sets $(2\pi m/\lambda_k)_{m\in\mathbb{Z}}$ for $k \geq 1$, $(\lambda_k)_{k\geq 1}$ being a Poisson-Dirichlet process of parameter θ. Now, X has exactly the distribution of the spectrum of iU, where U is the random operator associated to a virtual permutation following the Ewens measure of parameter θ.

Such a convergence of the renormalized eigenangles, when the dimension goes to infinity, also occurs for the Circular Unitary Ensemble, i.e. the Haar measure on the unitary group. The limiting point process is a determinantal sine-kernel process, and one may ask if it can be naturally viewed as the spectrum of a random operator. In Section 5, this question is discussed in detail.

4.4 The spectrum of randomized permutation matrices

The permutation matrices have the advantage that one can directly compute their spectrum in terms of their cycle structures. The behavior of the corresponding eigenvalues is very different from what one obtains for a general unitary matrix: for example, all the eigenvalues are roots of unity of finite order. However, this very strong property is relaxed when one takes matrices in a group which is larger than the symmetric group.

In our joint paper with Nikeghbali [42], we study random matrices which are contained in the wreath product of the symmetric group and the group $\mathbb{C}^*$, i.e. the group of matrices obtained from permutation matrices by changing the entries equal to one by arbitrary non-zero entries. More precisely, in [42], we introduce the following groups, for any $N \geq 1$:

- The group $\mathcal{G}(N)$ of matrices M such that there exists a permutation $\sigma \in \mathfrak{S}_N$, and $z_1, z_2, \dots, z_N \in \mathbb{C}^*$, such that $M_{jk} = z_j \mathbb{1}_{j=\sigma(k)}$ for all $j, k \in \{1, 2, \dots, N\}$.
- The group $\mathcal{H}(N)$ of matrices in $\mathcal{G}(N)$ whose non-zero entries $z_1, z_2, \dots, z_N$ are in the group $\mathbb{U}$ of complex numbers of modulus one.
- For all $k \geq 1$, the group $\mathcal{H}_k(N)$ of matrices in $\mathcal{H}(N)$ whose non-zero entries are roots of unity of order dividing k.

Then, we define a particular family of probability measures on each of these groups, for which the underlying permutation follows the Ewens measure of parameter θ for some $\theta > 0$, and the non-zero entries are i.i.d. random variables.

Definition 4.13 *Let $\theta > 0$, let k be a strictly positive integer, and let $\mathcal{L}$ be a probability distribution on $\mathbb{C}^*$ (resp. $\mathbb{U}$, $\mathbb{U}_k$). The probability measure $\mathbb{P}(N, \theta, \mathcal{L})$ on $\mathcal{G}(N)$ (resp. $\mathcal{H}(N)$, $\mathcal{H}_k(N)$) is the law of the matrix $M(\sigma, z_1, z_2, \dots, z_N)$, where:*

- *The permutation σ follows the Ewens measure of parameter θ on $\mathfrak{S}_N$.*
- *For all $j \in \{1, \dots, N\}$, z_j is a random variable following the probability law $\mathcal{L}$.*
- *The random permutation σ and the random variables $z_1, \dots, z_N$ are all independent.*
- *$M(\sigma, z_1, \dots, z_N)$ is the matrix $M \in \mathcal{G}(N)$ (resp. $\mathcal{H}(N)$, $\mathcal{H}_k(N)$) such that for all $j, k \in \{1, \dots, N\}$, $M_{jk} = z_j \mathbb{1}_{j=\sigma(k)}$.*

In this definition, as usually done in random matrix theory, we consider ensembles of random matrices of a given order N. As above, if we want to consider infinite-dimensional objects, and then obtain results of strong convergence when the dimension goes to infinity, we need to define all the finite-dimensional ensembles on a single probability space. This can be done by using virtual permutations, as follows:

Definition 4.14 *Let $\theta > 0$, and let $\mathcal{L}$ be a probability distribution on $\mathbb{C}^*$ (resp. $\mathbb{U}$, $\mathbb{U}_k$). The probability measure $\mathbb{P}(\infty, \theta, \mathcal{L})$, defined on the product of the probability spaces $\mathcal{G}(N)$ (resp. $\mathcal{H}(N)$, $\mathcal{H}_k(N)$) for $N \geq 1$, is the law of the sequence of random matrices $(M_N)_{N \geq 1}$, such that $M_N = M(\sigma_N, z_1, \dots, z_N)$, where:*

- *The sequence $(\sigma_N)_{N \geq 1}$ is a random virtual permutation following the Ewens measure of parameter θ.*
- *For all $j \geq 1$, z_j is a random variable following the distribution $\mathcal{L}$.*
- *The virtual permutation $(\sigma_N)_{N \geq 1}$ and the random variables $(z_j)_{j \geq 1}$ are independent.*

Notice that the eigenvalues of a matrix in $\mathcal{G}(N)$ can be easily computed in function of the corresponding cycle structure and the non-zero entries. More precisely, each cycle of length $m \geq 1$, for which the corresponding non-zero entries are $z_{j_1}, z_{j_2}, \dots, z_{j_m}$, gives m eigenvalues: the m-th roots of the product $z_{j_1} z_{j_2} \cdots z_{j_m}$, which form a regular polygon centered at the origin.

Now, let $N \geq 1$, $\theta > 0$, and suppose that $\mathcal{L}$ is a probability distribution on $\mathbb{C}^*$, which does not charge too much the elements with very large or very small modulus. For a random matrix following $\mathbb{P}(N, \theta, \mathcal{L})$, the large cycles give polygons with a large number of vertices, whose distance to the origin is given by the geometric mean of many i.i.d. random variables with law $\mathcal{L}$. By a geometric version of the law of large numbers, one can expect that this distance is close to the geometric mean R of the distribution $\mathcal{L}$. Since for N large, most of the eigenvalues correspond to large cycles, it is natural to expect that these eigenvalue are mostly concentrated around the circle of center 0 and radius R. The corresponding rigorous statement is the following:

Theorem 4.15 *Let $(M_N)_{N\geq 1}$ be a sequence of matrices following the distribution $\mathbb{P}(\infty, \theta, \mathcal{L})$ for some $\theta > 0$ and some probability measure $\mathcal{L}$ on $\mathbb{C}^*$. We suppose that for a random variable Z following the law $\mathcal{L}$, $\log(|Z|)$ is integrable, and we define $R := \exp(\mathbb{E}[\log(|Z|)])$. Then, almost surely, the probability measure $\mu(M_N)/N$, where $\mu(M_N)$ is the counting measure (with multiplicity) of the eigenvalues of M_N, converges weakly to the uniform distribution on the circle of center zero and radius R.*

Let us emphasize that in this result, the introduction of virtual permutations gives the possibility to obtain a strong convergence. Of course, this almost sure convergence implies the corresponding convergence in distribution for the finite-dimensional matrix ensembles, when the dimension goes to infinity.

Once we know that most of the eigenvalues are concentrated around a given circle, it is natural to ask about the behavior of the few eigenvalues which remain far from this circle. These eigenvalues are associated to the small cycles, and we know that asymptotically, the number of small cycles of a permutation following Ewens measure can be approximated in function of some independent Poisson random variables. More precisely, if $\sigma_N \in \mathfrak{S}_N$ is a permutation following Ewens measure of parameter $\theta > 0$, and if for $k \geq 1$, $a_{N,k}$ denotes the number of k-cycles in the permutation σ_N, then for $p \geq 1$, $(a_{N,1}, a_{N,2}, \dots, a_{N,p})$ converges in law to $(a_1, \dots, a_p)$, where $(a_k)_{k\geq 1}$ is a family of independent Poisson random variables, a_k having parameter θ/k. On the other hand, by looking carefully at the definitions, it is not difficult to check that for a random matrix M_N following $\mathbb{P}(N, \theta, \mathcal{L})$, the counting measure $\mu(M_N)$ of the eigenvalues of M_N can be given as follows:

$$\mu(M_N) = \sum_{k=1}^{\infty} \sum_{p=1}^{a_{N,k}} \sum_{\omega^k = T_{k,p}} \delta_\omega, \tag{4}$$

where for $k \geq 1$, $a_{N,k}$ is the number of k-cycles of the permutation induced by M_N, where for $k, p \geq 1$, the law of $T_{k,p}$ is the multiplicative convolution of k copies of $\mathcal{L}$, and where $(a_{N,k})_{k\geq 1}$ and the variables $T_{k,p}$, $k, p \geq 1$ are independent. In relation with (4), it is natural to introduce the following random measure:

$$\mu_\infty = \sum_{k=1}^{\infty} \sum_{p=1}^{a_k} \sum_{\omega^k = T_{k,p}} \delta_\omega$$

where the Poisson random variables $(a_k)_{k\geq 1}$ defined above are supposed to be independent of $T_{k,p}$, $k, p \geq 1$. One can then prove, under some technical assumptions, that the distribution of $\mu(M_N)$ converges to the distribution of μ_∞, if the measures are restricted to the points which are far from the circle considered in Theorem 4.15.

Theorem 4.16 *Let $(M_N)_{N\geq 1}$ be a sequence of matrices such that $M_N \in \mathcal{G}(N)$ follows the distribution $\mathbb{P}(N, \theta, \mathcal{L})$ for some $\theta > 0$ and some probability measure $\mathcal{L}$ on $\mathbb{C}^*$. We suppose that for a random* *variable Z following the law $\mathcal{L}$, $\log(|Z|)$ is square-integrable. Then, for all bounded, continuous functions f from $\mathbb{C}$ to $\mathbb{R}$, such that $f = 0$ on a neighborhood of the circle $|z| = R := \exp(\mathbb{E}[\log(|Z|)])$, f is almost surely integrable with respect to μ_∞, and the following convergence in distribution holds:*

$$\int_{\mathbb{C}} f\, d\mu(M_N) \underset{N\to\infty}{\longrightarrow} \int_{\mathbb{C}} f\, d\mu_\infty.$$

Remark 4.17 *For virtual permutations following the Ewens measure, the number of small cycles does not converge almost surely. Hence, it is not possible to prove a strong convergence in Theorem 4.16, even by introducing the measure $\mathbb{P}(\infty, \theta, \mathcal{L})$.*

The results stated until now are available only if the geometric law of large numbers applies under the measure $\mathcal{L}$. It is natural to ask what happens if this condition is not satisfied.

In this new setting, we are not able to obtain a general result, but we know how to deal with some particular cases, where the law $\mathcal{L}$ is constructed from a symmetric stable distribution.

Proposition 4.18 *Let $\rho > 0$, and let $\mathcal{L}$ be the law of $e^{i\Theta + \rho S_\alpha}$, where Θ is a uniform random variable on $[0, 2\pi)$ and S_α an independent standard symmetric stable random variable, with index $\alpha < 1$. For $\theta > 0$, let $(M_N)_{N\geq 1}$ be a sequence of random matrices such that M_N follows the distribution $\mathbb{P}(N, \theta, \mathcal{L})$. Then, the distribution of the random probability measure $\mu(M_N)/N$ converges to the law of the random measure $G_\theta \delta_0$, where G_θ is a beta random variable with parameters $(\theta/2, \theta/2)$, in the following sense: for all continuous functions f from $\mathbb{C}$ to $\mathbb{R}$, with compact support,*

$$\frac{1}{N} \int_{\mathbb{C}} f\, d\mu(M_N) \underset{N\to\infty}{\longrightarrow} G_\theta f(0)$$

in distribution.

The intuitive meaning of this result is that, since the stable random variables of index $\alpha < 1$ have heavy tail, the eigenvalues of M_N tend to be either close to zero or very large. The random variable G_θ represents the proportion of eigenvalues which remain small.

For the critical case $\alpha = 1$, corresponding to symmetric Cauchy random variables, the behavior of the eigenvalues is different: asymptotically, they concentrate around a family of circles centered at zero (corresponding to the cycle structure of the underlying random permutation), whose radii are given by exponentials of i.i.d., symmetric, Cauchy random variables. Notice that for $\alpha > 1$, the stable laws are integrable and then Theorem 4.15 applies.

The results we have previously stated are available for distributions $\mathcal{L}$ which are not a priori supposed to be carried by the unit circle $\mathbb{U}$. If we make this extra assumption, then the corresponding matrices are unitary, and one can study the behavior of their eigenangles, in particular in short scale. This problem is studied in the two last sections of [42].

More precisely, let $\mathcal{L}$ be a probability distribution on $\mathbb{U}$, and let $(M_N)_{N\geq 1}$ be a sequence of matrices following the distribution $\mathbb{P}(\infty, \theta, \mathcal{L})$ (in particular, M_N has law $\mathbb{P}(N, \theta, \mathcal{L})$ for all $N \geq 1$). For $N \geq 1$, we consider the random point measure $\tau_N(M_N)$ on $\mathbb{R}$ which charges the values $x \in \mathbb{R}$ (counted with multiplicity) such that $e^{2i\pi x/N}$ is an eigenvalue of M_N. The measure $\tau_N(M_N)$ is N-periodic, and $\tau_N(M_N)([0, N)) = N$, since M_N has N eigenvalues. In other words, the scaling of the eigenangles, which are multiplied by $N/2\pi$, is chosen to get a point process whose average spacing is one.

Now, for $N, m \geq 1$, let σ_N be the permutation induced by M_N, let $Z_{N,m} \in \mathbb{U}$ be the product of the non-zero entries of M_N which correspond to the m-th largest cycle of σ_N (if several cycles have the same length, they are ordered arbitrarily, if σ_N has less than m cycles, $Z_{N,m} = 1$), let $\ell_{N,m}$ be the length of this cycle ($\ell_{N,m} = 0$ if σ_N has less than m cycles), and let $y_{N,m} := \ell_{N,m}/N$. Then, the measure $\tau_N(M_N)$ is explicitly written as follows:

$$\tau_N(M_N) = \sum_{m=1}^{\infty} \mathbb{1}_{y_{N,m}>0} \sum_{k \in \mathbb{Z}} \delta_{(\gamma_{N,m}+k)/y_{N,m}}, \tag{5}$$

where δ_x denotes the Dirac measure at x, and where $\gamma_{N,m}$, defined modulo one, is $1/2\pi$ times the argument of $Z_{N,m}$.

For all $m \geq 1$, by the general properties of the virtual permutations (for example, see Tsilevich [66]), the renormalized cycle length $y_{N,m}$ converges almost surely, when N goes to infinity, to a limit y_m, and the sequence $(y_m)_{m\geq 1}$ is a Poisson-Dirichlet process of parameter θ. In the case where the distribution $\mathcal{L}$ is the Dirac measure at one, i.e. the matrices $(M_N)_{N\geq 1}$ are permutation matrices, the formula (5) is simplified by the fact that $\gamma_{N,m} = 0$, and one naturally expects a convergence of the measure $\tau_N(M_N)$ towards the following limit measure

$$\tau_\infty((M_N)_{N\geq 1}) := \sum_{m=1}^{\infty} \sum_{k \in \mathbb{Z}} \delta_{k/y_m}.$$

Note that this measure has an infinite mass at zero, which reflects the fact that the number of cycles of σ_N tends to infinity with N. The precise result we obtain is the following:

Proposition 4.19 *Let $(M_N)_{N\geq 1}$ be a sequence of matrices following the distribution $\mathbb{P}(\infty, \theta, \delta_1)$ for some $\theta > 0$. Then, the measure $\tau_N(M_N)$ converges a.s. to $\tau_\infty((M_N)_{N\geq 1})$, in the following sense: for all continuous functions from $\mathbb{R}$ to $\mathbb{R}_+$, with compact support:*

$$\int_{\mathbb{R}} f \, d\tau_N(M_N) \underset{N\to\infty}{\longrightarrow} \int_{\mathbb{R}} f \, d\tau_\infty((M_N)_{N\geq 1}),$$

almost surely.

In the case where $\mathcal{L}$ is not the Dirac measure at one, the shifts $\gamma_{N,m}$ introduced in (5) do not vanish in general, and one cannot expect that they converge almost surely. Hence, it is not possible to obtain an almost sure convergence of the measures $\tau_N(M_N)$ when N goes to infinity: however, a weak convergence is still possible. More precisely, the convergence of the renormalized cycle lengths $y_{N,m}$ holds for any choice of the distribution $\mathcal{L}$, and then one can expect a convergence in law of the random measures $\tau_N(M_N)$ if the distribution of $\gamma_{N,m}$ converges.

Intuitively, for N large, $Z_{N,m}$ is the product of many i.i.d. random variables on the unit circle, and then it is plausible, by some kind of mixing property, that $Z_{N,m}$ tends to a uniform distribution on the unit circle, implying a convergence of $\gamma_{N,m}$ to a uniform law on $[0, 1)$. This very informal argument is certainly not completely true: the case $\mathcal{L} = \delta_0$ stated just above is a counterexample, and more generally, if $\mathcal{L}$ is carried by $\mathbb{U}_k$ for some $k \geq 1$, then $Z_{N,m}$ is always a k-th root of unity and $\gamma_{N,m}$ is a multiple of $1/k$. However, up to a technical condition stated below, this case is essentially the only possible obstruction for a convergence of $\gamma_{N,m}$ to a uniform distribution. More precisely, the limiting measure of $\gamma_{N,m}$ is, under very general assumptions, given by the following definition:

Definition 4.20 *Let $\mathcal{L}$ be a distribution on the unit circle, and let k_0 be the infimum of the integers $k \geq 1$ such that $\mathcal{L}$ is carried by $\mathbb{U}_k$. Then, the probability measure $\mathcal{D}(\mathcal{L})$ on $[0, 1)$ is defined as follows:*

- *If $k_0 < \infty$, then $\mathcal{D}(\mathcal{L})$ is the uniform distribution on $\{0, 1/k_0, 2/k_0, \dots, (k_0-1)/k_0\}$.*
- *If $k_0 = \infty$, i.e. $\mathcal{L}$ is not carried by $\mathbb{U}_k$ for any $k \geq 1$, then $\mathcal{D}(\mathcal{L})$ is the uniform distribution on $[0,1)$.*

A more careful version of the mixing argument above suggests that the law of $\gamma_{N,m}$ converges to $\mathcal{D}(\mathcal{L})$, and that the measure $\tau_N(M_N)$ converges in distribution to the random measure

$$\tau_\infty := \sum_{m=1}^{\infty} \sum_{k \in \mathbb{Z}} \delta_{(\gamma_m + k)/y_m}, \tag{6}$$

where $(y_m)_{m\geq 1}$ is a Poisson-Dirichlet process of parameter θ, independent of the sequence $(\gamma_m)_{m\geq 1}$ of i.i.d. random variables on $[0,1)$ with law $\mathcal{D}(\mathcal{L})$. The following result gives a precise meaning to this intuition:

Proposition 4.21 *Let $(M_N)_{N\geq 1}$ be a sequence of random matrices, such that for all $N \geq 1$, M_N follows the distribution $\mathbb{P}(N, \theta, \mathcal{L})$, where $\theta > 0$ and where $\mathcal{L}$ is a probability measure on $\mathbb{U}$. We assume that $\mathcal{L}$ satisfies one of the two following conditions:*

- *$\mathcal{L}$ is carried by $\mathbb{U}_k$ for some $k \geq 1$.*
- *There exists $v > 1$ such that for $\epsilon > 0$ small enough, and for any arc $\mathcal{A}$ in $\mathbb{U}$ of size ϵ, $\mathcal{L}(\mathcal{A}) \leq |\log(\epsilon)|^{-v}$.*

Then, with the notation above, the random measure $\tau_N(M_N)$ converges in distribution to the measure τ_∞, in the following sense: for all continuous functions f from $\mathbb{R}$ to $\mathbb{R}_+$, with compact support,

- *If $f(0) > 0$ and $\mathcal{L}$ is carried by $\mathbb{U}_k$ for some $k \geq 1$, then*

$$\int_{\mathbb{R}} f \, d\tau_\infty = \infty$$

a.s., and for all $A > 0$,

$$\mathbb{P}\left[\int_{\mathbb{R}} f \, d\tau_N(M_N) \leq A\right] \underset{N\to\infty}{\longrightarrow} 0.$$

- *If $f(0) = 0$ or $\mathcal{L}$ is not carried by $\mathbb{U}_k$ for any $k \geq 1$, then*

$$\int_{\mathbb{R}} f \, d\tau_\infty < \infty,$$

and

$$\int_{\mathbb{R}} f \, d\tau_N(M_N) \underset{N\to\infty}{\longrightarrow} \int_{\mathbb{R}} f \, d\tau_\infty.$$

In this result, the distinction of two cases comes from the fact that $\tau_\infty(\{0\})$ is a.s. infinite if $\mathcal{L}$ is carried by $\mathbb{U}_k$ for some $k \geq 1$, and a.s. equal to zero otherwise.

A particularly simple case where Proposition 4.21 applies corresponds to a measure $\mathcal{L}$ equal to the uniform distribution on $\mathbb{U}$. In this case, and more generally in any situation for which $\mathcal{L}$ is not carried by $\mathbb{U}_k$ for any $k \geq 1$, the law of the measure τ_∞ is uniquely determined by the parameter θ, and in (6), the random variables $(\gamma_m)_{m\geq 1}$ are i.i.d., uniform on $[0,1)$.

In the case of the Circular Unitary Ensemble, we have seen that the point process of the renormalized eigenangles converges, in distribution, to a determinantal sine-kernel process. The point process induced by the measure τ_∞ is then the analogue of the sine-kernel process in the context of randomized permutation matrices. These two point processes have the following common properties:

- They have only simple points.
- Their distribution is invariant by translation.

However, there are important differences between the two processes. For example, since τ_∞ is obtained by taking unions of arithmetic progressions, the probability that there exist three points $x, y, z \in \mathbb{R}$ in τ_∞, such that $x + z = 2y$, is equal to one, whereas it is zero in the case of the sine-kernel process. For the same reason, the r-point correlation functions are not defined for τ_∞ as soon as $r \geq 3$. For $r = 1$, the average density is equal to one by translation invariance, and for $r = 2$, the correlation function ρ is well-defined and computed in [42]: one has $\rho(x, y) = \phi(x - y)$, where the function ϕ from $\mathbb{R}$ to $\mathbb{R}$ is given by

$$\phi(x) = \frac{\theta}{\theta + 1} + \frac{\theta}{x^2} \sum_{a \in \mathbb{N}, 0 < a \leq |x|} a \left(1 - \frac{a}{|x|}\right)^{\theta - 1}.$$

The correlation function ρ is always larger than or equal to $\theta/(\theta + 1)$: in particular, is does not tend to zero with the distance between the two points. In other words, there does not exist the same phenomenon of repulsion as in the unitary framework. On the other hand, one can check that $\rho(x, y)$ tends to one when $|x - y|$ goes to infinity: two points which are far from each other behave almost independently. In this point of the view, the situation is the same as for the sine-kernel process.

In [42], we also study the behavior of the smallest strictly positive point charged by the measure τ_∞. For the sine-kernel process, it is given in terms of a Painlevé differential equation. For τ_∞, we do not obtain a direct formula, but we get the following result:

Proposition 4.22 *For $x \geq 0$, let $G(x)$ be the probability that the point process τ_∞ has no point in the interval $(0, x)$, and for $x \in \mathbb{R}$, let us set*

$$H(x) := \mathbb{1}_{x>0} x^{\theta - 1} G(x).$$

Then H is integrable and satisfies the following equation:

$$xH(x) = \theta \int_0^1 (1 - y) H(x - y) dy.$$

Moreover, the Fourier transform $\widehat{H}$ of H satisfies the following:

$$\widehat{H}(\lambda) = \widehat{H}(0) \exp\left(-i\theta \int_0^\lambda \frac{1 - e^{-i\mu} - i\mu}{\mu^2} \, d\mu\right).$$

5 The framework of the unitary group

5.1 The extreme eigenvalues of generalized Cauchy ensembles

As written in Section 4, the Ewens measure on the symmetric group $\mathfrak{S}_N$, with parameter $\theta > 0$, is the measure giving to each permutation σ a probability proportional to $w(\sigma) := \theta^k$, where k is the number of cycles. This weight $w(\sigma)$ can be interpreted in the point of view of the random matrices: if M_σ denotes the matrix of σ, then $w(\sigma)$ is the limit of

$$[\det(x \operatorname{Id} - M_\sigma)]^{\log(\theta)/\log(x-1)}$$

when x tends to one from above.

This remark suggests to define an analog of the Ewens measure for the unitary matrices: for $N \geq 1$ and $s \in \mathbb{C}$ such that $\Re(s) > -1/2$, one considers a measure $\mu_{N,s}$ on the unitary group $U(N)$, whose density with respect to the Haar measure $\mu_{N,0}$ is given by

$$\frac{d\mu_{N,s}}{d\mu_{N,0}}(u) := C_{N,s} \det(\operatorname{Id} - u)^{\bar{s}} \det(\operatorname{Id} - u^{-1})^{s},$$

for $u \in U(N)$, $C_{N,s} > 0$ being a normalization constant. The complex powers are determined by assuming their continuity with respect to u and the restriction made on the real part of s is given in order to ensure the integrability of the density. Such samplings have first been introduced by Hua in [20], and have more recently been studied by Pickrell in [53] and [54], by Neretin in [49] and by Bourgade, Nikeghbali and Rouault in [12].

The punctured unit circle $\mathbb{U}\backslash\{1\}$ can be transformed to the real line $\mathbb{R}$ by the so-called *Cayley transform*, given by the bijective map:

$$z \mapsto i\,\frac{1+z}{1-z}.$$

This measure can be extended to a map from the space of unitary matrices whose eigenvalues are different from one, to the space of hermitian matrices.

The image, by the Cayley transform, of the measures introduced just above corresponds to the so-called *Generalized Cauchy Ensemble*. For $N \geq 1$ and $\Re(s) > -1/2$, this ensemble is given by the probability measure $\mathbb{P}_{N,s}$ on the space of N-dimensional hermitian matrices, defined by

$$\mathbb{P}_{N,s}(dX) = C'_{N,s} \det((\operatorname{Id} + iX)^{-s-N}) \det((\operatorname{Id} - iX)^{-\bar{s}-N}) \prod_{1 \leq j < k \leq N} d\Re(X_{j,k}) d\Im(X_{j,k}) \prod_{j=1}^{N} dX_{j,j},$$

where $C'_{N,s} > 0$ is a normalization constant. The distribution of the eigenvalues of a random matrix following the measure $\mathbb{P}_{N,s}$ is given, up to a normalization constant, by the following formula:

$$\prod_{1 \leq j < k \leq N} (x_k - x_j)^2 \prod_{1 \leq j \leq k} w_H(x_j)\, dx_j, \tag{7}$$

where the eigenvalues are $(x_j)_{1 \leq j \leq N}$, and where the weight w_H is given by

$$w_H(x) = (1 + ix)^{-s-N} (1 - ix)^{-\bar{s}-N}.$$

The repulsion between the eigenvalues of the matrix, which corresponds to the factors $(x_k - x_j)^2$, is involved in most of the classical ensembles of random hermitian or unitary matrices (e.g. Gaussian Unitary Ensemble and Circular Unitary Ensemble).

The point process whose density is given by (7) is determinantal an has been studied in detail by Borodin and Olshanski in [10], and by Forrester and Witte in [18] and [78]. By using the classical theory of orthogonal polynomials, applied to the particular case of the weight w_H, Borodin and Olshanski prove the following result:

Proposition 5.1 *For $1 \leq n \leq N$, the n-point correlation of the eigenvalue distribution given by (7) is expressed as follows:*

$$\rho_n^{N,s}(x_1, x_2, \ldots, x_n) = \det\left(K_{N,s}(x_j, x_k)\right)_{1 \leq j,k \leq n},$$

where the kernel $K_{N,s}$ is defined, for $x \neq y$, by

$$K_{N,s}(x,y) = C''_{N,s} \frac{p(x)q(y) - q(x)p(y)}{x-y} \sqrt{w_H(x)\, w_H(y)},$$

for

$$p(x) := (x-i)^N\, {}_2F_1\left[-N, s, 2\Re(s)+1; \frac{2}{1+ix}\right],$$

$$q(x) := (x-i)^{N-1}\, {}_2F_1\left[1-N, s+1, 2\Re(s)+2; \frac{2}{1+ix}\right],$$

and

$$C_{N,s} := \frac{2^{2\Re(s)}}{\pi} \frac{\Gamma(2\Re(s)+N+1)\Gamma(s+1)\Gamma(\bar{s}+1)}{\Gamma(N)\Gamma(2\Re(s)+1)\Gamma(2\Re(s)+2)}.$$

For $x = y$, the kernel is extended by continuity.

Here ${}_2F_1$ corresponds to the classical notation of the hypergeometric functions. Moreover, in [10], the authors give a scaling limit for the kernel $K_{N,s}$ when N goes to infinity. Namely, for fixed $s \in \mathbb{C}$ ($\Re(s) > -1/2$) and $x, y \in \mathbb{R}^*$,

$$N\, \mathrm{Sgn}(xy)^N\, K_{N,s}(Nx, Ny) \underset{N \to \infty}{\longrightarrow} K_{\infty,s}(x,y),$$

where for $x \neq y$, the kernel $K_{\infty,s}$ is given by

$$K_{\infty,s}(x,y) = \frac{1}{2\pi} \frac{\Gamma(s+1)\Gamma(\bar{s}+1)}{\Gamma(2\Re(s)+1)\Gamma(2\Re(s)+2)} \frac{P(x)Q(y) - Q(x)P(y)}{x-y}, \tag{8}$$

for

$$P(x) = |2/x|^{\Re(s)} e^{-i/x + \pi\Im(s)\,\mathrm{sgn}(x)/2}\, {}_1F_1\left[s, 2\Re(s)+1; \frac{2i}{x}\right] \tag{9}$$

and

$$Q(x) = (2/x)|2/x|^{\Re(s)} e^{-i/x + \pi\Im(s)\,\mathrm{sgn}(x)/2}\, {}_1F_1\left[s+1, 2\Re(s)+2; \frac{2i}{x}\right]. \tag{10}$$

Again, the kernel $K_{\infty,s}$ is extended by continuity for $x = y$.

This kernel defines a determinantal point process X_s, containing infinitely many points on the punctured real line $\mathbb{R}^*$. In the particular case $s = 0$, for which

$$K_{\infty,0}(x,y) = \frac{\sin[(1/y) - (1/x)]}{\pi(x-y)},$$

the image of X_s by the map $x \mapsto 1/\pi x$ is a determinantal sine-kernel process.

It is natural to expect that the convergence of the kernel $K_{N,s}$ to the kernel $K_{\infty,s}$ implies the convergence, in a sense which has to be made precise, of the eigenvalue point process of the Generalized Cauchy Ensemble towards the determinantal process X_s. In our joint work with Nikeghbali and Rubin [47], we study this problem in detail for the largest eigenvalue. We obtain the expected convergence, with an estimation of its rate:

Theorem 5.2 *For $N \geq 1$ and $\Re(s) > -1/2$, let $\lambda(N,s)$ be the largest eigenvalue of a random hermitian matrix following the distribution $\mathbb{P}_{N,s}$, and let $\lambda(s)$ be the largest point of the determinantal process X_s described above. Then, for all s, $\Re(s) > -1/2$, and for all $x_0 > 0$, there exists $C(x_0, s) > 0$ such that for all $N \geq 1$ and $x \geq x_0$,*

$$\left| \mathbb{P}\left(\frac{\lambda(N,s)}{N} \leq x\right) - \mathbb{P}(\lambda(s) \leq x) \right| \leq \frac{C(x_0,s)}{N}. \tag{11}$$

In particular,

$$\mathbb{P}\left(\frac{\lambda(N,s)}{N} \leq x\right) \underset{N\to\infty}{\longrightarrow} \mathbb{P}(\lambda(s) \leq x)$$

Notice the scaling factor $1/N$, which shows that the largest point of a Generalized Cauchy Ensemble has the same order of magnitude as its dimension. In [47], the proof of Theorem 5.2 uses some estimates of the Fredholm determinants giving the probabilities involved in (11). Our method is rather technical: in particular, we do some heavy computations on hypergeometric functions in order to get good majorizations of the kernel $K_{N,s}$ and its derivatives. However, our proof has the advantage to use only elementary tools, contrarily to the most common situation in random matrix theory.

Note that the probablity distribution of $\lambda(N,s)$ satisfies a differential equation, explicitly computed by Forrester and Witte in [18]:

Theorem 5.3 *Let σ be the function from $\mathbb{R}$ to $\mathbb{R}$, given by*

$$\sigma(t) := (1+t^2)\frac{d}{dt}\log \mathbb{P}(\lambda(N,s) \leq t).$$

Then, σ is well-defined, twice differentiable and satisfies the following equation, for all $t \in \mathbb{R}$:

$$(1+t^2)(\sigma''(t))^2 + 4(1+t^2)(\sigma'(t))^3 - 8t(\sigma'(t))^2\sigma(t) + 4(\sigma(t))^2(\sigma'(t) - (\Re(s))^2) + 8[t(\Re(s))^2 - \Re(s)\Im(s)$$
$$-N\Im(s)]\sigma(t)\sigma'(t) + 4[2t\Im(s)(N+\Re(s)) - (\Im(s))^2 - t^2(\Re(s))^2 + N(2\Re(s)+N)](\sigma'(t))^2 = 0.$$

Theorem 5.2 suggests that the distribution of $\lambda(s)$ should satisfy some kind of scaling limit of the differential equation given in Theorem 5.3. In [47], we prove that this intuition is good and we obtain the following result:

Theorem 5.4 *Let θ be the function from $\mathbb{R}_+^*$ to $\mathbb{R}$, given by*

$$\theta(\tau) = \tau \frac{d}{d\tau}\log \mathbb{P}(\lambda(s) \leq 1/\tau).$$

Then θ is well-defined, twice differentiable, and satisfies the following differential equation:

$$-\tau^2(\theta''(\tau))^2 = \left[2(\tau\theta'(\tau) - \theta(\tau)) + (\theta'(\tau))^2 + i(\bar{s}-s)\theta'(\tau)\right]^2 - (\theta'(\tau))^2(\theta'(\tau) - 2is)(\theta'(\tau) + 2i\bar{s}). \tag{12}$$

This theorem implies in particular the result of Jimbo, Miwa, Môri and Sato [22], which says that the probability distribution of the smallset positive point of a determinantal sine-kernel process satisfies the Painlevé V equation (12) for $s = 0$.

5.2 The space of virtual isometries

As proven by Hua [20] and Pickrell [53], and also stated by Borodin and Olshanski in [10], the measures $\mathbb{P}_{N,s}$ ($N \geq 1$, $\Re(s) > -1/2$) corresponding to the Generalized Cauchy Ensemble admit the following property of consistency: if for $N \geq 2$, a $N \times N$ random hermitian matrix follows the distribution $\mathbb{P}_{N,s}$, then its $(N-1) \times (N-1)$ upper left corner has law $\mathbb{P}_{N-1,s}$. This property of compatibility implies that for all $s \in \mathbb{C}$, $\Re(s) > -1/2$, the family of measures $(\mathbb{P}_{N,s})_{N \geq 1}$ induces a probability measure $\mathbb{P}_{\infty,s}$ on the space $\mathcal{H}$ of "infinite hermitian matrices", defined as follows: $\mathcal{H}$ is the set of families $(M_{j,k})_{j,k \geq 1}$ of complex numbers such that $M_{j,k} = \overline{M_{k,j}}$ for all integers $j, k \geq 1$. The measures $\mathbb{P}_{\infty,s}$ have been studied in detail in [10], where they are called *Hua-Pickrell measures*.

On the other hand, as we have seen before, it is natural to consider the Generalized Cauchy Ensembles as the images, by the Cayley transform, of the ensembles of unitary matrices studied by Bourgade, Nikeghbali and Rouault in [12]. The case where $s = 0$ is particularly simple in this point of view, since $\mathbb{P}_{N,0}$ is the image of the Haar measure on $U(N)$. A question which then naturally arises is the following: what becomes the compatibility property between the upper left corners of an infinite hermitian matrix when we look at their inverse images by the Cayley transform?

This question has been answered by Neretin in [49]: in the unitary point of view, the infinite hermitian matrices become the so-called *virtual isometries* (or *virtual rotations*), which are sequences $(u_N)_{N \geq 1}$ of unitary matrices, $u_N \in U(N)$, satisfying a relation of compatibility explicitly described in [49]. Now, the virtual permutations can also be identified to particular sequences of unitary matrices of increasing dimensions. Note that these two kinds of sequences are incompatible:

- If $(u_N)_{N \geq 1}$ is a virtual isometry in the sense of Neretin, then 1 is not an eigenvalue of u_N for any $N \geq 1$, since the inverse image of $\mathbb{R}$ by the Cayley transform is $\mathbb{U} \backslash \{1\}$.
- If $(u_N)_{N \geq 1}$ is a sequence of permutations matrices, then 1 is an eigenvalue of u_N for all $N \geq 1$.

However, the virtual isometries in the sense of Neretin and the virtual permutations can be proven to be particular cases of a more general notion of virtual isometries (or virtual rotations), which is the main topic of our joint work with Bourgade and Nikeghbali [11].

The first step of the construction given in [11] is to define, for all $N \geq M \geq 1$, a map $\pi_{N,M}$ from $U(N)$ to $U(M)$. This map can be described as follows: for all $u \in U(N)$, $\pi_{N,M}(u)$ is the unique matrix $v \in U(M)$ such that the rank of $u - \begin{pmatrix} v & 0 \\ 0 & \mathrm{Id_{N\text{-}M}} \end{pmatrix}$ is minimal. The uniqueness of v is proven in [11], and in the case where 1 is not an eigenvalue of u, the minimal rank is equal to $N - M$. Moreover, the maps $(\pi_{N,M})_{N \geq M \geq 1}$ satisfy the following projective property: for all $N \geq M \geq P \geq 1$, $\pi_{N,P} = \pi_{M,P} \circ \pi_{N,M}$. A virtual isometry is then defined as follows:

Definition 5.5 *A virtual isometry is a sequence $(u_N)_{N \geq 1}$ of unitary matrices, $u_N \in U(N)$, such that for all $N \geq 2$, $u_{N-1} = \pi_{N,N-1}(u_N)$.*

The projective property of $(\pi_{N,M})_{N \geq M \geq 1}$ implies that for any virtual isometry $(u_N)_{N \geq 1}$, and for all $N \geq M \geq 1$, $u_M = \pi_{N,M}(u_N)$. The link between the virtual isometries and the virtual permutations can be precisely stated as follows:

Proposition 5.6 *Let $(\sigma_N)_{N\geq 1}$ be a sequence of permutations ($\sigma_N \in \mathfrak{S}_N$) identified with the corresponding permutation matrices. Then, $(\sigma_N)_{N\geq 1}$ is a virtual permutation if and only if it is a virtual isometry.*

As we have seen in Section 4, the virtual permutations can be constructed as formal infinite products of transpositions, via the Chinese restaurant process (see Pitman [55], Chap. 3). This construction can be generalized to the unitary setting, replacing transpositions by the *complex reflections*, i.e. the unitary matrices $u \in U(N)$ such that $u - \mathrm{Id}$ has rank one. Notice that a transposition is a particular case of a reflection, if the permutations are identified with their matrices. Another property is important for our purpose: for all distinct vectors x, y on the unit sphere of $\mathbb{C}^N$, there exists a unique reflection $r \in U(N)$ such that $r(x) = y$. This property is used in the following result, generalizing the construction of the Chinese restaurant process:

Proposition 5.7 *Let $(x_N)_{N\geq 1}$ be a sequence of vectors, x_N lying on the complex unit sphere of $\mathbb{C}^N$. For $k \geq 1$, let $r_k \in U(k)$ be defined as follows:*

- *If x_k is not the last canonical basis vector e_k of $\mathbb{C}^k$, then r_k is the unique reflection such that $r_k(e_k) = x_k$.*
- *If $x_k = e_k$, then $r_k = \mathrm{Id}_k$.*

Then, there exists a unique virtual isometry $(u_N)_{N\geq 1}$ such that $u_N(e_N) = x_N$ for all $N \geq 1$: this isometry is given by

$$u_N = r_N \circ \begin{pmatrix} r_{N-1} & 0 \\ 0 & 1 \end{pmatrix} \circ \cdots \circ \begin{pmatrix} r_2 & 0 \\ 0 & \mathrm{Id}_{N-2} \end{pmatrix} \circ \begin{pmatrix} r_1 & 0 \\ 0 & \mathrm{Id}_{N-1} \end{pmatrix}.$$

Moreover, in the particular case where for all $N \geq 1$, x_N is the k_N-th canonical basis vector of $\mathbb{C}^N$ for $k_N \in \{1, 2, \dots, N\}$, $(u_N)_{N\geq 1}$ is identified with the virtual permutation $(\sigma_N)_{N\geq 1}$ constructed by the Chinese restaurant process, as in Section 4:

$$\sigma_N = \tau_{N,k_N} \circ \tau_{N-1,k_{N-1}} \circ \cdots \circ \tau_{2,k_2},$$

where $\tau_{j,k}$ is the identity for $j = k$, and the transposition (j, k) for $j \neq k$.

As we have seen before, the Generalized Cauchy Ensemble corresponds to a particular class of probability measures on the space of infinite hermitian matrices, which are called Hua-Pickrell measures in [10]. Applying the inverse Cayley transform gives a family of measures on the space of virtual isometries, which can also be called Hua-Pickrell measures. More precisely, the following result, stated in our paper [11], can be immediately proven by using the results obtained by Bourgade, Nikeghbali and Rouault in [12]:

Proposition 5.8 *Let $(x_N)_{N\geq 1}$ be a random sequence of independent vectors such that for all $N \geq 1$, the distribution of x_N is the h-sampling of the uniform law on the unit sphere of $\mathbb{C}^N$, where*

$$h(x) = (1 - \langle e_N | x \rangle)^{\bar{s}} \left(1 - \overline{\langle e_N | x \rangle}\right)^s.$$

Let $(u_N)_{N\geq 1}$ be the unique virtual isometry such that $u_N(e_N) = x_N$ for all $N \geq 1$. Then, the law of $(u_N)_{N\geq 1}$ is the Hua-Pickrell measure of parameter s, and for all $N \geq 1$, the distribution of u_N is the h-sampling of the Haar measure on $U(N)$, for

$$h(u) = \det(\mathrm{Id}_N - u)^{\bar{s}} \det(\mathrm{Id}_N - u^{-1})^s,$$

where the logarithm of $\mathrm{Id}_N - u$ *is taken in the unique way such that it is continuous on the connected set* $\{u \in U(n), \det(\mathrm{Id}_N - u) \neq 0\}$ *and real for* $u = -\mathrm{Id}_N$*.*

Remark 5.9 *In the particular case* $s = 0$*, and for all* $N \geq 1$*,* x_N *is uniform on the unit sphere of* $\mathbb{C}^N$ *and* u_N *follows the Haar measure on* $U(N)$*. Is it then natural to say that the law of* $(u_N)_{N\geq 1}$ *is the Haar measure on the space of virtual isometries.*

5.3 A determinantal sine-kernel process associated to virtual rotations

The probability measures on the space of infinite hermitian matrices which are invariant by conjugation have been classified by Olshanski and Vershik in [51]. In [10], Borodin and Olshanski study the particular case of the Hua-Pickrell measures, for which the following result holds:

Theorem 5.10 *Let* M *be a random infinite hermitian matrix following the Hua-Pickrell measure of parameter* $s \in \mathbb{C}$ $(\Re(s) > -1/2)$*. For* $N \geq 1$*, let* M_N *be the upper left corner of* M*, and for* $N \geq k \geq 1$*, let* $\lambda_k^+(N)$ *(resp.* $\lambda_k^-(N)$*) be the k-th largest (resp. smallest) eigenvalue of* M_N*. Then almost surely, for all* $k \geq 1$*,* $\lambda_k^+(N)$ *(resp.* $\lambda_k^-(N)$*) converges to a strictly positive (resp. strictly negative) limit* λ_k^+ *(resp.* λ_k^-*). Moreover, the point process* $\{\lambda_k^+, k \geq 1\} \cup \{\lambda_k^-, k \geq 1\}$ *is determinantal, and its kernel is given by the equations* (8)*,* (9) *and* (10)*.*

This statement can be translated in the unitary framework. In this case, one obtains, after a suitable rescaling, an almost sure convergence of the small eigenangles towards a limiting point process. This process, obtained by transforming the point process involved in Theorem 5.10 via the map $x \mapsto 1/\pi x$, is also determinantal and its kernel is explicitly described in [12]. In the particular case of the Haar measure, i.e. $s = 0$, one gets the following:

Theorem 5.11 *Let* $(u_N)_{N\geq 1}$ *be a random virtual isometry, following the Haar measure. For* $N \geq 1$*,* $k \geq 1$*, let* $\theta_k^{(N)}$ *be the k-th smallest strictly positive eigenangle of* u_N*, and let* $\theta_{1-k}^{(N)}$ *be the k-th largest nonnegative eigenangle of* u_N*. Then almost surely, for all* $k \in \mathbb{Z}$*,* $N\theta_k^{(N)}/2\pi$ *converges to a limit* x_k *when* N *goes to infinity, and the point process* $(x_k)_{k\in\mathbb{Z}}$ *is a determinantal sine-kernel process.*

This theorem can be easily deduced from the general results given by Olshanski and Vershik in [51], and from the study of the Hua-Pickrell measures made in [10]. However, the proof in [51] uses quite sophisticated tools, including representation theory. In [11], we give another proof of Theorem 5.11, more direct and purely probabilistic, using a recurrence relation between the characteristic polynomials of $(u_N)_{N\geq 1}$, and some martingale arguments. Another advantage of our proof is that it gives an estimate of the rate of convergence of the rescaled eigenvalues: there exists a universal constant $\epsilon > 0$ such that almost surely, for all $k \in \mathbb{Z}$,

$$N\theta_k^{(N)}/2\pi = x_k + O(N^{-\epsilon}).$$

It is well-known that the rescaled eigenangles of the Circular Unitary Ensemble weakly converge to a determinantal sine-kernel process, when the dimension goes to infinity. Theorem 5.11 gives a strong version of this classical result, and provides a deterministic link between the random virtual isometry $(u_N)_{N\geq 1}$ and the sine-kernel process $(x_k)_{k\in\mathbb{Z}}$. In this respect, the situation is similar to the setting of virtual permutations, for which there also exists a strong convergence of the rescaled eigenangles. More precisely, for a random virtual permutation $(\sigma_N)_{N\geq 1}$ following the Ewens measure of parameter $\theta > 0$, a result similar to Theorem 5.11 can be stated, and the limiting point process $(x_k)_{k\geq 1}$ is the union of the sets $(2\pi m/\lambda_k)_{m\in\mathbb{Z}}$ for $k \geq 1$, where $(\lambda_k)_{k\geq 1}$ is a Poisson-Dirichlet process of parameter θ.

As stated in Section 4, in the context of virtual permutations, the limiting point process $(x_k)_{k\in\mathbb{Z}}$ can be interpreted as the spectrum of an operator, which is naturally related to the corresponding virtual permutation $(\sigma_N)_{N\geq 1}$. In the context of Haar measure on virtual isometries, one may expect a similar situation, for which the sine-kernel process $(x_k)_{k\in\mathbb{Z}}$ is the spectrum of a random hermitian operator. However, we are still not able to construct such an operator in a fully satisfactory way.

6 Some other perspectives of research

A first direction of research we plan to explore is given by the problem suggested at the end of Section 5, i.e. the construction, for almost every virtual isometry (under Haar measure), of a suitable hermitian operator whose spectrum is given by the point process $(x_k)_{k\in\mathbb{Z}}$ of the limiting rescaled eigenangles. If we success in this construction, then the corresponding operator may satisfy some universal properties, which should improve our understanding of the universality of the sine-kernel process, still very incomplete. More precisely, the operator may represent, in a natural way, the typical microscopic behavior of a general random matrix around a point in the bulk of the spectrum, and it should play a role which can be compared to the role of the Brownian motion in the explanation of the universal properties of the gaussian distribution. Note that a stochastic operator whose eigenvalues form a determinantal Airy kernel point process (which appears in the limiting behavior of the Gaussian Unitary Ensemble, at the edge of the spectrum) has already been constructed by Ramírez, Rider and Virág [56]. It is perhaps possible to construct another random operator with the same distribution of the spectrum, for which one can see even more clearly the corresponding properties of universality. One may also be able to extend this kind of construction to more general point processes.

Another interest of a universal operator with sine-kernel process as its spectrum is that it can improve our understanding of the links between random matrices and other parts of mathematics, e.g. number theory. In 1973, Montgomery [34] made a conjecture about the asymptotic repartition of the zeros of the Riemann zeta function on the critical line, and Dyson remarked that the corresponding limit behavior is similar to what one obtains by considering the classical ensembles of random hermitian or unitary matrices. In other words, the zeros of the zeta function should locally behave like a determinantal sine-kernel process. On the other hand, if a remarkable set of complex numbers is studied, then it is particularly interesting to find a spectral interpretation of this set: in relation with the Riemann hypothesis, it has been conjectured, first by Hilbert and Pólya, that the nontrivial zeros of the function $t \mapsto \zeta(1/2 + it)$ are the eigenvalues of some remarkable hermitian operator. If such an operator exists (a construction has been suggested by Connes [14]), it should involve some kind of randomness, since its eigenvalues seem to behave like the spectrum of a large random matrix. Now, if we are able to construct a universal random hermitian operator whose spectrum is a sine-kernel process, then it may give some indication about the nature of the randomness involved in the hypothetical operator of Hilbert and Pólya.

It should also be interesting to construct, in a natural way, some random holomorphic functions whose zeros form a determinantal sine-kernel process: such functions may be used to model the asymptotic behavior of the Riemann zeta function. They may also be interpreted as an analog of the characteristic polynomial for virtual isometries, if we are still able to prove a result of strong convergence in this context. More generally, the conjectural link between random matrix theory and the Riemann zeta function suggests to compare this function to the characteristic polynomial of a random matrix.

Note that this conjectural link is supported by some similar results proven for zeta functions of function fields, which are constructed from the number of points of an algebraic curve on different extensions of a finite field. An analog of the Riemann hypothesis has been stated by Weil [72] and proven by Deligne (see [15] and [16]), and some other results about the repartition of the zeros have been obtained by Katz and Sarnak [23].

More generally, there are many links between probability theory and number theory, which have been much developed in the last years. For example, by working on the asymptotic behavior of the characteristic polynomial of random matrices on one hand, and by studying some problems in arithmetic on the other hand (number of prime factors of integers, zeta functions of function fields), Barbour, Jacod, Kowalski and Nikeghbali (see [4], [21], [29], [30]) have introduced a new notion of convergence of random variables, called *mod-* convergence*. A particular case of this notion can be stated as follows: a family of real random variables $(X_n)_{n\geq 1}$ converges in the *mod-gaussian* sense, if and only if there exists a sequence $(Y_n)_{n\geq 1}$ of gaussian variables such that for all $\lambda \in \mathbb{R}$,

$$\frac{\mathbb{E}[e^{i\lambda X_n}]}{\mathbb{E}[e^{i\lambda Y_n}]} \underset{n\to\infty}{\longrightarrow} \Phi(\lambda),$$

where Φ is a continuous function. This convergence can be considered to be a natural extension of the convergence in law, and it has many potential applications in probability and number theory. Since this notion is involved in the asymptotic behavior of the characteristic polynomial of random matrices, it is natural to expect that there are some deep links between mod-* convergence and the infinite-dimensional objects we are studying. These links should also be present in the conjectures one can make on the zeta or on the L-functions.

To conclude, the infinite-dimensional objects associated to random matrix models give a very promising topic of research, from which one can prove a number of new results improving our understanding of most of the classical results in random matrix theory. The area of research we are now exploring has also many links with other parts of mathematics, and even with theoretical physics: we have a good hope to find new connections in the near future of our research. This would give us the possibility to prove some results which were not expected until now.

References

[1] S. Albeverio and X.-Y. Zhou, *On the equality of two polymer measures in three dimensions*, Jour. Math. Sci. Univ. Tokyo **5** (1998), no. 3, 561–596.

[2] G.-W. Anderson, A. Guionnet, and O. Zeitouni, *An Introduction to Random Matrices*, Cambridge University Press, 2010.

[3] J. Azéma, P.-A. Meyer, and M. Yor, *Martingales relatives*, Séminaire de Probabilités, XXVI, Lecture Notes in Math., vol. 1526, Springer, Berlin, 1992, pp. 307–321.

[4] A.-D. Barbour, E. Kowalski, and A. Nikeghbali, *Mod-discrete expansions*, http://arxiv.org/pdf/0912.1886, 2009.

[5] M. Barlow, J. Pitman, and M. Yor, *On Walsh's Brownian motion*, Séminaire de Probabilités, XXIII, Lecture Notes in Math., vol. 1372, Springer, Berlin, 1989, pp. 275–293.

[6] O.-E. Barndorff-Nielsen and A. Shiryaev, *Change of Time and Change of Measure*, Advanced Series on Statistical Science and Applied Probability, vol. 13, Singapore: World Scientific, 2010.

[7] A. Bentata and M. Yor, *From Black-Scholes and Dupire formulae to last passage times of local martingales. Part A: The infinite time horizon*, 2008.

[8] K. Bichteler, *Stochastic integration and stochastic differential equations*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2002.

[9] E. Bolthausen, *On the construction of the three-dimensional polymer measure*, Probab. Theory Related Fields **97** (1993), no. 1-2, 81–101.

[10] A. Borodin and G. Olshanski, *Infinite random matrices and ergodic measures*, Comm. Math. Phys. **223** (2001), no. 1, 87–123.

[11] P. Bourgade, J. Najnudel, and A. Nikeghbali, *A unitary extension of virtual permutations*, http://arxiv.org/pdf/1102.2633v1, 2011.

[12] P. Bourgade, A. Nikeghbali, and A. Rouault, *Ewens measures on compact groups and hypergeometric kernels*, Séminaire de Probabilités, XLIII, Lecture Notes in Math., vol. 2006, Springer, 2011, pp. 351–377.

[13] P. Cheridito, A. Nikeghbali, and E. Platen, *Processes of the class sigma, last zero and drawdown processes*, http://arxiv.org/pdf/0910.5493, 2009.

[14] A. Connes, *Trace formula in noncommutative geometry and the zeros of the Riemann zeta function*, Selecta Mathematica, New Series **5** (1999), no. 1, 29–106.

[15] P. Deligne, *La conjecture de Weil, I*, Pub. Math. I.H.E.S. **43** (1974), 273–307.

[16] ______, *La conjecture de Weil, II*, Pub. Math. I.H.E.S. **52** (1980), 137–252.

[17] S.-F. Edwards, *The statistical mechanics of polymers with excluded volume*, Proc. Phys. Soc. **85** (1965), 613–624.

[18] P.-J. Forrester and N.-S. Witte, *Application of the τ-function theory of Painlevé equations to random matrices: P_VI, the JUE, CyUE, cJUE and scaled limits*, Nagoya Math. J. **174** (2002), 29–114.

[19] Y. Hu and M. Yor, *Asymptotic studies of Brownian functionals*, Random walks, Bolyai Soc. Math. Stud., vol. 9, János Bolyai Math. Soc., Budapest, 1999, pp. 187–217.

[20] L.-K. Hua, *Harmonic analysis of functions of several complex variables in the classical domains*, Transl. Math. Monographs, vol. 6, American Math. Society, 1963.

[21] J. Jacod, E. Kowalski, and A. Nikeghbali, *Mod-Gaussian convergence: new limit theorems in probability and number theory*, http://arxiv.org/pdf/0807.4739, 2008.

[22] M. Jimbo, T. Miwa, Y. Môri, and M. Sato, *Density matrix of an impenetrable Bose gas and the fifth Painlevé transcendent*, Physica **1D** (1980), 80–158.

[23] N.-M. Katz and P. Sarnak, *Zeroes of zeta functions and symmetry*, Bull. Amer. Math. Soc. **36** (1999), 1–26.

[24] S.-V. Kerov, G.-I. Olshanski, and A.-M. Vershik, *Harmonic analysis on the infinite symmetric group*, Comptes Rend. Acad. Sci. Paris **316** (1993), 773–778.

[25] J.-F.-C. Kingman, *Random discrete distribution*, J. Roy. Stat. Soc. B **37** (1975), 1–22.

[26] ______, *Random partitions in population genetics*, Proc. R. Soc. Lond. (A) **361** (1978), 1–20.

[27] ______, *The representation of partition structures*, J. London Math. Soc. (2) **18** (1978), 374–380.

[28] F.-B. Knight, *Essentials of Brownian Motion and Diffusion*, Math. Surveys and Monographs, vol. 18, American Math. Society, 1981.

[29] E. Kowalski and A. Nikeghbali, *Mod-Gaussian convergence and the value distribution of $\zeta(1/2+it)$ and related quantities*, http://arxiv.org/pdf/0912.3237, 2009.

[30] ______, *Mod-Poisson convergence in probability and number theory*, http://arxiv.org/pdf/0905.0318, 2009.

[31] J. Lörinczi, F. Hiroshima, and V. Betz, *Feynman-Kac-Type Theorems and Gibbs Measures on Path Space*, De Gruyter Studies in Mathematics, vol. 34, de Gruyter, 2011.

[32] D. Madan, B. Roynette, and M. Yor, *From Black-Scholes formula, to local times and last passage times for certain submartingales*, Prépublication IECN 2008/14.

[33] M.-L. Mehta, *Random matrices*, Pure and Applied Mathematics Series, vol. 142, Elsevier Academic Press, 2004, Third edition.

[34] H.-L. Montgomery, *The pair correlation of zeros of the zeta function*, Proc. of Symposia in Pure Math. **24** (1973), 181–193.

[35] J. Najnudel, *Pénalisations de l'araignée brownienne (Penalizations of Walsh Brownian motion)*, Annales de l'Institut Fourier **57** (2007), no. 4, 1063–1093.

[36] ______, *Temps locaux et pénalisations browniennes*, Ph.D. thesis, Université Paris VI, 2007.

[37] ______, *Penalisations of the Brownian motion with a functional of its local times*, Stochastic Process. Appl. **118** (2008), 1407–1433.

[38] ______, *Construction of an Edwards' probability measure on $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$*, Ann. Probab. **38** (2010), no. 6, 2295–2321.

[39] J. Najnudel and A. Nikeghbali, *On penalisation results related with a remarkable class of submartingales*, http://arxiv.org/pdf/0911.4365, 2009.

[40] ———, *On some properties of a universal sigma-finite measure associated with a remarkable class of submartingales*, http://arxiv.org/pdf/0911.2571, to appear in Pub. of the Res. Instit. of Math. Sci. (Kyoto University), 2009.

[41] ———, *On some universal σ-finite measures and some extensions of Doob's optional stopping theorem*, http://arxiv.org/pdf/0906.1782, 2009.

[42] ———, *The distribution of eigenvalues of randomized permutation matrices*, http://arxiv.org/pdf/1005.0402, 2010.

[43] ———, *A new construction of the σ-finite measures associated with submartingales of class* (Σ), Comptes Rend. Acad. Sci. Paris **348** (2010), 311–316.

[44] ———, *On a flow of operators associated to virtual permutations*, http://arxiv.org/pdf/1008.4972, 2010.

[45] ———, *A remarkable σ-finite measure associated with last passage times and penalisation problems*, Contemporary Quantitative Finance, Essays in Honour of Eckhard Platen, Springer, 2010, pp. 77–98.

[46] ———, *A new kind of augmentation of filtrations*, ESAIM: Probability and Statistics **15** (2011), S39–S57.

[47] J. Najnudel, A. Nikeghbali, and F. Rubin, *Scaled limit and rate of convergence for the largest eigenvalue from the generalized Cauchy random matrix ensemble*, Jour. of Stat. Phys. **137** (2009), no. 2, 373–406.

[48] J. Najnudel, B. Roynette, and M. Yor, *A Global View of Brownian Penalisations*, MSJ Memoirs, vol. 19, Mathematical Society of Japan, Tokyo, 2009.

[49] Y.-A. Neretin, *Hua type integrals over unitary groups and over projective limits of unitary groups*, Duke Math. Jour. **114** (2002), 239–266.

[50] A. Nikeghbali, *A class of remarkable submartingales*, Stochastic Process. Appl. **116** (2006), no. 6, 917–938.

[51] G. Olshanski and A. Vershik, *Ergodic unitarily invariant measures on the space of infinite Hermitian matrices*, Amer. Math. Soc. Trans. **175** (1996), 137–175.

[52] K.-R. Parthasarathy, *Probability measures on metric spaces*, New York: Academic Press, 1967.

[53] D. Pickrell, *Measures on infinite-dimensional Grassmann manifolds*, Jour. Func. Anal. **70** (1987), no. 2, 323–356.

[54] ———, *Mackey analysis of infinite classical motion groups*, Pacific J. Math. **150** (1991), 139–166.

[55] J. Pitman, *Combinatorial Stochastic Processes*, Lecture Notes in Math., vol. 1875, Springer-Verlag, Berlin, 2006.

[56] J. Ramirez, B. Valkó, and B. Virág, *Beta ensembles, stochastic Airy spectrum, and a diffusion*, http://arxiv.org/pdf/math/0607331, 2006.

[57] J. Rosen, *A local time approach to the self-intersection of Brownian paths in space*, Comm. Math. Phys. **88** (1983), no. 3, 327–338.

[58] ———, *Tanaka's formula and renormalization for intersections of planar Brownian motion*, Ann. Probab. **14** (1986), no. 4, 1245–1251.

[59] B. Roynette, P. Vallois, and M. Yor, *Limiting laws for long Brownian bridges perturbed by their one-sided maximum. III*, Period. Math. Hungar. **50** (2005), no. 1-2, 247–280.

[60] ———, *Limiting laws associated with Brownian motion perturbed by its maximum, minimum and local time. II*, Studia Sci. Math. Hungar. **43** (2006), no. 3, 295–360.

[61] ———, *Limiting laws associated with Brownian motion perturbed by normalized exponential weights. I*, Studia Sci. Math. Hungar. **43** (2006), no. 2, 171–246.

[62] ———, *Some penalisations of the Wiener measure*, Japan Jour. of Math. **1** (2006), no. 1, 263–290.

[63] ———, *Penalisations of multidimensional Brownian motion, VI*, ESAIM: Probability and Statistics **13** (2009), 152–180.

[64] B. Roynette and M. Yor, *Penalising Brownian Paths*, Lecture Notes in Math., vol. 1969, Springer, 2009.

[65] P. Salminen, P. Vallois, and M. Yor, *On the excursion theory for linear diffusions*, Japan Jour. of Math. **2** (2007), 91–127.

[66] N.-V. Tsilevich, *Distribution of cycle lengths of infinite permutations*, Jour. of Math. Sciences **87** (1997), no. 6, 4072–4081.

[67] ———, *Stationary measures on the space of virtual permutations for an action of the infinite symmetric group*, 1998.

[68] R. van der Hofstad, F. den Hollander, and W. König, *Central limit theorem for the Edwards model*, Ann. Probab. **25** (1997), no. 2, 573–597.

[69] ———, *Large deviations for the one-dimensional Edwards model*, Ann. Probab. **31** (2003), 2003–2039.

[70] S.-R.-S. Varadhan, *Appendix to: Euclidian quantum field theory, by K. Symanzik*, Local Quantum Theory, Jost R. (Ed.), Academic Press, New York, 1969.

[71] J.-B. Walsh, *A diffusion with a discontinuous local time*, Temps locaux, Astérisque **52-53** (1978), 37–45.

[72] A. Weil, *Number of solutions of equations in a finite field*, Bull. Amer. Math. Soc. **55** (1952), 497–508.

[73] M.-J. Westwater, *On Edwards' model for long polymer chains*, Comm. Math. Phys. **72** (1980), no. 2, 131–174.

[74] ———, *On Edwards' model for polymer chains. III. Borel summability*, Comm. Math. Phys. **84** (1982), no. 4, 459–470.

[75] ———, *On Edwards' model for polymer chains*, Trends and developments in the eighties, World Sci. Publishing, Singapore, 1985, pp. 384–404.

[76] E.-P. Wigner, *On the statistical distribution of the widths and spacing of nuclear resonance levels*, Proc. Cambridge Phil. Soc. **47** (1951), 790–798.

[77] J. Wishart, *The generalized product moment distribution in samples from a normal multivariate population*, Biometrika A **20** (1928), 32–43.

[78] N.-S. Witte and P.-J. Forrester, *Gap probabilities in the finite and scaled Cauchy random matrix ensembles*, Nonlinearity **13** (2000), 1965–1986.

[79] M. Yor, *Les inégalités de sous-martingales, comme conséquences de la relation de domination*, Stochastics **3** (1979), no. 1, 1–15.

[80] X.-Y. Zhou, *The intersection local time of Westwater process*, Probab. Theory Related Fields **91** (1992), 375–397.