UNIVERSITÉ DE GENÈVE
Section de Physique

FACULTÉ DES SCIENCES
Professeur J.-P. ECKMANN

Comportement Asymptotique d'Équations à Dérivées Partielles Stochastiques

THÈSE

présentée à la Faculté des Sciences
de l'Université de Genève
pour obtenir le grade de Docteur ès Sciences,
mention physique

par

Martin HAIRER

Autriche

Thèse No 3XXX

GENÈVE
Atelier de reproduction de la Section de Physique
2001

Remerciements

Lorsqu'arrive le moment de tirer un bilan de trois années passées à travailler sur une thèse, il convient de s'interroger sur les personnes qui nous ont aidées directement ou indirecement dans ce parcours. Une telle démarche comporte toujours le risque d'un oubli involontaire et je ne ferai certainement pas exception à cette règle. Je demande donc en tout premier lieu aux personnes que j'aurai oublié de citer de bien vouloir me pardonner et de mettre un tel oubli sur le compte des erreurs humaines.

La personne qui a sans conteste joué le rôle majeur dans l'élaboration de ce travail est mon directeur de thèse, Jean-Pierre Eckmann. Il a su rester disponible à tout moment et a toujours pleinement assumé les diverses fonctions de son rôle avec compétence et sérieux, mais non sans humour. Son approche très directe de la science et son aptitude à saisir l'essentiel d'un problème auront, je l'espère, une influence durable sur mon travail.

Parmi toutes les autres personnes qui ont accepté de partager avec moi leurs connaissances et leur vision de la science, je voudrais citer Mathieu Baillif, Dirk Blömker, Stella Brassesco, Sandra Cerrai, David Cimasoni, Pierre Collet, Thierry Gallay, Tarik Garidi, Ernst Hairer, André Henriques, Markus Kunze, Olivier Lévêque, Xue-Mei Li, Daniel Matthey, Jonathan Mattingly, Marius Mantoiu, Nicolas Musolino, Claude-Alain Pillet, Luc Rey-Bellet, Marco Romito, Jacques Rougemont, Alain Schenkel, Armen Shirikyan, André Stefanov, Larry Thomas, Guillaume van Baalen, Gerhard Wanner, Peter Wittwer, Lai-Sang Young et Emmanuel Zabey.

Un travail de thèse ne se résume jamais à la seule démarche scientifique. C'est pourquoi je voudrais également remercier ici ma famille ainsi que mes amis. J'espère qu'ils se reconnaîtront tous sans que je doive les citer individuellement par peur d'une omission inévitable.

Je remercie aussi Donald Knuth et Leslie Lamport pour avoir développé TeX et LaTeX respectivement. Je ne sais pas où j'en serais dans la rédaction de ce travail sans ces outils aussi indispensables que robustes.

Enfin, je voudrais exprimer ma gratitude aux membres du jury, Gérard Ben Arous, David Elworthy et Peter Wittwer d'avoir accepté d'endosser cette responsabilité.

Table des Matières

I. Introduction

Ce mémoire est consacré à l'étude asymptotique (à grand temps) des solutions d'équations à dérivées partielles paraboliques perturbées par une force aléatoire. Le problème principal que nous abordons est la démonstration de l'existence et surtout de l'unicité d'un état stationnaire pour certaines classes d'équations. Le prototype des problèmes que nous considérons est l'équation de Ginzburg-Landau, donnée par

$$\partial_t u = \partial_x^2 u + u - u^3 , \qquad u(x,0) = u_0(x) , \tag{0.1}$$

où $u(x,t)$ est une fonction réelle et périodique de période $2L$ en x. On peut considérer (0.1) comme une équation d'évolution dans un espace de Hilbert abstrait $\mathcal{H}$, par exemple $\mathcal{H} = \mathrm{L}^2([-L,L],\mathbf{R})$. Il découle alors de résultats bien connus [Lun95] que les solutions de (0.1) définissent un semiflot $\{\varphi_t\}_{t\ge 0}$ sur $\mathcal{H}$, via la formule $u(x,t) = \varphi_t(u_0)(x)$.

Des équations du type (0.1) (comme par exemple aussi l'équation de Swift-Hohenberg qui peut être traîtée de la même manière) servent à décrire une multitude de problèmes physiques. Quelques exemples sont donnés par la croissance d'interfaces, les flots de Couette-Taylor ou encore l'évolution d'un échantillon dans lequel coexistent deux phases stables comme par exemple un système ferromagnétique. Dans un monde idéalisé sans bruit, l'équation déterministe (0.1) donnerait une description correcte des phénomènes observés. Une formulation plus proche de la nature est obtenue en ajoutant un terme de bruit. Il existe plusieurs causes physiques donnant lieu à un tel bruit.

- Une équation du type (0.1) est généralement vue comme limite hydrodynamique d'une dynamique microscopique sous-jacente. Tout système réel est fini et comporte donc des fluctuations dans les variables macroscopiques dues à cette dynamique microscopique. Ainsi, il est possible de dériver la version stochastique de (0.1) comme limite macroscopique de la dynamique de Glauber (voir par exemple l'article de revue [GLP99]).
- L'interaction d'un système avec un environnement aléatoire donne également lieu à des fluctuations dans les observables macroscopiques.

D'un point de vue aussi bien mathématique que physique, il est également intéressant d'étudier quelles caractéristiques de (0.1) sont préservées sous des perturbations stochastiques et lesquelles ne le sont pas.

Avant de discuter de l'effet d'un terme aléatoire sur cette équation, cherchons à comprendre sa dynamique déterministe. Pour une exposition détaillée de cette dynamique, nous renvoyons le lecteur aux travaux [CP89, CP90, ER98, Rou99]. Calculons d'abord les points fixes de (0.1). Il s'agit donc de trouver des solutions $2L$-périodiques à l'équation

$$\partial_x^2 u + u - u^3 = 0 . \tag{0.2}$$

Si l'on interprète la variable u comme une position q et la variable x comme un temps, on voit que l'équation (0.2) est équivalente au système Hamiltonien

$$\dot q = \partial_p H(p,q) , \qquad \dot p = -\partial_q H(p,q) , \tag{0.3}$$

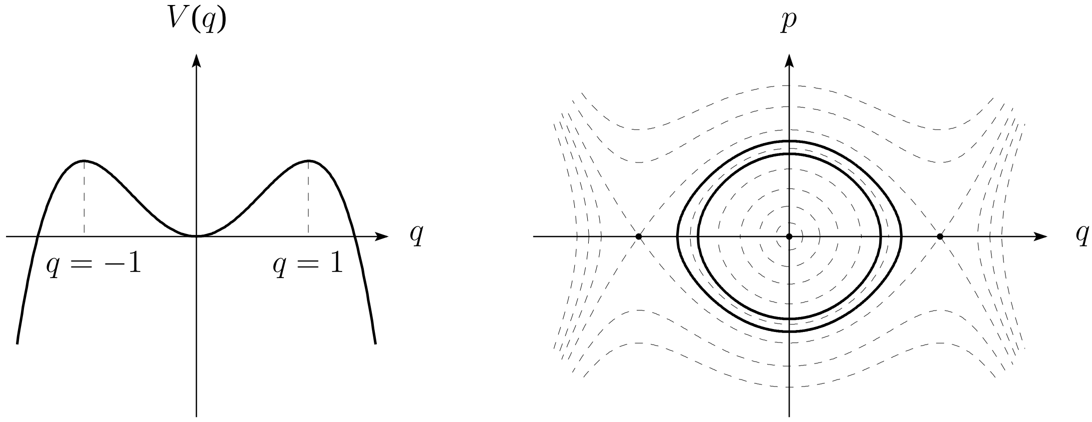

Figure 1: Potentiel et lignes de niveau pour (0.3).

avec

$$H(p,q) = \frac{p^2}{2} + V(q)\,, \qquad V(q) = -\frac{1}{4}(1-q^2)^2\;.$$

La Figure 1 ci-dessus montre la forme du potentiel V, ainsi que les lignes de niveau de H dans l'espace (p,q). Les lignes dessinées en gras correspondent aux solutions périodiques de période $2L$ pour $L = 8$.

En plus des trois solutions triviales $u \equiv 0$ et $u \equiv \pm 1$, il existe dans ce cas deux familles de points fixes paramétrisées par leur phase. Une étude de stabilité permet de voir que les points fixes $u \equiv \pm 1$ sont linéairement stables, alors que les autres points fixes possèdent des variétés instables de dimension finie (dépendant de la valeur de L), voir par exemple [CP90]. L'ensemble de ces points fixes et de leurs variétés instables définit l'*attracteur* $\mathcal{A}$ de (0.1), en d'autres termes $\mathcal{A}$ est un ensemble compact, invariant sous le semiflot φ_t et qui attire chaque ensemble borné suffisamment grand.

Cet attracteur $\mathcal{A}$ caractérise une partie du comportement asymptotique des solutions dans le sens que toutes les solutions finiront par se trouver arbitrairement proches de $\mathcal{A}$ lorsqu'on fait tendre le temps t vers l'infini. (Dans l'exemple considéré, on peut montrer que toutes solutions convergent finalement vers un des points fixes du système.) Néanmoins, même si on peut définir des attracteurs stochastiques [CDF97], la notion d'attracteur (ou de point fixe) n'est pas très bien adaptée à l'étude de systèmes perturbés par du bruit, surtout si l'on s'intéresse à leurs propriétés statistiques, plutôt qu'au comportement d'une trajectoire particulière.

En effet, étant donnée une observable G du système (c'est-à-dire une fonction mesurable et bornée $G : \mathcal{H} \to \mathbf{R}$), on s'intéresse souvent à la convergence des *moyennes empiriques* données par

$$\langle G \rangle = \lim_{T\to\infty} \frac{1}{T} \int_0^T (G \circ \varphi_t)(u_0)\, dt\;. \tag{0.4}$$

(On prendra l'espérance de l'expression de droite dans un contexte probabiliste.) Une telle limite n'existe pas forcément et, même si elle existe, elle peut bien sûr dépendre de la condition initiale u_0. Dans notre cas, par exemple, on aura pour des observables continues $\langle G \rangle = G(u_f)$, où u_f est le point fixe vers lequel la solution converge. Dans certains cas néanmoins, le système perd la mémoire de sa condition initiale sous l'influence d'un bruit extérieur ou d'une chaoticité intrinsèque. Dans ce cas, il arrive que la limite (0.4) existe et soit la même pour “la plupart” des conditions initiales u_0. On peut alors trouver une mesure μ sur $\mathcal{H}$ telle que, pour la plupart des conditions initiales,

$$\langle G \rangle = \int_{\mathcal{H}} G(u)\,\mu(du)\;. \tag{0.5}$$

Si une telle mesure μ existe, elle sera une mesure invariante pour (0.1).

De manière générale, une *mesure invariante* pour (0.1) est une mesure de Borel μ sur $\mathcal{H}$ qui reste inchangée lorsqu’on la transporte avec le semiflot φ_t. En d’autres termes, une mesure μ est invariante lorsque

$$\mu(A) = \mu(\varphi_t^{-1}(A))\;,$$

pour tous les temps $t \geq 0$ et pour tous les ensemblesBoréliens $A \subset \mathcal{H}$. Il ressort de la définition des moyennes empiriques que s’il existe une mesure μ satisfaisant (0.5), alors μ est une mesure invariante. En effet, notant par χ_A la fonction caractéristique d’un ensemble A, on a

$$\begin{aligned}\mu(\varphi_t^{-1}(A)) &= \int_{\mathcal{H}} \chi_{\varphi_t^{-1}(A)}\mu(du) = \langle \chi_{\varphi_t^{-1}(A)}\rangle = \lim_{T\to\infty}\frac{1}{T}\int_0^T \chi_{\varphi_t^{-1}(A)}(\varphi_s(u_0))\,ds \\ &= \lim_{T\to\infty}\frac{1}{T}\int_0^T \chi_A(\varphi_{t+s}(u_0))\,ds = \lim_{T\to\infty}\frac{1}{T}\int_t^{T+t}\chi_A(\varphi_s(u_0))\,ds \\ &= \langle \chi_A\rangle = \mu(A)\;.\end{aligned} \tag{0.6}$$

On voit donc que l’étude des mesures invariantes d’un système est primordiale dans la caractérisation de son comportement asymptotique.

Il ressort de la discussion sur la dynamique du système déterministe qu’il existe beaucoup de mesures invariantes différentes pour (0.1). Il suffit de prendre par exemple une mesure de Dirac concentrée sur un des points fixes de l’équation. On peut se demander si, parmi toutes ces mesures invariantes, il en existe une qui soit plus “naturelle” que les autres. Une manière de caractériser une mesure invariante naturelle est de demander qu’elle soit stable sous l’addition d’une perturbation stochastique au système. Dans le cas de certains systèmes chaotiques, les mesures SRB (Sinai, Ruelle, Bowen) possèdent justement cette propriété [Col98]. Il est donc intéressant de se demander combien de bruit il faut ajouter à un système pour qu’il ne possède plus qu’une seule mesure invariante.

L’avantage de considérer une équation “bruitée” est que, dans de nombreux cas de figure, son comportement asymptotique est beaucoup plus simple à décrire. En effet, nous verrons dans les chapitres suivants que, même sous l’addition de “peu” de bruit, les solutions de la version stochastique de (0.1) tendent vers une unique mesure invariante. Une question intéressante et qui reste ouverte est de savoir s’il existe un moyen de faire tendre le bruit vers 0 qui permette de prouver que la suite de mesures invariantes ainsi obtenue possède une limite et pas seulement des points d’accumulation. Cette limite serait alors un candidat naturel au titre de “mesure SRB” pour un tel système.

1 Présentation du Modèle et Formulation du Problème

Dans cette section, nous allons formuler plus précisément de quelle manière un terme stochastique est ajouté à (0.1). Nous considérons l'équation donnée formellement par

$$\partial_t u = \partial_x^2 u + u - u^3 + \sum_{i=1}^{\infty} q_i e_i \partial_t w_i \;, \qquad u(x,0) = u_0(x) \;. \tag{1.1}$$

Dans cette équation, les q_i sont des nombres positifs ou nuls uniformément bornés, $\{e_i\}_{i=1}^{\infty}$ est une base orthonormée de $\mathcal{H}$ qui diagonalise l'opérateur linéaire ∂_x^2 et les w_i sont des mouvements Browniens indépendants, de manière à ce que l'expression $\partial_t w_i$ dénote un bruit blanc.

Une équation du type (1.1) est habituellement écrite sous la forme abstraite

$$du = Au\,dt + F(u)\,dt + Q\,dW(t) \;, \qquad u(0) = u_0 \;. \tag{1.2}$$

Ici, A dénote l'opérateur linéaire ∂_x^2 de domaine $\mathcal{D}(A)$, F est l'opérateur nonlinéaire $u \mapsto u - u^3$ de domaine $\mathcal{D}(F)$, Q est l'opérateur linéaire borné donné par $Qe_i = q_i e_i$ et W dénote un processus de Wiener cylindrique sur $\mathcal{H}$, formellement donné par $W(t) = \sum_{i=1}^{\infty} e_i\, w_i(t)$. Nous désignons par $(\Omega, \mathscr{F}, \mathbf{P})$ l'espace de probabilité sous-jacent à W.

Nous appelons solution de (1.2) un processus stochastique $u(t)$ à valeurs dans $\mathcal{H}$ tel que $u(t) \in \mathcal{D}(F)$ pour $t > 0$ et

$$u(t) = e^{At} u_0 + \int_0^t e^{A(t-s)} F(u(s))\,ds + \int_0^t e^{A(t-s)} Q\,dW(s) \;. \tag{1.3}$$

Pour une définition rigoureuse de l'intégrale stochastique apparaissant dans (1.3), voir par exemple [DPZ92b]. Nous dirons que la solution de (1.3) définit un flot stochastique si les applications

$$\begin{aligned} \varphi_t(\,\cdot\,,\omega) : \mathcal{H} &\to \mathcal{H} \\ u_0 &\mapsto u(t,\omega) \end{aligned}$$

sont continues pour $\mathbf{P}$-presque tout $\omega \in \Omega$. Le théorème d'existence suivant peut être dérivé facilement des résultats présentés dans [DPZ96]:

Théorème 1 *L'équation (1.1) possède une unique solution qui définit un flot stochastique φ_t.*

Ce théorème nous permet de définir, à l'aide du flot stochastique φ_t, une évolution sur les observables G, ainsi qu'une évolution duale sur les mesures μ par les formules:

$$(\mathcal{P}_t G)(u) = \mathbf{E}(G \circ \varphi_t)(u) \;, \qquad u \in \mathcal{H} \;, \tag{1.4a}$$

$$(\mathcal{P}_t^* \mu)(B) = \mathbf{E}(\mu \circ \varphi_t^{-1})(B) \;, \qquad B \subset \mathcal{H} \;. \tag{1.4b}$$

Avec ces définitions, une mesure invariante pour le problème (1.2) est simplement un point fixe de $\mathcal{P}_t^*$. Le problème que nous abordons dans ce travail est l'existence et l'unicité d'une telle mesure invariante.

Dans le cadre du problème que nous venons de décrire, l'existence d'une mesure invariante est relativement aisée à obtenir. En effet, par le théorème d'existence de Krylov-Bogolyubov

[DPZ96], il suffit de montrer que, pour une mesure μ donnée, la suite des moyennes empiriques données par

$$\frac{1}{T}\int_0^T \mathcal{P}_t^* \mu \, dt \tag{1.5}$$

possède au moins un point d'accumulation lorsque $T \to \infty$. Ceci découle des propriétés de régularisation de l'opérateur e^{At}. Ainsi, on peut montrer que les solutions de (1.1) appartiennent presque sûrement à un espace de Sobolev $\mathcal{W}$ tel que l'inclusion $\mathcal{W} \subset \mathcal{H}$ est compacte. De plus, il est possible d'obtenir des estimations uniformes en temps sur les moments de la norme de la solution dans $\mathcal{W}$. Par conséquent, la suite des moyennes empiriques est tendue et possède donc au moins un point d'accumulation (dans la topologie faible-*) [Bil68].

Néanmoins, nous verrons au Chapitre II une situation dans laquelle la preuve de l'existence d'une mesure invariante n'est pas aussi simple. C'est le cas notamment lorsque la variable spatiale x n'est pas restreinte à un intervalle mais peut prendre ses valeurs sur toute la droite réelle. En effet, l'opérateur e^{At} n'est alors plus compact, ce qui complique l'argumentation ci-dessus. Dans la suite de cette introduction, nous discuterons principalement le problème de l'unicité de la mesure invariante pour des équations du type (1.2).

2 Unicité de la Mesure Invariante – Techniques de Démonstration

Avant d'énoncer quelques résultats obtenus dans ce travail, nous donnons un bref aperçu historique des méthodes connues et de leur champ d'application. Cet historique ne prétend pas être exhaustif, mais il nous semble qu'il couvre les méthodes dont l'impact a été le plus important dans la compréhension du problème. Pour un aperçu des principaux travaux couvrant les deux premières méthodes présentées ci-après, le lecteur se référera à l'excellent travail de revue de Maslovski et Seidler [MS99]. Un exposé détaillé des résultats obtenus, ainsi que de leurs démonstrations peut être trouvé dans la monographie [DPZ96].

2.1 La méthode de la dissipativité

Cette méthode s'applique à des situations où le problème linéaire

$$\dot{u} = Au + F(u) \tag{2.1}$$

associé à (1.2) possède de bonnes propriétés de dissipation. Une condition typique consiste à imposer l'existence d'une constante $\alpha > 0$ telle que, pour toutes les fonctions u et v appartenant aux domaines de A et de F, on ait l'inégalité

$$\langle Au - Av, u - v\rangle + \langle F(u) - F(v), u - v\rangle \le \alpha \|u - v\|^2 \; . \tag{2.2}$$

Cette condition implique que le système déterministe possède un seul point fixe et que toutes les solutions s'en approchent avec un taux exponentiel. On peut alors construire une suite u_T de variables aléatoires obtenues en évaluant au temps $t = 0$ la solution du système (1.2), avec pour condition initiale $u(-T) = u_0$. Cette suite de variables aléatoires converge dans $\mathrm{L}^2(\Omega, \mathbf{P})$ vers une vaiable aléatoire u_∞, dont on peut montrer qu'elle ne dépend pas de u_0. Cette convergence est exponentielle avec un taux α, comme on peut le voir intuitivement de la condition (2.2). La loi de u_∞ est la mesure invariante recherchée.

Remarquons encore que la méthode de la dissipativité ne dépend que très peu de la nature du bruit. Elle n'utilise d'ailleurs pas du tout le bruit pour obtenir l'unicité de la mesure invariante.

2.2 La méthode du recouvrement

Contrairement à la méthode de la dissipativité, la méthode du recouvrement utilise le bruit de manière cruciale. Dénotons par $P_t(u,\cdot\,)$ la famille de mesures sur $\mathcal{H}$ donnant les probabilités de transition à temps t pour les solutions de (1.2), c'est-à-dire que

$$P_t(u,B) = \mathbf{P}(\varphi_t(u) \in B)\;.$$

La méthode du recouvrement est basée principalement sur la constatation suivante:

Proposition 2 *Supposons qu'il existe un temps $T > 0$ et une mesure positive (non-nulle) δ telle que $P_t(u,B) \ge \delta$ pour tous les $u \in \mathcal{H}$. Alors, l'opérateur $\mathcal{P}_T^*$ défini dans (1.4b) est une contraction sur les mesures de probabilité dans la norme de variation totale.*

La démonstration de cet énoncé est élémentaire. Prenons deux mesures de probabilité μ et ν mutuellement sigulières. Nous avons alors $\mathcal{P}_T^*\mu \ge \delta$ et $\mathcal{P}_T^*\nu \ge \delta$. Il y a donc un recouvrement de masse $\|\delta\|_{\mathrm{TV}}$ entre ces deux mesures. Ce recouvrement s'annule lorsque l'on prend la différence et l'on a ainsi

$$\|\mathcal{P}_T^*\mu - \mathcal{P}_T^*\nu\|_{\mathrm{TV}} \le (1 - \|\delta\|_{\mathrm{TV}})\|\mu-\nu\|_{\mathrm{TV}}\;.$$

Cette formule est aisément généralisable au cas de deux mesures quelconques. Ceci implique qu'il existe une et une seule mesure de probabilité invariante pour (1.2) et que la convergence des probabilités de transition vers cette mesure invariante est exponentielle dans la norme de la variation totale. En général, un système donné ne montrera pas un recouvrement aussi uniforme que ce que nous supposons dans la Proposition 2, mais l'idée sera la même.

La technique habituelle afin d'obtenir une estimation sur le recouvrement est de vérifier les trois propriétés suivantes pour les solutions de (1.2).

(*a*) La dynamique est fortement Feller, c'est-à-dire que l'opérateur $\mathcal{P}_t$ défini en (1.4a) envoie les fonctions bornées mesurables sur des fonctions bornées continues.

(*b*) La dynamique est topologiquement irréductible, c'est-à-dire que les probabilités de transition satisfont $P_t(u,B) > 0$ pour tout $u \in \mathcal{H}$ et pour tout ouvert $B \subset \mathcal{H}$.

(*c*) Il existe un compact K qui attire les solutions, dans le sens où les temps de retour vers K et les temps d'entrée dans K sont bien contrôlés.

Les deux premières propriétés suffisent déjà, par le théorème de Doob [Doo48], à garantir l'unicité de la mesure invariante (si elle existe). La difficulté technique principale consiste à démontrer que la dynamique est fortement Feller. Cette difficulté est résolue en dimension finie par le critère de Hörmander [Hör67, Hör85, Mal78, Str86, Nor86]. Ce critère dit que, pour que la dynamique générée par les solutions de l'équation de Stratanovitch

$$dx = f_0(x)\,dt + \sum_{i=1}^m f_i(x)\circ d\omega_i\;, \qquad x \in \mathbf{R}^n\;, \tag{2.3}$$

soit fortement Feller, il est suffisant que les champs de vecteurs

$$f_i\;,\quad [f_i,f_j]\;,\quad [[f_i,f_j],f_k]\;,\ldots \qquad i=1,\ldots,m\;,\quad j,k=0,\ldots,m\;,$$

engendrent tout $\mathbf{R}^n$ en chaque point. Dans cette expression $[\cdot,\cdot]$ désigne le crochet de Lie entre deux champs de vecteurs (qui est égal au commutateur, si on interprète les champs de vecteurs comme des opérateurs différentiels). Malheureusement, il n'existe pas encore de critère général équivalent en dimension infinie. Un outil très utile en dimension infinie est la formule de Bismut-Elworthy [Bis84, EL94]. Celle-ci permet d'exprimer la dérivée de $\mathcal{P}_t G$ en terme de G et du flot φ_t de la manière suivante:

$$(D\mathcal{P}_t G)(u)h = \frac{1}{t}\mathbf{E}\Big((G\circ\varphi_t)(u)\int_0^t \Big\langle Q^{-1}(D\varphi_s)(u)h, dW(s)\Big\rangle\Big) . \tag{2.4}$$

Néanmoins, afin de pouvoir utiliser la formule (2.4), il faut que l'image de l'opérateur Q contienne l'image de $D\varphi_s$. Cette formule n'est donc applicable telle quelle pratiquement que dans des situations où l'image de Q est dense dans $\mathcal{H}$.

2.3 La méthode du couplage

De manière générale, si l'on se donne une mesure de probabilité μ sur un espace M et une mesure de probabilité ν sur un espace N, un *couplage* pour la paire (μ,ν) est la donnée d'une mesure de probabilité P sur $M \times N$, telle que les marginales de P (donc les projections de P sur M et N) sont précisément μ et ν. Ce que l'on appelle un couplage pour (1.2) est donc la donnée d'un processus stochastique $(u(t),v(t))$, tel que u et v pris séparément sont des solutions de (1.2) (avec pour conditions initiales u_0 et v_0). S'il est possible de construire ce couplage de manière à ce que la distance entre u et v tende vers zéro pour de grands temps, l'unicité de la mesure invariante suit (voir Section 3.2 ci-après).

La méthode du couplage et son application aux problèmes d'ergodicité et d'unicité de l'état stationnaire ont une longue histoire. Les premières applications semblent remonter aussi loin que Doeblin [Doe38] dans l'étude de chaînes de Markov ayant un nombre fini d'états. Dans un contexte et une formulation un peu plus proches de ce que nous abordons dans ce mémoire, cette technique a été utilisée par Vaseršteĭn et Dobrušin [Vas69, Dob71] dans l'étude de systèmes de spins. Ce domaine d'application a donné lieu à une multitude de travaux; pour un travail de revue de cette époque, nous renvoyons le lecteur à [Lig77].

A notre connaissance, la méthode du couplage a pour la première fois été appliquée dans le cadre des équations à dérivées partielles stochastiques dans [Mue93]. L'idée était alors de choisir le couplage de manière à ce que, lorsque u s'approche de v, les deux processus s'attirent et finalement se “collent” l'un à l'autre. Une fois qu'ils sont collés dans une région spatiale, ils y restent collés pour les temps futurs. Plus précisément, le couplage est fait de manière à ce que le processus $\varrho = |u - v|$ se comporte comme les solutions de

$$d\varrho = \Delta\varrho\, dt + \varrho^{1/2}\, dW(t) .$$

Il est bien connu que les solutions de cette équations atteignent $\varrho \equiv 0$ en temps fini avec probabilité 1. Cette technique à également été appliquée avec succès à l'étude de de la propagation de fronts pour l'équation de Kolmogorov-Petrovskii-Piscuinov avec bruit [MS95c].

Beaucoup plus récemment, un autre type de construction pour la méthode du couplage est apparu dans une série d'articles sur les propriétés ergodiques de l'équation de Navier-Stokes bidimensionelle (voir [KS01, Mat01, MY01] et les références données à la Section V.6). L'idée qui apparaît dans ces travaux est de décomposer l'espace $\mathcal{H}$ en une partie "stable" $\mathcal{H}_s$ et une partie "instable" $\mathcal{H}_u$ et de supposer que le bruit agit sur la partie instable, c'est-à-dire que l'image de Q est égale à (ou du moins contient) $\mathcal{H}_u$. On supposera de plus que la partie linéaire A laisse invariante cette décomposition. L'équation (1.2) devient alors

$$du_u = A_u u_u \, dt + F_u(u_u, u_s) \, dt + Q \, dW(t) \; , \tag{2.5a}$$

$$du_s = A_s u_s \, dt + F_s(u_u, u_s) \, dt \; . \tag{2.5b}$$

Dans cette décomposition, on choisira la partie stable $\mathcal{H}_s$ de manière à ce que, étant donnée une trajectoire u_u, deux solutions quelconques de l'équation (2.5b) avec des conditions initiales différentes convergent toujours l'une vers l'autre.

D'autre part, comme le bruit agit de façon non-dégénérée sur la partie instable (2.5a), il y a un moyen de construire un couplage $(u(t), v(t))$ avec $t \in [0,1]$, tel que $u_u(1) = v_u(1)$ avec une probabilité non-nulle, quelles que soient les conditions initiales $u(0)$ et $v(0)$. Ces deux constatations permettent alors de construire un couplage $(u(t), v(t))$ (cette fois avec le temps $t \in [0, \infty)$), tel qu'il existe un temps aléatoire τ avec la propriété que $u_u(t) = v_u(t)$ pour tous les temps $t \ge \tau$. Si on peut montrer que τ est presque sûrement fini, l'unicité de la mesure en découle. On peut également obtenir des estimations sur la vitesse de convergence vers cette mesure invariante si l'on sait estimer τ, ainsi que le taux de stabilité de (2.5b).

Dans ce mémoire, nous appliquerons aux chapitres III et IV la méthode du recouvrement à une situation où l'image de Q n'est pas dense dans $\mathcal{H}$, en combinant un argument à la Hörmander avec une version modifiée de la formule de Bismut-Elworthy. Nous généraliserons également au Chapitre V la méthode du couplage à des situations où le bruit agit de façon dégénérée sur la partie instable de l'équation. Une présentation plus détaillée de ces résultats est l'objet de la section suivante.

3 Unicité de la Mesure Invariante – Résultats Obtenus

3.1 Méthode du recouvrement pour des situations dégénérées

Dans cette partie du mémoire, qui fera l'objet des chapitres III et IV, nous considérons l'équation de Ginzburg-Landau stochastique donnée par

$$\partial_t u = \partial_x^2 u + u - u^3 + \sum_{i=1}^{\infty} q_i e_i \partial_t w_i \; , \qquad u(x,0) = u_0(x) \; , \tag{3.1}$$

avec $x \in [-L, L]$ et des conditions aux bords périodiques pour u. Pour des raisons techniques, nous considérons cette équation non pas dans l'espace $\mathrm{L}^2([-L,L])$, mais dans l'espace de Sobolev $\mathcal{H} = \mathcal{W}^{(1,2)}([-L,L])$ des fonctions périodiques dans L^2 et à dérivée dans L^2. Nous dénotons de plus par $\mathsf{P}_t(u, \cdot\,)$ les probabilités de transition sur $\mathcal{H}$ induites par la solution de (3.1).

Nous supposons qu'il existe des constantes c_1, c_2 et k_*, ainsi que des exposants α et β satisfaisant

$$\alpha \ge 4 \qquad \text{et} \qquad \alpha - 1/4 < \beta \le \alpha \; , \tag{3.2}$$

tels que

$$c_1 k^{-\alpha} \le q_k \le c_2 k^{-\beta} \qquad \text{pour } k \ge k_*. \tag{3.3}$$

Remarquons qu'aucune supposition n'est faite sur la taille de k_*, ni sur les valeurs de q_k pour $k < k_*$. Ces valeurs de q_k peuvent par exemple être choisies nulles. Notre résultat est alors le suivant.

Théorème 3 *Supposons que les valeurs q_k satisfassent (3.3). Alors, l'équation (3.1) possède une unique mesure invariante μ_* sur $\mathcal{H}$ et il existe des constantes $C, \gamma > 0$ telles que*

$$\|\mathsf{P}_t(u,\cdot\,) - \mu_*\|_{\mathrm{TV}} \le Ce^{-\gamma t}\;,$$

pour tout $t \ge 0$ et pour tout $u \in \mathcal{H}$.

Voici les idées principales permettant de démontrer ce théorème. La démarche générale consiste à vérifier les propriétés (*a*), (*b*) et (*c*) de la Section 2.2. La difficulté technique principale consiste à montrer que la dynamique est fortement Feller. L'irréductibilité topologique (*b*) peut être montrée de la manière suivante. On se donne une condition initiale $u_i \in \mathcal{H}$, un temps $\tau > 0$ et une "cible" $u_f \in \mathcal{H}$. On peut alors montrer que pour tout $\varepsilon > 0$, il existe une fonction lisse $F\colon [0,\tau] \to \mathcal{H}$, telle que la solution $u(t)$ de l'équation

$$\dot u = \Delta u + u - u^3 + F(t)\;,\qquad u(0) = u_i\;,$$

satisfait $\|u(\tau) - u_f\| \le \varepsilon$. On montre ceci "brutalement" en construisant explicitement la fonction F. On a ainsi construit pour chaque cible une réalisation du bruit qui amène les solutions arbitrairement près de la cible en un temps donné. Il est intuitivement clair que ceci est suffisant pour avoir l'irréductibilité topologique. L'égalité entre le support des probabilités de transition et l'ensemble atteignable du problème de contrôle associé a été montré pour la première fois en dimension finie dans [SV72].

La propriété (*c*) découle immédiatement de la dissipativité très forte de ce système. Le terme $-u^3$ oblige les solutions à rester relativement près de l'origine et le terme Δu les régularise, fournissant la compacité cherchée.

Il reste donc à montrer la propriété (*a*), à savoir que la dynamique est fortement Feller. Nous nous contenterons dans cette introduction à montrer que l'équation (3.1) satisfait formellement la condition de Hörmander. Afin de simplifier les notations, nous nous restreignons au cas de fonctions paires et réelles $u(x) = u(-x) = \bar u(x)$. Dans ce cas, les coefficients de Fourier de u satisfont $u_k = u_{-k} = \bar u_k$. L'équation (3.1) peut alors s'écrire comme

$$du_k = (1-k^2)\,u_k\,dt - \sum_{\ell+m+n=k} u_\ell u_m u_n\,dt + q_k\,d\omega_k\;,\qquad k,\ell,m,n \in \mathbf{Z}\;.$$

Le terme important dans cette expression est celui correspondant à la nonlinéarité cubique. En effet, on voit que ce terme est le seul qui couple les différents modes de Fourier entre eux. Si on interprète ce terme comme un champ de vecteurs (de dimension infinie), on peut l'écrire comme

$$f_0 = \sum_{\ell,m,n} u_\ell u_m u_n\,\partial_{\ell+m+n}\;,$$

où ∂_k désigne l'opérateur de dérivation dans la direction u_k. Les champs de vecteurs décrivant le bruit sont donnés par $g_k = q_k\,\partial_k$. Ces champs peuvent éventuellement être nuls pour $|k| < k_*$.

Un calcul élémentaire donne alors

$$\Big[g_\ell,[g_m,[g_n,f_0]]\Big] \propto q_\ell q_m q_n \sum_{k=\pm\ell\pm m\pm n} \partial_k \;.$$

Comme tout nombre entre $-k_*$ et $+k_*$ peut s'écrire comme une somme de trois nombres, on voit que, du moins au niveau formel, trois commutateurs suffisent pour engendrer tout l'espace.

3.2 Généralisation de la méthode du couplage

Dans cette partie du mémoire, qui fera l'object du Chapitre V, nous décrirons une manière de construire un couplage (u,v) tel que $\|u(t)-v(t)\|$ converge exponentiellement vers 0 lorsque $t\to\infty$. Une telle construction implique alors immédiatement l'unicité de la mesure invariante, ainsi que la convergence exponentielle des probabilités de transition vers cette mesure invariante.

Notre construction diffère de celle décrite à la Section 2.3 par le fait que nous remplaçons la condition que les parties "instables" u_u et v_u se rencontrent en un temps fini par celle, plus faible, qu'elles convergent (exponentiellement) l'une vers l'autre. Cette condition permet d'étendre la méthode du couplage à certaines situations où le bruit agit seulement de manière dégénérée sur la partie instable de l'équation et où les méthodes de couplage décrites précédemment font défaut. Dans cette introduction, nous considérons le modèle donné par le système d'équations suivant:

$$\begin{aligned} du_0 &= (a^2u_0 + u_1 - u_0^3)\,dt + d\omega\;,\\ \dot u_k &= (a^2-k^2)u_k + u_{k-1} + u_{k+1} - u_k^3\;, \qquad k=1,2,\ldots \end{aligned} \tag{3.4}$$

Ici, $a\in\mathbf{R}$ dénote une constante que l'on peut choisir arbitrairement et ω est un mouvement Brownien unidimensionnel. Nous choisissons de considérer l'équation (3.4) dans l'espace de Hilbert $\mathcal{H}=\ell^2$. On peut voir (3.4) comme un modèle simplifé pour des équations à dérivées partielles du type Ginzburg-Landau. Remarquons encore que le bruit ω n'agit que sur le mode 0, alors que tous les modes jusqu'à $k=|a|$ sont linéairement instables. Néanmoins, chaque mode est couplé à ses voisins, ce qui a pour conséquence une transmission du bruit à tous les modes du système. C'est ce couplage que nous exploiterons par la suite pour notre construction. Comme précédemment, nous dénotons par $\mathsf{P}_t(u,\cdot\,)$ les probabilités de transition sur $\mathcal{H}$ induites par la solution de (3.4). Notre résultat est alors le suivant.

Théorème 4 *L'équation (3.4) possède une unique mesure invariante μ_* sur $\mathcal{H}$ et il existe des constantes $C,\gamma>0$ telles que*

$$\|\mathsf{P}_t(u,\cdot\,)-\mu_*\|_{\mathrm{L}} \le Ce^{-\gamma t}\;,$$

pour tout $t\ge 0$ et pour tout $u\in\mathcal{H}$.

La norme $\|\cdot\|_{\mathrm{L}}$ apparaissant dans cette équation est la norme duale à la norme de Lipschitz, c'est-à-dire que

$$\|\mu-\nu\|_{\mathrm{L}} = \sup\Big\{\int_{\mathcal{H}} G(u)\,(\mu-\nu)(du)\;\Big|\;|G(u)|\le 1\;,\;|G(u)-G(v)|\le\|u-v\|\;,\;\forall u,v\in\mathcal{H}\Big\}\;.$$

Cette norme définit une topologie plus faible que la norme de la variation totale apparaissant dans la section précédente. Un exemple qui permet de voir la différence entre ces deux normes est la suite de mesures sur **R** donnée par

$$\mu_n = \delta_{1/n} \;, \qquad \mu_\infty = \delta_0 \;,$$

où δ_x dénote la mesure de Dirac concentrée au point x. On a alors $\|\mu_n - \mu_\infty\|_{\mathrm{L}} = 1/n$, mais $\|\mu_n - \mu_\infty\|_{\mathrm{TV}} = 2$.

Nous donnons maintenant l'idée principale permettant de démontrer le Théorème 4. Notre technique consiste à construire un couplage (u, v) pour (3.4) tel que $\|u - v\| \to 0$ exponentiellement à grand temps. Supposons qu'un tel couplage existe et que l'on ait une estimation du type

$$\mathbf{P}\Big(\|u(t) - v(t)\| > C_1 \varepsilon^{-\gamma_1 t}\Big) \le C_2 e^{-\gamma_2 t} \;. \tag{3.5}$$

Par la définition de la norme de Lipschitz, on voit que l'estimation ci-dessus implique pour les probabilités de transition que

$$\|\mathsf{P}_t(u_0, \cdot\,) - \mathsf{P}_t(v_0, \cdot\,)\|_{\mathrm{L}} \le C_1 \varepsilon^{-\gamma_1 t} + 2 C_2 e^{-\gamma_2 t} \;.$$

Les probabilités de transition pour deux conditions initiales différentes convergent donc l'une vers l'autre. Si on a de plus une certaine uniformité des constantes C_1 et C_2 sur un ensemble qui attire les solutions, on peut montrer que les probabilités de transition $\mathsf{P}_n(u_0, \cdot\,)$ forment une suite de Cauchy et possèdent donc une limite μ_*. Il est simple de vérifier que μ_* est une mesure invariante.

Il reste donc à construire le couplage en question. Soit $(\Omega, \mathscr{F}, \mathsf{W})$ l'espace de probabilité sous-jacent à un processus de Wiener. Si $(\omega, \tilde\omega)$ est une variable aléatoire à valeurs dans $\Omega \times \Omega$ distribuée selon une mesure P telle que les marginales de P sur chacune des deux copies Ω sont égales à W, un couplage pour (3.4) est donné par les solutions de

$$\begin{aligned} du_0 &= (a^2 u_0 + u_1 - u_0^3)\, dt + d\omega \;, \\ \dot u_k &= (a^2 - k^2) u_k + u_{k-1} + u_{k+1} - u_k^3 \;, \\ dv_0 &= (a^2 v_0 + v_1 - v_0^3)\, dt + d\tilde\omega \;, \\ \dot v_k &= (a^2 - k^2) v_k + v_{k-1} + v_{k+1} - v_k^3 \;, \qquad k = 1, 2, \ldots \end{aligned}$$

L'idée principale dans la construction de P est de considérer l'équation

$$\begin{aligned} du_0 &= (a^2 u_0 + u_1 - u_0^3)\, dt + d\omega \;, \\ \dot u_k &= (a^2 - k^2) u_k + u_{k-1} + u_{k+1} - u_k^3 \;, \\ dv_0 &= (a^2 v_0 + v_1 - v_0^3)\, dt + G(u, v)\, dt + d\omega \;, \\ \dot v_k &= (a^2 - k^2) v_k + v_{k-1} + v_{k+1} - v_k^3 \;, \qquad k = 1, 2, \ldots \end{aligned} \tag{3.6}$$

avec pour G une fonction telle que les solutions de (3.6) satisfont une estimation du type (3.5). Cette équation n'est pas encore un couplage pour (3.4), puisque le processus

$$\tilde\omega(t) = \omega(t) + \int_0^t G(u(s), v(s))\, ds$$

n'est pas distribué selon la mesure de Wiener. Il est néanmoins possible de construire un couplage tel qu'il existe un temps aléatoire τ ayant la propriété que le processus (u, v) satisfait l'équation (3.6) pour des temps $t > \tau$. On peut également trouver des estimations sur la taille de τ.

Il ne reste donc plus qu'à construire la fonction G. On considère pour ceci le processus donné par $\varrho = v - u$. On a alors (si on définit $\varrho_{-1} = 0$):

$$\dot{\varrho}_k = (a^2 - k^2)\varrho_k + \varrho_{k+1} + \varrho_{k-1} - \varrho_k(u_k^2 + u_k v_k + v_k^2) + \delta_{k,0}\, G(u, v) \; . \tag{3.7}$$

On voit bien que, grâce au terme en $-k^2$, il existe une valeur $\tilde{k}$ telle que l'équation ci-dessus est stable pour $k > \tilde{k}$. Si nous trouvons un G tel que les modes avec $k \leq \tilde{k}$ tendent vers 0, les autres vont donc suivre automatiquement. La construction que nous utilisons ensuite pour trouver G est similaire à une construction utilisée dans [EPR99b] pour contrôler une chaîne d'oscillateurs couplés par ses extrémités à des réservoirs thermiques. Les équations (3.7) exhibent un couplage par plus proches voisins. Comme ce couplage est non-dégénéré on voit que, si l'on obtient d'une manière ou d'une autre le contrôle sur $\varrho_{\tilde{k}-1}$, on peut à travers ce couplage forcer $\varrho_{\tilde{k}}$ à tendre vers 0. Par le même raisonnement, on peut obtenir le contrôle sur $\varrho_{\tilde{k}-1}$ si on a le contrôle sur $\varrho_{\tilde{k}-2}$, etc. Ce raisonnement peut être poursuivi en descendant les modes jusqu'à arriver au mode 0, sur lequel on a justement un bon contrôle via la fonction G. On peut donc de cette manière faire tendre $\varrho_{\tilde{k}}$ vers 0.

Enfin, il est possible de démontrer, en regardant en détail la construction ainsi obtenue, que non seulement $\varrho_{\tilde{k}}$ tend vers 0, mais également tous les autres modes. Ceci achève la démonstration du Théorème 4.

4 Conclusions et Perspectives

Dans ce travail, nous avons généralisé les méthodes du recouvrement et du couplage, afin de pouvoir les appliquer dans des situations qui n'ont jusqu'alors pas pu être étudiées. Nous avons fait la plupart de nos estimations sur l'exemple concret de l'équation de Ginzburg-Landau stochastique, mais une partie des résultats se généralisent également à d'autres situations (voir les Exemples à la fin du Chapitre V). Il reste en tous cas trois questions ouvertes qui nous semblent intéressantes:

- Quel est le nombre minimal de modes qu'il faut forcer pour que l'équation de Ginzburg-Landau stochastique (réelle ou complexe) sur un intervalle ne possède qu'une seule mesure invariante?
- La suite de mesures invariantes pour (3.1) obtenue en faisant tendre k_* vers l'infini possède-t-elle une limite?
- Est-ce que la mesure invariante pour l'équation de Ginzburg-Landau en domaine infini est unique si une bande de fréquences seulement est forcée par le bruit?

Dans les deux cas, la partie déterministe de l'équation peut être remplacée par n'importe quelle autre équation à dérivées partielles comportant une certaine instabilité (comme par exemple l'équation de Navier-Stokes en deux dimensions).

La suite de ce travail (en anglais) est structurée de la manière suivante. Au Chapitre II, nous considérons l'équation de Ginzburg-Landau stochastique sur la droite réelle et nous montrons que l'on peut obtenir l'existence d'une mesure invariante à l'aide des propriétés analytiques de ses solutions. Au Chapitre III, nous considérons l'équation de Ginzburg-Landau en domaine borné avec un bruit n'agissant que sur les hautes fréquences. Nous montrons que sa dynamique est fortement Feller et topologiquement irréductible, ce qui implique l'unicité de sa mesure invariante. Au Chapitre IV, nous montrons que la convergence vers cette mesure invariante s'effectue de façon exponentielle, en utilisant la méthode du recouvrement. Au Chapitre V enfin, nous utilisons la méthode du couplage pour démontrer l'unicité de la mesure invariante dans des situations très dégénérées où le bruit est de dimension finie et agit seulement indirectement sur les modes déterminants du système.

Les chapitres II et III sont des reproductions des articles [EH01a] et [EH01b] respectivement. Ces deux publications ont été réalisées en collaboration avec J.-P. Eckmann.

II. Invariant Measures for Stochastic PDE's in Unbounded Domains

Abstract

We study stochastically forced semilinear parabolic PDE's of the Ginzburg-Landau type. The class of forcings considered are white noises in time and colored smooth noises in space. Existence of the dynamics in L^∞, as well as existence of an invariant measure are proven. We also show that the solutions are with high probability analytic in a strip around the real axis and give estimates on the width of that strip.

1 Introduction

We consider the stochastic partial differential equation (SPDE) given by

$$\begin{aligned} du_\xi(t) &= \Delta u_\xi(t)\,dt + (1 - |u_\xi(t)|^2)u_\xi(t)\,dt + Q\,dW(t)\;, \\ u_\xi(0) &= \xi\;, \qquad \xi \in \mathrm{L}^\infty(\mathbf{R})\;. \end{aligned} \tag{SGL}$$

In this equation, $dW(t)$ denotes the canonical cylindrical Wiener process on the Hilbert space $\mathrm{L}^2(\mathbf{R}, dx)$, *i.e.* we have the formal expression

$$\mathbf{E}(dW(s,x)\,dW(t,y)) = \delta(s-t)\delta(x-y)\,ds\,dt\;.$$

Think for the moment of $u_\xi(t)$ as a distribution on the real line. We will introduce later the space of functions in which (SGL) makes sense. The symbol Q denotes a bounded operator of the type $Qf = \varphi_1 \star (\varphi_2\, f)$ where $\hat\varphi_1$, the Fourier transform of φ_1, is some positive $\mathcal{C}_0^\infty$ function and φ_2 is some smooth function that decays sufficiently fast at infinity to be square-integrable. In fact, we will assume for convenience that there are constants $c > 0$ and $\beta > 0$ such that

$$|\varphi_2(x)| \le \frac{c}{\langle\!\langle x\rangle\!\rangle^{1/2+\beta}}\;, \qquad \langle\!\langle x\rangle\!\rangle \equiv \sqrt{1+x^2}\;. \tag{1.1}$$

The space in which we show the existence of the solutions is $C_u(\mathbf{R})$, the Banach space of complex-valued uniformly continuous functions. The reason of this choice is that we want to work in a translational invariant space which is big enough to contain the interesting part of the dynamics of the deterministic part of the equation, *i.e.* the three fixed points 0 and ± 1, as well as various kinds of fronts and waves. The meaning of the assumptions on φ_1 and φ_2 is the following.

- The noise does not shake the solution too badly at infinity (in the space variable x). If it did, the solution would not stay in L^∞.
- The noise is smooth in x (it is even analytic), so it will not lead to irregular functions in x-space. This assumption is crucial for our existence theorem concerning the invariant measure.

For convenience, we write (SGL) as

$$\begin{aligned} du_\xi(t) &= (Lu_\xi(t) + F(u_\xi(t)))\,dt + Q\,dW(t)\;, \\ L = \Delta - 1\;, &\quad (F(u))(x) = u(x) + (1 - |u(x)|^2)u(x)\;. \end{aligned} \tag{1.2}$$

This is also to emphasize that our proofs apply in fact to a much larger class of SPDE's of the form (1.2). For example, all our results apply to the stochastically perturbed Swift-Hohenberg equation

$$du_\xi(t) = (1-\Delta)^2 u_\xi(t)\,dt + (1 - |u_\xi(t)|^2)u_\xi(t)\,dt + Q\,dW(t)\;,$$

but one has to be more careful in the computations, since one does not know an explicit formula for the kernel of the linear semigroup. It is also possible to replace the nonlinearity by some slightly more complicated expression of $u(t)$.

For any Banach space $\mathscr{B}$, a $\mathscr{B}$-valued stochastic process $u_\xi(t)$ is called a *mild solution* of (1.2) with initial condition ξ, if it satisfies the associated integral equation

$$u_\xi(t) = e^{Lt}\xi + \int_0^t e^{L(t-s)}F(u_\xi(s))\,ds + \int_0^t e^{L(t-s)}Q\,dW(s)\;, \tag{1.3}$$

in the sense that every term defines a stochastic process on $\mathscr{B}$ and that the equality holds almost surely with respect to the probability measure on the abstract probability space underlying the Wiener process. The initial condition does not have to belong to $\mathscr{B}$, provided $e^{Lt}\xi \in \mathscr{B}$ for all times $t > 0$.

To a Markovian solution, we can associate (under suitable conditions) the *transition semigroup* $\mathcal{P}_t$ defined on and into the set of bounded Borel functions $\varphi : \mathscr{B} \to \mathbf{C}$ by

$$(\mathcal{P}_t\varphi)(\xi) = \int_{\mathscr{B}} \varphi(\eta)\mathbf{P}(u_\xi(t) \in d\eta)\;. \tag{1.4}$$

Its dual semigroup $\mathcal{P}_t^*$ is defined on and into the set of Borel probability measures ν on $\mathscr{B}$ by

$$(\mathcal{P}_t^*\nu)(\Gamma) = \int_{\mathscr{B}} \mathbf{P}(u_\xi(t) \in \Gamma)\,\nu(d\xi)\;, \tag{1.5}$$

where Γ is a $\mathscr{B}$-Borel set. If the existence of the solutions is shown for initial conditions in a larger Banach space $\mathscr{B}'$ in which $\mathscr{B}$ is continuously embedded, $\mathcal{P}_t^*$ can be extended to a map from the $\mathscr{B}'$-Borel probability measures into the $\mathscr{B}$-Borel probability measures.

An *invariant measure* for (1.2) is a probability measure on $\mathscr{B}$ which is a fixed point for $\mathcal{P}_t^*$. If $\mathscr{T}$ is a weaker topology on $\mathscr{B}$, we can under appropriate conditions extend $\mathcal{P}_t^*$ by (1.5) to a mapping from the $\mathscr{T}$-Borel probability measures into themselves. In the case of $\mathrm{L}^\infty(\mathbf{R})$, we may for example consider a "weighted topology" $\mathscr{T}_\varrho$ induced by some weighted norm $\|\varrho\cdot\|_\infty$.

If we take $\varphi_2(x) = 1$, it is known (we refer to [DPZ96] for details) that (1.2) possesses a mild solution in $\mathrm{L}^p(\mathbf{R}, \varrho(x)\,dx)$ for a weight function ϱ that decays at infinity. Our choice for Q makes it possible to work in flat spaces, since the noise is damped at infinity. In fact, we will show that, for every initial condition $u_0 \in \mathrm{L}^\infty(\mathbf{R})$, (1.2) possesses a mild solution in $C_u(\mathbf{R})$, the space of bounded uniformly continuous functions on $\mathbf{R}$. This leads to slight technical difficulties since neither $\mathrm{L}^\infty(\mathbf{R})$ nor $C_u(\mathbf{R})$ are separable Banach spaces, and thus standard existence theorems do not apply.

After proving the existence of the solutions, we will be concerned with their regularity. We prove that with high probability the solution $u_\xi(t)$ of (SGL) for a fixed time is analytic in a strip around the real axis. We will also derive estimates on the width of that strip. These estimates will finally allow to show the existence of an invariant measure for $\mathcal{P}_t^*$, provided we equip $C_u(\mathbf{R})$ with a slightly weaker topology. The existence of an invariant measure is not a trivial result since

a. The linear semigroup of (SGL) is not made of compact operators in $C_u(\mathbf{R})$.

b. The deterministic equation is not strictly dissipative, in the sense that there is not a unique fixed point that attracts every solution.

c. The deterministic equation is of the gradient type, but the operator Q is not invertible, so we can not make the *a priori* guess that the invariant measure is some Gibbs measure.

The results we found in the literature about the existence of invariant measures for infinite-dimensional stochastic differential equations (see *e.g.* [JLM85, DPZ92a, DPZ96, BKL00a] and references therein) usually assume that the converse of either *a.*, *b.* or *c.* holds. The main result of this paper is the following.

Theorem 1.1 *There exist slowly decaying weight functions ϱ such that the extension of $\mathcal{P}_t^*$ to the $\mathscr{T}_\varrho$-Borel probability measures is well-defined and admits a fixed point.*

Remark 1.2 The hypotheses of this theorem have been made with the following future project in mind. We hope to prove that the measure found in Theorem 1.1 is *unique*. The basic idea is to apply the methods of [EPR99b] to the context of SPDE's to show uniqueness of the measure by the tools of control theory. In this context, it is interesting if the noise drives the system only in the dissipative range, namely in a *finite* interval of frequencies which need not contain the unstable modes of the deterministic Ginzburg-Landau equation. In particular, such forces do *not* have invertible covariances and hence methods such as those found in [DPZ96] do not apply.

This is also the reason why the setting considered in this paper imposes $\hat{\varphi}_1$ to have compact support, although the extension to exponentially decaying functions would have been easy.

The next sections will be organized as follows. In Section 2, we give detailed bounds on the stochastic convolution, *i.e.* on the evolution of the noise under the action of the semigroup generated by L. In Section 3 we then prove the existence of a unique solution for (1.2) and derive an *a priori* estimate on its amplitude. Section 4 is devoted to the study of the analyticity properties of the solution. In Section 5, we finally show the existence of an invariant measure for the dynamics, *i.e.* we prove Theorem 1.1 which will be restated as Theorem 5.4. The appendix gives conditions under which one can prove the existence of a global strong solution to a class of semilinear PDE's in a Banach space.

1.1 Definitions and notations

Consider the sets $\mathscr{A}_\eta$ of functions that are analytic and uniformly bounded in an open strip of width 2η centered around the real axis. They are Banach spaces with respect to the norms

$$\|f\|_{\eta,\infty} \equiv \sup_{z\,:\,|\mathrm{Im} z|<\eta} |f(z)| \ .$$

Fix $T > 0$. We define $\mathscr{B}_T$ as the Banach space of functions $f(t,x)$ with $t \in (0,T]$ and $x \in \mathbf{R}$ such that for fixed $t > 0$, $f(t,\cdot)$ is analytic and bounded in the strip $\{z = x + iy \,|\, |y| < \sqrt{t}\}$. We equip $\mathscr{B}_T$ with the norm

$$\|\!|f|\!\|_T \equiv \sup_{t\in(0,T]} \|\!|f(t,z)|\!\|_{\sqrt{t},\infty} \;.$$

In the sequel we denote by $\|\cdot\|_p$ the norm of $\mathrm{L}^p(\mathbf{R}, dx)$. For M a metric space and $\mathscr{B}$ a Banach space, the symbol $C_b(M,\mathscr{B})$ (resp. $C_u(M,\mathscr{B})$) stands for the Banach space of bounded (uniformly) continuous functions $M \to \mathscr{B}$ endowed with the usual sup norm. If $\mathscr{B} = \mathbf{C}$, it is usually suppressed in the notation. Moreover, the symbol C denotes a constant which is independent of the running parameters and which may change from one line to the other (even inside the same equation).

The symbol $\mathscr{L}(X)$ denotes the probability law of a random variable X. The symbol $\mathcal{B}(M,r)$ denotes the open ball of radius r centered at the origin of a metric vector space M.

2 The Stochastic Convolution

This section is devoted to the detailed study of the properties of the stochastic process obtained by letting the semigroup generated by L act on the noise.

2.1 Basic properties

Let us denote by $(\Omega, \mathscr{F}, \mathbf{P})$ the underlying probability space for the cylindrical Wiener process dW, and by $\mathbf{E}$ the expectation in Ω. We define the stochastic convolution

$$W_L(t,\omega) = \int_0^t e^{L(t-s)} Q\, dW(s,\omega)\;, \quad \omega \in \Omega\;. \tag{2.1}$$

The argument ω will be suppressed during the major part of the discussion. For a discussion on the definition of the stochastic integral in infinite-dimensional Banach spaces, we refer to [DPZ92b]. Notice that since $\hat\varphi_1$ has compact support, we can find a $\mathcal{C}_0^\infty$ function $\hat\psi$ such that $\hat\psi(x) = 1$ for $x \in \operatorname{supp}\hat\varphi$. We define $\tilde Q f = \psi \star f$ and fix a constant R such that

$$\operatorname{supp}\hat\varphi \subset \operatorname{supp}\hat\psi \subset \{x \in \mathbf{R} \,|\, |x| \le R\}\;. \tag{2.2}$$

We have of course $\tilde Q Q = Q$. An important consequence of this property is

Lemma 2.1 *Fix $\eta > 0$ and $\alpha < 1/2$. Then there exists a version of W_L with α-Hölder continuous sample paths in $\mathscr{A}_\eta$. Furthermore, for every $T > 0$, the mapping*

$$\begin{aligned} W_L^\eta : \Omega &\to C_b([0,T], \mathscr{A}_\eta)\;,\\ \omega &\mapsto W_L(\cdot,\omega)\;, \end{aligned} \tag{2.3}$$

is measurable with respect to the Borel σ-field generated by the strong topology on the space $C_b([0,T],\mathscr{A}_\eta)$.

Remark 2.2 The meaning of the word "version" is that the process constructed here differs from (2.1) only on a set of **P**-measure 0. We will in the sequel not make any distinction between both processes.

Proof of Lemma 2.1. We first notice that $W_L(t)$ has an α-Hölder continuous version in $\mathrm{L}^2(\mathbf{R})$. This is a consequence of the fact that the Hilbert-Schmidt norm in $\mathrm{L}^2(\mathbf{R})$ of $\exp(Lt)Q$ is bounded by $e^{-t}\|\varphi_1\|_2\|\varphi_2\|_2$. Since $\mathrm{L}^2(\mathbf{R})$ is separable, the mapping

$$\begin{aligned} W_L : \Omega &\to C_b([0,T], \mathrm{L}^2(\mathbf{R})) , \\ \omega &\mapsto W_L(\cdot, \omega) , \end{aligned}$$

is measurable [DPZ92b, Prop 3.17]. Since L and $\tilde{Q}$ commute, we can write

$$W_L(t,\omega) = \int_0^t \tilde{Q}^2 e^{L(t-s)} Q \, dW(s,\omega) = \tilde{Q}^2 W_L(t,\omega) , \tag{2.4}$$

where we used [DPZ96, Prop. 4.15] to commute the operator and the integral. We will show that $\tilde{Q}^2$ defines a bounded continuous linear operator from $\mathrm{L}^2(\mathbf{R})$ into $\mathscr{A}_\eta$. The claim then follows if we define the map $W_L^\eta = \tilde{Q}_\eta^2 \circ W_L$, where we denote by $\tilde{Q}_\eta^2$ the operator constructed in an obvious way from $\tilde{Q}^2$ as a map from $C_b([0,T], \mathrm{L}^2(\mathbf{R}))$ into $C_b([0,T], \mathscr{A}_\eta)$.

Notice first that if $f \in \mathrm{L}^2(\mathbf{R})$, we have by the Young inequality $\tilde{Q}f \in \mathrm{L}^\infty(\mathbf{R})$ and the estimate

$$\|\tilde{Q}f\|_\infty \le \|\psi\|_2 \|f\|_2 \tag{2.5}$$

holds. Take now $f \in \mathrm{L}^\infty(\mathbf{R})$. Since $\tilde{Q}$ maps any measurable function onto an entire analytic function, $\tilde{Q}f(z)$ has a meaning for every $z \in \mathbf{C}$. We have for any $x \in \mathbf{R}$

$$|(\tilde{Q}f)(x+i\eta)| = \Big| \int_R \psi(x+i\eta-y) f(y) \, dy \Big| . \tag{2.6}$$

By assumption, the Fourier transform of ψ belongs to $\mathcal{C}_0^\infty$. We know that such functions enjoy the property – see *e.g.* [RS80] – that for each $N > 0$ there exists a constant C_N such that

$$|\psi(x+i\eta)| \le \frac{C_N e^{R|\eta|}}{(1+x^2+\eta^2)^N} ,$$

where the constant R is defined in (2.2). We thus have the estimate

$$\begin{aligned} |(\tilde{Q}f)(x+i\eta)| &\le \|f\|_\infty \int_{\mathbf{R}} |\psi(x+i\eta-y)| \, dy \\ &\le C e^{R|\eta|} \|f\|_\infty , \end{aligned} \tag{2.7}$$

and thus

$$|\!|\!|\tilde{Q}f|\!|\!|_{\eta,\infty} \le C e^{R|\eta|} \|f\|_\infty . \tag{2.8}$$

Collecting (2.5) and (2.8) proves the claim. □

Remark 2.3 As an evident corollary of the proof of the lemma, note that $W_L(t) \in \mathcal{D}(L)$ for all times $t \geq 0$ and that the mapping

$$\begin{aligned} W_L : \Omega &\to C_b([0,T], \mathcal{D}(L)) \;, \\ \omega &\mapsto W_L(\cdot, \omega) \;, \end{aligned} \tag{2.9}$$

has the same properties as the mapping W_L^η if we equip $\mathcal{D}(L)$ with the graph norm. In particular, W_L has almost surely α-Hölder continuous sample paths in $\mathcal{D}(L)$.

We will now give more precise bounds on the magnitude of the process W_L. Our main tool will be the so-called "factorization formula" which will allow to get uniform bounds over some finite time interval.

2.2 Factorization of the stochastic convolution

We define, for $\delta \in (0, 1/2)$,

$$\begin{aligned} Y_{L,\delta}(t) &= \int_0^t (t-s)^{-\delta} e^{L(t-s)} Q \, dW(s) \;, \\ (G_\delta \Psi)(t) &= \int_0^t (t-s)^{\delta-1} e^{L(t-s)} \Psi(s) \, ds \;. \end{aligned}$$

Notice that we can show by the same arguments as in Lemma 2.1 that the process $Y_{L,\delta}(t)$ has a version which takes values in $\mathscr{A}_\eta$. Thus, in particular the expression $Y_{L,\delta}(t,x)$ is a well-defined complex-valued random variable. A corollary of the stochastic Fubini theorem (sometimes referred to as the "factorization formula" [DPZ92b]) shows that

$$W_L(t) = \frac{\sin \pi\delta}{\pi} (G_\delta Y_{L,\delta})(t) \;. \tag{2.10}$$

Before we start to estimate $\|W_L(t)\|_\infty$, we state without proof the following trivial consequence of the Young inequality:

Lemma 2.4 *Denote by g_t the heat kernel and choose $p > 1$. Then there exists a constant c depending on p such that*

$$\|g_t \star f\|_\infty \leq c t^{-1/(2p)} \|f\|_p \;, \tag{2.11}$$

holds for every $f \in \mathrm{L}^p(\mathbf{R})$.

We have, using (2.10), Lemma 2.4, and the Hölder inequality,

$$\begin{aligned} \|W_L(t)\|_\infty &\leq C \int_0^t (t-s)^{\delta-1} e^{-(t-s)} \|g_{t-s} \star Y_{L,\delta}(s)\|_\infty \, ds \\ &\leq C \int_0^t (t-s)^{\delta-1-1/(2p)} \|Y_{L,\delta}(s)\|_p \, ds \\ &\leq C \Big(\int_0^t (t-s)^{q(\delta-1-1/(2p))} \, ds \Big)^{1/q} \Big(\int_0^t \|Y_{L,\delta}(s)\|_p^p \, ds \Big)^{1/p} \;, \end{aligned}$$

where q is chosen such that $p^{-1} + q^{-1} = 1$. It is easy to check that the first integral converges when

$$p > \frac{3}{2\delta} \;. \tag{2.12}$$

In that case, we have

$$\|W_L(t)\|_\infty^p \le Ct^\gamma \int_0^t \|Y_{L,\delta}(s)\|_p^p \, ds \;, \qquad \gamma = p\delta - \frac{3}{2} \;. \tag{2.13}$$

So it remains to estimate $\|Y_{L,\delta}(t)\|_p$.

2.3 Estimate on the process $Y_{L,\delta}(t)$

This subsection is devoted to the proof of the following lemma.

Lemma 2.5 *Let $Y_{L,\delta}$ be as above and choose $p \ge 2$ and $\delta \in (0, 1/2)$. There exists a constant c depending on δ, p, φ_1 and φ_2 but independent of t such that $\mathbf{E}\|Y_{L,\delta}(t)\|_p^p \le c$.*

Remember that the convolution of two decaying functions decays like the one that decays slower at infinity:

Lemma 2.6 *Let f and g be two positive even functions which are integrable and monotone decreasing between 0 and ∞. Then the estimate*

$$|(f \star g)(x)| \le |f(x/2)| \, \|g\|_1 + |g(x/2)| \, \|f\|_1$$

holds.

Proof. Assume $x \ge 0$ (the case $x < 0$ can be treated in a similar way) and define $I_x = (x/2, 3x/2)$. We can decompose the convolution as

$$\begin{aligned} |(f \star g)(x)| &\le \int_{I_x} |f(y - x)g(y)| \, dy + \int_{\mathbf{R}\setminus I_x} |f(y - x)g(y)| \, dy \\ &\le |g(x/2)| \int_{\mathbf{R}} |f(y)| \, dy + |f(x/2)| \int_{\mathbf{R}} |g(y)| \, dy \;, \end{aligned}$$

which proves the assertion. □

Proof of Lemma 2.5. We use the formal expansion

$$dW(x,t) = \sum_{j=1}^\infty e_j(x) \, dw_j(t) \;,$$

where the e_i form an orthonormal basis of $\mathrm{L}^2(\mathbf{R}, dx)$ (say the eigenfunctions of the harmonic oscillator) and the dw_i are independent Wiener increments. We also denote by T_x the translation operator $(T_x f)(y) = f(y - x)$. We then have

$$\begin{aligned} \mathbf{E}|Y_{L,\delta}(t,x)|^2 &= \mathbf{E}\Big|\int_0^t \sum_{j=1}^\infty (t-s)^{-\delta} e^{-(t-s)} (g_{t-s} \star \varphi_1 \star (\varphi_2 \, e_j))(x) \, dw_j(s)\Big|^2 \\ &= \int_0^t \sum_{j=1}^\infty (t-s)^{-2\delta} e^{-2(t-s)} |(g_{t-s} \star \varphi_1 \star (\varphi_2 \, e_j))(x)|^2 \, ds \\ &= \int_0^t (t-s)^{-2\delta} e^{-2(t-s)} \sum_{j=1}^\infty |\langle \varphi_2 \, T_x(g_{t-s} \star \varphi_1), e_j \rangle|^2 \, ds \end{aligned}$$

$$= \int_0^t s^{-2\delta} e^{-2s} \|\varphi_2\, T_x(g_s \star \varphi_1)\|_2^2 \, ds \;.$$

An explicit computation shows the equality

$$\|\varphi_2\, T_x(g_s \star \varphi_1)\|_2^2 = (\varphi_2^2 \star (g_s \star \varphi_1)^2)(x) \;.$$

Using Lemma 2.6, the fact that $\varphi_1(x) \le C_N \langle\!\langle x \rangle\!\rangle^{-N}$ for every N, and the well-known inequality $|g_s \star \varphi_1|(x) \le \|\varphi_1\|_\infty$, we get the estimate

$$(g_s \star \varphi_1)^2(x) \le C\Big(\frac{e^{-x^2/(16s)}}{\langle\!\langle s \rangle\!\rangle} + \frac{1}{\langle\!\langle x \rangle\!\rangle^N}\Big) \;.$$

Using again Lemma 2.6 and (1.1), we get

$$\|\varphi_2\, T_x(g_s \star \varphi_1)\|_2^2 \le C\Big(\frac{e^{-x^2/(64s)}}{\langle\!\langle s \rangle\!\rangle} + \frac{1}{\langle\!\langle x \rangle\!\rangle^{1+2\beta}}\Big) \;.$$

It is now an easy exercise to show that

$$\sup_{s>0} \|\varphi_2\, T_x(g_s \star \varphi_1)\|_2^2 \le C\Big(\frac{1}{\langle\!\langle x \rangle\!\rangle^2} + \frac{1}{\langle\!\langle x \rangle\!\rangle^{1+2\beta}}\Big) \;.$$

Defining $\beta' = \min\{1/2, \beta\}$, and using $\langle\!\langle x \rangle\!\rangle \ge 1$, we have

$$\mathbf{E}|Y_{L,\delta}(t,x)|^2 \le C \langle\!\langle x \rangle\!\rangle^{-1-2\beta'} \int_0^t s^{-2\delta} e^{-2s} \, ds \le C \langle\!\langle x \rangle\!\rangle^{-1-2\beta'} \;.$$

Since $Y_{L,\delta}(t,x)$ is a Gaussian random variable, this implies, for $p \ge 2$

$$\begin{aligned} \mathbf{E}\|Y_{L,\delta}(t)\|_p^p &= \int_{\mathbf{R}} \mathbf{E}|Y_{L,\delta}(t,x)|^p \, dx \le C \int_{\mathbf{R}} (\mathbf{E}|Y_{L,\delta}(t,x)|^2)^{p/2} \, dx \\ &\le C \int_{\mathbf{R}} \frac{1}{\langle\!\langle x \rangle\!\rangle^{p/2+\beta' p}} \, dx \le C \;. \end{aligned} \tag{2.14}$$

This proves the assertion. □

As a corollary of Lemma 2.5, we have the following estimate on the process $W_L(t)$.

Corollary 2.7 *For any $p \ge 2$, there is a constant $C > 0$ such that $\mathbf{E}\|W_L(t)\|_\infty^p \le C$ for all times $t \ge 0$.*

Proof. Using again the equality $W_L(t) = \tilde{Q} W_L(t)$, we notice that it is enough to have an estimate on $\mathbf{E}\|W_L(t)\|_p^p$. This can be done by retracing the proof of Lemma 2.5 with δ replaced by 0. □

We have now collected all the necessary tools to obtain the main result of this section.

Theorem 2.8 *For every $\varepsilon > 0$, there are constants $C, R > 0$ depending only on the choices of φ_1, φ_2 and ε such that the estimate*

$$\mathbf{E}|||W_L|||_T \le Ce^{R\sqrt{T}}T^{1/2-\varepsilon}$$

holds. □

Proof. The estimate

$$|||W_L|||_T \le Ce^{R\sqrt{T}} \sup_{t\in(0,T]} \|W_L(t)\|_\infty \;, \tag{2.15}$$

holds as a consequence of Eqs. (2.6) and (2.8). We thus need an estimate on $\|W_L(t)\|_\infty$ which is uniform on some time interval. This is achieved by combining Lemma 2.5 with Eq. (2.13). Let us first choose a constant $\delta > 1/2$, but very close to $1/2$ and then a (big) constant p such that $p > \max\{2, 3/(2\delta)\}$. Since $\sup_{t\in(0,T]} \|W_L(t)\|_\infty$ is a positive random variable, we have

$$\begin{aligned}
\mathbf{E}\Big(\sup_{t\in(0,T]} \|W_L(t)\|_\infty\Big) &\le C\Big(\mathbf{E}(\sup_{t\in(0,T]}\|W_L(t)\|_\infty)^p\Big)^{1/p} \\
&= C\Big(\mathbf{E}(\sup_{t\in(0,T]}\|W_L(t)\|_\infty^p)\Big)^{1/p} \\
&\le C\Big(T^\gamma \int_0^T \mathbf{E}\|Y_{L,\delta}(s)\|_p^p\,ds\Big)^{1/p} \\
&\le CT^{(\gamma+1)/p} \le CT^{\delta-1/(2p)} \;.
\end{aligned} \tag{2.16}$$

The exponent $\delta - 1/(2p)$ can be brought arbitrarily close to $1/2$. This, together with the previous estimate (2.15), proves the claim. □

We have now the necessary tools to prove the existence of a unique solution to the SPDE (1.2).

3 Existence of the Solutions

Throughout this section, we denote by $\mathscr{B}$ the Banach space $C_u(\mathbf{R})$ of bounded uniformly continuous complex-valued functions on the real line endowed with the norm $\|\cdot\|_\infty$. The reason why we can not use a standard existence theorem is that $\mathscr{B}$ is not separable. Nevertheless, the outline of our proof is quite similar to the proofs one can find in [DPZ92b]. The technique is to solve (1.2) pathwise and then to show that the result yields a well-defined stochastic process on $\mathscr{B}$ which is a mild solution to the considered problem. In order to prepare the existence proof for solutions of (1.2), we study the dynamics of the *deterministic* equation

$$\dot{X}_\xi(W,t) = LX_\xi(W,t) + F(X_\xi(W,t) + W(t)) \;, \quad X_\xi(W,0) = \xi. \tag{3.1}$$

In this equation, $\xi \in \mathrm{L}^\infty(\mathbf{R})$ is an arbitrary initial condition and $W \in C_b([0,T], \mathscr{A}_\eta)$ is an arbitrary noise function with $W(0) = 0$ and $\eta > 0$ fixed. For the moment, we choose an arbitrary time $T > 0$ and study the solutions up to time T. The reason why we study (3.1) is that if X_ξ is a solution of (3.1), then $Y_\xi(t) = X_\xi(t) + W(t)$ is a solution of

$$\dot{Y}_\xi(t) = LY_\xi(t) + F(Y_\xi(t)) + \dot{W}(t) \;, \quad Y_\xi(0) = \xi,$$

provided $W : [0,T] \to \mathscr{A}_\eta$ is a differentiable function. Because of the dissipativity of F, we will show that (3.1) possesses a unique bounded and continuous solution in $\mathscr{B}$ for all times $t \in (0,T]$. Consider the map

$$\begin{aligned} S_\xi^T : C_b([0,T],\mathscr{A}_\eta) &\to C_b((0,T],\mathscr{B}) , \\ W(\cdot) &\mapsto X_\xi(W,\cdot) , \end{aligned}$$

that associates to every noise function W and every initial condition $\xi \in \mathrm{L}^\infty(\mathbf{R})$ the solution of (3.1). (We do not show explicitly the value of η in the notations, since the map S_ξ^T is in an obvious sense independent of η.) We have the following result.

Lemma 3.1 *The map* $(\xi, W) \mapsto S_\xi^T(W)$ *is locally Lipschitz continuous in both arguments. Furthermore, the estimates*

$$\|S_\xi^T(W)\| \le \max\{\|\xi\|_\infty, C(1+\|W\|^3)\} , \tag{3.2a}$$

$$\|S_\xi^T(W) - S_\zeta^T(W)\| \le e^T \|\xi - \zeta\|_\infty , \tag{3.2b}$$

hold.

Proof. The proof relies on the results of Appendix A. As a first step, we verify that the assumptions of Theorem A.2 are satisfied with $F_t(x) = F(x + W(t))$. It is well-known [Lun95] that **A1** is satisfied for the Laplacean and thus for L. Using the easy-to-check inequality

$$|(a-b) + \alpha(a|a|^2 - b|b|^2)| \ge |a-b|\Big(1 + \alpha\frac{|a|^2+|b|^2}{2}\Big) ,$$

which holds for any $a, b \in \mathbf{C}$ and $\alpha \ge 0$, it is also straightforward to check that the mapping $L + F_t$ is κ-quasi dissipative for all times with $\kappa = 1$ and therefore **A2** holds. Assumption **A3** can be checked in a similar way. To check **A4**, notice that by Cauchy's integral representation theorem, $\mathscr{A}_\eta \subset \mathcal{D}(L)$, and so F_t maps $\mathcal{D}(L)$ into itself. Furthermore, it is easy to check the inequality

$$\|\partial_x v\|_\infty^2 \le C\|v\|_\infty \|\partial_x^2 v\|_\infty , \qquad v \in \mathcal{D}(L) . \tag{3.3}$$

We leave it to the reader to verify, with the help of (3.3), that **A4** is indeed satisfied. It is clear by the continuity of $W(\cdot)$ that **A5** holds as well, so we are allowed to use Theorem A.2.

We will show that (3.2) holds for arbitrary initial conditions in $\mathcal{D}(L)$. To show that they also hold for arbitrary initial conditions in $\mathrm{L}^\infty(\mathbf{R})$, we can apply arguments similar to what is done at the end of the proof of Theorem A.2.

Until the end of the proof, we will always omit the subscript ∞ in the norms. Denote by $X(t)$ the solution of (3.1). Since $X(t)$ is strongly differentiable by Theorem A.2, the left lower Dini derivative $D_-\|X(t)\|$ satisfies by (A.2)

$$\begin{aligned} D_-\|X(t)\| &\le \liminf_{h\to 0^+} h^{-1}(\|X(t)\| - \|X(t) - hLX(t) - hF_t(X(t))\|) \\ &\le -\|X(t)\| + C(1 + \|W(t)\|^3) , \end{aligned} \tag{3.4}$$

where the last inequality is easily obtained by inspection, absorbing the linear instability into the strongly dissipative term $-X(t)|X(t)+W(t)|^2$. The estimate (3.2a) follows immediately from a standard theorem about differential inequalities [Wal64].

Inequality (3.2b) is an immediate consequence of Theorem A.2.

It remains to show that $S_\xi^T(W)$ is a locally Lipschitz continuous function of W. We call $X(t)$ and $\tilde{X}(t)$ the solutions of (3.1) with noise functions W and V respectively. We also denote by F_t^W and F_t^V the corresponding nonlinearities. In a similar way as above, we obtain the inequality

$$\begin{aligned} D_-\|X(t)-\tilde{X}(t)\| \le \|X(t)-\tilde{X}(t)\| &+ \frac{\|(F_t^W-F_t^V)(X(t))\|}{2} \\ &+ \frac{\|(F_t^W-F_t^V)(\tilde{X}(t))\|}{2} \;. \end{aligned}$$

The claim now follows from the estimate

$$\|(F_t^W-F_t^V)(x)\| \le C\|W-V\|(1+\|x\|^2+\|W\|^2+\|V\|^2) \;,$$

and from the *a priori* estimate (3.2a) on the norms of $X(t)$ and $\tilde{X}(t)$. □

Before we state the existence theorem, let us define the following.

Definition 3.2 A transition semigroup $\mathcal{P}_t$ on a Banach space $\mathscr{B}$ has the *weak Feller* property if $\mathcal{P}_t\varphi \in C_u(\mathscr{B})$ for every $\varphi \in C_u(\mathscr{B})$.

Theorem 3.3 *For every initial condition in* $\mathrm{L}^\infty(\mathbf{R})$*, the SPDE defined by (SGL) possesses a unique continuous mild solution in* $\mathscr{B}$ *for all times. The solution is Markov, its transition semigroup is well-defined and weak Feller and its sample paths are almost surely* α*-Hölder continuous for every* $\alpha < 1/2$*.*

Proof. The main work for the proof was done in Lemma 3.1. Recall the definition (2.3) of the mapping W_L^η that associates to every element of Ω a continuous noise function in $\mathscr{A}_\eta$. Since $\mathscr{A}_\eta$ is continuously embedded in $\mathscr{B}$, we can *define* the random variable

$$\begin{aligned} u_\xi^T : \Omega &\to C_b((0,T],\mathscr{B}) \;, \\ \omega &\mapsto (S_\xi^T \circ W_L^\eta)(\omega) + W_L^\eta(\omega) \;, \end{aligned}$$

for some $\eta > 0$ and some $T > 0$. This allows to define the stochastic process

$$\begin{aligned} u_\xi(t) : \Omega &\to \mathscr{B} \;, \\ \omega &\mapsto (u_\xi^T(\omega))(t) \;, \end{aligned}$$

for some $T > t$. It is clear by the uniqueness of the solutions to the deterministic equation (3.1) that this expression is well-defined, *i.e.* does not depend on the particular choice of T. It is also independent of the choice of η. Since W_L^η is measurable and S_ξ^T is continuous, u_ξ is a well-defined stochastic process with values in $\mathscr{B}$. It is immediate from the definitions of W_L^η and S_ξ^T that u_ξ is indeed a mild solution to (SGL). The Markov property follows from the construction and the Markov property of W_L.

To show that the transition semigroup is well-defined, it suffices by Fubini's theorem to show that the function

$$P_{\xi,t}(\Gamma) = \mathbf{P}(u_\xi(t) \in \Gamma) = \int_\Omega \chi_\Gamma(u_\xi(t,\omega))\,\mathbf{P}(d\omega)\;,$$

is measurable as a function of ξ for every $\mathscr{B}$-Borel set Γ and every $t \geq 0$. This is (again by Fubini's theorem) an immediate consequence of the measurability of W_L and the joint continuity of $S^\eta_\xi(W)$.

The weak Feller property is an immediate consequence of (3.2b), since

$$|(\mathcal{P}_t\varphi)(\xi) - (\mathcal{P}_t\varphi)(\zeta)| \leq \int_\Omega |\varphi(u_\xi(t,\omega)) - \varphi(u_\zeta(t,\omega))|\,\mathbf{P}(d\omega)\;.$$

Now choose $\varepsilon > 0$. Since $\varphi \in C_u(\mathscr{B})$, there exists $\delta > 0$ such that $|\varphi(x) - \varphi(y)| < \varepsilon$ for $\|x - y\| < \delta$. It suffices to choose ξ close enough to ζ such that $\|u_\xi(t,\omega) - u_\zeta(t,\omega)\| \leq e^t\|\xi - \zeta\| < \delta$ holds.

The α-Hölder continuity of the sample paths is a consequence of the strong differentiability (and thus local Lipschitz continuity) of the solutions of (3.1) and of the almost sure α-Hölder continuity of the sample paths of W_L. □

We now show that the solution of (1.2) not only exists in $C_b(\mathbf{R})$ but also stays bounded in probability. In fact we have

Lemma 3.4 *Let $u_\xi(t)$ be the solution of (1.2) constructed above with $\xi \in \mathrm{L}^\infty(\mathbf{R})$. There exist a time $T^* > 0$ depending on ξ and a constant $C > 0$ such that $\mathbf{E}\|u(t)\|_\infty \leq C$ for every time $t > T^*$.*

Proof. From (3.4), we obtain the estimate

$$\|u(t) - W_L(t)\|_\infty \leq e^{-t}\|\xi\|_\infty + C\int_0^t e^{-(t-s)}(1 + \|W_L(s)\|_\infty)^3\,ds\;.$$

This yields immediately

$$\sup_{t>T}\mathbf{E}\|u(t)\|_\infty \leq e^{-T}\|\xi\|_\infty + C\sup_{s>0}\mathbf{E}(1 + \|W_L(s)\|_\infty + \|W_L(s)\|^3_\infty)\;.$$

The claim follows now easily from Corollary 2.7. □

4 Analyticity of the Solutions

Our first step towards the existence proof for an invariant measure consists in proving that the solution of (SGL) constructed in Section 3 lies for all times in some suitable space of analytic functions. More precisely, we show that there is a (small) time T such that the solution of (SGL) up to time T belongs to $\mathscr{B}_T$. (Recall the definition of $\mathscr{B}_T$ given in Subsection 1.1.) The proof is inspired by that of [Col94] for the deterministic case, making use of the estimates of the preceding sections, in particular of Theorem 2.8.

We split the evolution into a linear part and the remaining nonlinearity. Recall the definitions

$$L = \Delta - 1 \quad \text{and} \quad F(u)(x) = u(x)(2 - |u(x)|^2) \; .$$

Throughout this section, we assume that $u(t)$ is a stochastic process solving (SGL) in the mild sense, *i.e.* there exists a $\xi \in \mathrm{L}^\infty(\mathbf{R})$ such that $u(t)$ satisfies (1.3). Such a process exists and is unique (given ξ) by Theorem 3.3.

For given functions $g \in \mathrm{L}^\infty(\mathbf{R})$ and $h \in \mathscr{B}_T$, we define the map $\mathcal{M}_{g,h} : \mathscr{B}_T \to \mathscr{B}_T$ as

$$\begin{aligned}(\mathcal{M}_{g,h}(f))(t) &= h(t) + e^{Lt} g + \int_0^t e^{L\tau} F(f(t-\tau))\, d\tau \\ &\equiv h(t) + (\mathcal{L}g)(t) + (\mathcal{N}f)(t) \; .\end{aligned} \tag{4.1}$$

Until the end of this proof, we write $\|\cdot\|$ instead of $\|\cdot\|_T$. It is possible to show – see [Col94] – that $\mathcal{M}_{g,h}$ is always well-defined on $\mathscr{B}_T$ and that there are constants k_1, k_2, k_3 such that

$$\begin{aligned}\|\mathcal{L}g\| &\le k_1 \|g\|_\infty \; , \\ \|\mathcal{N}f\| &\le k_2 T \|f\|^3 \; , \\ \|\mathcal{M}_{g,h} f_1 - \mathcal{M}_{g,h} f_2\| &\le k_3 T (1 + \|f_1\| + \|f_2\|)^2 \|f_1 - f_2\| \; .\end{aligned}$$

We now show that $u(t) \in \mathscr{A}_\eta$ with high probability for some $\eta > 0$. The precise statement of the result is

Theorem 4.1 *For any $\varepsilon > 0$ there are constants $\eta, \tilde{T}, C > 0$ such that $\mathbf{P}(u(t) \in \mathcal{B}(\mathscr{A}_\eta, C)) > 1 - \varepsilon$ for every time $t > \tilde{T}$.*

Proof. We fix $\tilde{T}$ bigger than the value T^* we found in Lemma 3.4, say $\tilde{T} = T^* + 1$. We also fix some time $T < 1$ to be chosen later and we choose an arbitrary time $t > \tilde{T}$. We show that with high probability, the solution $u(t - T + \cdot)$ belongs to $\mathscr{B}_T$. To begin, we take $g = u(t-T)$ and, for $s > 0$, we define

$$h(s) = \int_{t-T}^{t-T+s} e^{L(t-T+s-\sigma)} Q \, dW(\sigma) \; .$$

Since the Wiener increments are identically distributed independent random variables, it is clear that $\mathscr{L}(h(s)) = \mathscr{L}(W_L(s))$. In particular, Theorem 2.8 ensures the existence of a constant C_1 such that $\mathbf{E}\|h\| \le C_1$. By Lemma 3.4, there exists another constant C_2 such that $\mathbf{E}\|g\|_\infty < C_2$. Since the solution is Markovian, g and h are independent random variables and we have

$$\begin{aligned}\mathbf{P}\Big(\|g\|_\infty < \frac{2C_2}{\varepsilon} \quad \text{and} \quad \|h\| < \frac{2C_1}{\varepsilon}\Big) &= \mathbf{P}\Big(\|g\|_\infty < \frac{2C_2}{\varepsilon}\Big)\mathbf{P}\Big(\|h\| < \frac{2C_1}{\varepsilon}\Big) \\ &> (1 - \varepsilon/2)^2 > 1 - \varepsilon \; .\end{aligned}$$

From now on we assume that the above event is satisfied. Thus there is a constant $C_3 \approx \mathcal{O}(1/\varepsilon)$ such that

$$\|\mathcal{M}_{g,h} f\| \le C_3 + k_2 T \|f\|^3 \; .$$

If we impose now $T < 1/(8 k_2 C_3^2)$, we see that $\mathcal{M}_{g,h}$ maps the ball of radius $2C_3$ centered at 0 into itself. If we also impose the condition

$$T < \frac{1}{k_3 (1 + 4C_3)^2} \; ,$$

we see that $\mathcal{M}_{g,h}$ is a contraction on that ball. This, together with the uniqueness of the solutions of (SGL), proves the claim. It moreover shows that the width η of analyticity behaves asymptotically like $\eta \approx \mathcal{O}(\varepsilon)$. □

The above theorem tells us the probability for the solution to be analytic in a strip at a fixed time. Another property of interest is the behavior of the individual sample paths. We will show that any given sample path is always analytic with probability 1. Recall that $\mathscr{F}$ denotes the σ-field of the probability space underlying the cylindrical Wiener process.

Proposition 4.2 *There is an event $\Gamma \in \mathscr{F}$ with* $\mathbf{P}(\Gamma) = 1$ *such that for every $\xi \in \mathrm{L}^\infty(\mathbf{R})$, every $\omega \in \Gamma$, and every positive time $t > 0$, there exists a strictly positive value $\eta(t) > 0$ such that $u_\xi(t,\omega) \in \mathscr{A}_\eta$.*

Proof. Define for each integer n the set Γ_n as

$$\Gamma_n = \{w \in \Omega \,|\, W_L(\cdot,\omega) \in C([0,n],\mathscr{A}_n)\} \;.$$

We have $\mathbf{P}(\Gamma_n) = 1$ for all n by Lemma 2.1. By σ-completeness, $\Gamma = \bigcap_{n>0} \Gamma_n$ belongs to $\mathscr{F}$ and $\mathbf{P}(\Gamma) = 1$. We claim that Γ is the right event.

By the construction of Γ, the sample paths $u_\xi(\cdot,\omega)$ and $W_L(\cdot,\omega)$ are continuous and thus bounded on every finite time interval. Furthermore, $W_L(t,\omega) \in \mathscr{A}_\eta$ for every time and every positive η. The claim now follows easily from the proof of Theorem 4.1. □

5 Existence of an Invariant Measure

We can now turn to the proof of Theorem 1.1. We first define the set of weight functions $\mathscr{W}$ as the set of all functions $\varrho : \mathbf{R} \to \mathbf{R}$ which satisfy

a. The function $\varrho(x)$ is bounded, two times continuously differentiable and strictly positive.

b. For every $\varepsilon > 0$ there exists $x_\varepsilon > 0$ such that $|\varrho(x)| \le \varepsilon$ if $|x| \ge x_\varepsilon$.

c. There exist constants c_1 and c_2 such that

$$\Bigl|\frac{\partial_x \varrho(x)}{\varrho(x)}\Bigr| \le c_1 \qquad \text{and} \qquad \Bigl|\frac{\partial_x^2 \varrho(x)}{\varrho(x)}\Bigr| \le c_2 \;, \tag{5.1}$$

for all $x \in \mathbf{R}$.

Remark 5.1 The meaning of the expression “slowly decaying” used in Theorem 1.1 becomes clear from the following statement, the verification of which we leave to the reader. For *every* strictly positive decreasing sequence $\{x_n\}_{n=0}^\infty$ satisfying $\lim_{n\to\infty} x_n = 0$ and such that x_n/x_{n+1} remains bounded, it is possible to construct a function $\varrho \in \mathscr{W}$ such that $\varrho(n) = x_{|n|}$ for every $n \in \mathbf{Z}$. In particular, x_n may decay as slowly as $1/\log(\log(\ldots\log(C+n)\ldots))$, but is not allowed to decay faster than exponentially.

For every $\varrho \in \mathscr{W}$, we define the weighted norm

$$\|f\|_\varrho = \|\varrho f\|_\infty \;.$$

We can now consider the topological vector space $\mathscr{B}_\varrho$ which is equal as a set to $\mathscr{B} = C_u(\mathbf{R})$, but endowed with the (slightly weaker) topology induced by the norm $\|\cdot\|_\varrho$. The space $\mathscr{B}_\varrho$ is a metric space, but it is neither complete nor separable. Since the topology of $\mathscr{B}_\varrho$ is weaker than that of the original space $\mathscr{B}$, every $\mathscr{B}_\varrho$-Borel set is also a $\mathscr{B}$-Borel set and every probability measure on $\mathscr{B}$ can be restricted to a probability measure on $\mathscr{B}_\varrho$. Let us show that we can define consistently a transition semigroup $\mathcal{P}^*_{t,\varrho}$ acting on and into the set of $\mathscr{B}_\varrho$-Borel probability measures. We have

Proposition 5.2 *For every $\varrho \in \mathscr{W}$, the transition semigroup $\mathcal{P}^*_t$ associated to (SGL) can be extended to a transition semigroup $\mathcal{P}^*_{t,\varrho}$ such that (1.5) holds for every $\mathscr{B}_\varrho$-Borel set Γ. Furthermore, the transition semigroup $\mathcal{P}^*_{t,\varrho}$ is weak Feller.*

In order to prove this proposition, we will show the Lipschitz continuous dependence of the solutions on the initial conditions in the new topology. For this, we need (see Appendix A for the definition of a dissipative mapping in a Banach space):

Lemma 5.3 *The operator Δ is quasi dissipative with respect to the norm $\|\cdot\|_\varrho$.*

Proof. We have the equality

$$\varrho\Delta u = \Delta(\varrho u) - \frac{\Delta\varrho}{\varrho}(\varrho u) + 2\frac{\nabla\varrho}{\varrho}\nabla(\varrho u) - 2\Big|\frac{\nabla\varrho}{\varrho}\Big|^2(\varrho u)\ .$$

The claim follows from (5.1) and the fact that Δ and ∇ are dissipative operators with respect to $\|\cdot\|_\infty$. □

Proof of Proposition 5.2. Using Lemma 5.3, it is easy to check that the operator $L + F_t$ is, for all times and for a $\kappa \in \mathbf{R}$, κ-quasi dissipative with respect to the norm $\|\cdot\|_\varrho$. This yields as in Lemma 3.1 the estimate

$$\|S^T_\xi(W) - S^T_\eta(W)\|_\varrho \le e^{\kappa T}\|\xi - \eta\|_\varrho\ .$$

Using this estimate, we can retrace the arguments exposed in the proof of Theorem 3.3 to show that $\mathcal{P}^*_{t,\varrho}$ is well-defined and weak Feller. □

This construction is reminiscent of what was done in [MS95a, FLS96] to construct an attractor for the deterministic case. They also introduce a weighted topology on $\mathrm{L}^\infty(\mathbf{R})$ to overcome the fact that the attractor of the deterministic Ginzburg-Landau equation is not compact. Our result is the following.

Theorem 5.4 *For every $\varrho \in \mathscr{W}$, there exists a $\mathscr{B}_\varrho$-Borel probability measure μ_ϱ which is invariant for the transition semigroup $\mathcal{P}^*_{\varrho,t}$.*

The proof follows from a standard tightness argument. The main point is to notice that the unit ball of $\mathscr{A}_\eta$ is compact in $\mathscr{B}_\varrho$ for any weight function $\varrho \in \mathscr{W}$. We formulate this as a lemma.

Lemma 5.5 *The unit ball of $\mathscr{A}_\eta$ is a compact subset of $\mathscr{B}_\varrho$ for every $\varrho \in \mathscr{W}$.*

Proof. Since $\mathscr{B}_\varrho$ is a metric space, compact sets coincide with sequentially compact sets, see [Köt83]. We use the latter characterization. Choose a sequence $\mathcal{F} = \{f_n\}_{n=1}^\infty$ of functions in $\mathscr{A}_\eta$ with $|||f_n|||_{\eta,\infty} \le 1$ for all n. It is a standard theorem of complex analysis [Die68] that if $\mathcal{D} \subset \mathbf{C}$ is open and $\mathcal{F}$ is a family of analytic functions uniformly bounded on $\mathcal{D}$, then for every compact domain $K \subset \mathcal{D}$ there is a subsequence of $\mathcal{F}$ that converges uniformly on K to an analytic limit.

We define the subsequences $\mathcal{F}_n$ inductively by the following construction. First we choose $\mathcal{F}_{-1} = \mathcal{F}$. Then we consider the compact sets $\mathcal{D}_n = [-n, n]$ and we define $\mathcal{F}_n$ as a subsequence of $\mathcal{F}_{n-1}$ that converges uniformly on $\mathcal{D}_n$. Call $\hat{f}_n$ the resulting limit function on $\mathcal{D}_n$. We now define a global limit function $\hat{f}_\infty$ by $\hat{f}_\infty(x) = \hat{f}_n(x)$ if $x \in \mathcal{D}_n$. This procedure is well-defined since different $\hat{f}_n$ must by construction coincide on the intersection of their domains.

It remains now to exhibit a subsequence of $\mathcal{F}$ that converges to $\hat{f}_\infty$ in the topology of $\mathscr{B}_\varrho$. For every $n \ge 1$, choose $g_n \in \mathcal{F}_n$ such that $|g_n(z) - f_n(z)| < 1/n$ for $z \in \mathcal{D}_n$. The g_n form a subsequence of $\mathcal{F}$. We have moreover

$$\|g_n - \hat{f}_\infty\|_\varrho \le \|g_n - \hat{f}_N\|_\varrho + \|\hat{f}_N - \hat{f}_\infty\|_\varrho \le \frac{\|\varrho\|_\infty}{N} + 4 \sup_{|x|\ge N} |\varrho(x)| \; .$$

By hypotheses *a.* and *b.* on ϱ, this expression tends to 0 as N tends to ∞. □

Remark 5.6 By the compatibility of the various topologies with the linear structures, every bounded closed subset of $\mathscr{A}_\eta$ is compact as a subset of $\mathscr{B}_\varrho$.

Proof of Theorem 5.4. We choose an initial condition $\xi \in \mathrm{L}^\infty(\mathbf{R})$ and consider the family of $\mathscr{B}_\varrho$-Borel probability measures given by

$$\mu_t = \frac{1}{t}\int_0^t \mathcal{P}^*_{\varrho,t}(\delta_\xi)\, dt \; .$$

Fix now an arbitrary $\varepsilon > 0$. By Theorem 4.1 there exist $\eta, C, T > 0$ such that $\mu_t(\mathcal{B}(\mathscr{A}_\eta, C)) > 1 - \varepsilon$ for every $t > T$. Since $\mathcal{B}(\mathscr{A}_\eta, C)$ is compact in $\mathscr{B}_\varrho$ by Lemma 5.5, the family $\{\mu_t\}_{t>T}$ is tight and thus contains a weakly convergent subsequence by Prohorov's theorem. Denote by μ_ϱ the limit measure. Remember that a Borel probability measure on a metric space M is uniquely determined by its values on $C_u(M)$ [Bil68]. The weak Feller property of $\mathcal{P}^*_{t,\varrho}$ is thus sufficient to retrace the proof of the Krylov-Bogoluboff existence theorem [BK37, DPZ96], which states that μ_ϱ is invariant for $\mathcal{P}^*_{\varrho,t}$. □

A Dissipative Maps

This appendix will first give a short caracterization of dissipative maps in Banach spaces. We will then prove a global existence theorem for the solutions of non-autonomous semilinear PDE's with a dissipative nonlinearity.

Definition A.1 Given a Banach space $\mathscr{B}$ and a map $F : \mathcal{D}(F) \subset \mathscr{B} \to \mathscr{B}$, one says [DPZ92b] that F is *dissipative* if

$$\|x - y\| \le \|x - y - \alpha(F(x) - F(y))\| \; , \tag{A.1}$$

holds for every $x, y \in \mathcal{D}(F)$ and every $\alpha > 0$. If there exists a $\kappa \in \mathbf{R}$ such that $x \mapsto F(x) - \kappa x$ is dissipative, we say that F is κ-quasi dissipative (or quasi dissipative for short).

In the following, $u : (0,\infty) \to \mathscr{B}$ denotes a differentiable map. The function $\|u(\cdot)\|$ is of course continuous and its left-handed lower Dini derivative satisfies the inequality

$$\begin{aligned} D_-\|u(t)\| &= \liminf_{h\to 0^+} \frac{\|u(t)\| - \|u(t-h)\|}{h} \\ &\le \liminf_{h\to 0^+} \Big(\frac{\|u(t)\| - \|u(t) - h\dot{u}(t)\|}{h} + \frac{\|u(t-h) - u(t) + h\dot{u}(t)\|}{h} \Big) \\ &= \liminf_{h\to 0^+} \frac{\|u(t)\| - \|u(t) - h\dot{u}(t)\|}{h} \;. \end{aligned} \tag{A.2}$$

This estimate allows to get easily very useful estimates on the norm of the solutions of dissipative differential equations. For example, if $\dot{u}(t) = F(u(t))$ holds for all times and F is κ-quasi dissipative, then the estimate

$$\|u(t)\| \le e^{\kappa t} |\, \|u(0)\| - \|F(0)\|\, | + \|F(0)\| \tag{A.3}$$

holds as a consequence of a standard theorem about differential inequalities [Wal64].

We will now use standard techniques to prove a global existence theorem for the Cauchy problem

$$\dot{X}_\xi(t) = LX_\xi(t) + F_t(X_\xi(t)) \;, \quad X_\xi(0) = \xi \;, \tag{A.4}$$

and the associated integral equation

$$X_\xi(t) = e^{Lt}\xi + \int_0^t e^{L(t-s)} F_s(X_\xi(s))\, ds \;, \tag{A.5}$$

in a Banach space $\mathscr{B}$. We do *not* require that the domain of L be dense in $\mathscr{B}$. Let us denote by $\overline{\mathcal{D}(L)}$ the Banach space obtained by closing the domain of L in $\mathscr{B}$. Since, by assumption **A1** below, L is chosen to be closed, we can equip $\mathcal{D}(L)$ with the graph norm $\|x\|_L = \|x\| + \|Lx\|$ to obtain a Banach space. Our assumptions on L and F_t will be the following.

A1. The operator L is sectorial in the sense that its resolvent set contains the complement of a sector in the complex plane and that its resolvent satisfies the usual bounds [Lun95, Def 2.0.1].

This assumption implies [Lun95] that L generates an analytic semigroup $S(t)$ which is strongly continuous on $\overline{\mathcal{D}(L)}$ and maps $\mathscr{B}$ into $\mathcal{D}(L^k)$ for any $k \ge 0$. Furthermore, a bound of the form $\|S(t)\| \le Me^{\Omega t}$ holds. We will assume without loss of generality that $M \le 1$ and $\Omega = 0$. The latter assumption can be made since a constant can always be added to the nonlinear part. The former assumption is only made for convenience to simplify the notations. All the results also hold for $M > 1$. Another useful property of $S(t)$ is that there exists a constant c such that $\|S(t)\xi\|_L \le ct^{-1}\|\xi\|$ for $\xi \in \mathscr{B}$ and $t > 0$.

A2. There exist a positive time T and a real constant κ such that the mapping $x \mapsto Lx + F_t(x)$ is κ-quasi dissipative for all times $t \in [0,T]$.

This assumption will ensure the existence of the solutions up to the time T, which may be infinite.

A3. The function F_t is everywhere defined and there exist continuous increasing functions $a, \tilde{a} : \mathbf{R}_+ \to \mathbf{R}_+$ such that

$$\begin{aligned} \|F_t(x)\| &\leq a(\|x\|) , \\ \|F_t(x) - F_t(y)\| &\leq \|x - y\| \cdot \tilde{a}(\|x\| + \|y\|) , \end{aligned} \tag{A.6}$$

holds for every $x, y \in \mathscr{B}$ and for every $t \in [0, T]$.

A4. The map F_t maps $\mathcal{D}(L)$ into $\mathcal{D}(L)$ for all times and there exist continuous at most polynomially growing functions $b, \tilde{b} : \mathbf{R}_+ \to \mathbf{R}_+$ such that

$$\begin{aligned} \|F_t(x)\|_L &\leq b(\|x\|_L) , \\ \|F_t(x) - F_t(y)\|_L &\leq \|x - y\|_L \cdot \tilde{b}(\|x\|_L + \|y\|_L) , \end{aligned} \tag{A.7}$$

holds for every $x, y \in \mathcal{D}(L)$ and for every $t \in [0, T]$.

A5. The mapping $t \mapsto F_t(x)$ is continuous as a mapping $[0, T] \to \mathscr{B}$ for every $x \in \mathscr{B}$, and as a mapping $[0, T] \to \mathcal{D}(L)$ for every $x \in \mathcal{D}(L)$.

These assumptions allow us to show the existence of the solutions of (A.4) in the mild sense for any initial condition $\xi \in \mathscr{B}$ and in the strict sense for $\xi \in \mathcal{D}(L)$. Furthermore, we show that for any initial condition $\xi \in \mathscr{B}$, the solution lies in $\mathcal{D}(L)$ after an infinitesimal amount of time. Similar results can be found in the literature (see *e.g.* [Lun95, Hen81] and references therein), but with slightly different assumptions. The present result has by no means the pretention to generality but is tailored to fit our needs. Since the proof is not excessively long, we give it here for the sake of completeness.

Theorem A.2 *Assume* **A1–A5** *hold and choose* $\xi \in \mathscr{B}$. *Then there exists a unique function* $X_\xi : [0, T] \to \mathscr{B}$ *solving (A.5) for* $t \in [0, T]$. *The solutions satisfy* $\|X_\xi(t) - X_\eta(t)\| \leq e^{\kappa t}\|\xi - \eta\|$ *for all times. Furthermore,* $t \mapsto X_\xi(t)$ *is differentiable for* $t > 0$, $X_\xi(t) \in \mathcal{D}(L)$ *and its derivative satisfies (A.4).*

Proof. Assume first that the initial condition ξ belongs to $\mathcal{D}(L)$. We denote by $\mathscr{B}_{L,T}$ the Banach space $C([0, T], \mathcal{D}(L))$ with the usual sup norm. We show the local existence of a classical solution to (A.4) in $\mathscr{B}_{L,T}$ by a standard contraction argument. Choose $T_0 > 0$ and define the map $\mathcal{M}_\xi : \mathscr{B}_{L,T_0} \to \mathscr{B}_{L,T_0}$ by

$$(\mathcal{M}_\xi f)(t) = S(t)\xi + \int_0^t S(t - s)F_s(f(s))\, ds .$$

It is clear by **A1**, **A3**, **A4** and **A5** that $\mathcal{M}_\xi$ is well-defined and that the bounds

$$\|\mathcal{M}_\xi f\| \leq \|\xi\|_L + T_0 b(\|f\|) , \tag{A.8a}$$

$$\|\mathcal{M}_\xi f - \mathcal{M}_\xi g\| \leq T_0\|f - g\| \cdot \tilde{b}(\|f\| + \|g\|) , \tag{A.8b}$$

$$\|\mathcal{M}_\xi f - \mathcal{M}_\zeta f\| \leq \|\xi - \zeta\|_L , \tag{A.8c}$$

hold. It is clearly enough to take T_0 small enough, for example

$$T_0 < \min\Big\{ \frac{\|\xi\|_L}{b(2\|\xi\|_L)}\,,\ \frac{1}{\tilde b(4\|\xi\|_L)} \Big\}\;, \tag{A.9}$$

to find a contraction in the ball $\mathcal{B}(\mathscr{B}_{L,T_0}, 2\|\xi\|_L)$. Thus $\mathcal{M}_\xi$ possesses a unique fixed point X_ξ in $\mathscr{B}_{L,T_0}$. By [Lun95, Lem. 4.1.6], X_ξ is strongly differentiable in $\mathscr{B}$ and its derivative satisfies (A.4).

Using (A.2) and **A2**, we see immediately that for any $\xi, \zeta \in \mathcal{D}(L)$ and $t > 0$ such that the strong solutions X_ζ and X_ξ exist up to time t, the estimates

$$\begin{aligned} \|X_\xi(t)\| &\le |\|\xi\| - a(0)|e^{\kappa t} + a(0)\;, \\ \|X_\zeta(t) - X_\xi(t)\| &\le e^{\kappa t}\|\zeta - \xi\|\;, \end{aligned} \tag{A.10}$$

hold. The global existence of the solution now follows by iterating the above arguments, using (A.10) to ensure the non-explosion of the solutions. We leave it to the reader to verify that one can indeed continue the solutions up to the time T.

We next now show that for any initial condition $\xi \in \mathscr{B}$, the solution of (A.5) exists locally and lies in $\mathcal{D}(L)$ for positive times. We define $\mathcal{M}_\xi$ as above, but replace the space $\mathscr{B}_{L,T_0}$ by the larger space $\bar{\mathscr{B}}_{L,T_0}$ given by the measurable functions $f : (0, T_0] \to \mathcal{D}(L)$ with finite norm

$$|\!|\!| f |\!|\!| = \sup_{t\in(0,T_0]} \|tf(t)\|_L + \sup_{t\in(0,T_0]} \|f(t)\|\;.$$

We first show that $\mathcal{M}_\xi$ is well-defined on $\bar{\mathscr{B}}_{L,T_0}$. Choose $f \in \bar{\mathscr{B}}_{L,T_0}$. It is easy to check that, by **A3**, $\|(\mathcal{M}_\xi f)(t)\| \le \|\eta\| + T_0 a(|\!|\!| f |\!|\!|)$. By **A4**, we can choose n such that b and $\tilde b$ grow slower than $(1 + x)^n$. We also choose an exponent $N > n$ and choose $T_0 < 1$. We have, by the remark following **A1**, the estimate

$$\begin{aligned} \|t(\mathcal{M}_\xi f)(t)\|_L &\le \|tS(t)\xi\|_L + \int_0^{t-t^N} \Big\|tS(t-s)F_s(f(s))\Big\|_L\,ds \\ &\quad + \int_{t-t^N}^t \Big\|tS(t-s)F_s(f(s))\Big\|_L\,ds \\ &\le c\|\xi\| + \int_0^{t-t^N} \frac{ct}{t-s} a(\|f(s)\|)\,ds + \int_{t-t^N}^t tb(\|f(s)\|_L)\,ds \\ &\le c\|\xi\| + C_1 t\ln(t) a(|\!|\!| f |\!|\!|) + C_2 t^{N+1}\Big(1 + \frac{|\!|\!| f |\!|\!|}{t}\Big)^n\;. \end{aligned}$$

A similar estimate holds for $|\!|\!| \mathcal{M}_\xi f - \mathcal{M}_\xi g |\!|\!|$. Since $N > n$, there exists a function χ such that estimates of the type

$$|\!|\!| \mathcal{M}_\xi f |\!|\!| \le \sqrt{T_0}\chi(|\!|\!| f |\!|\!|) \quad \text{and} \quad |\!|\!| \mathcal{M}_\xi f - \mathcal{M}_\xi g |\!|\!| \le \sqrt{T_0} |\!|\!| f - g |\!|\!| \chi(|\!|\!| f |\!|\!| + |\!|\!| g |\!|\!|)$$

hold. It follows that T_0 can be chosen sufficiently small to make $\mathcal{M}_\xi$ a contraction on some ball of $\bar{\mathscr{B}}_{L,T_0}$, and so the fixed point of $\mathcal{M}_\xi$ takes its values in $\mathcal{D}(L)$.

In order to complete the proof of the theorem, it remains to show that (A.10) holds for arbitrary intial conditions. We again consider the same mapping $\mathcal{M}_\xi$, but this time on the space

$C_b((0,T_0],\mathscr{B})$. It is straightforward to check, using the assumptions, that bounds similar to (A.8), but with $\|\cdot\|_L$ replaced by $\|\cdot\|$ and b, $\tilde{b}$ replaced by a, $\tilde{a}$ hold. We notice that, by (A.8a), we can, for arbitrary $\varepsilon > 0$, choose δ so small that $\|u_\eta(\delta)\| \le (1+\varepsilon)\|\eta\|$. Since $u(\delta) \in \mathcal{D}(L)$, this gives the estimate $\|u_\eta(t)\| \le |(1+\varepsilon)\|\eta\| - a(0)|e^{\kappa(t-\delta)} + a(0)$, holding for every $\varepsilon > 0$. By using (A.8b) and a similar argument, we can show that $\|u_\eta(t) - u_\xi(t)\| \le e^{\kappa(t-\delta)}(1+\varepsilon)\|\eta - \xi\|$ holds and thus (A.10) is true for $\eta, \xi \in \mathscr{B}$. □

Acknowledgements

We would like to thank Marius Mantoiou, Radu Purice, Guillaume van Baalen and Peter Wittwer for helpful discussions. This work was partially supported by the Fonds National Suisse.

III. Uniqueness of the Invariant Measure for a Stochastic PDE Driven by Degenerate Noise

Abstract

We consider the stochastic Ginzburg-Landau equation in a bounded domain. We assume the stochastic forcing acts only on high spatial frequencies. The low-lying frequencies are then only connected to this forcing through the non-linear (cubic) term of the Ginzburg-Landau equation. Under these assumptions, we show that the stochastic PDE has a *unique* invariant measure. The techniques of proof combine a controllability argument for the low-lying frequencies with an infinite dimensional version of the Malliavin calculus to show positivity and regularity of the invariant measure. This then implies the uniqueness of that measure.

1 Introduction

In this paper, we study a stochastic variant of the Ginzburg-Landau equation on a finite domain with periodic boundary conditions. The deterministic equation is

$$\dot{u} = \Delta u + u - u^3 \;, \quad u(0) = u^{(0)} \in \mathcal{H} \;, \tag{1.1}$$

where $\mathcal{H}$ is the real Hilbert space $\mathcal{W}^1_{\text{per}}([-\pi,\pi])$, *i.e.*, the closure of the space of smooth periodic functions $u : [-\pi,\pi] \to \mathbf{R}$ equipped with the norm

$$\|u\|^2 = \int_{-\pi}^{\pi} (|u(x)|^2 + |u'(x)|^2)\, dx \;.$$

(The restriction to the interval $[-\pi,\pi]$ is irrelevant since other lengths of intervals can be obtained by scaling space, time and amplitude u in (1.1).) While we work exclusively with the real Ginzburg-Landau equation (1.1) our methods generalize immediately to the complex Ginzburg-Landau equation

$$\dot{u} = (1 + ia)\Delta u + u - (1 + ib)|u|^2 u \;, \quad a, b \in \mathbf{R} \;, \tag{1.2}$$

which has a more interesting dynamics than (1.1). But the notational details are slightly more involved because of the complex values of u and so we stick with (1.1).

While a lot is known about existence and regularity of solutions of (1.1) or (1.2), only very little information has been obtained about the attractor of such systems, and in particular, nothing seems to be known about invariant measures on the attractor.

On the other hand, when (1.1) is replaced by a stochastic differential equation, more can be said about the invariant measure, see [DPZ96] and references therein. Since the problem (1.1) involves only functions with periodic boundary conditions, it can be rewritten in terms of the Fourier series for u:

$$u(x,t) \;=\; \sum_{k\in\mathbf{Z}} e^{ikx} u_k(t) \;, \quad u_k \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikx} u(x)\, dx \;.$$

We call k the momenta, u_k the modes, and, since $u(x,t)$ is real we must always have $u_k(t) = \bar{u}_{-k}(t)$, where $\bar{z}$ is the complex conjugate of z. With these notations (1.1) takes the form

$$\dot{u}_k \;=\; (1-k^2)u_k - \sum_{k_1+k_2+k_3=k} u_{k_1}u_{k_2}u_{k_3}\;,$$

for all $k \in \mathbf{Z}$ and the initial condition satisfies $\{(1+|k|)u_k(0)\} \in \ell^2$. In the sequel, we will use the symbol $\mathcal{H}$ indifferently for the space $\mathcal{W}^1_{\text{per}}([-\pi,\pi])$ and for its counterpart in Fourier space. In the earlier literature on uniqueness of the invariant measure for stochastic differential equations, see the recent review [MS95a], the authors are mostly interested in systems where each of the u_k is forced by some external noise term. The main aim of our work is to study forcing by noise which *acts only on the high-frequency part* of u, namely on the u_k with $|k| \geq k_*$ for some finite $k_* \in \mathbf{N}$. The low-frequency amplitudes u_k with $|k| < k_*$ are then only *indirectly* forced through the noise, namely through the nonlinear coupling of the modes. In this respect, our approach is reminiscent of the work done on thermally driven chains in [EPR99a, EPR99b, EH00], where the chains were only stochastically driven at the ends.

In the context of our problem, the *existence* of an invariant measure is a classical result for the noise we consider [DPZ96], and the main novelty of our paper is a proof of *uniqueness* of that measure. To prove uniqueness we begin by proving controllability of the equations, *i.e.*, to show that the high-frequency noise together with non-linear coupling effectively drives the low-frequency modes. Using this, we then use Malliavin calculus in infinite dimensions, to show regularity of the transition probabilities. This then implies uniqueness of the invariant measure.

We will study the system of equations

$$du_k = -k^2 u_k\,dt + (u_k - (u^3)_k)\,dt + \frac{q_k}{\sqrt{4\pi(1+k^2)}}\,dw_k(t)\;, \tag{1.3}$$

with $u \in \mathcal{H}$. The above equations hold for $k \in \mathbf{Z}$, and it is always understood that

$$(u^3)_k = \sum_{\substack{k_1+k_2+k_3=k\\ k_1,k_2,k_3\in\mathbf{Z}}} u_{k_1}u_{k_2}u_{k_3}\;, \tag{1.4}$$

with $u_{-k} = \bar{u}_k$. *To avoid inessential notational problems we will work with even periodic functions, so that* $u_k = u_{-k} \in \mathbf{R}$. We will work with the basis

$$e_k(x) \;=\; \frac{1}{\sqrt{\pi(1+k^2)}}\cos(kx)\;. \tag{1.5}$$

Note that this basis is orthonormal w.r.t. the scalar product in $\mathcal{H}$, but the u_k are actually given by $u_k = (4\pi(1+k^2))^{-1/2}\langle u, e_k\rangle$. (We choose this to make the cubic term (1.4) look simple.)

The noise is supposed to act only on the high frequencies, but there we need it to be strong enough in the following way. Let $a_k = k^2+1$. Then we require that there exist constants $c_1, c_2 > 0$ such that for $k \geq k_*$,

$$c_1 a_k^{-\alpha} \leq q_k \leq c_2 a_k^{-\beta}\;, \qquad \alpha \geq 2\;, \quad \alpha - 1/8 < \beta \leq \alpha\;. \tag{1.6}$$

These conditions imply

$$\sum_{k=0}^{\infty}(1+k^{4\alpha-3/2})q_k^2 < \infty\;,$$

$$\sup_{k \geq k_*} k^{-2\alpha} q_k^{-1} < \infty \; .$$

We formulate the problem in a more general setting: Let $F(u)$ be a polynomial of odd degree with negative leading coefficient. Let A be the operator of multiplication by $1 + k^2$ and let Q be the operator of multiplication by q_k. Then (1.3) is of the form

$$d\Phi^t \;=\; -A\Phi^t \, dt + F(\Phi^t) \, dt + Q \, dW(t) \; , \tag{1.7}$$

where $dW(t) = \sum_{k=0}^{\infty} e_k dw_k(t)$ is the cylindrical Wiener process on $\mathcal{H}$ with the w_k mutually independent real Brownian motions.[1] We define $\Phi^t(\xi)$ as the solution of (1.7) with initial condition $\Phi^0(\xi) = \xi$. Clearly, the conditions on Q can be formulated as

$$\|A^{\alpha-3/8} Q\|_{\text{HS}} < \infty \; , \tag{1.8a}$$

$$q_k^{-1} k^{-2\alpha} \text{ is bounded for } k \geq k_* \; , \tag{1.8b}$$

where $\| \cdot \|_{\text{HS}}$ is the Hilbert-Schmidt norm on $\mathcal{H}$. Note that for each k, (1.3) is obtained by multiplying (1.7) by $(4\pi(1 + k^2))^{-1/2} \langle \cdot, e_k \rangle$.

Important Remark. The crucial aspect of our conditions is the possibility of choosing $q_k = 0$ for all $k < k_*$, *i.e.*, the noise drives only the high frequencies. But we also allow any of the q_k with $k < k_*$ to be different from 0, which corresponds to long wavelength forcing. Furthermore, as we are allowing α to be arbitrarily large, this means that the forcing at high frequencies has an amplitude which can decay like any power. The point of this paper is to show that these conditions are sufficient to ensure the existence of a unique invariant measure for (1.7).

Theorem 1.1 *The process (1.7) has a unique invariant Borel measure on* $\mathcal{H}$.

There are two main steps in the proof of Theorem 1.1. First, the nature of the nonlinearity F implies that the modes with $k \geq k_*$ couple in such a way to those with $k < k_*$ as to allow *controllability*. Intuitively, this means that any point in phase space can be reached to arbitrary precision in any given time, by a suitable choice of the high-frequency controls.

Second, verifying a Hörmander-like condition, we show that a version of the Malliavin calculus can be implemented in our infinite-dimensional context. This will be the hard part of our study, and the main result of that part is a proof that the strong Feller property holds. This means that for any measurable function $\varphi \in \mathcal{B}_b(\mathcal{H})$, the function

$$(\mathcal{P}^t \varphi)(\xi) \;\equiv\; \mathbf{E}\Big((\varphi \circ \Phi^t)(\xi)\Big) \tag{1.9}$$

is *continuous*.[2] We show this by proving that a cutoff version of (1.7) (modifying the dynamics at large amplitudes by a parameter ϱ) makes $\mathcal{P}_\varrho^t \varphi$ a *differentiable* map.

The interest in such highly degenerate stochastic PDE's is related to questions in hydrodynamics where one would ask how "energy" is transferred from high to low frequency modes,

[1]It is convenient to have, in the case of (1.3), $A = 1 - \Delta$ and $F(u) = 2u - u^3$ rather than $A = -1 - \Delta$ and $F(u) = -u^3$.

[2]Throughout the paper, **E** denotes expectation and **P** denotes probability for the random variables.

and vice versa when only some of the modes are driven. This could then shed some light on the entropy-enstrophy problem in the (driven) Navier-Stokes equation.

To end this introduction, we will try to compare the results of our paper to current work of others. These groups consider the 2-D Navier Stokes equation without deterministic external forces, also in bounded domains. In these equations, any initial condition eventually converges to zero, as long as there is no stochastic forcing. First there is earlier work by Flandoli-Maslowski [FM95] dealing with noise whose amplitude is bounded below by $|k|^{-c}$. In the work of Bricmont, Kupiainen and Lefevere [BKL00c, BKL00a], the stochastic forcing acts on modes with low k, and they get uniqueness of the invariant measure and analyticity, with probability 1. Furthermore, they obtain exponential convergence to the stationary measure. In the work of Kuksin and Shirikyan [KS00] the bounded noise is quite general, acts on low-lying Fourier modes, and acts at definite times with "noise-less" intervals in-between. Again, the invariant measure is unique. It is supported by $\mathcal{C}^\infty$ functions, is mixing and has a Gibbs property. In the work of [EMS01], a result similar to [BKL00a] is shown.

The main difference between those results and the present paper is our control of a situation which is already unstable at the deterministic level. Thus, in this sense, it comes closer to a description of a deterministically turbulent fluid (*e.g.*, obtained by an external force). On the other hand, in our work, we need to actually force all high spatial frequencies. Perhaps, this could be eliminated by a combination with ideas from the papers above.

2 Some Preliminaries on the Dynamics

Here, we summarize some facts about deterministic and stochastic GL equations from the literature which we need to get started.

We will consider the dynamics on the following space:

Definition 2.1 We define $\mathcal{H}$ as the subspace of even functions in $\mathcal{W}^1_{\text{per}}([-\pi,\pi])$. The norm on $\mathcal{H}$ will be denoted by $\|\cdot\|$, and the scalar product by $\langle\cdot,\cdot\rangle$.

We consider first the deterministic equation

$$\dot{u} = \Delta u + u - u^3 , \quad u(0) = u^{(0)} \in \mathcal{H} , \tag{2.1}$$

Due to its dissipative character the solutions are, for positive times, analytic in a strip around the real axis. More precisely, denote by $\|\cdot\|_{\mathscr{A}_\eta}$ the norm

$$\|f\|_{\mathscr{A}_\eta} = \sup_{|\mathrm{Im}z|\le\eta} |f(z)| ,$$

and by $\mathscr{A}_\eta$ the corresponding Banach space of analytic functions. Then the following result holds.

Lemma 2.2 *For every initial value $u^{(0)} \in \mathcal{H}$, there exist a time T and a constant C such that for $0 < t \le T$, the solution $u(t,u^{(0)})$ of (2.1) belongs to $\mathscr{A}_{\sqrt{t}}$ and satisfies $\|u(t,u^{(0)})\|_{\mathscr{A}_{\sqrt{t}}} \le C$.*

Proof. The statement is proven in [Col94] for the case of the infinite line. Since the periodic functions form an invariant subspace under the evolution, the result applies to our case. □

We next collect some useful results for the stochastic equation (1.7):

Proposition 2.3 *For every $t > 0$ and every $p \geq 1$ the solution of (1.7) with initial condition $\Phi^0(\xi) = \xi \in \mathcal{H}$ exists in $\mathcal{H}$ up to time t. It defines by (1.9) a Markovian transition semigroup on $\mathcal{H}$. One has the bound*

$$\mathbf{E}\Big(\sup_{s\in[0,t]} \|\Phi^s(\xi)\|^p\Big) \leq C_{t,p}(1+\|\xi\|)^p \;.$$

Furthermore, the process (1.7) has an invariant measure.

These results are well-known and in Section 8.6 we sketch where to find them in the literature.

3 Controllability

In this section we show the "approximate controllability" of (1.3). The control problem under consideration is

$$\dot u = \Delta u + u - u^3 + Q\, f(t)\;, \qquad u(0) = u^{(\mathrm{i})} \in \mathcal{H}\;, \tag{3.1}$$

where f is the control. Using Fourier series' and the hypotheses on Q, we see that by choosing $f_k \equiv 0$ for $|k| < k_*$, (3.1) can be brought to the form

$$\dot u_k = \begin{cases} -k^2 u_k + u_k - \sum\limits_{\ell+m+n=k} u_\ell u_m u_n + \dfrac{q_k}{\sqrt{4\pi(1+k^2)}} f_k(t)\;, & |k| \geq k_*\;, \\ -k^2 u_k + u_k - \sum\limits_{\ell+m+n=k} u_\ell u_m u_n\;, & |k| < k_*\;, \end{cases} \tag{3.2}$$

with $\{u_k\} \in \mathcal{H}$ and $t \mapsto \{f_k(t)\} \in \mathrm{L}^\infty([0,\tau],\mathcal{H})$. We will refer in the sequel to $\{u_k\}_{|k|<k_*}$ as the *low-frequency modes* and to $\{u_k\}_{|k|\geq k_*}$ as the *high-frequency modes*. We also introduce the projectors Π_{L} and Π_{H} which project onto the low (resp. high) frequency modes. Let $\mathcal{H}_{\mathrm{L}}$ and $\mathcal{H}_{\mathrm{H}}$ denote the ranges of Π_{L} and Π_{H} respectively. Clearly $\mathcal{H}_{\mathrm{L}}$ is finite dimensional, whereas $\mathcal{H}_{\mathrm{H}}$ is a separable Hilbert space.

The main result of this section is approximate controllability in the following sense:

Theorem 3.1 *For every time $\tau > 0$ the following is true: For every $u^{(\mathrm{i})}, u^{(\mathrm{f})} \in \mathcal{H}$ and every $\varepsilon > 0$, there exists a control $f \in \mathrm{L}^\infty([0,\tau],\mathcal{H})$ such that the solution $u(t)$ of (3.1) with $u(0) = u^{(\mathrm{i})}$ satisfies $\|u(\tau) - u^{(\mathrm{f})}\| \leq \varepsilon$.*

Proof. The construction of the control proceeds in 4 different phases, of which the third is the actual controlling of the low-frequency part by the high-frequency controls. In the construction, we will encounter a time $\tau(R,\varepsilon')$ which depends on the norm R of $u^{(\mathrm{f})}$ and some precision ε'. Given this function, we split the given time τ as $\tau = \sum_{i=1}^4 \tau_i$, with $\tau_4 \leq \tau(\|u^{(\mathrm{f})}\|, \varepsilon/2)$ and all $\tau_i > 0$. We will use the cumulated times $t_j = \sum_{i=1}^j \tau_i$.

Step 1. In this step we choose $f \equiv 0$, and we define $u^{(1)} = u(t_1)$, where $t \mapsto u(t)$ is the solution of (3.1) with initial condition $u(0) = u^{(\mathrm{i})}$. Since there is no control, we really have (2.1) and hence, by Lemma 2.2, we see that $u^{(1)} \in \mathscr{A}_\eta$ for some $\eta > 0$.

Step 2. We will construct a smooth control $f : [t_1, t_2] \to \mathcal{H}$ such that $u^{(2)} = u(t_2)$ satisfies:

$$\Pi_{\mathrm{H}} u^{(2)} = 0 .$$

In other words, in this step, we drive the high-frequency part to 0. To construct f, we choose a $\mathcal{C}^\infty$ function $\varphi : [t_1, t_2] \to \mathbf{R}$, interpolating between 1 and 0 with vanishing derivatives at the ends. Define $u_{\mathrm{H}}(t) = \varphi(t)\Pi_{\mathrm{H}} u^{(1)}$ for $t \in [t_1, t_2]$. This will be the evolution of the high-frequency part. We next define the low-frequency part $u_{\mathrm{L}} = u_{\mathrm{L}}(t)$ as the solution of the ordinary differential equation

$$\dot u_{\mathrm{L}} = \Delta u_{\mathrm{L}} + u_{\mathrm{L}} - \Pi_{\mathrm{L}}\big((u_{\mathrm{L}} + u_{\mathrm{H}})^3\big) ,$$

with $u_{\mathrm{L}}(t_1) = \Pi_{\mathrm{L}} u^{(1)}$. We then set $u(t) = u_{\mathrm{L}}(t) \oplus u_{\mathrm{H}}(t)$ and substitute into (3.1) which we need to solve for the control $Qf(t)$ for $t \in [t_1, t_2]$.

Since $u_{\mathrm{L}}(t) \oplus u_{\mathrm{H}}(t)$ as constructed above is in $\mathscr{A}_\eta$ and since $Qf = \dot u - \Delta u - u + u^3$, and Δ maps $\mathscr{A}_\eta$ to $\mathscr{A}_{\eta/2}$ we conclude that $Qf \in \mathscr{A}_{\eta/2}$. By construction, the components q_k of Q decay polynomially with k and do not vanish for $k \ge k_*$. Therefore, Q^{-1} is a bounded operator from $\mathscr{A}_{\eta/2} \cap \mathcal{H}_{\mathrm{H}}$ to $\mathcal{H}_{\mathrm{H}}$. Thus, we can solve for f in this step.

Step 3. As mentioned before, this step really exploits the coupling between high and low frequencies. Here, we start from $u^{(2)}$ at time t_2 and we want to reach $\Pi_{\mathrm{L}} u^{(\mathrm{f})}$ at time t_3. In fact, we will instead reach a point $u^{(3)}$ with $\|\Pi_{\mathrm{L}} u^{(3)} - \Pi_{\mathrm{L}} u^{(\mathrm{f})}\| < \varepsilon/2$.

The idea is to choose for every low frequency $|k| < k_*$ a set of three[3] high frequencies that will be used to control u_k. To simplify matters we will assume (without loss of generality) that $k_* > 2$:

Definition 3.2 We define for every k with $0 \le k < k_*$ the set $\mathscr{I}_k$ by

$$\mathscr{I}_k = \{10^{k_*+k} + k\ ,\ 2 \cdot 10^{k_*+k}, 3 \cdot 10^{k_*+k}\} .$$

We also define $\mathscr{I}_{\mathrm{L}}^0 = \{k\ :\ 0 \le k < k_*\}$ and

$$\mathscr{I} = \mathscr{I}_{\mathrm{L}}^0 \cup \Big(\bigcup_{0 \le k < k_*} \mathscr{I}_k \Big) .$$

Lemma 3.3 *The sets defined above have the following properties:*

(A) *Let $\mathscr{I}_k = \{k_1, k_2, k_3\}$. Then, of the six sums $\pm k_1 \pm k_2 \pm k_3$ exactly one equals k and one equals $-k$. All others have modulus larger than k_*.*

(B) *The sets $\mathscr{I}_k$ and $\mathscr{I}_{\mathrm{L}}^0$ are all mutually disjoint.*

(C) *Let S be a collection of three indices in $\mathscr{I}$, $S = \{k_1, k_2, k_3\}$. If any of the sums $\pm k_1 \pm k_2 \pm k_3$ adds up to k with $|k| < k_*$ then either $S = \mathscr{I}_k$ or $S \subset \mathscr{I}_{\mathrm{L}}^0$ or S is of the form $S = \{k, k', k'\}$.*

Remark 3.4 At the end of this section, we indicate how this construction generalizes to the complex Ginzburg-Landau equation.

[3] The number 3 is the highest power of the nonlinearity F in the GL equation.

Proof. The claims (A) and (B) are obvious from the definition of $\mathscr{I}_k$. To prove (C) let $S = \{k_1, k_2, k_3\}$. If $S \subset \mathscr{I}_{\rm L}^0$, we are done. Otherwise, at least one of the k_i is an element of an $\mathscr{I}_\ell$ for some $\ell = 0, \dots, k_* - 1$. Clearly, if the two others are in $\mathscr{I}_{\rm L}^0$, none of the sums have modulus less than k_*. If a second k_j is in $\mathscr{I}_{\ell'}$ with $\ell' \neq \ell$ then again none of the 6 sums can lead to a modulus less than k_*. Finally if k_j is in $\mathscr{I}_\ell$ then either all 3 are in $\mathscr{I}_\ell$ and we are done, or $k_i = k_j$ and thus $S = \{k, k', k'\}$. We have covered all cases and the proof of the lemma is complete. □

We are going to construct a control which, in addition to driving the low frequency part as indicated, also implies $u_k(t) \equiv 0$ for $k \notin \mathscr{I}$ for $t \in [t_2, t_3]$. By the conditions on $\mathscr{I}$, the low-frequency part of (3.2) is for $0 < k < k_*$ equal to (having chosen the controls equal to 0 for $k < k_*$):

$$\dot u_k = \Big(1 - k^2 - 6 \sum_{n \in \mathscr{I} \setminus \mathscr{I}_{\rm L}^0} |u_n|^2\Big) u_k - \sum_{\substack{\pm\ell\pm m\pm n = k \\ \{\ell,m,n\} \subset \mathscr{I}_{\rm L}^0}} u_\ell u_m u_n - 6 \prod_{n \in \mathscr{I}_k} u_n \;. \tag{3.3}$$

When $k = 0$, the last term in (3.3) is replaced by $-12 \prod_{n \in \mathscr{I}_0} u_n$. This identity exploits the relations $u_{-n} = u_n$. To simplify the combinatorial problem, we choose the controls of the 3 amplitudes u_n with $n \in \mathscr{I}_k$ in such a way that these u_n are all equal to a fixed real function $z_k(t)$ which we will determine below. With this particular choice, (3.3) reduces for $0 < k < k_*$ to

$$0 = -\dot u_k + \Big(1 - k^2 - 18 \sum_{0 \le n < k_*} |z_n|^2\Big) u_k - \big((\Pi_{\rm L} u)^3\big)_k - 6 z_k^3 \;. \tag{3.4}$$

For $k = 0$ the last term is $-12 z_0^3$. We claim that for every path $\gamma \in \mathcal{C}^\infty([t_2, t_3]; \mathcal{H}_{\rm L})$ and every $\varepsilon > 0$, *we can find a set of bounded functions* $t \mapsto z_k(t)$ *such that the solution of (3.4) shadows* γ *at a distance at most* ε.

To prove this statement, consider the map $F : \mathbf{R}^{k_*} \to \mathbf{R}^{k_*}$ of the form (obtained when substituting the path γ into (3.4))

$$F : \begin{pmatrix} z_0 \\ z_1 \\ \vdots \\ z_{k_*-1} \end{pmatrix} \mapsto \begin{pmatrix} F_0(z) \\ F_1(z) \\ \vdots \\ F_{k_*-1}(z) \end{pmatrix} = \begin{pmatrix} 2z_0^3 \\ z_1^3 \\ \vdots \\ z_{k_*-1}^3 \end{pmatrix} + \begin{pmatrix} \mathcal{P}_0(z) \\ \mathcal{P}_1(z) \\ \vdots \\ \mathcal{P}_{k_*-1}(z) \end{pmatrix} ,$$

where the $\mathcal{P}_m$ are polynomials of degree at most 2. We want to find a solution to $F = 0$. The F_m form a Gröbner basis for the ideal of the ring of polynomials they generate. As an immediate consequence, the equation $F(z) = 0$ possesses exactly 3^{k_*} complex solutions, if they are counted with multiplicities [MS95b]. Since the coefficients of the $\mathcal{P}_m$ are real this implies that there exists at least one real solution.

Having found a (possibly discontinuous) solution for the z_k, we find nearby smooth functions $\tilde z_k$ with the following properties:

– The equation (3.4) with $\tilde z_k$ replacing z_k and initial condition $u_k(t_2) = u_k^{(2)}$ leads to a solution u with $\|u(t_3) - \Pi_{\rm L} u^{(\rm f)}\| \le \varepsilon/2$.

– One has $\tilde z_k(t_3) = 0$.

Having found the $\tilde{z}_k$ we construct the f_k in such a way that for $n \in \mathscr{I}_k$ one has $u_n(t) = \tilde{z}_k(t)$. Finally, for $k \notin \mathscr{I}$ we choose the controls in such a way that $u_k(t) \equiv 0$ for $t \in [t_2, t_3]$. We define $u^{(3)}$ as the solution obtained in this way for $t = t_3$.

Step 4. Starting from $u^{(3)}$ we want to reach $u^{(\mathrm{f})}$. Note that $u^{(3)}$ is in $\mathscr{A}_\eta$ (for every $\eta > 0$) since it has only a finite number of non-vanishing modes. By construction we also have $\|\Pi_{\mathrm{L}} u^{(3)} - \Pi_{\mathrm{L}} u^{(\mathrm{f})}\| \le \varepsilon/2$. We only need to adapt the high frequency part without moving the low-frequency part too much.

Since $\mathscr{A}_\eta$ is dense in $\mathcal{H}$, there is a $u^{(4)} \in \mathscr{A}_\eta$ with $\|u^{(4)} - u^{(\mathrm{f})}\| \le \varepsilon/4$. By the reasoning of Step 2 there is for every $\tau' > 0$ a control for which $\Pi_{\mathrm{H}} u(t_3 + \tau') = \Pi_{\mathrm{H}} u^{(4)}$ when starting from $u(t_3) = u^{(3)}$. Given ε there is a τ_* such that if $\tau' < \tau_*$ then $\|\Pi_{\mathrm{L}} u(t_3 + \tau') - \Pi_{\mathrm{L}} u(t_3)\| < \varepsilon/4$. *This τ_* depends only on $\|u^{(\mathrm{f})}\|$ and ε,* as can be seen from the following argument: Since $\Pi_{\mathrm{H}} u^{(3)} = 0$, we can choose the controls in such a way that $\|\Pi_{\mathrm{H}} u(t_3 + t)\|$ is an increasing function of t and is therefore bounded by $\|\Pi_{\mathrm{H}} u^{(\mathrm{f})}\|$. The equation for the low-frequency part is then a finite dimensional ODE in which all high-frequency contributions can be bounded in terms of $R = \|u^{(\mathrm{f})}\|$.

Combining the estimates we see that

$$
\begin{aligned}
\|u(t_4) - u^{(\mathrm{f})}\| &= \|\Pi_{\mathrm{L}}(u(t_4) - u^{(\mathrm{f})})\| + \|\Pi_{\mathrm{H}}(u(t_4) - u^{(\mathrm{f})})\| \\
&\le \|\Pi_{\mathrm{L}}(u(t_4) - u(t_3))\| + \|\Pi_{\mathrm{L}}(u(t_3) - u^{(\mathrm{f})})\| \\
&\quad + \|\Pi_{\mathrm{H}}(u^{(4)} - u^{(\mathrm{f})})\| \le \varepsilon \,.
\end{aligned}
$$

The proof of Theorem 3.1 is complete. □

3.1 The combinatorics for the complex Ginzburg-Landau equation

We sketch here those aspects of the combinatorics which change for the complex Ginzburg-Landau equation. In this case, both the real and the imaginary parts of u_n and u_{-n} are independent. Thus, we would need a noise which acts on each of the real and imaginary components of u_n and of u_{-n} independently *i.e.*, *four* components per $n > 0$ and *two* for $n = 0$. A possible definition of $\mathscr{I}_k$ for $|k| < k_*$ is:

$$
\mathscr{I}_k = \begin{cases} \{10^{k_*+2k} + k, 2 \cdot 10^{k_*+2k}, -3 \cdot 10^{k_*+2k}\} & \text{for } k \ge 0, \\ \{10^{k_*+2|k|+1} - |k|, 2 \cdot 10^{k_*+2|k|+1}, -3 \cdot 10^{k_*+2|k|+1}\} & \text{for } k < 0. \end{cases}
$$

We also define $\mathscr{I}_{\mathrm{L}}^0 = \{k \;:\; |k| < k_*\}$ and

$$
\mathscr{I} = \mathscr{I}_{\mathrm{L}}^0 \cup \Big(\bigcup_{|k|<k_*} \mathscr{I}_k \Big) \,.
$$

The analog of Lemma 3.3 is

Lemma 3.5 *The sets defined above have the following properties:*

(A) *Let $\mathscr{I}_k = \{k_1, k_2, k_3\}$. Then, the sum $k_1 + k_2 + k_3$ equals k.*

(B) *The sets $\mathscr{I}_k$ and $\mathscr{I}_{\mathrm{L}}^0$ are all mutually disjoint.*

(C) *Let S be a collection of three indices in $\mathscr{I}$, $S = \{k_1, k_2, k_3\}$. If the sum $k_1+k_2+k_3$ equals k with $|k| < k_*$ then either $S = \mathscr{I}_k$ or $S \subset \mathscr{I}_{\rm L}^0$ or S is of the form $S = \{k, k', -k'\}$.*

Finally, the analog of (3.4) is for $|k| < k_*$:

$$0 = -\dot{u}_k + (1 - (1 + ia)k^2)u_k - (1 + ib)\Big((\Pi_{\rm L} u\, |\Pi_{\rm L} u|^2)_k + 6z_k^3\Big) \ .$$

Apart from these combinatorial changes the complex Ginzburg-Landau equation is treated like the real one.

4 Strong Feller Property and Proof of Theorem 1.1

The aim of this section is to show the strong Feller property of the process defined by (1.3) resp. (1.7).

Theorem 4.1 *The Markov semigroup $\mathcal{P}^t$ defined in (1.9) is strong Feller.*

Proof of Theorem 1.1. This proof follows a well-known strategy, see *e.g.*, [DPZ96]. First of all, there is at least one invariant measure for the process (1.7), since for a problem in a finite domain, the semigroup $t \mapsto e^{-At}$ is compact, and therefore [DPZ96, Theorem 6.3.5] applies.

By the controllability Theorem 3.1, we deduce, see [DPZ96, Theorem 7.4.1], that the transition probability from any point in $\mathcal{H}$ to any open set in $\mathcal{H}$ cannot vanish, *i.e.*, the Markov process is irreducible. Furthermore, by Theorem 4.1 the process is strong Feller. By a classical result of Khas'minskiĭ, this implies that $\mathcal{P}^t$ is regular. Therefore we can use Doob's theorem [DPZ96, pp.42–43] to conclude that the invariant measure is unique. This completes the proof of Theorem 1.1. □

Before we start with the proof of Theorem 4.1, we explain our strategy. Because of the polynomial nature of the nonlinearity in (1.3), the natural bounds diverge with some power of the norm of the initial data. On the other hand, the nonlinearity is strongly dissipative at large amplitudes. Therefore we introduce a cutoff version of the dynamics beyond some fixed amplitude and then take the limit in which this cutoff goes to infinity. We seem to need such a technique to get the bounds (5.11) and (5.12).

The precise definition of the cutoff version F_ϱ of F is:

$$F_\varrho(x) \;=\; (1 - \chi(\|x\|/(3\varrho)))F(x) \ ,$$

where χ is a smooth, non-negative function satisfying

$$\chi(z) \;=\; \begin{cases} 1 & \text{if } z > 2, \\ 0 & \text{if } z < 1. \end{cases}$$

Similarly, we define

$$Q_\varrho(x) \;=\; Q + \chi(\|x\|/\varrho)\Pi_{k_*} \ , \tag{4.1}$$

where Π_{k_*} is the projection onto the frequencies below k_*.

Remark 4.2 These cutoffs have the following effect as a function of $\|x\|$:

- When $\|x\| \le \varrho$ then $Q_\varrho(x) = Q$ and $F_\varrho(x) = F(x)$.
- When $\varrho < \|x\| \le 2\varrho$ then $Q_\varrho(x)$ depends on x and $F_\varrho(x) = F(x)$.
- When $2\varrho < \|x\| \le 6\varrho$ then all Fourier components of $Q_\varrho(x)$ *including the ones below* k_* are non-zero and $F_\varrho(x)$ is proportional to a $F(x)$ times a factor ≤ 1.
- When $6\varrho < \|x\|$ then all Fourier components of $Q_\varrho(x)$ *including the ones below* k_* are non-zero and $F_\varrho(x) = 0$.

At high amplitudes, the nonlinearity is truncated to 0. Thus, the Hörmander condition cannot be satisfied there unless the diffusion process is non-degenerate. We achieve this non-degeneracy by extending the stochastic forcing to *all* degrees of freedom when $\|x\|$ is large.

Instead of (1.7) we then consider the modified problem

$$d\Phi^t_\varrho = -A\Phi^t_\varrho\, dt + (F_\varrho \circ \Phi^t_\varrho)\, dt + (Q_\varrho \circ \Phi^t_\varrho)\, dW(t)\ , \tag{4.2}$$

with $\Phi^0_\varrho(\xi) = \xi \in \mathcal{H}$. Note that the cutoffs are chosen in such a way that the dynamics of $\Phi^t_\varrho(\xi)$ *coincides* with that of $\Phi^t(\xi)$ as long as $\|\Phi^t(\xi)\| < \varrho$. We will show that the solution of (4.2) defines a Markov semigroup

$$\mathcal{P}^t_\varrho\varphi(\xi) = \mathbf{E}(\varphi \circ \Phi^t_\varrho)(\xi)\ ,$$

with the following smoothing property:

Theorem 4.3 *There exist exponents $\mu, \nu > 0$, and for all $\varrho > 0$ there is a constant C_ϱ such that for every $\varphi \in \mathcal{B}_b(\mathcal{H})$, for every $t > 0$ and for every $\xi \in \mathcal{H}$, the function $\mathcal{P}^t_\varrho\varphi$ is differentiable and its derivative satisfies*

$$\|D\mathcal{P}^t_\varrho\varphi(\xi)\| \le C_\varrho(1 + t^{-\mu})(1 + \|\xi\|^\nu)\|\varphi\|_{\mathrm{L}^\infty}\ . \tag{4.3}$$

Using this theorem, the proof of Theorem 4.1 follows from a limiting argument.

Proof of Theorem 4.1. Choose $x \in \mathcal{H}$, $t > 0$, and $\varepsilon > 0$. We denote by $\mathcal{B}$ the ball of radius $2\|x\|$ centered around the origin in $\mathcal{H}$. Using Proposition 2.3 we can find a sufficiently large constant $\varrho = \varrho(x, t, \varepsilon)$ such that for every $y \in \mathcal{B}$, the inequality

$$\mathbf{P}\Big(\sup_{s\in[0,t]} \|\Phi^s(y)\| > \varrho\Big) \le \frac{\varepsilon}{8}$$

holds. Choose $\varphi \in \mathcal{B}_b(\mathcal{H})$ with $\|\varphi\|_{\mathrm{L}^\infty} \le 1$. We have by the triangle inequality

$$\begin{aligned}|\mathcal{P}^t\varphi(x) - \mathcal{P}^t\varphi(y)| \le |\mathcal{P}^t\varphi(x) - \mathcal{P}^t_\varrho\varphi(x)| + |\mathcal{P}^t_\varrho\varphi(x) - \mathcal{P}^t_\varrho\varphi(y)|\\ + |\mathcal{P}^t\varphi(y) - \mathcal{P}^t_\varrho\varphi(y)|\ .\end{aligned}$$

Since the dynamics of the cutoff equation and the dynamics of the original equation coincide on the ball of radius ϱ, we can write, for every $z \in \mathcal{B}$,

$$|\mathcal{P}^t\varphi(z) - \mathcal{P}^t_\varrho\varphi(z)| = \mathbf{E}|(\varphi \circ \Phi^t)(z) - (\varphi \circ \Phi^t_\varrho)(z)|$$

$$\leq 2\|\varphi\|_{L^\infty}\ \mathbf{P}\Big(\sup_{s\in[0,t]}\|\Phi^s(z)\| > \varrho\Big) \leq \frac{\varepsilon}{4}\ .$$

This implies that

$$|\mathcal{P}^t\varphi(x) - \mathcal{P}^t\varphi(y)| \leq \frac{\varepsilon}{2} + |\mathcal{P}^t_\varrho\varphi(x) - \mathcal{P}^t_\varrho\varphi(y)|\ .$$

By Theorem 4.3 we see that if y is sufficiently close to x then

$$|\mathcal{P}^t_\varrho\varphi(x) - \mathcal{P}^t_\varrho\varphi(y)| \leq \frac{\varepsilon}{2}\ .$$

Since ε is arbitrary we conclude that $\mathcal{P}^t\varphi$ is continuous when $\|\varphi\|_{L^\infty} \leq 1$. The generalization to any value of $\|\varphi\|_{L^\infty}$ follows by linearity in φ. The proof of Theorem 4.1 is complete. □

5 Regularity of the Cutoff Process

In this section, we start the proof of Theorem 4.3. If the cutoff problem were finite dimensional, a result like Theorem 4.3 could be derived easily using, *e.g.*, the works of Hörmander [Hör67, Hör85], Malliavin [Mal78], Stroock [Str86], or Norris [Nor86]. In the present infinite-dimensional context we need to modify the corresponding techniques, but the general idea retained is Norris'. The main idea will be to treat the (infinite number of) high-frequency modes by a method which is an extension of [DPZ96, Cer99], while the low-frequency part is handled by a variant of the Malliavin calculus adapted from [Nor86]. It is at the juncture of these two techniques that we need a cutoff in the nonlinearity.

5.1 Splitting and interpolation spaces

Throughout the remainder of this paper, we will again denote by $\mathcal{H}_L$ and $\mathcal{H}_H$ the spaces corresponding to the low (resp. high)-frequency parts. We slightly change the meaning of "low-frequency" by including in the low-frequency part all those frequencies that are driven by the noise which are in $\mathscr{I}$ as defined in Definition 3.2. More precisely, the low-frequency part is now $\{k\ :\ |k| \leq L-1\}$, where $L = \max\{k\ :\ k \in \mathscr{I}\} + 1$. Note that L is *finite*.

Since $A = 1 - \Delta$ is diagonal with respect to this splitting, we can define its low (resp. high)-frequency parts A_L and A_H as operators on $\mathcal{H}_L$ and $\mathcal{H}_H$. From now on, L will always denote the dimension of $\mathcal{H}_L$, which will therefore be identified with $\mathbf{R}^L$.[4] We also allow ourselves to switch freely between equivalent norms on $\mathbf{R}^L$, when deriving the various bounds.

In the sequel, we will always use the notations D_L and D_H to denote the derivatives with respect to $\mathcal{H}_L$ (resp. $\mathcal{H}_H$) of a differentiable function defined on $\mathcal{H}$. The words "derivative" and "differentiable" will always be understood in the strong sense, *i.e.*, if $f : \mathscr{B}_1 \to \mathscr{B}_2$ with $\mathscr{B}_1$ and $\mathscr{B}_2$ some Banach spaces, then $Df : \mathscr{B}_1 \to \mathscr{L}(\mathscr{B}_1, \mathscr{B}_2)$, *i.e.*, it is bounded from $\mathscr{B}_1$ to $\mathscr{B}_2$.

We introduce the interpolation spaces $\mathcal{H}^\gamma$ (for every $\gamma \geq 0$) defined as being equal to the domain of A^γ equipped with the graph norm

$$\|x\|^2_\gamma = \|A^\gamma x\|^2 = \|(1-\Delta)^\gamma x\|^2\ .$$

[4]The choice of L above is dictated by the desire to obtain a dimension equal to L and not $L+1$.

Clearly, the $\mathcal{H}^\gamma$ are Hilbert spaces and we have the inclusions

$$\mathcal{H}^\gamma \subset \mathcal{H}^\delta \quad \text{if} \quad \gamma \geq \delta \;.$$

Note that in usual conventions, $\mathcal{H}^\gamma$ would be the Sobolev space of index $2\gamma + 1$. Our motivation for using non-standard notation comes from the fact that our basic space is that with *one* derivative, which we call $\mathcal{H}$, and that γ measures additional smoothness in terms of powers of the generator of the linear part.

5.2 Proof of Theorem 4.3

The proof of Theorem 4.3 is based on Proposition 5.1 and Proposition 5.2 which we now state.

Proposition 5.1 *Assume that the noise satisfies condition (1.6). Then (4.2) defines a stochastic flow Φ^t_ϱ on $\mathcal{H}$ with the following properties which hold for any $p \geq 1$:*

(A) *If $\xi \in \mathcal{H}^\gamma$ with some γ satisfying $0 \leq \gamma \leq \alpha$, the solution of (4.2) stays in $\mathcal{H}^\gamma$, with a bound*

$$\mathbf{E}\Big(\sup_{0<t<T} \|\Phi^t_\varrho(\xi)\|^p_\gamma\Big) \leq C_{T,p,\varrho}(1 + \|\xi\|_\gamma)^p \;, \tag{5.1a}$$

for every $T > 0$. If $\gamma \geq 1$ the solution exists in the strong sense in $\mathcal{H}$.

(B) *The quantity $\Phi^t_\varrho(\xi)$ is in $\mathcal{H}^\alpha$ with probability 1 for every time $t > 0$ and every $\xi \in \mathcal{H}$. Furthermore, for every $T > 0$ there is a constant $C_{T,p,\varrho}$ for which*

$$\mathbf{E}\Big(\sup_{0<t<T} t^{\alpha p}\|\Phi^t_\varrho(\xi)\|^p_\alpha\Big) \leq C_{T,p,\varrho}(1 + \|\xi\|)^p \;. \tag{5.1b}$$

(C) *The mapping $\xi \mapsto \Phi^t_\varrho(\xi)$ (for ω and t fixed) has a.s. bounded partial derivatives with respect to ξ. Furthermore, we have for every $\xi, h \in \mathcal{H}$ the bound*

$$\mathbf{E}\Big(\sup_{0<t<T} \|(D\Phi^t_\varrho(\xi))h\|^p\Big) \leq C_{T,p,\varrho}\|h\|^p \;, \tag{5.1c}$$

for every $T > 0$.

(D) *For every $h \in \mathcal{H}$ and $\xi \in \mathcal{H}^\alpha$, the quantity $(D\Phi^t_\varrho(\xi))h$ is in $\mathcal{H}^\alpha$ with probability 1 for every $t > 0$. Furthermore, for a ν depending only on α the bound*

$$\mathbf{E}\Big(\sup_{0<t<T} t^{\alpha p}\|(D\Phi^t_\varrho(\xi))h\|^p_\alpha\Big) \leq C_{T,p,\varrho}(1 + \|\xi\|_\alpha)^{\nu p}\|h\|^p \;, \tag{5.1d}$$

holds for every $T > 0$.

(E) *For every $\xi \in \mathcal{H}^\gamma$ with $\gamma \leq \alpha$, we have the small-time estimate*

$$\mathbf{E}\Big(\sup_{0<t<\varepsilon} \|\Phi^t_\varrho(\xi) - e^{-At}\xi\|^p_\gamma\Big) \leq C_{T,p,\varrho}(1 + \|\xi\|_\gamma)^p\varepsilon^{p/16} \;, \tag{5.1e}$$

which holds for every $\varepsilon \in (0, T]$ and every $T > 0$.

This proposition will be proved in Section 8.4.

Proposition 5.2 *There exist exponents $\mu_*, \nu_* > 0$ such that for every $\varphi \in \mathcal{C}_b^2(\mathcal{H})$, every $\xi \in \mathcal{H}^\alpha$ and every $t > 0$,*

$$\|D\mathcal{P}_\varrho^t \varphi(\xi)\| \leq C_\varrho(1 + t^{-\mu_*})(1 + \|\xi\|_\alpha^{\nu_*})\|\varphi\|_{\mathrm{L}^\infty} . \tag{5.2}$$

Proof of Theorem 4.3. Note first that for all $\tau > 0$, one has $\|\mathcal{P}_\varrho^\tau \varphi\|_{\mathrm{L}^\infty} \leq \|\varphi\|_{\mathrm{L}^\infty}$. Furthermore, for $\tau > 1$,

$$\|D\mathcal{P}_\varrho^\tau \varphi(\xi)\| = \|D(\mathcal{P}_\varrho^1(\mathcal{P}_\varrho^{\tau-1}\varphi))(\xi)\| .$$

Therefore, if we can show (4.3) for $t \leq 1$, then we find for any $\tau > 1$:

$$\|D\mathcal{P}_\varrho^\tau \varphi(\xi)\| \leq 2C_\varrho(1 + \|\xi\|^\nu)\|\mathcal{P}_\varrho^{\tau-1}\varphi\|_{\mathrm{L}^\infty} \leq 2C_\varrho(1 + \|\xi\|^\nu)\|\varphi\|_{\mathrm{L}^\infty} .$$

In view of the above, it clearly suffices to show Theorem 4.3 for $t \in (0, 1]$.

We first prove the bound for the case $\varphi \in \mathcal{C}_b^2(\mathcal{H})$. Let $h \in \mathcal{H}$. Using the definition (1.9) of $\mathcal{P}_\varrho^t \varphi$ and the Markov property of the flow we write

$$\begin{aligned} \|D\mathcal{P}_\varrho^{2t}\varphi(\xi)h\| &= \|D\mathbf{E}(\mathcal{P}_\varrho^t\varphi \circ \Phi_\varrho^t)(\xi)h\| = \Big\|\mathbf{E}\Big((D\mathcal{P}_\varrho^t\varphi \circ \Phi_\varrho^t)(\xi)D\Phi_\varrho^t(\xi)h\Big)\Big\| \\ &\leq \sqrt{\mathbf{E}\|(D\mathcal{P}_\varrho^t\varphi \circ \Phi_\varrho^t)(\xi)\|^2}\sqrt{\mathbf{E}\|D\Phi_\varrho^t(\xi)h\|^2} . \end{aligned}$$

Bounding the first square root by Proposition 5.2 and then applying Proposition 5.1 (B–C), (with $T = 1$) we get a bound

$$\begin{aligned} \|D\mathcal{P}_\varrho^{2t}\varphi(\xi)h\| &\leq C_\varrho\|\varphi\|_{\mathrm{L}^\infty}(1 + t^{-\mu_*})\sqrt{\mathbf{E}(1 + \|\Phi_\varrho^t(\xi)\|_\alpha^{\nu_*})^2}\sqrt{\mathbf{E}\|D\Phi_\varrho^t(\xi)h\|^2} \\ &\leq C_\varrho\|\varphi\|_{\mathrm{L}^\infty}(1 + t^{-\mu_*})t^{-\alpha\nu_*}(1 + \|\xi\|)^{\nu_*}\|h\| . \end{aligned}$$

Choosing $\mu = \mu_* + \alpha\nu_*$ and $\nu = \nu_*$ we find (4.3) in the case when $\varphi \in \mathcal{C}_b^2(\mathcal{H})$. The method of extension to arbitrary $\varphi \in \mathcal{B}_b(\mathcal{H})$ can be found in [DPZ96, Lemma 7.1.5]. The proof of Theorem 4.3 is complete. □

5.3 Smoothing properties of the transition semigroup

In this subsection we prove the smoothing bound Proposition 5.2. Thus, we will no longer be interested in smoothing in position space as shown in Proposition 5.1 but in smoothing properties of the transition semigroup associated to (4.2).

Important remark. In this section and up to Section 8.6 we always tacitly assume that we are considering the cutoff equation (4.2) and we will omit the index ϱ.

Thus, we will write Eq.(4.2) as

$$d\Phi^t = -A\Phi^t\,dt + (F \circ \Phi^t)\,dt + (Q \circ \Phi^t)\,dW(t) . \tag{5.3}$$

The solution of (5.3) generates a semigroup on the space $\mathcal{B}_b(\mathcal{H})$ of bounded Borel functions over $\mathcal{H} = \mathcal{H}_{\mathrm{L}} \oplus \mathcal{H}_{\mathrm{H}}$ by

$$\mathcal{P}^t\varphi = \mathbf{E}(\varphi \circ \Phi^t) , \qquad \varphi \in \mathcal{B}_b(\mathcal{H}) .$$

Our goal will be to show that the mixing properties of the nonlinearity are strong enough to make $\mathcal{P}^t\varphi$ differentiable, even if φ is only measurable.

We will need a separate treatment of the high and low frequencies, and so we reformulate (5.3) as

$$d\Phi^t_{\mathrm{L}} = -A_{\mathrm{L}}\Phi^t_{\mathrm{L}}\,dt + (F_{\mathrm{L}}\circ\Phi^t)\,dt + (Q_{\mathrm{L}}\circ\Phi^t)\,dW_{\mathrm{L}}(t)\;, \qquad \Phi^t_{\mathrm{L}}\in\mathcal{H}_{\mathrm{L}}\;, \tag{5.4a}$$

$$d\Phi^t_{\mathrm{H}} = -A_{\mathrm{H}}\Phi^t_{\mathrm{H}}\,dt + (F_{\mathrm{H}}\circ\Phi^t)\,dt + Q_{\mathrm{H}}\,dW_{\mathrm{H}}(t)\;, \qquad \Phi^t_{\mathrm{H}}\in\mathcal{H}_{\mathrm{H}}\;, \tag{5.4b}$$

where $\mathcal{H}_{\mathrm{L}}$ and $\mathcal{H}_{\mathrm{H}}$ are defined in Section 5.1 and the cutoff version of Q was defined in (4.1). Note that $Q_{\mathrm{H}}(\Phi^t(\xi))$ is independent of ξ and t by construction, which is why we can use Q_{H} in (5.4b).

The proof of Proposition 5.2 is based on the following two results dealing with the low-frequency part and the cross-terms between low and high frequencies, respectively.

Proposition 5.3 *There exist exponents $\mu,\nu>0$ such that for every $\varphi\in\mathcal{C}^2_b(\mathcal{H})$, every $\xi\in\mathcal{H}^\alpha$ and every $T>0$, one has*

$$\Big\|\mathbf{E}\Big((D_{\mathrm{L}}\varphi\circ\Phi^t)(\xi)(D_{\mathrm{L}}\Phi^t_{\mathrm{L}})(\xi)\Big)\Big\| \le C_T\,t^{-\mu}(1+\|\xi\|^\nu_\alpha)\|\varphi\|_{\mathrm{L}^\infty}\;,$$

for all $t\in(0,T]$.[5]

Lemma 5.4 *For every $T>0$ and every $p\ge1$, there is a constant $C_{T,p}>0$ such that for every $t\le T$, one has the estimates (valid for $h_{\mathrm{L}}\in\mathcal{H}_{\mathrm{L}}$ and $h_{\mathrm{H}}\in\mathcal{H}_{\mathrm{H}}$):*

$$\mathbf{E}\sup_{0<s<t}\|(D_{\mathrm{L}}\Phi^s_{\mathrm{H}})(\xi)h_{\mathrm{L}}\|^p \le C_{T,p}\,t^p\|h_{\mathrm{L}}\|^p\;, \tag{5.5a}$$

$$\mathbf{E}\sup_{0<s<t}\|(D_{\mathrm{H}}\Phi^s_{\mathrm{L}})(\xi)h_{\mathrm{H}}\|^p \le C_{T,p}\,t^{p/4}\|h_{\mathrm{H}}\|^p\;. \tag{5.5b}$$

These bounds are independent of $\xi\in\mathcal{H}$.

Remark 5.5 In the absence of the cutoff ϱ one can prove inequalities like (5.5), but with an additional factor of $(1+\|\xi\|^2)^p$ on the right. This is not good enough for our strategy and is the reason for introducing a cutoff.

The proof of Proposition 5.3 will be given in Section 6 and the proof of Lemma 5.4 will be given in Section 8.5.

Proof of Proposition 5.2. As in the proof of Theorem 4.3, it suffices to consider times $t\le T$, where T is any (small) positive constant. The proof will be performed in the spirit of [DPZ96] and [Cer99], using a modified version of the Bismut-Elworthy formula. Take a function $\varphi\in\mathcal{C}^2_b(\mathcal{H})$. We consider Q_{L} and Q_{H} as acting on and into $\mathcal{H}_{\mathrm{L}}$ and $\mathcal{H}_{\mathrm{H}}$ respectively. It is possible to write as a consequence of Itô's formula:

$$\begin{aligned}(\varphi\circ\Phi^t)(\xi) &= \mathcal{P}^t\varphi(\xi) + \int_0^t((D\mathcal{P}^{t-s}\varphi)\circ\Phi^s)(\xi)\,(Q\circ\Phi^s)(\xi)\,dW(s)\\ &= \mathcal{P}^t\varphi(\xi) + \int_0^t((D_{\mathrm{L}}\mathcal{P}^{t-s}\varphi)\circ\Phi^s)(\xi)\,(Q_{\mathrm{L}}\circ\Phi^s)(\xi)\,dW_{\mathrm{L}}(s)\end{aligned}$$

[5] Recall that not only the flow, but for example also the constant C_T depends on ϱ.

$$+\int_0^t ((D_{\rm H}\mathcal{P}^{t-s}\varphi)\circ\Phi^s)(\xi)\,Q_{\rm H}\,dW_{\rm H}(s)\;. \tag{5.6}$$

Choose some $h\in\mathcal{H}_{\rm H}$. By Proposition 5.1 (D), $(D_{\rm H}\Phi_{\rm H}^t)(\xi)h$ is in $\mathcal{H}^\alpha$ for positive times and is bounded by (5.1d). Using condition (1.8b) we see that $Q_{\rm H}^{-1}$ maps to $\mathcal{H}_{\rm H}$ and so we can multiply both sides of (5.6) by

$$\int_{t/4}^{3t/4}\Big\langle Q_{\rm H}^{-1}(D_{\rm H}\Phi_{\rm H}^s)(\xi)h\,,\,dW_{\rm H}(s)\Big\rangle\;,$$

where the scalar product is taken in $\mathcal{H}_{\rm H}$. Taking expectations on both sides, the first two terms on the right vanish because $dW_{\rm L}$ and $dW_{\rm H}$ are independent and of mean zero. Thus, we get

$$\begin{aligned}\mathbf{E}\Big((\varphi\circ\Phi^t)(\xi)\int_{t/4}^{3t/4}\Big\langle Q_{\rm H}^{-1}(D_{\rm H}\Phi_{\rm H}^s)(\xi)h\,,\,dW_{\rm H}(s)\Big\rangle\Big)\\ =\mathbf{E}\int_{t/4}^{3t/4}((D_{\rm H}\mathcal{P}^{t-s}\varphi)\circ\Phi^s)(\xi)\,(D_{\rm H}\Phi_{\rm H}^s)(\xi)h\,ds\;,\end{aligned} \tag{5.7}$$

We add to both sides of (5.7) the term

$$\mathbf{E}\int_{t/4}^{3t/4}((D_{\rm L}\mathcal{P}^{t-s}\varphi)\circ\Phi^s)(\xi)\,(D_{\rm H}\Phi_{\rm L}^s)(\xi)h\,ds\;,$$

and note that the r.h.s. can be rewritten as

$$\int_{t/4}^{3t/4}D_{\rm H}\mathbf{E}((\mathcal{P}^{t-s}\varphi)\circ\Phi^s)(\xi)h\,ds=\frac{t}{2}D_{\rm H}\mathbf{E}(\varphi\circ\Phi^t)(\xi)h\;,$$

since by the Markov property, $\mathbf{E}(\mathcal{P}^{t-s}\varphi\circ\Phi^s)(\xi)=\mathbf{E}(\varphi\circ\Phi^t)(\xi)$. Therefore, (5.7) leads to

$$\begin{aligned}(D_{\rm H}\mathcal{P}^t\varphi)(\xi)h=\frac{2}{t}\mathbf{E}\Big((\varphi\circ\Phi^t)(\xi)\int_{t/4}^{3t/4}\langle Q_{\rm H}^{-1}(D_{\rm H}\Phi_{\rm H}^s)(\xi)h,dW_{\rm H}(s)\rangle\Big)\\ +\frac{2}{t}\mathbf{E}\int_{t/4}^{3t/4}((D_{\rm L}\mathcal{P}^{t-s}\varphi)\circ\Phi^s)(\xi)\,(D_{\rm H}\Phi_{\rm L}^s)(\xi)h\,ds\;.\end{aligned} \tag{5.8}$$

For the low-frequency part, we use the equality

$$\begin{aligned}(D_{\rm L}\mathcal{P}^t\varphi)(\xi)=\mathbf{E}\Big((D_{\rm L}\mathcal{P}^{t/2}\varphi\circ\Phi^{t/2})(\xi)(D_{\rm L}\Phi_{\rm L}^{t/2})(\xi)\Big)\\ +\mathbf{E}\Big((D_{\rm H}\mathcal{P}^{t/2}\varphi\circ\Phi^{t/2})(\xi)(D_{\rm L}\Phi_{\rm H}^{t/2})(\xi)\Big)\;.\end{aligned} \tag{5.9}$$

We introduce the Banach spaces $\mathscr{B}_{T,\mu_*,\nu_*}$ of measurable functions $f:(0,T)\times\mathcal{H}^\alpha\to\mathcal{H}$, for which

$$|||f|||_{T,\mu_*,\nu_*}\equiv\sup_{0<t<T}\sup_{\xi\in\mathcal{H}^\alpha}\frac{t^{\mu_*}\|f(t,\xi)\|}{1+\|\xi\|_\alpha^{\nu_*}} \tag{5.10}$$

is finite. Recall that we consider here only times smaller than the (small) time $T\in(0,1]$ which we will fix below. Choose μ_* as the maximum of the constants α and the μ appearing in Proposition 5.3. Similarly ν_* is the maximum of the ν of Proposition 5.1 (D) and the one in Proposition 5.3.

We will construct a $T > 0$ such that $f_\varphi : (t,\xi) \mapsto (D\mathcal{P}^t\varphi)(\xi)$ belongs to $\mathscr{B}_{T,\mu_*,\nu_*}$ and that $|||f_\varphi|||_{T,\mu_*,\nu_*} \le C\|\varphi\|_{\mathrm{L}^\infty}$, thus proving Proposition 5.2. The fact that $f_\varphi \in \mathscr{B}_{T,\mu_*,\nu_*}$ for every T if $\varphi \in \mathcal{C}_b^2(\mathcal{H})$ is shown in [DPZ92b, Theorem 9.17], so we only have to show the bound on its norm.

The following inequalities are obtained by applying to (5.8) in order the Cauchy-Schwarz inequality and the definition (5.10), then (1.8b), (5.1d), and again Cauchy-Schwarz. The last inequality is obtained by applying (5.1a) and (5.1c). This yields for $h \in \mathcal{H}_\mathrm{H}$:

$$\begin{aligned}
|(D_\mathrm{H}\mathcal{P}^t\varphi)(\xi)h| &\le \|\varphi\|_{\mathrm{L}^\infty}\frac{2}{t}\left(\mathbf{E}\int_{t/4}^{3t/4}\|Q_\mathrm{H}^{-1}(D_\mathrm{H}\Phi_\mathrm{H}^s)(\xi)h\|^2\,ds\right)^{1/2} \\
&\quad + \frac{2}{t}|||f_\varphi|||_{t,\mu_*,\nu_*}\,\mathbf{E}\int_{t/4}^{3t/4}\frac{1+\|\Phi^s(\xi)\|_\alpha^{\nu_*}}{(t-s)^{\mu_*}}\|(D_\mathrm{H}\Phi_\mathrm{L}^s)(\xi)h\|\,ds \\
&\le Ct^{-\alpha}\|\varphi\|_{\mathrm{L}^\infty}(1+\|\xi\|_\alpha^{\nu_*})\|h\| \qquad (5.11)\\
&\quad + Ct^{-\mu_*}|||f_\varphi|||_{t,\mu_*,\nu_*}\left(\mathbf{E}\sup_{s\in[\frac{t}{4},\frac{3t}{4}]}(1+\|\Phi^s(\xi)\|_\alpha^{\nu_*})^2\right)^{1/2} \\
&\qquad \times\left(\mathbf{E}\sup_{s\in[\frac{t}{4},\frac{3t}{4}]}\|(D_\mathrm{H}\Phi_\mathrm{L}^s)(\xi)h\|^2\right)^{1/2} \\
&\le Ct^{-\alpha}\|\varphi\|_{\mathrm{L}^\infty}(1+\|\xi\|_\alpha^{\nu_*})\|h\| + Ct^{-\mu_*+1/4}|||f_\varphi|||_{t,\mu_*,\nu_*}(1+\|\xi\|_\alpha^{\nu_*})\|h\|\;.
\end{aligned}$$

Note that this is the place where the lower bound (1.8b) on the noise is really used.

For the low-frequency part Eq.(5.9) we use first Proposition 5.3, $\|\mathcal{P}^{t/2}\varphi\|_{\mathrm{L}^\infty} \le \|\varphi\|_{\mathrm{L}^\infty}$, and the definition (5.10), then Cauchy-Schwarz, and finally (5.5a) and (5.1b). This leads for $h \in \mathcal{H}_\mathrm{L}$ to:

$$\begin{aligned}
|(D_\mathrm{L}\mathcal{P}^t\varphi)(\xi)h| &\le Ct^{-\mu_*}\|\varphi\|_{\mathrm{L}^\infty}(1+\|\xi\|_\alpha^{\nu_*})\,\|h\| \\
&\quad + Ct^{-\mu_*}|||f_\varphi|||_{t,\mu_*,\nu_*}\mathbf{E}\Big((1+\|\Phi^{t/2}(\xi)\|_\alpha^{\nu_*})\|(D_\mathrm{L}\Phi_\mathrm{H}^{t/2})(\xi)h\|\Big) \\
&\le Ct^{-\mu_*}\|\varphi\|_{\mathrm{L}^\infty}(1+\|\xi\|_\alpha^{\nu_*})\,\|h\| \qquad (5.12)\\
&\quad + Ct^{-\mu_*}|||f_\varphi|||_{t,\mu_*,\nu_*}\sqrt{\mathbf{E}(1+\|\Phi^{t/2}(\xi)\|_\alpha^{\nu_*})^2}\sqrt{\mathbf{E}\|(D_\mathrm{L}\Phi_\mathrm{H}^{t/2})(\xi)h\|^2} \\
&\le Ct^{-\mu_*}\|\varphi\|_{\mathrm{L}^\infty}(1+\|\xi\|_\alpha^{\nu_*})\,\|h\| + Ct^{-\mu_*+1}|||f_\varphi|||_{t,\mu_*,\nu_*}(1+\|\xi\|_\alpha^{\nu_*})\,\|h\|\;.
\end{aligned}$$

Combining the above expressions we get for every $T \in (0,1]$ a bound of the type

$$|||f_\varphi|||_{T,\mu_*,\nu_*} \le C_1\|\varphi\|_{\mathrm{L}^\infty} + C_2T^{1/4}|||f_\varphi|||_{T,\mu_*,\nu_*}\;.$$

Our final choice of T is now $T^{1/4} = \min\{1, 1/(2C_2)\}$, and we find

$$|||f_\varphi|||_{T,\mu_*,\nu_*} \le C\|\varphi\|_{\mathrm{L}^\infty}\;. \qquad (5.13)$$

Since $f_\varphi(t,\xi) = (D\mathcal{P}^t\varphi)(\xi)$, inspection of (5.10) shows that (5.13) is equivalent to (5.2). The proof of Proposition 5.2 is complete. □

6 Malliavin Calculus

To prove Proposition 5.3 we will apply a modification of Norris' version of the Malliavin calculus. This modification takes into account some new features which are necessary due to our splitting of the problem in high and low frequencies (which in turn was done to deal with the infinite dimensional nature of the problem).

Consider first the deterministic PDE for a flow:

$$\frac{d\Psi^t(\xi)}{dt} = -A\Psi^t(\xi) + (F \circ \Psi^t)(\xi) . \tag{6.1}$$

This is really an abstract reformulation for the flow defined by the GL equation, and ξ belongs to a space $\mathcal{H}$, which for our problem is a suitable Sobolev space. The linear operator A is chosen as $1 - \Delta$, while the non-linear term F corresponds to $2u - u^3$ in the GL equation. Below, we will work with approximations to the GL equation, and all we need to know is that $A : \mathcal{H} \to \mathcal{H}$ is the generator of a strongly continuous semigroup, and F will be seen to be bounded with bounded derivatives.

For each fixed $\xi \in \mathcal{H}$ we consider the following stochastic variant of (6.1):

$$d\Psi^t(\xi) = -A\Psi^t(\xi)\,dt + (F \circ \Psi^t)(\xi)\,dt + (Q \circ \Psi^t)(\xi)\,dW(t) . \tag{6.2}$$

with initial condition $\Psi^0(\xi) = \xi$. Furthermore, W is the cylindrical Wiener process on a separable Hilbert space $\mathcal{W}$ and Q is a strongly differentiable map from $\mathcal{H}$ to $\mathscr{L}^2(\mathcal{W}, \mathcal{H})$, the space of bounded linear Hilbert-Schmidt operators from $\mathcal{W}$ to $\mathcal{H}$.

We next introduce the notion of directional derivative (in the direction of the noise) and the reader familiar with this concept can pass directly to (6.3). To understand this concept consider first the case of a function $t \mapsto v_i^t \in \mathcal{W}$. Then the variation $\mathscr{D}_{v_i}\Psi^t$ of Ψ^t in the direction v_i is obtained by replacing $dW(t)$ by $dW(t) + \varepsilon v_i^t\,dt$ and it satisfies the equation

$$\begin{aligned} d\mathscr{D}_{v_i}\Psi^t \;=\; & (-A\mathscr{D}_{v_i}\Psi^t + (DF \circ \Psi^t)\mathscr{D}_{v_i}\Psi^t)\,dt + ((DQ \circ \Psi^t)\mathscr{D}_{v_i}\Psi^t)\,dW(t) \\ & + (Q \circ \Psi^t)v_i^t\,dt . \end{aligned}$$

Intuitively, the first line comes from varying Ψ^t with respect to the noise and the second comes from varying the noise itself.

We will need a finite number L of directional derivatives, and so we introduce some more general notation. We combine L vectors v_i as used above into a matrix called v which is an element of $\Omega \times [0, \infty) \to \mathcal{W}^L$. We identify $\mathcal{W}^L$ with $\mathscr{L}(\mathbf{R}^L, \mathcal{W})$. Note that we now allow v to depend on Ω, and to make things work, we require v to be a predictable stochastic process, *i.e.*, v^t only depends on the noise before time t. The stochastic process $G_v^t \in \mathcal{H}^L$ (corresponding to $\mathscr{D}_v\Psi^t$) is then defined as the solution of the equation

$$\begin{aligned} dG_v^t h &= \Big(-AG_v^t + (DF \circ \Psi^t)G_v^t + (Q \circ \Psi^t)v^t\Big)h\,dt \\ &\qquad + \Big((DQ \circ \Psi^t)\,G_v^t h\Big)\,dW(t) , \\ G_v^0 &= 0 , \end{aligned} \tag{6.3}$$

which has to hold for all $h \in \mathbf{R}^L$.

Having given the detailed definition of G_v^t, we will denote it henceforth by the more suggestive

$$G_v^t(\xi) = \mathscr{D}_v \Psi^t(\xi) ,$$

to make clear that it is a directional derivative. We use the notation $\mathscr{D}_v$ to distinguish this derivative from the derivative D with respect to the initial condition ξ.

For (6.2) and (6.3) to make sense, two assumptions on F, Q and v are needed:

A1 $F : \mathcal{H} \to \mathcal{H}$ and $Q : \mathcal{H} \to \mathscr{L}^2(\mathcal{W}, \mathcal{H})$ are of at most linear growth and have bounded first and second derivatives.

A2 The stochastic process $t \mapsto v^t$ is predictable, has a continuous version, and satisfies

$$\mathbf{E}\Big(\sup_{s\in[0,t]} \|v^s\|^p\Big) < \infty ,$$

for every $t > 0$ and every $p \geq 1$. (The norm being the norm of $\mathcal{W}^L$.)

It is easy to see that these conditions imply the hypotheses of Theorem 8.9 for the problems (6.2) and (6.3). Therefore G_v^t is a well-defined strongly Markovian stochastic process.

With these notations one has the well-known Bismut integration by parts formula [Nor86].

Proposition 6.1 *Let Ψ^t and $\mathscr{D}_v\Psi^t$ be defined as above and assume* **A1** *and* **A2** *are satisfied. Let $\mathcal{B} \subset \mathcal{H}$ be an open subset of $\mathcal{H}$ such that $\Psi^t \in \mathcal{B}$ almost surely and let $\varphi : \mathcal{B} \to \mathbf{R}$ be a differentiable function such that*

$$\mathbf{E}\|\varphi(\Psi^t)\|^2 + \mathbf{E}\|D\varphi(\Psi^t)\|^2 < \infty .$$

Then we have for every $h \in \mathbf{R}^L$ the following identity in $\mathbf{R}$:

$$\mathbf{E}(D\varphi(\Psi^t)\mathscr{D}_v\Psi^t h) = \mathbf{E}\Big(\varphi(\Psi^t)\int_0^t \langle v^s h, dW(s)\rangle\Big) , \qquad (6.4)$$

where $\langle\cdot,\cdot\rangle$ is the scalar product of $\mathcal{W}$.

Remark 6.2 The Eq.(6.4) is useful because it relates the expectation of $D\varphi$ to that of φ. In order to fully exploit (6.4) we will need to get rid of the factor $\mathscr{D}_v\Psi^t$. This will be possible by a clever choice of v. This procedure is explained for example in [Nor86] but we will need a new variant of his results because of the high-frequency part. In the sequel, we will proceed in two steps. *We need only bounds on $D_{\rm L}\varphi$, since the smoothness of the high-frequency part follows by other means.* Thus, it *suffices* to construct $\mathscr{D}_v\Psi^t$ in such a way that $\Pi_{\rm L}\mathscr{D}_v\Psi^t$ is invertible, where $\Pi_{\rm L}$ is the orthogonal projection onto $\mathcal{H}_{\rm L}$. The construction of $\Pi_{\rm L}\mathscr{D}_v\Psi^t$ follows closely the presentation of [Nor86]. However, we also want $\Pi_{\rm H}\mathscr{D}_v\Psi^t = 0$ and this elimination of the high-frequency part seems to be new.

Proof. The finite dimensional case is stated (with slightly different assumptions on the function F) in [Nor86]. The extension to the infinite-dimensional setting can be done without major difficulty. By **A1–A2** and Theorem 8.9, we ensure that all the expressions appearing in the proof and the statement are well-defined. By **A2**, we can use Itô's formula to ensure the validity of the assumptions for the infinite-dimensional version of Girsanov's theorem [DPZ96]. □

6.1 The construction of v

In order to use Proposition 6.1 we will construct $v = (v_{\rm L}, v_{\rm H})$ in such a way that the high-frequency part of $\mathscr{D}_v\Phi^t = (\mathscr{D}_v\Phi^t_{\rm L}, \mathscr{D}_v\Phi^t_{\rm H})$ vanishes. This construction is new and will be explained in detail in this subsection.

Notation. The equations which follow are quite involved. To keep the notation at a reasonable level without sacrificing precision we will adopt the following conventions:

$$(D_{\rm L}F_{\rm L})^t \equiv (D_{\rm L}F_{\rm L}) \circ \Phi^t \;,$$
$$(D_{\rm L}Q_{\rm L})^t \equiv (D_{\rm L}Q_{\rm L}) \circ \Phi^t \;,$$

and similarly for other derivatives of the Q and the F. Furthermore, the reader should note that $D_{\rm L}Q_{\rm L}$ is a linear map from $\mathcal{H}_{\rm L}$ to the linear maps $\mathcal{H}_{\rm L} \to \mathcal{H}_{\rm L}$ and therefore, below, $(D_{\rm L}Q_{\rm L})h$ with $h \in \mathcal{H}_{\rm L}$ is a linear map $\mathcal{H}_{\rm L} \to \mathcal{H}_{\rm L}$. The dimension of $\mathcal{H}_{\rm L}$ is $L < \infty$.

Inspired by [Nor86], we define the $L \times L$ matrix-valued stochastic processes $U^t_{\rm L}$ and $V^t_{\rm L}$ by the following SDE's, which must hold for every $h \in \mathcal{H}_{\rm L}$:

$$\begin{aligned} dU^t_{\rm L}h &= -A_{\rm L}U^t_{\rm L}h\,dt + (D_{\rm L}F_{\rm L})^t U^t_{\rm L}h\,dt + ((D_{\rm L}Q_{\rm L})^t U^t_{\rm L}h)\,dW_{\rm L}(t) \;, \\ U^0_{\rm L} &= I \in \mathscr{L}(\mathcal{H}_{\rm L}, \mathcal{H}_{\rm L}) \;, \end{aligned} \tag{6.5a}$$

$$\begin{aligned} dV^t_{\rm L}h &= V^t_{\rm L}A_{\rm L}h\,dt - V^t_{\rm L}(D_{\rm L}F_{\rm L})^t h\,dt - V^t_{\rm L}((D_{\rm L}Q_{\rm L})^t h)\,dW_{\rm L}(t) \\ &\quad + \sum_{i=0}^{L-1} V^t_{\rm L}\Big((D_{\rm L}Q_{\rm L})^t((D_{\rm L}Q_{\rm L})^t h)e_i\Big)e_i\,dt \;, \\ V^0_{\rm L} &= I \in \mathscr{L}(\mathcal{H}_{\rm L}, \mathcal{H}_{\rm L}) \;. \end{aligned} \tag{6.5b}$$

The last term in the definition of $V^t_{\rm L}$ will be written as

$$\sum_{i=0}^{L-1} V^t_{\rm L}((D_{\rm L}Q^i_{\rm L})^t)^2 h\,dt \;,$$

where $Q^i_{\rm L}$ is the $i^{\rm th}$ column of the matrix $Q_{\rm L}$.

For small times, the process $U^t_{\rm L}$ is an approximation to the partial Jacobian $D_{\rm L}\Phi^t_{\rm L}$, and $V^t_{\rm L}$ is an approximation to its inverse.

We first make sure that the objects in (6.5) are well-defined. The following lemma summarizes the properties of $U_{\rm L}$ and $V_{\rm L}$ which we need later.

Lemma 6.3 *The processes $U^t_{\rm L}$ and $V^t_{\rm L}$ satisfy the following bounds. For every $p \ge 1$ and all $T > 0$ there is a constant $C_{T,p,\varrho}$ independent of the initial data (for Φ^t) such that*

$$\mathbf{E} \sup_{t\in[0,T]} (\|U^t_{\rm L}\|^p + \|V^t_{\rm L}\|^p) \le C_{T,p,\varrho} \;, \tag{6.6a}$$

$$\mathbf{E}(\sup_{t\in[0,\varepsilon]} \|V^t_{\rm L} - I\|^p) \le C_{T,p,\varrho}\varepsilon^{p/2} \;, \tag{6.6b}$$

for all $\varepsilon < T$. Furthermore, $V_{\rm L}$ is the inverse of $U_{\rm L}$ in the sense that $V^t_{\rm L} = (U^t_{\rm L})^{-1}$ almost surely

Proof. The bound (6.6a) is a straightforward application of Theorem 8.9 whose conditions are easily checked. (Note that we are here in a finite-dimensional, linear setting.) To prove (6.6b), note that I is the initial condition for $V_{\rm L}$. One writes (6.5b) in its integral form and then the result follows by applying (6.6a). The last statement can be shown easily by applying Itô's formula to the product $V_{\rm L}^t U_{\rm L}^t$. (In fact, the definition of $V_{\rm L}$ was precisely made with this in mind.) □

We continue with the construction of v. Since A and Q are diagonal with respect to the splitting $\mathcal{H} = \mathcal{H}_{\rm L} \oplus \mathcal{H}_{\rm H}$, we can write (6.3) as

$$\begin{aligned} d\,\mathscr{D}_v\Phi_{\rm L}^t = \Big(&-A_{\rm L}\,\mathscr{D}_v\Phi_{\rm L}^t + (D_{\rm L}F_{\rm L})^t\,\mathscr{D}_v\Phi_{\rm L}^t \\ &+ (D_{\rm H}F_{\rm L})^t\,\mathscr{D}_v\Phi_{\rm H}^t + Q_{\rm L}^t v_{\rm L}^t\Big)\,dt \\ &+ \Big((D_{\rm L}Q_{\rm L})^t\,\mathscr{D}_v\Phi_{\rm L}^t\Big)\,dW_{\rm L}(t) \\ &+ \Big((D_{\rm H}Q_{\rm L})^t\,\mathscr{D}_v\Phi_{\rm H}^t\Big)\,dW_{\rm L}(t)\,, \end{aligned} \tag{6.7a}$$

$$\begin{aligned} d\,\mathscr{D}_v\Phi_{\rm H}^t = \Big(&-A_{\rm H}\,\mathscr{D}_v\Phi_{\rm H}^t + (D_{\rm L}F_{\rm H})^t\,\mathscr{D}_v\Phi_{\rm L}^t \\ &+ (D_{\rm H}F_{\rm H})^t\,\mathscr{D}_v\Phi_{\rm H}^t + Q_{\rm H}v_{\rm H}^t\Big)\,dt\;, \end{aligned} \tag{6.7b}$$

with zero initial condition. Since we want to consider derivatives with respect to the low-frequency part, we would like to define (implicitly) $v_{\rm H}^t$ as

$$v_{\rm H}^t = -Q_{\rm H}^{-1}(D_{\rm L}F_{\rm H})^t\,\mathscr{D}_v\Phi_{\rm L}^t\;.$$

In this way, the solution of (6.7b) would be $\mathscr{D}_v\Phi_{\rm H}^t \equiv 0$. We next would define the "directions" $v_{\rm L}$ and $v_{\rm H}$ by

$$\begin{aligned} v_{\rm L}^t &= (V_{\rm L}^t\,Q_{\rm L}^t)^*\;, \\ v_{\rm H}^t &= -Q_{\rm H}^{-1}(D_{\rm L}F_{\rm H})^t\,\mathscr{D}_v\Phi_{\rm L}^t\;, \end{aligned} \tag{6.8}$$

where $\mathscr{D}_v\Phi_{\rm L}^t$ is the solution to (6.7a) with $\mathscr{D}_v\Phi_{\rm H}^t$ replaced by 0 and $v_{\rm L}$ replaced by $(V_{\rm L}^t\,Q_{\rm L}^t)^*$. Here, X^* denotes the transpose of the real matrix X.

The implict problem (6.8) can be somewhat simplified by the following device: Since we are constructing a solution of (6.7) whose high-frequency part is going to vanish, we consider instead the simpler equation for $y^t \in \mathscr{L}(\mathcal{H}_{\rm L}, \mathcal{H}_{\rm L})$:

$$dy^t = \Big(-A_{\rm L}y^t + (D_{\rm L}F_{\rm L})^t y^t + Q_{\rm L}^t(V_{\rm L}^t Q_{\rm L}^t)^*\Big)\,dt + ((D_{\rm L}Q_{\rm L})^t y^t)\,dW_{\rm L}(t)\;, \tag{6.9}$$

with $y^0 = 0$, and where we use again the notation $F^t = F \circ \Phi^t$, and similar notation for Q.

The verification that (6.9) is well-defined and can be bounded is again a consequence of Theorem 8.9 and is left to the reader. *Given the solution of (6.9) we proceed to make our definitive choice of $v_{\rm L}^t$ and $v_{\rm H}^t$:*

Definition 6.4 Given an initial condition $\xi \in \mathcal{H}^\alpha$ (for Φ^t) and a cutoff $\varrho < \infty$ we define $v^t = v_{\rm L}^t \oplus v_{\rm H}^t$ by

$$\begin{aligned} v_{\rm L}^t &\equiv \quad (V_{\rm L}^t\,Q_{\rm L}^t)^* &&= \quad (V_{\rm L}^t\,(Q_{\rm L}\circ\Phi^t))^*\;, \\ v_{\rm H}^t &\equiv -Q_{\rm H}^{-1}(D_{\rm L}F_{\rm H})^t\,y^t &&= -Q_{\rm H}^{-1}((D_{\rm L}F_{\rm H})\circ\Phi^t)\,y^t\;, \end{aligned} \tag{6.10}$$

where Φ^t solves (5.3), $V_{\rm L}^t$ is the solution of (6.5b), and y^t solves (6.9).

Lemma 6.5 *The process v^t satisfies for all $p \geq 1$ and all $t > 0$:*

$$\mathbf{E}\Big(\sup_{s\in[0,t]} \|v^s\|^p\Big) < C_{t,p,\varrho}(1+\|\xi\|_\alpha)^p ,$$

i.e., it satisfies assumption **A2** *of Proposition 6.1.*

Proof. By Proposition 5.1 (B), Φ^t is in $\mathcal{H}^\alpha$ for all $t \geq 0$. In Lemma 8.1 **P6**, it will be checked that $D_{\rm L}F_{\rm H}$ maps $\mathcal{H}^\alpha$ into $\mathscr{L}(\mathcal{H}_{\rm L}, \mathcal{H}^\alpha \cap \mathcal{H}_{\rm H})$ and that this map has linear growth. By the *lower bound* (1.6) on the amplitudes q_k, we see that $Q_{\rm H}^{-1}$ is bounded from $\mathcal{H}^\alpha \cap \mathcal{H}_{\rm H}$ to $\mathcal{H}_{\rm H}$ and thus the assertion follows. □

We now verify that $\mathscr{D}_v\Phi^t_{\rm H} \equiv 0$. Indeed, consider the equations (6.7). This is a system for two unknowns, $Y^t = \mathscr{D}_v\Phi^t_{\rm L}$ and $X^t = \mathscr{D}_v\Phi^t_{\rm H}$. For our choice of $v^t_{\rm L}$ and $v^t_{\rm H}$ this system takes the form

$$\begin{aligned} d\,Y^t = \Big(&-A_{\rm L}\,Y^t + (D_{\rm L}F_{\rm L})^t\,Y^t \\ &+ (D_{\rm H}F_{\rm L})^t\,X^t + Q^t_{\rm L}(V^t_{\rm L}Q^t_{\rm L})^*\Big)\,dt \\ &+ \Big((D_{\rm L}Q_{\rm L})^t\,Y^t\Big)\,dW_{\rm L}(t) \\ &+ \Big((D_{\rm H}Q_{\rm L})^t\,X^t\Big)\,dW_{\rm L}(t)\,, \end{aligned} \tag{6.11a}$$

$$\begin{aligned} d\,X^t = \Big(&-A_{\rm H}\,X^t + (D_{\rm L}F_{\rm H})^t\,Y^t \\ &+ (D_{\rm H}F_{\rm H})^t\,X^t - (D_{\rm L}F_{\rm H})^t y^t\Big)\,dt\ . \end{aligned} \tag{6.11b}$$

By inspection, we see that $X^t \equiv 0$ and

$$dY^t = (-A_{\rm L}Y^t + (D_{\rm L}F_{\rm L})^tY^t)\,dt + ((D_{\rm L}Q_{\rm L})^t\,Y^t)\,dW_{\rm L}(t) + Q^t_{\rm L}(V^t_{\rm L}Q^t_{\rm L})^*\,dt \tag{6.12}$$

solve the problem, *i.e.*, $Y^t = y^t$, by the construction of y^t. Applying the Itô formula to the product $V^t_{\rm L}Y^t$ and using Eqs.(6.5b) and (6.12), we see immediately that we have defined $Y^t = \mathscr{D}_v\Phi^t_{\rm L}$ in such a way that

$$d(V^t_{\rm L}\mathscr{D}_v\Phi^t_{\rm L}) = V^t_{\rm L}Q^t_{\rm L}(Q^t_{\rm L})^*(V^t_{\rm L})^*\,dt\ ,$$

because all other terms cancel. Thus we finally have shown

Theorem 6.6 *Given an initial condition $\xi \in \mathcal{H}^\alpha$ (for Φ^t) and a cutoff $\varrho < \infty$, the following is true: If v^t is given by Definition 6.4 then*

$$\begin{aligned} &\mathscr{D}_v\Phi^t_{\rm L} = U^t_{\rm L}\int_0^t V^s_{\rm L}((Q_{\rm L}Q^*_{\rm L})\circ\Phi^s)(V^s_{\rm L})^*\,ds \equiv U^t_{\rm L}C^t_{\rm L}\ , \\ &\mathscr{D}_v\Phi^t_{\rm H} \equiv 0\ . \end{aligned} \tag{6.13}$$

Definition 6.7 We will call the matrix $C^t_{\rm L}$ the *partial Malliavin matrix* of our system.

7 The Partial Malliavin Matrix

In this section, we estimate the partial Malliavin matrix $C_{\rm L}^t$ from below. We fix some time $t > 0$ and denote by $\mathcal{S}^L$ the unit sphere in $\mathbf{R}^L$. Our bound is

Theorem 7.1 *There are constants $\mu, \nu \geq 0$ such that for every $T > 0$ and every $p \geq 1$ there is a $C_{T,p,\varrho}$ such that for all initial conditions $\xi \in \mathcal{H}^\alpha$ for the flow Φ^t and all $t < T$, one has*

$$\mathbf{E}\Big((\det C_{\rm L}^t)^{-p}\Big) \leq C_{T,p,\varrho} t^{-\mu p}(1 + \|\xi\|_\alpha)^{\nu p} \ .$$

Corollary 7.2 *There are constants $\mu, \nu \geq 0$ such that for every $T > 0$ and every $p \geq 1$ there is a $C_{T,p,\varrho}$ such that for all initial conditions $\xi \in \mathcal{H}^\alpha$ for the flow Φ^t and all $t < T$, one has, with v given by Definition 6.4:*

$$\mathbf{E}\|(\mathscr{D}_v \Phi_{\rm L}^t)^{-p}\| \leq C_{T,p,\varrho} t^{-\mu p}(1 + \|\xi\|_\alpha)^{\nu p} \ .$$

This corollary follows from $(\mathscr{D}_v \Phi_{\rm L}^t)^{-1} = (C_{\rm L}^t)^{-1} V_{\rm L}^t$ and Eq.(6.6a).

As a first step, we formulate a bound from which Theorem 7.1 follows easily.

Theorem 7.3 *There are a $\mu > 0$ and a $\nu > 0$ such that for every $p \geq 1$, every $t < T$ and every $\xi \in \mathcal{H}^2$, one has*

$$\mathbf{P}\Big(\inf_{h \in \mathcal{S}^L} \int_0^t \|Q_{\rm L}^s (V_{\rm L}^s)^* h\|^2 \, ds < \varepsilon\Big) \ \leq \ C_{T,p,\varrho} \varepsilon^p t^{-\mu p}(1 + \|\xi\|_2)^{\nu p} \ ,$$

with $C_{T,p,\varrho}$ independent of ξ.

Proof of Theorem 7.1. Note that $\int_0^t \|Q_{\rm L}^s (V_{\rm L}^s)^* h\|^2 \, ds$ is, by Eq.(6.13), nothing but the quantity $\langle h, C_{\rm L}^t h\rangle$. Then, Theorem 7.1 follows at once. □

The proof of Theorem 7.3 is largely inspired from [Nor86, Sect. 4], but we need some new features to deal with the infinite dimensional high-frequency part. This will take up the next three subsections.

Our proof needs a modification of the Lie brackets considered when we study the Hörmander condition. We explain first these identities in a finite dimensional setting.

7.1 Finite dimensional case

Throughout this subsection we assume that both $\mathcal{H}_{\rm L}$ and $\mathcal{H}_{\rm H}$ are finite dimensional and we denote by N the dimension of $\mathcal{H}$. The function Q maps $\mathcal{H}$ to $\mathscr{L}(\mathcal{H}, \mathcal{H})$, and we denote by $Q_i : \mathcal{H} \to \mathcal{H}$ its $i^{\rm th}$ column ($i = 0, \ldots, N-1$).[6] Finally, $\widehat{F}$ is the drift (in this section, we absorb the linear part of the SDE into $\widehat{F} = -A + F$, to simplify the expressions). The equation for Φ^t is

$$\Phi^t(\xi) = \xi + \int_0^t (\widehat{F} \circ \Phi^s)(\xi) \, ds + \int_0^t \sum_{i=0}^{N-1} (Q_i \circ \Phi^s)(\xi) \, dw_i(s) \ .$$

[6] There is a slight ambiguity of notation here, since Q_i really means $Q_{\varrho,i}$ which is not the same as Q_ϱ.

Let $K : \mathcal{H} \to \mathcal{H}_{\rm L}$ be a smooth function whose derivatives are all bounded and define $K^t = K \circ \Phi^t$, $\widehat{F}^t = \widehat{F} \circ \Phi^t$, and $Q_i^t = Q_i \circ \Phi^t$. We then have by Itô's formula

$$dK^t = (DK)^t \widehat{F}^t\,dt + \sum_{i=0}^{N-1} (DK)^t Q_i^t\,dw_i(t) + \tfrac{1}{2} \sum_{i=0}^{N-1} (D^2K)^t(Q_i^t; Q_i^t)\,dt\;. \tag{7.1}$$

We next rewrite the equation (6.5) for $V_{\rm L}^t$ as:

$$dV_{\rm L}^t = -V_{\rm L}^t(D_{\rm L}\widehat{F}_{\rm L})^t\,dt - \sum_{i=0}^{L-1} V_{\rm L}^t(D_{\rm L}Q_i)^t\,dw_i(t) + \sum_{i=0}^{L-1} V_{\rm L}^t((D_{\rm L}Q_i)^t)^2\,dt\;.$$

By Itô's formula, we have therefore the following equation for the product $V_{\rm L}^t K^t$:

$$\begin{aligned} d(V_{\rm L}^tK^t) = -V_{\rm L}^t(D_{\rm L}\widehat{F}_{\rm L})^t K^t\,dt &- V_{\rm L}^t \sum_{i=0}^{L-1}(D_{\rm L}Q_i)^tK^t\,dw_i(t) \\ &+ V_{\rm L}^t \sum_{i=0}^{L-1}((D_{\rm L}Q_i)^t)^2 K^t\,dt + V_{\rm L}^t(DK)^t\widehat{F}^t\,dt \\ &+ V_{\rm L}^t \sum_{i=0}^{N-1}(DK)^tQ_i^t\,dw_i(t) \\ &+ \tfrac{1}{2}V_{\rm L}^t \sum_{i=0}^{N-1}(D^2K)^t(Q_i^t;Q_i^t)\,dt \\ &- V_{\rm L}^t \sum_{i=0}^{L-1}(D_{\rm L}Q_i)^t(DK)^tQ_i^t\,dt\;. \end{aligned} \tag{7.2}$$

By construction, $D_{\rm L}Q_i = 0$ for $i \geq L$ and therefore we can extend all the sums above to $N-1$.

The following definition is useful to simplify (7.2). Let $A : \mathcal{H} \to \mathcal{H}$ and $B : \mathcal{H} \to \mathcal{H}_{\rm L}$ be two functions with continuous bounded derivatives. We define the *projected Lie bracket* $[A,B]_{\rm L} : \mathcal{H} \to \mathcal{H}_{\rm L}$ by

$$[A,B]_{\rm L}(x) = \Pi_{\rm L}[A,B](x) = (DB(x))A(x) - (D_{\rm L}A_{\rm L}(x))B(x)\;.$$

A straightforward calculation then leads to

$$\begin{aligned} d(V_{\rm L}^tK^t) = V_{\rm L}^t\Big([\widehat{F},K]_{\rm L}^t &+ \tfrac{1}{2}\sum_{i=0}^{N-1}[Q_i,[Q_i,K]_{\rm L}]_{\rm L}^t\Big)\,dt \\ &+ V_{\rm L}^t\sum_{i=0}^{N-1}[Q_i,K]_{\rm L}^t\;\;dw_i(t) \\ &+ \tfrac{1}{2}V_{\rm L}^t\sum_{i=0}^{N-1}\Big(((D_{\rm L}Q_i)^t)^2K^t - (DK)^t(DQ_i)^tQ_i^t \\ &\qquad\qquad + (DD_{\rm L}Q_i)^t(Q_i^t;K^t)\Big)\,dt\;. \end{aligned} \tag{7.3}$$

Note next that for $i < L$, both K and Q_i map to $\mathcal{H}_{\rm L}$ and therefore $DD_{\rm L}Q_i(Q_i;K)$ equals $D_{\rm L}^2Q_i(Q_i;K)$ when $i < L$ and is 0 otherwise. Similarly, $(DK)(DQ_i)Q_i$ equals $(DK)(D_{\rm L}Q_i)Q_i$ when $i < L$ and vanishes otherwise. Thus, the last sum in (7.3) only extends to $L-1$.

In order to simplify (7.3) further, we define the vector field $\widetilde{F} : \mathcal{H} \to \mathcal{H}$ by

$$\widetilde{F} = \widehat{F} - \tfrac{1}{2} \sum_{i=0}^{L-1} (D_{\rm L} Q_i) Q_i \;.$$

Then we get

$$d(V_{\rm L}^t K^t) = V_{\rm L}^t \Big([\widetilde{F}, K]_{\rm L}^t + \tfrac{1}{2} \sum_{i=0}^{N-1} [Q_i, [Q_i, K]_{\rm L}]_{\rm L}^t \Big)\, dt + V_{\rm L}^t \sum_{i=0}^{N-1} [Q_i, K]_{\rm L}^t \, dw_i(t) \;.$$

This is very similar to [Nor86, p. 128], who uses conventional Lie brackets instead of $[\cdot,\cdot]_{\rm L}$.

7.2 Infinite dimensional case

In this case, some additional care is needed when we transcribe (7.1). The problem is that the stochastic flow Φ^t solves (5.4) in the mild sense but not in the strong sense. Nevertheless, this technical difficulty will be circumvented by choosing the initial condition in $\mathcal{H}^\alpha$. We have indeed by Proposition 5.1 (A) that if the initial condition is in $\mathcal{H}^\gamma$ with $\gamma \in [1, \alpha]$, then the solution of (5.4) is in the same space. Thus, Proposition 5.1 allows us to use Itô's formula also in the infinite dimensional case.

For any two Banach (or Hilbert) spaces $\mathscr{B}_1$, $\mathscr{B}_2$, we denote by $P(\mathscr{B}_1, \mathscr{B}_2)$ the set of all $\mathcal{C}^\infty$ functions $\mathscr{B}_1 \to \mathscr{B}_2$, which are polynomially bounded together with all their derivatives. Let $K \in P(\mathcal{H}, \mathcal{H}_{\rm L})$ and $X \in P(\mathcal{H}, \mathcal{H})$. We define as above $[X, K]_{\rm L} \in P(\mathcal{H}, \mathcal{H}_{\rm L})$ by

$$[X, K]_{\rm L}(x) = (DK(x))X(x) - (D_{\rm L} X_{\rm L}(x))K(x) \;.$$

Furthermore, we define $[A, K]_{\rm L} \in P(\mathcal{D}(A), \mathcal{H}_{\rm L})$ by the corresponding formula, *i.e.*,

$$[A, K]_{\rm L}(x) = (DK(x))Ax - A_{\rm L} K(x) \;,$$

where $A = 1 - \Delta$. Notice that if K is a constant vector field, *i.e.*, $DK = 0$, then $[A, K]_{\rm L}$ extends uniquely to an element of $P(\mathcal{H}, \mathcal{H}_{\rm L})$.

We choose again the basis $\{e_i\}_{i=0}^\infty$ of Fourier modes in $\mathcal{H}$ (see Eq.(1.5)) and define $dw_i(t) = \langle e_i, dW(t)\rangle$. We also define the stochastic process $K^t(\xi) = (K \circ \Phi^t)(\xi)$ and

$$\widetilde{F} = F - \tfrac{1}{2} \sum_{i=0}^{L-1} (D_{\rm L} Q_i) Q_i \;,$$

where $Q_i(x) = Q(x)e_i$. Then one has

Proposition 7.4 *Let $\xi \in \mathcal{H}^1$ and $K \in P(\mathcal{H}, \mathcal{H}_{\rm L})$. Then the equality*

$$\begin{aligned} V_{\rm L}^t(\xi) K^t(\xi) = K(\xi) &+ \int_0^t V_{\rm L}^s(\xi) \sum_{i=0}^\infty [Q_i, K]_{\rm L}^s(\xi)\, dw_i(s) \\ &+ \int_0^t V_{\rm L}^s(\xi) \Big(-[A, K]_{\rm L}^s(\xi) + [\widetilde{F}, K]_{\rm L}^s(\xi) \Big)\, ds \\ &+ \tfrac{1}{2} \int_0^t V_{\rm L}^s(\xi) \sum_{i=0}^\infty [Q_i, [Q_i, K]_{\rm L}]_{\rm L}^s(\xi)\, ds \;, \end{aligned}$$

holds almost surely. Furthermore, the same equality holds if $\xi \in \mathcal{H}^2$ and $K \in P(\mathcal{H}^1, \mathcal{H}_{\rm L})$.

Note that by $[A,K]^s_{\mathrm{L}}(\xi)$ we mean $\Big(DK(\Phi^s(\xi))\Big)(A\Phi^s(\xi)) - A_{\mathrm{L}}K(\Phi^s(\xi))$.

Proof. This follows as in the finite dimensional case by Itô's formula. □

7.3 The restricted Hörmander condition

The condition for having appropriate mixing properties is the following Hörmander-like condition.

Definition 7.5 Let $\mathcal{K} = \{K^{(i)}\}_{i=1}^M$ be a collection of functions in $P(\mathcal{H},\mathcal{H}_{\mathrm{L}})$. We say that $\mathcal{K}$ satisfies the *restricted Hörmander condition* if there exist constants $\delta, R > 0$ such that for every $h \in \mathcal{H}_{\mathrm{L}}$ and every $y \in \mathcal{H}$ one has

$$\sup_{K\in\mathcal{K}} \inf_{\|x-y\|\le R} \langle h, K(x)\rangle^2 \ge \delta\|h\|^2 . \tag{7.4}$$

We now construct the set $\mathcal{K}$ for our problem. We define the operator

$$[X^0, \cdot\,]_{\mathrm{L}} : P(\mathcal{H}^\gamma, \mathcal{H}_{\mathrm{L}}) \to P(\mathcal{H}^{\gamma+1}, \mathcal{H}_{\mathrm{L}})$$

by

$$[X^0,K]_{\mathrm{L}} = -[A,K]_{\mathrm{L}} + [F,K]_{\mathrm{L}} + \tfrac12 \sum_{i=0}^{\infty} [Q_i,[Q_i,K]_{\mathrm{L}}]_{\mathrm{L}} - \tfrac12 \sum_{i=0}^{L-1} [(D_{\mathrm{L}}Q_i)Q_i, K]_{\mathrm{L}} .$$

This is a well-defined operation since Q is Hilbert-Schmidt and DQ is finite rank and we can write

$$\sum_{i=0}^{\infty} [Q_i,[Q_i,K]_{\mathrm{L}}]_{\mathrm{L}} = \sum_{i=0}^{\infty} (D^2K)(Q_i;Q_i) + r ,$$

with r a finite sum of bounded terms.

Definition 7.6 We define

- $\mathcal{K}_0 = \{Q_i, \text{ with } i = 0,\dots,L-1\}$,
- $\mathcal{K}_1 = \{[X^0,Q_i]_{\mathrm{L}}, \text{ with } i = k_*,\dots,L-1\}$,
- $\mathcal{K}_\ell = \{[Q_i,K]_{\mathrm{L}}, \text{ with } K\in\mathcal{K}_{\ell-1}$ and $i = k_*,\dots,L-1\}$, when $\ell > 1$.

Finally,

$$\mathcal{K} = \mathcal{K}_0 \cup \cdots \cup \mathcal{K}_3 .^7$$

Remark 7.7 Since for $i \ge k_*$ the Q_i are constant vector fields, the quantity $[X^0,K]$ is in $P(\mathcal{H},\mathcal{H}_{\mathrm{L}})$ and not only in $P(\mathcal{H}^1,\mathcal{H}_{\mathrm{L}})$. Furthermore, if $K\in\mathcal{K}$ then D^jK is bounded for all $j \ge 0$.

We have

Theorem 7.8 *The set $\mathcal{K}$ constructed above satisfies the restricted Hörmander condition for the cutoff GL equation if ϱ is chosen sufficiently large. Furthermore, the inequality (7.4) holds for $R = \varrho/2$. Finally, $\delta > \delta_0 > 0$ for all sufficiently large ϱ.*

[7] The number 3 is the power 3 in u^3.

Proof. The basic idea of the proof is as follows: The leading term of F is the cubic term u^m with $m = 3$. Clearly, if i_1, i_2, i_3 are any 3 modes, we find

$$[e_{i_1}, [e_{i_2}, [u \mapsto u^3, e_{i_3}]_{\rm L}]_{\rm L}]_{\rm L} = \sum_{k=\pm i_1 \pm i_2 \pm i_3} C_k \Pi_{\rm L} e_k \;, \tag{7.5}$$

where the e_ℓ are the basis vectors of $\mathcal{H}$ defined in (1.5), and the C_k are *non-zero* combinatorial constants. By Lemma 3.3 the following is true: For every choice of a fixed k the three numbers i_1, i_2, and i_3 of $\mathscr{I}_k$ satisfy

– For $j = 1, 2, 3$ one has $i_j \in \{k_*, \ldots, L-1\}$.
– If $|k| < k_*$ exactly one of the six sums $\pm i_1 \pm i_2 \pm i_3$ lies in the set $\{0, \ldots, k_* - 1\}$ and exactly one lies in $\{-(k_* - 1), \ldots, 0\}$.

In particular, the expression (7.5) does not depend on u. If instead of u^3 we take a lower power, the triple commutator will vanish.

The basic idea has to be slightly modified because of the cutoff ϱ. First of all, the constant R in the definition of the Hörmander condition is set to $R = \varrho/2$. Consider first the case where $\|x\| \geq 5\varrho/2$. In that case we see from (4.1) that the $Q_{\varrho,i}$, viewed as vector fields, are of the form

$$Q_{\varrho,i}(x) \;=\; \begin{cases} (q_i + 1)e_i, & \text{if } i < k_*, \\ q_i e_i, & \text{if } i \geq k_*. \end{cases}$$

Since these vectors span a basis of $\mathcal{H}_{\rm L}$ the inequality (7.4) follows in this case (already by choosing only $K \in \mathcal{K}_0$).

Consider next the more delicate case when $\|x\| \leq 5\varrho/2$.

Lemma 7.9 *For all $\|x\| \leq 3\varrho$ one has for $\{i_1, i_2, i_3\} = \mathscr{I}_k$ the identity*

$$[e_{i_1}, [e_{i_2}, [X^0, e_{i_3}]_{\rm L}]_{\rm L}]_{\rm L}(x) = \sum_{k=\pm i_1 \pm i_2 \pm i_3} C_k \Pi_{\rm L} e_k + r_\varrho(x) \;, \tag{7.6}$$

where r_ϱ satisfies a bound

$$\|r_\varrho(x)\| \;\leq\; C\varrho^{-1} \;,$$

with the constant C independent of x and of $k < k_*$.

Proof. In $[X^0, \cdot]_{\rm L}$ there are 4 terms. The first, A, leads successively to $[A, e_{i_3}]_{\rm L} = (1 + i_3^2)e_{i_3}$, which is constant, and hence the Lie bracket with e_{i_2} vanishes. The second term contains the non-linear interaction F_ϱ. Since $\|x\| \leq 3\varrho$ one has $F_\varrho(x) = F(x)$. Thus, (7.5) yields the leading term of (7.6). The two remaining terms will contribute to $r_\varrho(x)$. We just discuss the first one. We have, using (4.1),

$$[Q_{\varrho,i}, e_{i_3}]_{\rm L}(x) = -DQ_{\varrho,i}(x)e_{i_3} \;=\; -\frac{1}{\varrho}\chi'(\|x\|/\varrho)\frac{\langle x, e_{i_3}\rangle}{\|x\|}\Pi_{k_*} e_i \;.$$

This gives clearly a bound of order ϱ^{-1} for this Lie bracket, and the further ones are handled in the same way. □

We continue the proof of Theorem 7.8. When $k < k_*$, we consider the elements of $\mathcal{K}_3$. They are of the form

$$[Q_{\varrho,i_1},[Q_{\varrho,i_2},[X^0,Q_{\varrho,i_3}]_{\rm L}]_{\rm L}]_{\rm L}(x) = q_{i_1}q_{i_2}q_{i_3}\Big(\sum_{k=\pm i_1\pm i_2\pm i_3} C_k\Pi_{\rm L}e_k + r_\varrho(x)\Big)\;.$$

Thus, for $\varrho=\infty$ these vectors together with the Q_i with $i\in\{k_*,\ldots,L-1\}$ span $\mathcal{H}_{\rm L}$ (independently of y with $\|y\|\le 3\varrho$) and therefore (7.5) holds in this case, if $\|x\|\le 5\varrho/2$ and $R=\varrho/2$. The assertion for finite, but large enough ϱ follows immediately by a perturbation argument. This completes the case of $\|x\|\le 5\varrho/2$ and hence the proof of Theorem 7.8. □

Proof of Theorem 7.3. The proof is very similar to the one in [Nor86], but we have to keep track of the x,t-dependence of the estimates. First of all, choose $x\in\mathcal{H}^2$ and $t\in(0,t_0]$.

From now on, we will use the notation $\mathcal{O}(\nu)$ as a shortcut for $C(1+\|x\|_2^\nu)$, where the constant C may depend on t_0 and p, but is independent of x and t. Denote by R the constant found in Theorem 7.8 and define the subset $\mathcal{B}_x$ of $\mathcal{H}^2$ by

$$\mathcal{B}_x = \{y\in\mathcal{H}^2 \;:\; \|y-x\|\le R \;\text{ and }\; \|y\|_\gamma \le \|x\|_\gamma + 1 \text{ for } \gamma=1,2\}\;.$$

We also denote by $\mathcal{B}(I)$ a ball of (small) radius $\mathcal{O}(1/L)$ centered at the identity in the space of all $L\times L$ matrices. (Recall that L is the dimension of $\mathcal{H}_{\rm L}$, and that $K\in\mathcal{K}$ maps to $\mathcal{H}_{\rm L}$.) We then have a bound of the type

$$\sup_{y\in\mathcal{B}_x}\sup_{K\in\mathcal{K}}\sum_{i=0}^\infty \|[Q_i,K]_{\rm L}(y)\|^2 \le \mathcal{O}(0)\;. \tag{7.7}$$

This is a consequence of the fact that QQ^* is trace class and thus the sum converges and its principal term is equal to

$$\begin{aligned}\mathrm{Tr}(Q^*(y)\,(DK)^*(y)\,(DK)(y)\,Q(y))&\\ &= \mathrm{Tr}((DK)(y)Q(y)\,Q^*(y)\,(DK)^*(y))\\ &= \sum_{i=0}^{L-1}\|Q^*(y)\,(DK)^*(y)e_i\|^2 \le C_\varrho\;.\end{aligned}$$

The last inequality follows from Remark 7.7. The other terms form a finite sum containing derivatives of the Q_i and are bounded in a similar way.

We have furthermore bounds of the type

$$\begin{aligned}\sup_{y\in\mathcal{B}_x}\sup_{K\in\mathcal{K}}\|[X^0,K]_{\rm L}(y)\|^2 &\le \mathcal{O}(\nu)\;,\\ \sup_{y\in\mathcal{B}_x}\sup_{K\in\mathcal{K}}\|[X^0,[X^0,K]_{\rm L}]_{\rm L}(y)\|^2 &\le \mathcal{O}(\nu)\;,\\ \sup_{y\in\mathcal{B}_x}\sup_{K\in\mathcal{K}}\sum_{i=0}^\infty\|[Q_i,[X^0,K]_{\rm L}]_{\rm L}(y)\|^2 &\le \mathcal{O}(\nu)\;,\end{aligned} \tag{7.8}$$

where $\nu=1$.

Let $\mathcal{S}_{\rm L}$ be the unit sphere in $\mathcal{H}_{\rm L}$. By the assumptions on $\mathcal{K}$ and the choice of $\mathcal{B}(I)$ we see that:

(A) *For every $h_0 \in \mathcal{S}_{\rm L}$, there exist a $K \in \mathcal{K}$ and a neighborhood $\mathcal{N}$ of h_0 in $\mathcal{S}_{\rm L}$ such that*

$$\inf_{y \in \mathcal{B}_x} \inf_{V \in \mathcal{B}(I)} \inf_{h \in \mathcal{N}} \langle VK(y), h\rangle^2 \geq \frac{\delta}{2} \;,$$

with δ the constant appearing in (7.4).

Next, we define a stopping time τ by $\tau = \min\{t, \tau_1, \tau_2\}$ with

$$\begin{aligned}\tau_1 &= \inf\{s \geq 0 \;:\; \Phi^s(x) \notin \mathcal{B}_x\} \;,\\ \tau_2 &= \inf\{s \geq 0 \;:\; V_{\rm L}^s(x) \notin \mathcal{B}(I)\} \;,\\ t &< T \text{ as chosen in the statement of Theorem 7.3} \;.\end{aligned}$$

It follows easily from Proposition 5.1 (E) that the probability of τ_1 being small (meaning that in the sequel we will always assume $\varepsilon \leq 1$) can be bounded by

$$\mathbf{P}(\tau_1 < \varepsilon) \leq C_p (1 + \|x\|_2)^{16p} \varepsilon^p \;,$$

with C_p independent of x. Similarly, using Lemma 6.3, we see that

$$\mathbf{P}(\tau_2 < \varepsilon) \leq C_p \varepsilon^p \;.$$

Observing that $\mathbf{P}(t < \varepsilon) < t^{-p}\varepsilon^p$ and combining this with the two estimates, we get for every $p \geq 1$:

$$\mathbf{P}(\tau < \varepsilon) \leq \mathcal{O}(16p) t^{-p} \varepsilon^p \;.$$

From this and (A) we deduce

(B) *for every $h_0 \in \mathcal{S}_{\rm L}$ there exist a $K \in \mathcal{K}$ and a neighborhood $\mathcal{N}$ of h_0 in $\mathcal{S}_{\rm L}$ such that for $\varepsilon < 1$,*

$$\sup_{h \in \mathcal{N}} \mathbf{P}\Big(\int_0^\tau \langle V_{\rm L}^s(x) K^s(x), h\rangle^2 \, ds \leq \varepsilon\Big) \leq \mathbf{P}(\tau < 2\varepsilon/\delta) \leq \mathcal{O}(16p) t^{-p} \varepsilon^p \;,$$

with δ the constant appearing in (7.4).

Following [Nor86], we will show below that (B) implies:

(C) *for every $h_0 \in \mathcal{S}_{\rm L}$ there exist an $i \in \{k_*, \ldots, L-1\}$, a neighborhood $\mathcal{N}$ of h_0 in $\mathcal{S}_{\rm L}$ and constants $\nu, \mu > 0$ such that for $\varepsilon < 1$ and $p > 1$ one has*

$$\sup_{h \in \mathcal{N}} \mathbf{P}\Big(\int_0^\tau \langle V_{\rm L}^s(x) Q_i^s(x), h\rangle^2 \, ds \leq \varepsilon\Big) \leq \mathcal{O}(\nu p) t^{-\mu p} \varepsilon^p \;.$$

Remark 7.10 Note that for small $\|x\|$, $Q_i(x) = Q_{i,\varrho}(x)$ may be 0 when $i < k_*$, but the point is that then we can find another i for which the inequality holds.

By a straightforward argument, given in detail in [Nor86, p. 127], one concludes that (C) implies Theorem 7.3.

It thus only remains to show that (B) implies (C). We follow closely Norris and choose a $K \in \mathcal{K}$ such that (B) holds. If K happens to be in $\mathcal{K}_0$ then it is equal to a Q_i, and thus we already have (C). Otherwise, assume $K \in \mathcal{K}_j$ with $j \geq 1$. Then we use a Martingale inequality.

Lemma 7.11 *Let $\mathcal{H}$ be a separable Hilbert space and $W(t)$ be the cylindrical Wiener process on $\mathcal{H}$. Let β^t be a real-valued predictable process and γ^t, ζ^t be predictable $\mathcal{H}$-valued processes. Define*

$$a^t = a^0 + \int_0^t \beta^s\,ds + \int_0^t \langle \gamma^s, dW(s)\rangle\ ,$$
$$b^t = b^0 + \int_0^t a^s\,ds + \int_0^t \langle \zeta^s, dW(s)\rangle\ .$$

Suppose $\tau \le t_0$ is a bounded stopping time such that for some constant $C_0 < \infty$ we have

$$\sup_{0<s<\tau} \{|\beta^s|,\ |a^s|,\ \|\zeta^s\|,\ \|\gamma^s\|\} \le C_0\ .$$

Then, for every $p > 1$, there exists a constant C_{p,t_0} such that

$$\mathbf{P}\Big(\int_0^\tau (b^s)^2\,ds < \varepsilon^{20} \quad\text{and}\quad \int_0^\tau (|a^s|^2 + \|\zeta^s\|^2)\,ds \ge \varepsilon\Big) \le C_{p,t_0}\,(1+C_0^6)^p\varepsilon^p\ ,$$

for every $\varepsilon \le 1$.

Proof. The proof is given in [Nor86], but without the explicit dependence on C_0. If we follow his proof carefully we get an estimate of the type

$$\mathbf{P}\Big(\int_0^\tau (b^s)^2\,ds < \varepsilon^{10}\ \text{and}\ \int_0^\tau (|a^s|^2 + \|\zeta^s\|^2)\,ds \ge (1+C_0^3)\varepsilon\Big) \le C_1\,(1+C_0^{12})^p\varepsilon^p\ .$$

Replacing ε by ε^2 and making the assumption $\varepsilon < 1/(1+C_0^3)$, we recover our statement. The statement is trivial for $\varepsilon > 1/(1+C_0^3)$, since any probability is always smaller than 1. □

We apply this inequality as follows: Define, for $K_0 \in \mathcal{K}$,

$$a^t(x) = \langle V_{\rm L}^t([X^0, K_0]_{\rm L}^t)(x), h\rangle\ ,$$
$$b^t(x) = \langle V_{\rm L}^t K_0^t(x), h\rangle\ ,$$
$$\beta^t(x) = \langle V_{\rm L}^t([X^0, [X^0, K_0]_{\rm L}]_{\rm L}^t)(x), h\rangle\ ,$$
$$(\gamma^t)^i(x) = \langle V_{\rm L}^t([Q_i, [X^0, K_0]_{\rm L}]_{\rm L}^t)(x), h\rangle\ ,$$
$$(\zeta^t)^i(x) = \langle V_{\rm L}^t([Q_i, K_0]_{\rm L}^t)(x), h\rangle\ .$$

In this expression, $\zeta^t(x) \in \mathcal{H}$, $(\zeta^t)^i(x) = \langle \zeta^t(x), e_i\rangle$ and similarly for the γ^t. It is clear by Proposition 7.4, Eq.(7.7), and Eq.(7.8) that the assumptions of Lemma 7.11 are satisfied with $C_0 = \mathcal{O}(\nu)$ for some $\nu > 0$.

We continue the proof that (B) implies (C) in the case when $K \in \mathcal{K}_j$, with $j = 1$. Then, by the construction of $\mathcal{K}_j$ with $j \ge 1$, there is a $K_0 \in \mathcal{K}_{j-1}$ such that we have either $K = [Q_i, K_0]_{\rm L}$ for some $i \in \{k_*, \dots, L-1\}$, or $K = [X^0, K_0]_{\rm L}$. In fact, for $j = 1$ only the second case occurs and $K_0 = Q_i$ for some i, but we are already preparing an inductive step. Applying Lemma 7.11, we have for every $\varepsilon \le 1$:

$$\mathbf{P}\Big(\int_0^\tau \langle V_{\rm L}^s K_0^s(x), h\rangle^2\,ds < \varepsilon \quad\text{and}\quad \int_0^\tau \Big(\langle V_{\rm L}^s [X^0, K_0]_{\rm L}^s(x), h\rangle^2$$

$$+ \sum_{i=0}^{\infty} \langle V_{\rm L}^s [Q_i, K_0]_{\rm L}^s(x), h\rangle^2 \Big)\, ds \geq \varepsilon^{1/20} \Big) \leq \mathcal{O}(6\nu p)\varepsilon^{p/20} \;.$$

Since the second integral above is always larger than $\int_0^\tau \langle V_{\rm L}^t K^t(x), h\rangle^2\, dt$, the probability for it to be smaller than $\varepsilon^{1/20}$ is, by (B), bounded by $\mathcal{O}(16p) t^{-p}\varepsilon^{p/20}$. This implies (replacing ν by $\max\{6\nu, 16\}$) that

$$\mathbf{P}\Big(\int_0^\tau \langle V_{\rm L}^s K_0^s(x), h\rangle^2\, ds < \varepsilon\Big) \leq \mathcal{O}(\nu p) t^{-p}\varepsilon^{p/20} \;.$$

Since for $j = 1$ we have $K_0 = Q_i$ with $i \in \{k_*, \ldots, L-1\}$, we have shown (C) in this case. The above reasoning is repeated for $j = 2$ and $j = 3$, by iterating the above argument. For example, if $K = [Q_{i_1}, [X^0, Q_{i_2}]_{\rm L}]_{\rm L}$,with $i_1, i_2 \in \{k_*, \ldots, L-1\}$, we apply Lemma 7.11 twice, showing the first time that $\langle [X^0, Q_{i_2}]_{\rm L}, h\rangle^2$ is unlikely to be small and then again to show that $\langle Q_{i_2}, h\rangle^2$ is also unlikely to be small (with other powers of ε), which is what we wanted. Finally, since every K used in (B) is in $\mathcal{K}$, at most 3 such invocations of Lemma 7.11 will be sufficient to conclude that (C) holds. The proof of Theorem 7.3 is complete. □

7.4 Estimates on the low-frequency derivatives (Proof of Proposition 5.3)

Having proven the crucial bound Theorem 7.1 on the reduced Malliavin matrix, we can now proceed to prove Proposition 5.3, *i.e.*, the smoothing properties of the dynamics in the low-frequency part. For convenience, we restate it here.

Proposition 7.12 *There exist exponents $\mu, \nu > 0$ such that for every $\varphi \in \mathcal{C}_b^2(\mathcal{H})$, every $\xi \in \mathcal{H}^\alpha$ and every $T > 0$, one has*

$$\Big\|\mathbf{E}\Big((D_{\rm L}\varphi \circ \Phi^t)(\xi)(D_{\rm L}\Phi_{\rm L}^t)(\xi)\Big)\Big\| \leq C_T\, t^{-\mu}(1 + \|\xi\|_\alpha^\nu)\|\varphi\|_{\rm L^\infty} \;, \tag{7.9}$$

for all $t \in (0, T]$.

Proof. The proof will use the integration by parts formula (6.4) together with Theorem 7.1. Fix $\xi \in \mathcal{H}^\alpha$ and $t > 0$. In this proof, we omit the argument ξ to gain legibility, but it will be understood that the formulas do generally only hold if evaluated at some $\xi \in \mathcal{H}^\alpha$. We extend our phase space to include $D_{\rm L}\Phi^t$, $V_{\rm L}^t$ and $\mathscr{D}_v\Phi_{\rm L}^t$. We define a new stochastic process Ψ^t by

$$\Psi^t = (\Phi^t, \mathscr{D}_v\Phi_{\rm L}^t, D_{\rm L}\Phi^t, V_{\rm L}^t) \in \widetilde{\mathcal{H}} = \mathcal{H} \oplus \mathbf{R}^{L\cdot L} \oplus \mathcal{H}^L \oplus \mathbf{R}^{L\cdot L} \;.$$

Applying the definitions of these processes, we see that Ψ^t is defined by the autonomous SDE given by

$$\begin{aligned}
d\Phi^t &= -A\Phi^t\, dt + F(\Phi^t)\, dt + Q(\Phi^t)\, dW(t) \;,\\
dD_{\rm L}\Phi^t &= -AD_{\rm L}\Phi^t\, dt + DF(\Phi^t)D_{\rm L}\Phi^t\, dt + DQ(\Phi^t)D_{\rm L}\Phi^t\, dW(t) \;,\\
dV_{\rm L}^t &= V_{\rm L}^t A_{\rm L}\, dt - V_{\rm L}^t\, D_{\rm L}F_{\rm L}(\Phi^t)\, dt - V_{\rm L}^t D_{\rm L}Q_{\rm L}(\Phi^t)\, dW_{\rm L}(t)\\
&\quad + V_{\rm L}^t \sum_{i=0}^{L-1} (D_{\rm L}Q_{\rm L}^i(\Phi^t))^2\, dt \;,\\
d\mathscr{D}_v\Phi_{\rm L}^t &= -A_{\rm L}\mathscr{D}_v\Phi_{\rm L}^t\, dt + D_{\rm L}F_{\rm L}(\Phi^t)\mathscr{D}_v\Phi_{\rm L}^t\, dt + Q_{\rm L}(\Phi^t)^2(V_{\rm L}^t)^*\, dt
\end{aligned}$$

$$+ D_{\rm L}Q_{\rm L}(\Phi^t)\,\mathscr{D}_v\Phi^t_{\rm L}\,dW_{\rm L}(t)\;.$$

This expression will be written in the short form

$$d\Psi^t = -\widetilde{A}\Psi^t\,dt + \widetilde{F}(\Psi^t)\,dt + \widetilde{Q}(\Psi^t)\,dW(t)\;,$$

with $\Psi^t \in \widetilde{\mathcal{H}}$ and $dW(t)$ the cylindrical Wiener process on $\mathcal{H}$. It can easily be verified that this equation satisfies assumption **A1** of Proposition 6.1. We consider again the stochastic process $v^t \in \mathcal{H}$ defined in (6.10). It is clear from Lemma 6.5 that v^t satisfies **A2**. With this particular choice of v, the first component of $\mathscr{D}_v\Psi^t$ (the one in $\mathcal{H}$) is equal to $\mathscr{D}_v\Phi^t_{\rm L}\oplus 0$.

We choose a function $\varphi \in \mathcal{C}^2_b(\mathcal{H})$ and fix two indices $i,k \in \{0,\ldots,L-1\}$. Define $\tilde\varphi_{i,k} : \widetilde{\mathcal{H}} \to \mathbf{R}$ by

$$\tilde\varphi_{i,k}(\Psi^t) = \sum_{j=0}^{L-1}\varphi(\Phi^t)((\mathscr{D}_v\Phi^t_{\rm L})^{-1})_{i,j}(D_{\rm L}\Phi^t_{\rm L})_{j,k}\;,$$

where the inverse has to be understood as the inverse of a square matrix. By Theorem 7.1, $\tilde\varphi_{i,k}$ satisfies the assumptions of Proposition 6.1. A simple computation gives for every $h \in \mathbf{R}^L$ the identity:

$$\begin{aligned} D\tilde\varphi_{i,k}(\Psi^t)\mathscr{D}_v\Psi^t h &= D_{\rm L}\varphi(\Phi^t)(\mathscr{D}_v\Phi^t_{\rm L}h)((\mathscr{D}_v\Phi^t_{\rm L})^{-1})_{i,j}(D_{\rm L}\Phi^t_{\rm L})_{j,k} \\ &\quad + \varphi(\Phi^t)((\mathscr{D}_v\Phi^t_{\rm L})^{-1}(\mathscr{D}^2_v\Phi^t_{\rm L}h)(\mathscr{D}_v\Phi^t_{\rm L})^{-1})_{i,j}(D_{\rm L}\Phi^t_{\rm L})_{j,k} \\ &\quad + \varphi(\Phi^t)((\mathscr{D}_v D_{\rm L}\Phi^t_{\rm L})h)_{i,j}((\mathscr{D}_v\Phi^t_{\rm L})^{-1})_{j,k}\;, \end{aligned} \tag{7.10}$$

where summation over j is implicit. We now apply the integration by parts formula in the form of Proposition 6.1. This gives the identity

$$\mathbf{E}(D\tilde\varphi_{i,k}(\Psi^t)\mathscr{D}_v\Psi^t h) = \mathbf{E}\Big(\tilde\varphi_{i,k}(\Psi^t)\int_0^t\langle v^s h, dW(s)\rangle\Big)\;.$$

Substituting (7.10), we find

$$\begin{aligned} \mathbf{E}\Big(D_{\rm L}\varphi(\Phi^t)(\mathscr{D}_v\Phi^t_{\rm L}h)((\mathscr{D}_v\Phi^t_{\rm L})^{-1})_{i,j}(D_{\rm L}\Phi^t_{\rm L})_{j,k}\Big) &= \\ &- \mathbf{E}\Big(\varphi(\Phi^t)((\mathscr{D}_v\Phi^t_{\rm L})^{-1}(\mathscr{D}^2_v\Phi^t_{\rm L}h)(\mathscr{D}_v\Phi^t_{\rm L})^{-1})_{i,j}(D_{\rm L}\Phi^t_{\rm L})_{j,k}\Big) \\ &- \mathbf{E}\Big(\varphi(\Phi^t)((\mathscr{D}_v D_{\rm L}\Phi^t_{\rm L})h)_{i,j}((\mathscr{D}_v\Phi^t_{\rm L})^{-1})_{j,k}\Big) \\ &+ \mathbf{E}\Big(\varphi(\Phi^t)((\mathscr{D}_v\Phi^t_{\rm L})^{-1})_{i,j}(D_{\rm L}\Phi^t_{\rm L})_{j,k}\int_0^t\langle v^s h, dW(s)\rangle\Big)\;. \end{aligned}$$

The summation over the index j is implicit in every term. We now choose $h = e_i$ and sum over the index i. The left-hand side is then equal to

$$\mathbf{E}\Big((D_{\rm L}\varphi(\Phi^t))D_{\rm L}\Phi^t_{\rm L}e_k\Big)\;,$$

which is precisely the expression we want to bound. The right-hand side can be bounded in terms of $\|\varphi\|_{\rm L^\infty}$ and of $\mathbf{E}((\mathscr{D}_v\Phi^t_{\rm L})^{-4})$ (at worst). The other factors are all given by components of $\mathscr{D}_v\Psi^t$ and can therefore be bounded by means of Theorem 8.9. Therefore, (7.9) follows. The proof of Proposition 7.12 is complete. □

8 Existence Theorems

In this section, we prove existence theorems for several PDE's and SDE's, in particular we prove Proposition 5.1 and Lemma 5.4. Much of the material here relies on well-known techniques, but we include the details for completeness.

We consider again the problem

$$d\Phi^t = -A\Phi^t\,dt + F(\Phi^t)\,dt + Q(\Phi^t)\,dW(t)\;, \tag{8.1}$$

with $\Phi^0 = \xi$ given. The initial condition ξ will be taken in one of the Hilbert spaces $\mathcal{H}^\gamma$. We will show that, after some time, the solution lies in some smaller Hilbert space. Note that we are working here with the *cutoff* equations, but we omit the index ϱ.

We will of course require that all stochastic processes are predictable. This means that if we write $\mathrm{L}^p(\Omega, \mathscr{Y})$, with $\mathscr{Y}$ some Banach space of functions of the interval $[0, T]$, we really mean that the only functions we consider are those that are measurable with respect to the predictable σ-field when considered as functions over $\Omega \times [0, T]$.

We first state precisely what is known about the ingredients of (8.1).

Lemma 8.1 *The following properties hold for A, F and Q.*

P1 *The space $\mathcal{H}$ is a real separable Hilbert space and $A : \mathcal{D}(A) \to \mathcal{H}$ is a self-adjoint strictly positive operator.*

P2 *The map $F : \mathcal{H} \to \mathcal{H}$ has bounded derivatives of all orders.*

P3 *For every $\gamma \ge 0$, F maps $\mathcal{H}^\gamma$ into itself. Furthermore, there exists a constant $n > 0$ independent of γ and constants $C_{F,\gamma}$ such that F satisfies the bounds*

$$\|F(x)\|_\gamma \le C_{F,\gamma}(1 + \|x\|_\gamma)\;, \tag{8.2a}$$

$$\|F(x) - F(y)\|_\gamma \le C_{F,\gamma}\|x - y\|_\gamma(1 + \|x\|_\gamma + \|y\|_\gamma)^n\;, \tag{8.2b}$$

for all x and y in $\mathcal{H}^\gamma$.

P4 *There exists an $\alpha > 0$ such that for every $x, x_1, x_2 \in \mathcal{H}$ the map $Q : \mathcal{H} \to \mathscr{L}(\mathcal{H}, \mathcal{H})$ satisfies*

$$\|A^{\alpha-3/8}Q(x)\|_{\mathrm{HS}} \le C\;, \quad \|A^{\alpha-3/8}(Q(x_1) - Q(x_2))\|_{\mathrm{HS}} \le C\|x_1 - x_2\|\;,$$

where $\|\cdot\|_{\mathrm{HS}}$ denotes the Hilbert-Schmidt norm in $\mathcal{H}$.

P5 *The derivative of Q satisfies*

$$\|A^\alpha(DQ(x))h\|_{\mathrm{HS}} \le C\|h\|\;, \tag{8.3}$$

for every $x, h \in \mathcal{H}$.

P6 *The derivative of F satisfies*

$$\|(DF(x))y\|_\gamma \le C(1 + \|x\|_\gamma)\|y\|_\gamma\;,$$

for every $x, y \in \mathcal{H}^\gamma$.

Proof. The points **P1**, **P2** are obvious. The point **P4** follows from the definition (1.6) of Q and the construction of Q_ϱ in (4.1). To prove **P3**, recall that the map $F = F_\varrho$ of the GL equation is of the type

$$F_\varrho(u) = \chi(\|u\|/(3\varrho))P(u) ,$$

with P some polynomial and $\chi \in \mathcal{C}_0^\infty(\mathbf{R})$. The key point is to notice that the estimate

$$\|uv\|_\gamma \le C_\gamma(\|u\|\,\|v\|_\gamma + \|u\|_\gamma\|v\|)$$

holds for every $\gamma \ge 0$, where uv denotes the multiplication of two functions. In particular, we have

$$\|u^n\|_\gamma \le C\|u\|_\gamma\|u\|^{n-1} ,$$

which, together with the fact that χ has compact support, shows (8.2a). This also shows that the derivatives of F in $\mathcal{H}^\gamma$ are polynomially bounded and so (8.2b) holds. **P6** follows by the same argument.

The point **P5** immediately follows from the fact that the image of the operator $(DQ(x))h$ is contained in $\mathcal{H}_\mathrm{L}$ for every $x, h \in \mathcal{H}$. □

Remark 8.2 The condition **P1** implies that e^{-At} is an analytic semigroup of contraction operators on $\mathcal{H}$. We will use repeatedly the bound

$$\|e^{-At}x\|_\gamma \;\le\; C_\gamma t^{-\gamma}\|x\| \;.$$

We begin the study of (8.1) by considering the equation for the mild solution

$$\begin{aligned}\Psi(t,\xi,\omega) \;=\; e^{-At}\xi + \int_0^t e^{-A(t-s)}F(\Psi(s,\xi,\omega))\,ds \\ + \int_0^t e^{-A(t-s)}Q(\Psi(s,\xi,\omega))\,dW(s,\omega) \;.\end{aligned} \tag{8.4}$$

The study of this equation is in several steps. We will consider first the noise term, then the equation for a fixed instance of ω, and finally prove existence and bounds.

We need some more notation:

Definition 8.3 Let $\mathcal{H}^\alpha$ be as above the domain of A^α with the graph norm. We fix, once and for all, a maximal time T. We denote by $\mathcal{H}^\alpha_T$ the space $\mathcal{C}([0,T],\mathcal{H}^\alpha)$ equipped with the norm

$$\|y\|_{\mathcal{H}^\alpha_T} = \sup_{t\in[0,T]} \|y(t)\|_\alpha \;.$$

We write $\mathcal{H}_T$ instead of $\mathcal{H}^0_T$.

8.1 The noise term

Let $y \in \mathrm{L}^p(\Omega,\mathcal{H}_T)$. (One should think of y as being $y(t) = \Phi^t$.) The noise term in (8.4) will be studied as a function on $\mathrm{L}^p(\Omega,\mathcal{H}_T)$. It is given by the function Z defined as

$$(Z(y))(\omega) = t \mapsto \int_0^t e^{-A(t-s)}Q(y(\omega)(s))\,dW(s,\omega) \;. \tag{8.5}$$

We will show that $Z(y)$ is in $\mathrm{L}^p(\Omega, \mathcal{H}_T^\alpha)$ when y is in $\mathrm{L}^p(\Omega, \mathcal{H}_T)$. The natural norm here is the L^p norm defined by

$$|||Z(y)|||_{\mathcal{H}_T^\alpha, p} = \left(\mathbf{E}_\omega \sup_{t\in[0,T]} \|(Z(y))_t(\omega)\|_\alpha^p\right)^{1/p} .$$

Proposition 8.4 *Let $\mathcal{H}$, A and Q be as above and assume* **P1** *and* **P4** *are satisfied. Then, for every $p \geq 1$ and every $T < T_0$ one has*

$$|||Z(y)|||_{\mathcal{H}_T^\alpha, p} \;\leq\; C_{T_0} T^{p/16} . \tag{8.6}$$

Proof. Choose an element $y \in \mathrm{L}^p(\Omega, \mathcal{H}_T)$. In the sequel, we will consider y as a function over $[0, T] \times \Omega$ and we will not write explicitly the dependence on Ω.

In order to get bounds on Z, we use the factorization formula and the Young inequality. Choose $\gamma \in (1/p, 1/8)$. The factorization formula [DPZ92b] then gives the equality

$$(Z(y))(t) = C\int_0^t (t-s)^{\gamma-1} e^{-A(t-s)} \int_0^s (s-r)^{-\gamma} e^{-A(s-r)} Q(y(r))\,dW(r)\,ds .$$

Since A commutes with e^{-At}, the Hölder inequality leads to

$$\begin{aligned} &\|(Z(y))(t)\|_\alpha^p \\ &= C\Big\|\int_0^t (t-s)^{\gamma-1} e^{-A(t-s)} \int_0^s (s-r)^{-\gamma} A^\alpha e^{-A(s-r)} Q(y(r))\,dW(r)\,ds\Big\|^p \\ &\leq Ct^\nu \int_0^t \Big\|\int_0^s (s-r)^{-\gamma} A^\alpha e^{-A(s-r)} Q(y(r))\,dW(r)\Big\|^p\,ds , \end{aligned} \tag{8.7}$$

with $\nu = (p\gamma - 1)/(p-1)$. For the next bound we need the following result:

Lemma 8.5 *[DPZ92b, Thm. 7.2]. Let $r \mapsto \Psi^r$ be an arbitrary predictable $\mathscr{L}^2(\mathcal{H})$-valued process. Then, for every $p \geq 2$, there exists a constant C such that*

$$\mathbf{E}\Big(\Big\|\int_0^s \Psi^r\,dW(r)\Big\|^p\Big) \leq C\mathbf{E}\Big(\int_0^s \|\Psi^r\|_{\mathrm{HS}}^2\,dr\Big)^{p/2} .$$

This lemma, the Young inequality applied to (8.7), and **P4** above imply

$$\begin{aligned} |||Z(y)|||_{\mathcal{H}_T^\alpha, p}^p &= \mathbf{E}\Big(\sup_{0\leq t\leq T}\Big\|\int_0^t A^\alpha e^{-A(t-s)} Q(y(s))\,dW(s)\Big\|^p\Big) \\ &\leq CT^\nu \mathbf{E}\int_0^T \Big\|\int_0^s (s-r)^{-\gamma} A^\alpha e^{-A(s-r)} Q(y(r))\,dW(r)\Big\|^p\,ds \\ &\leq CT^\nu \mathbf{E}\int_0^T \Big(\int_0^s (s-r)^{-2\gamma} \|A^\alpha e^{-A(s-r)} Q(y(r))\|_{\mathrm{HS}}^2\,dr\Big)^{p/2}\,ds \\ &\leq CT^\nu \mathbf{E}\int_0^T \Big(\int_0^s (s-r)^{-2\gamma} \|A^{3/8} e^{-A(s-r)}\|^2 \|A^{\alpha-3/8} Q(y(r))\|_{\mathrm{HS}}^2\,dr\Big)^{p/2}\,ds \\ &\leq CT^\nu \mathbf{E}\int_0^T \Big(\int_0^s (s-r)^{-2\gamma-3/4} \|A^{\alpha-3/8} Q(y(r))\|_{\mathrm{HS}}^2\,dr\Big)^{p/2}\,ds \end{aligned}$$

$$\begin{aligned}&\le CT^\nu\Big(\int_0^T s^{-2\gamma-3/4}\,ds\Big)^{p/2}\mathbf{E}\int_0^T \|A^{\alpha-3/8}Q(y(s))\|_{\rm HS}^p\,ds\\ &\le CT^{1+\nu}\Big(\int_0^T s^{-2\gamma-3/4}\,ds\Big)^{p/2}\;,\end{aligned}\tag{8.8}$$

provided $\gamma < 1/8$. We choose $\gamma = 1/16$ (which thus imposes the condition $p > 16$), and we find

$$|\!|\!|Z(y)|\!|\!|^p_{\mathcal{H}^\alpha_T,p} \le CT_0^{1+\nu}T^{p/16}\;.$$

Thus, we have shown (8.6) for $p > 16$. Since we are working in a probability space the case of $p \ge 1$ follows. This completes the proof of Proposition 8.4. □

8.2 A deterministic problem

The next step in our study of (8.4) is the analysis of the problem for a *fixed* instance of the noise ω. Then (8.4) is of the form

$$h(t,\xi,z) = e^{-At}\xi + \int_0^t e^{-A(t-s)}F(h(s,\xi,z))\,ds + z(t)\;,$$

where we assume that $z \in \mathcal{H}^\alpha_T$. One should think of this as an instance of $Z(\Phi)$, but at this point of our proof, the necessary bounds are not yet available.

We find it more convenient to study instead of h the quantity g defined by $g(t,\xi,z) = h(t,\xi,z) - z(t)$. Then g satisfies

$$g(t,\xi,z) = e^{-At}\xi + \int_0^t e^{-A(t-s)}F(g(s,\xi,z) + z(s))\,ds\;. \tag{8.9}$$

We consider the solution (assuming it exists) as a map from the initial condition ξ and the deterministic noise term z. More precisely, we define

$$G(\xi,z)_t \;=\; g(t,\xi,z)\;.$$

This is a map defined on $\mathcal{H} \times \mathcal{H}^\alpha_T$. Clearly, (8.9) reads:

$$G(\xi,z)_t \;=\; e^{-At}\xi + \int_0^t e^{-A(t-s)}F(G(\xi,z)_s + z(s))\,ds\;. \tag{8.10}$$

To formulate the bounds on G, we need some more spaces that take into account the regularizing effect of the semigroup $t \mapsto e^{-At}$.

Definition 8.6 For $\gamma \ge 0$ the spaces $\mathcal{G}^\gamma_T$ are defined as the closures of $\mathcal{C}([0,T],\mathcal{H}^\gamma)$ under the norm

$$\|y\|_{\mathcal{G}^\gamma_T} = \sup_{t\in(0,T]} t^\gamma\|y(t)\|_\gamma + \sup_{t\in[0,T]}\|y(t)\|\;.$$

Note that

$$\|y\|_{\mathcal{G}^\gamma_T} \;\le\; C_{\gamma,T}\|y\|_{\mathcal{H}^\gamma_T}\;.$$

With these definitions, one has:

Proposition 8.7 *Assume the conditions* **P1–P4** *are satisfied. Assume* $\xi \in \mathcal{H}$ *and* $z \in \mathcal{H}_T^\alpha$*. Then, there exists a map* $G : \mathcal{H} \times \mathcal{H}_T^\alpha \to \mathcal{H}_T$ *solving (8.10). One has the following bounds:*

(A) *If* $\xi \in \mathcal{H}^\gamma$ *with* $\gamma \le \alpha$ *one has for every* $T > 0$ *the bound*

$$\|G(\xi, z)\|_{\mathcal{H}_T^\gamma} \le C_T(1 + \|\xi\|_\gamma + \|z\|_{\mathcal{H}_T^\gamma}) . \tag{8.11}$$

(B) *If* $\xi \in \mathcal{H}$ *one has for every* $T > 0$ *the bound*

$$\|G(\xi, z)\|_{\mathcal{G}_T^\alpha} \le C_T(1 + \|\xi\| + \|z\|_{\mathcal{H}_T^\alpha}) . \tag{8.12}$$

Before we start with the proof proper we note the following regularizing bound: Define

$$(\mathcal{N} f)(t) \;=\; \int_0^t e^{-A(t-s)} f(s)\, ds \; . \tag{8.13}$$

Then one has:

Lemma 8.8 *For every* $\varepsilon \in [0, 1)$ *and every* $\gamma > \varepsilon$ *there is a constant* $C_{\varepsilon,\gamma}$ *such that*

$$\|\mathcal{N} f\|_{\mathcal{G}_T^\gamma} \;\le\; C_{\varepsilon,\gamma} T \|f\|_{\mathcal{G}_T^{\gamma-\varepsilon}} \;,$$

for all $f \in \mathcal{G}_T^{\gamma-\varepsilon}$.

Proof. We start with

$$\begin{aligned}
\|(\mathcal{N} f)(t)\|_\gamma \;&\le\; \int_0^{t/2} \|A^\gamma e^{-A(t-s)} f(s)\|\, ds + \int_{t/2}^t \|A^\varepsilon e^{-A(t-s)} A^{\gamma-\varepsilon} f(s)\|\, ds \\
&\le\; \int_0^{t/2} (t-s)^{-\gamma} \|f(s)\|\, ds + \int_{t/2}^t (t-s)^{-\varepsilon} \|f(s)\|_{\gamma-\varepsilon}\, ds \\
&\le\; \int_0^{t/2} (t-s)^{-\gamma} \|f\|_{\mathcal{G}_T^{\gamma-\varepsilon}}\, ds + \int_{t/2}^t (t-s)^{-\varepsilon} s^{\varepsilon-\gamma} \|f\|_{\mathcal{G}_T^{\gamma-\varepsilon}}\, ds \\
&\le\; C t^{1-\gamma} \|f\|_{\mathcal{G}_T^{\gamma-\varepsilon}} + C t^{1-\varepsilon} t^{\varepsilon-\gamma} \|f\|_{\mathcal{G}_T^{\gamma-\varepsilon}} \; .
\end{aligned}$$

Therefore, $t^\gamma \|(\mathcal{N} f)(t)\|_\gamma \le CT \|f\|_{\mathcal{G}_T^{\gamma-\varepsilon}}$. Similarly, we have

$$\|(\mathcal{N} f)(t)\| \;\le\; \int_0^t \|e^{-A(t-s)} f(s)\|\, ds \;\le\; Ct \|f\|_{\mathcal{G}_T^{\gamma-\varepsilon}} \; .$$

Combining the two inequalities, the result follows. □

Proof of Proposition 8.7. We first choose an initial condition $\xi \in \mathcal{H}^\gamma$ and a function $z \in \mathcal{H}_T^\gamma$. The local existence of the solutions in $\mathcal{H}^\gamma$ is a well-known result. Thus there exists, for a possibly small time $\widetilde{T} > 0$, a function $u \in \mathcal{C}([0, \widetilde{T}], \mathcal{H}^\gamma)$ satisfying

$$u(t) = e^{-At}\xi + \int_0^t e^{-A(t-s)} F(u(s) + z(s))\, ds \; .$$

In order to get an *a priori* bound on $\|u(t)\|_\gamma$ we use assumption **P3** and find

$$\begin{aligned}\|u(t)\|_\gamma &\leq \|\xi\|_\gamma + C_{F,\gamma}\int_0^t (1+\|u(s)+z(s)\|_\gamma)\,ds\\ &\leq C(1+\|\xi\|_\gamma + \|z\|_{\mathcal{H}_T^\gamma}) + C_{F,\gamma}\int_0^t \|u(s)\|_\gamma\,ds\;.\end{aligned}$$

By Gronwall's lemma we get for $t < T$,

$$\|u(t)\|_\gamma \leq C_T(1+\|\xi\|_\gamma + \|z\|_{\mathcal{H}_T^\gamma})\;. \tag{8.14}$$

Note that (8.14) tells us that if the initial condition ξ is in $\mathcal{H}^\gamma$ and if z is in $\mathcal{H}_T^\gamma$, then $u(t)$ is, for small enough t, again in $\mathcal{H}^\gamma$ with the above bound. Therefore, we can iterate the above reasoning and show the global existence of the solutions up to time T, with bounds. Thus, G is well-defined and satisfies the bound (8.11).

We turn to the proof of the estimate (8.12). Define for $z \in \mathcal{H}_T$ the map $\mathcal{M}_z$ by

$$(\mathcal{M}_z(x))(t) = e^{-At}\xi + \int_0^t e^{-A(t-s)}F(x(s)+z(s))\,ds\;. \tag{8.15}$$

Taking $\xi \in \mathcal{H}$ we see from (8.14) with $\gamma = 0$ that there exists a fixed point u of $\mathcal{M}_z$ which satisfies

$$\|u\|_{\mathcal{H}_T} \;=\; \sup_{t\in[0,T]}\|u(t)\| \;\leq\; C_T(1+\|\xi\|+\|z\|_{\mathcal{H}_T})\;.$$

Assume next that $z \in \mathcal{H}_T^\alpha$ and hence *a fortiori* $z \in \mathcal{G}_T^\alpha$. Then, by **P3** one has

$$\|F(x+z)\|_{\mathcal{G}_T^\gamma} \;\leq\; C(1+\|x\|_{\mathcal{G}_T^\gamma} + \|z\|_{\mathcal{G}_T^\gamma})\;.$$

Since u is a fixed point and (8.15) contains a term of the form of (8.13) we can apply Lemma 8.8 and obtain for every $\gamma \leq \alpha$ and $\varepsilon \in [0,1)$:

$$\begin{aligned}\|u\|_{\mathcal{G}_T^{\gamma+\varepsilon}} \;&=\; \|\mathcal{M}_z(u)\|_{\mathcal{G}_T^{\gamma+\varepsilon}} \;\leq\; C\|\xi\| + CT\|F(u+z)\|_{\mathcal{G}_T^\gamma}\\ &\leq\; C\|\xi\| + C_T(1+\|u\|_{\mathcal{G}_T^\gamma} + \|z\|_{\mathcal{G}_T^\gamma})\;.\end{aligned} \tag{8.16}$$

Thus, as long as $\|z\|_{\mathcal{G}_T^\gamma}$ is finite, we can apply repeatedly (8.16) until reaching $\gamma = \alpha$, and this proves (8.12). The proof of Proposition 8.7 is complete. □

8.3 Stochastic differential equations in Hilbert spaces

Before we can start with the final steps of the proof of Proposition 5.1 we state in the next subsection a general existence theorem for stochastic differential equations in Hilbert spaces. The symbol $\mathcal{H}$ denotes a separable Hilbert space. We are interested in solutions to the SDE

$$dX^t = (-AX^t + N(t,\omega,X^t) + M^t)\,dt + B(t,\omega,X^t)\,dW(t)\;, \tag{8.17}$$

where $W(t)$ is the cylindrical Wiener process on a separable Hilbert space $\mathcal{H}_0$. We assume $B(t,\omega,X^t) : \mathcal{H}_0 \to \mathcal{H}$ is Hilbert-Schmidt. We will denote by Ω the underlying probability space and by $\{\mathscr{F}_t\}_{t\geq 0}$ the associated filtration.

The exact conditions spell out as follows:

C1 The operator $A : \mathcal{D}(A) \to \mathcal{H}$ is the generator of a strongly continuous semigroup in $\mathcal{H}$.

C2 There exists a constant $C > 0$ such that for arbitrary $x, y \in \mathcal{H}$, $t \geq 0$ and $\omega \in \Omega$ the estimates

$$\|N(t,\omega,x) - N(t,\omega,y)\| + \|B(t,\omega,x) - B(t,\omega,y)\|_{\mathrm{HS}} \leq C\|x-y\| \;,$$
$$\|N(t,\omega,x)\|^2 + \|B(t,\omega,x)\|_{\mathrm{HS}}^2 \leq C^2(1+\|x\|^2) \;,$$

hold.

C3 For every $x, h \in \mathcal{H}$ and $h_0 \in \mathcal{H}_0$, the stochastic processes $\langle N(\cdot,\cdot,x), h\rangle$ and $\langle B(\cdot,\cdot,x)h_0, h\rangle$ are predictable.

C4 The $\mathcal{H}$-valued stochastic process M^t is predictable, has continuous sample paths, and satisfies

$$\sup_{t\in[0,T]} \mathbf{E}\, \|M^t\|^p < \infty \;,$$

for every $T > 0$ and every $p \geq 1$.

C5 For arbitrary $t > 0$ and $\omega \in \Omega$, the maps $x \mapsto N(t,\omega,x)$ and $x \mapsto B(t,\omega,x)$ are twice continuously differentiable with their derivatives bounded by a constant independent of t, x and ω.

We have the following existence theorem.

Theorem 8.9 *Assume that $\xi \in \mathcal{H}$ and that* **C1 – C4** *are satisfied.*

- *For any $T > 0$, there exists a mild solution X_ξ^t of (8.17) with $X_\xi^0 = \xi$. This solution is unique among the $\mathcal{H}$-valued processes satisfying*

$$\mathbf{P}\biggl(\int_0^T \|X_\xi^t\|^2\, dt < \infty\biggr) \;=\; 1 \;.$$

  Furthermore, X_ξ has a continuous version and is strongly Markov.

- *For every $p \geq 1$ and $T > 0$, there exists a constant $C_{p,T}$ such that*

$$\mathbf{E}(\sup_{t\in[0,T]} \|X_\xi^t\|^p) \leq C_{p,T}(1 + \|\xi\|^p) \;. \tag{8.18}$$

- *If, in addition,* **C5** *is satisfied, the mapping $\xi \mapsto X_\xi^t(\omega)$ has a.s. bounded partial derivatives with respect to the initial condition ξ. These derivatives satisfy the SDE's obtained by formally differentiating (8.17) with respect to X.*

Proof. The proof of this theorem for the case $M^t \equiv 0$ can be found in [DPZ96]. The same proof carries through for the case of non-vanishing M^t satisfying **C4**. □

8.4 Bounds on the cutoff dynamics (Proof of Proposition 5.1)

With the tools from stochastic analysis in place, we can now prove Proposition 5.1. We start with the

Proof of (A). In this case we identify the equation (8.17) with (4.2) and apply Theorem 8.9. The condition **C1** of Theorem 8.9 is obviously true, and the condition **C3** is redundant in this case. The condition **C2** is satisfied because F and Q of (8.17) satisfy **P2–P4**. Therefore, (8.18) holds and hence we have shown (5.1a) for the case of $\gamma = 0$. In particular, Φ_ϱ^t exists and satisfies

$$\begin{aligned}\Phi_\varrho^t(\xi,\omega) \;=\; e^{-At}\xi + \int_0^t e^{-A(t-s)} F(\Phi_\varrho^s(\xi,\omega))\,ds \\ + \int_0^t e^{-A(t-s)} Q(\Phi_\varrho^s(\xi,\omega))\,dW(s)\;.\end{aligned} \tag{8.19}$$

We can extend (5.1a) to arbitrary $\gamma \le \alpha$ as follows. We set as in (8.5),

$$(Z(\Phi_\varrho))_t(\omega) \;=\; \int_0^t e^{-A(t-s)} Q(\Phi_\varrho^s(\xi,\omega))\,dW(s)\;. \tag{8.20}$$

By Proposition 8.4, we find that for all $p \ge 1$ one has

$$\Bigl(\mathbf{E}_\omega \sup_{t\in[0,T]} \|(Z(\Phi_\varrho))_t(\omega)\|_\alpha^p\Bigr)^{1/p} < C_{T,p} \tag{8.21}$$

for all ξ. From this, we conclude that, almost surely,

$$\sup_{t\in[0,T]} \|(Z(\Phi_\varrho))_t(\omega)\|_\alpha < \infty\;. \tag{8.22}$$

Subtracting (8.20) from (8.19) we get

$$\begin{aligned}\Phi_\varrho^t(\xi,\omega) - (Z(\Phi_\varrho))_t(\omega) \;&=\; e^{-At}\xi + \int_0^t e^{-A(t-s)} F(\Phi_\varrho^s(\xi,\omega))\,ds \\ =\; e^{-At}\xi + \int_0^t &e^{-A(t-s)} F\Bigl(\Phi_\varrho^s(\xi,\omega) - (Z(\Phi_\varrho))_s(\omega) + (Z(\Phi_\varrho))_s(\omega)\Bigr)\,ds\;.\end{aligned} \tag{8.23}$$

Comparing (8.23) with (8.10) we see that, a.s.,

$$\Phi_\varrho^t(\xi,\omega) - (Z(\Phi_\varrho))_t(\omega) = G(\xi, Z(\Phi_\varrho(\xi,\cdot))(\omega))\;.$$

We now use z as a shorthand:

$$z(t) \;=\; \Bigl(Z(\Phi_\varrho(\xi,\cdot))\Bigr)_t(\omega)\;.$$

Assume now $\xi \in \mathcal{H}^\gamma$. Note that by (8.22), $z(t)$ is in $\mathcal{H}^\alpha$. If $\gamma \le \alpha$, we can apply Proposition 8.7 and from (8.11) we conclude that almost surely,

$$\sup_{t\in[0,T]} \|G(\xi,z)\|_\gamma \le C_T(1 + \|\xi\|_\gamma + \sup_{t\in[0,T]} \|z\|_\gamma)\;.$$

Finally, since $\gamma \le \alpha$, we find

$$\mathbf{E}\Bigl(\sup_{t\in[0,T]} \|\Phi_\varrho^t(\xi)\|_\gamma^p\Bigr) \;\le\; C\mathbf{E}\Bigl(\sup_{t\in[0,T]} \|G(\xi,z)_t\|_\gamma^p\Bigr) + C\mathbf{E}\Bigl(\sup_{t\in[0,T]} \|z(t)\|_\gamma^p\Bigr)$$

$$\begin{aligned}&\leq C_{T,p}(1 + \|\xi\|_\gamma)^p + C\mathbf{E}\Big(\sup_{t\in[0,T]} \|z(t)\|_\gamma^p\Big)\\&\leq C_{T,p}(1 + \|\xi\|_\gamma)^p\;,\end{aligned}\tag{8.24}$$

where we applied (8.21) to get the last inequality. Thus, we have shown (5.1a) for all $\gamma \le \alpha$. The fact that the solution is strong if $\gamma \ge 1$ is an immediate consequence of [Lun95, Lemma 4.1.6] and [DPZ92b, Thm. 5.29].

Proof of (B). This bound can be shown in a similar way, using the bound (8.12) of Proposition 8.7: Take $\xi \in \mathcal{H}$. By the above, we know that there exists a solution to (8.19) satisfying the bound (5.1b) with $\gamma = 0$. We define $z(t)$ and $G(\xi, z)_t$ as above. But now we apply the bound (8.12) of Proposition 8.7 and we conclude that almost surely,

$$\sup_{t\in[0,T]} t^\alpha \|G(\xi, z)\|_\alpha \le C_T(1 + \|\xi\| + \sup_{t\in[0,T]} \|z\|_\alpha)\;.$$

Following a procedure similar to (8.24), we conclude that (5.1b) holds.

Proof of (C). The existence of the partial derivatives follows from Theorem 8.9. To show the bound, choose $\xi \in \mathcal{H}$ and $h \in \mathcal{H}$ with $\|h\| = 1$, and define the process $\Psi^t = (D\Phi^t_\varrho(\xi))h$. It is by Theorem 8.9 a mild solution to the equation

$$d\Psi^t = -A\Psi^t\,dt + \Big((DF \circ \Phi^t_\varrho)(\xi)\Psi^t\Big)\,dt + \Big((DQ \circ \Phi^t_\varrho)(\xi)\Psi^t\Big)\,dW(t)\;.\tag{8.25}$$

By **P3** and **P5**, this equation satisfies conditions **C1–C3** of Theorem 8.9, so we can apply it to get the desired bound (5.1c). (The constant term drops since the problem is linear in h.)

Proof of (D). Choose $h \in \mathcal{H}$ and $\xi \in \mathcal{H}^\alpha$ and define as above $\Psi^t = (D\Phi^t_\varrho(\xi))h$, which is the mild solution to (8.25) with initial condition h. We write this as

$$\begin{aligned}\Psi^t &= e^{-At}h + \int_0^t e^{-A(t-s)}\Big((DF \circ \Phi^s_\varrho)(\xi)\Psi^s\Big)\,ds\\&\quad+ \int_0^t e^{-A(t-s)}\Big((DQ \circ \Phi^s_\varrho)(\xi)\Psi^s\Big)\,dW(s)\\&\equiv S_1^t + S_2^t + S_3^t\;.\end{aligned}$$

The term S_1^t satisfies

$$\sup_{t\in(0,T]} t^\alpha \|S_1^t\|_\alpha \le C_T\|h\|\;.\tag{8.26}$$

The term S_3^t is very similar to what is found in (8.5), with $Q(y(s))$ replaced by $(DQ \circ \Phi^s_\varrho)\Psi^s$. Repeating the steps of (8.8) for a sufficiently large p, we obtain now with $\gamma = \frac{1}{4}$, some $\mu > 0$ and writing $X^s = (DQ \circ \Phi^s_\varrho)(\xi)\Psi^s$:

$$\begin{aligned}\mathbf{E}\sup_{t\in[0,T]} \|S_3^t\|_\alpha^p &= \mathbf{E}\Big(\sup_{0\le t\le T}\Big\|\int_0^t A^\alpha e^{-A(t-s)}X^s\,dW(s)\Big\|^p\Big)\\&\le CT^\mu\mathbf{E}\int_0^T\Big\|\int_0^s (s-r)^{-\gamma}A^\alpha e^{-A(s-r)}X^r\,dW(r)\Big\|^p\,ds\end{aligned}$$

$$\leq CT^{\mu}\mathbf{E}\int_0^T\Big(\int_0^s(s-r)^{-2\gamma}\|A^{\alpha}e^{-A(s-r)}X^r\|_{\text{HS}}^2\,dr\Big)^{p/2}\,ds$$
$$\leq CT^{\mu}\mathbf{E}\int_0^T\Big(\int_0^s(s-r)^{-2\gamma}\|A^{\alpha}X^r\|_{\text{HS}}^2\,dr\Big)^{p/2}\,ds$$
$$\leq CT^{\mu}\Big(\int_0^T s^{-2\gamma}\,ds\Big)^{p/2}\mathbf{E}\int_0^T\|A^{\alpha}X^s\|_{\text{HS}}^p\,ds$$
$$\leq CT^{\mu+p/4}\mathbf{E}\int_0^T\|A^{\alpha}(DQ\circ\Phi_{\varrho}^s)(\xi)\Psi^s\|_{\text{HS}}^p\,ds\;.$$

We now use **P5**, *i.e.*, (8.3) and then (5.1c) and get

$$\mathbf{E}\sup_{t\in[0,T]}\|S_3^t\|_{\alpha}^p\;\leq\;CT^{\mu+p/4}\mathbf{E}\int_0^T\|\Psi^s\|^p\,ds\;\leq\;CT^{\mu+p/4+1}\|h\|^p\;. \tag{8.27}$$

To treat the term S_2^t, we fix a realization $\omega\in\Omega$ of the noise and use Lemma 8.8. This gives for $\varepsilon\in[0,1)$ the bound

$$\sup_{t\in(0,T]}t^{\gamma}\|S_2^t\|_{\gamma}\leq CT\sup_{t\in(0,T]}t^{\gamma-\varepsilon}\|(DF\circ\Phi_{\varrho}^t)(\xi)\Psi^t\|_{\gamma-\varepsilon}\;.$$

By **P6**, this leads to the bound, a.s.,

$$\sup_{t\in(0,T]}t^{\gamma}\|S_2^t\|_{\gamma}\leq C_T\Big(1+\sup_{t\in(0,T]}\|\Phi_{\varrho}^t(\xi)\|_{\gamma-\varepsilon}\Big)\sup_{t\in(0,T]}t^{\gamma-\varepsilon}\|\Psi^t\|_{\gamma-\varepsilon}\;.$$

Taking expectations we have

$$\mathbf{E}\sup_{t\in(0,T]}t^{\gamma p}\|S_2^t\|_{\gamma}^p\leq C_T^p\mathbf{E}\bigg(\Big(1+\sup_{t\in(0,T]}\|\Phi_{\varrho}^t(\xi)\|_{\gamma-\varepsilon}\Big)^p\sup_{t\in(0,T]}t^{(\gamma-\varepsilon)p}\|\Psi^t\|_{\gamma-\varepsilon}^p\bigg)\;.$$

By the Schwarz inequality and (5.1a) we get

$$\mathbf{E}\sup_{t\in(0,T]}t^{\gamma p}\|S_2^t\|_{\gamma}^p\leq C_{T,p}(1+\|\xi\|_{\gamma-\varepsilon}^p)\Big(\mathbf{E}\sup_{t\in(0,T]}t^{(\gamma-\varepsilon)2p}\|\Psi^t\|_{\gamma-\varepsilon}^{2p}\Big)^{1/2}\;. \tag{8.28}$$

Since $\Psi^t=(D\Phi_{\varrho}^t(\xi))h=S_1^t+S_2^t+S_3^t$, combining (8.26)–(8.28) leads to

$$\mathbf{E}\sup_{t\in(0,T]}t^{\gamma p}\|(D\Phi_{\varrho}^t(\xi))h\|_{\gamma}^p$$
$$\leq C_{T,p}\|h\|^p+C_{T,p}(1+\|\xi\|_{\gamma-\varepsilon}^p)\Big(\mathbf{E}\sup_{t\in(0,T]}t^{(\gamma-\varepsilon)2p}\|(D\Phi_{\varrho}^t(\xi))h\|_{\gamma-\varepsilon}^{2p}\Big)^{1/2}\;.$$

Thus, we have gained ε in regularity. Choosing $\varepsilon=\frac{1}{2}$ and iterating sufficiently many times we obtain (5.1d) for sufficiently large p. The general case then follows from the Hölder inequality.

Proof of (E). We estimate this expression by

$$\|\Phi_{\varrho}^t(\xi)-e^{-At}\xi\|_{\gamma}\leq\int_0^t\|F(\Phi_{\varrho}^s(\xi))\|_{\gamma}\,ds+\Big\|\int_0^t e^{-A(t-s)}Q(\Phi_{\varrho}^s(\xi))\,dW(s)\Big\|_{\gamma}\;.$$

The first term can be bounded by combining (5.1b) and **P3**. The second term is bounded by Proposition 8.4.

The proof Proposition 5.1 is complete.

8.5 Bounds on the off-diagonal terms

Here, we prove Lemma 5.4. This is very similar to the proof of (D) of Proposition 5.1.

Proof. We fix $T > 0$ and $p \geq 1$. We start with (5.5b). Recall that here we do not write the cutoff ϱ. We choose $h \in \mathcal{H}_{\mathrm{H}}$ and $\xi \in \mathcal{H}$. The equation for $\Psi^s = (D_{\mathrm{H}}\Phi_{\mathrm{L}}^s(\xi))h$ is :

$$\begin{aligned}\Psi^s &= \int_0^s e^{-A(s-s')}\Big((DF_{\mathrm{L}} \circ \Phi_\varrho^{s'})(\xi)(D_{\mathrm{H}}\Phi^{s'}(\xi))h\Big)\,ds' \\ &\quad + \int_0^s e^{-A(s-s')}\Big((DQ_{\mathrm{L}} \circ \Phi_\varrho^{s'})(\xi)(D_{\mathrm{H}}\Phi^{s'}(\xi))h\Big)\,dW(s') \\ &\equiv R_1^s + R_2^s\ .\end{aligned}$$

Since $DF = DF_\varrho$ is bounded we get

$$\|R_1^s\| \;\leq\; C\int_0^s \|(D_{\mathrm{H}}\Phi^{s'}(\xi))h\|\,ds' \;\leq\; Cs \sup_{s'\in[0,s]} \|(D_{\mathrm{H}}\Phi^{s'}(\xi))h\|\ .$$

Using (5.1c), this leads to

$$\mathbf{E}\sup_{s\in[0,t]} \|R_1^s\|^p \;\leq\; C^p t^p \mathbf{E}\sup_{s\in[0,t]} \|(D_{\mathrm{H}}\Phi^s(\xi))h\|^p \;\leq\; C_{T,p}t^p\|h\|^p\ .$$

The term R_2^s is bounded exactly as in (8.27). Combining the bounds, (5.5b) follows.

Since Q_{H} is constant, see (4.1), we get for $\Psi^s = (D_{\mathrm{L}}\Phi_{\mathrm{H}}^s(\xi))h$ and $h \in \mathcal{H}_{\mathrm{L}}$:

$$\Psi^s = \int_0^s e^{-A(s-s')}\Big((DF_{\mathrm{H}} \circ \Phi_\varrho^{s'})(\xi)(D_{\mathrm{L}}\Phi^{s'}(\xi))h\Big)\,ds'\ .$$

This is bounded like R_1^s and leads to (5.5a). This completes the proof of Lemma 5.4. □

8.6 Proof of Proposition 2.3

Here we point out where to find the general results on (1.7) which we stated in Proposition 2.3. Note that these are bounds on the flow *without* cutoff ϱ.

Proof of Proposition 2.3. There are many ways to prove this. To make things simple, without getting the best estimate possible, we note that a bound in L^∞ can be found in [Cer99, Prop. 3.2]. To get from L^∞ to $\mathcal{H}$, we note that $\xi \in \mathcal{H}$ and we use (1.7) in its integral form. The term $e^{-At}\xi$ is bounded in $\mathcal{H}$, while the non-linear term $\int_0^t e^{-A(t-s)}F(\Phi^s(\xi))\,ds$ can be bounded by using a version of Lemma 8.8. Finally, the noise term is bounded by Proposition 8.4.

Furthermore, because of the compactness of the semigroup generated by A, it is possible to show [DPZ96, Thm. 6.3.5] that an invariant measure exists. □

Acknowledgements

We thank L. Rey-Bellet and G. Ben-Arous for helpful discussions. We also thank the referee for insightful remarks which clarified some ambiguities of an earlier version. This research was partially supported by the Fonds National Suisse.

IV. Exponential Mixing for a Stochastic PDE Driven by Degenerate Noise

Abstract

We study stochastic partial differential equations of the reaction-diffusion type. We show that, even if the forcing is very degenerate (*i.e.*has not full rank), one has exponential convergence towards the invariant measure. The convergence takes place in the topology induced by a weighted variation norm and uses a kind of (uniform) Doeblin condition.

1 Model and Result

We consider the stochastic partial differential equation given by

$$du = \partial_\xi^2 u \, dt - \mathsf{P}(u) \, dt + Q \, dW(t) \;, \qquad u \in W^{(1,2)}_{\text{per}}([0,1]) \;. \tag{SGL}$$

In this equation, P is a polynomial of odd degree with positive leading coefficient and deg $\mathsf{P} \geq 3$, dW is the cylindrical Wiener process on $\mathcal{H} \equiv W^{(1,2)}_{\text{per}}([0,1])$, and $Q : \mathcal{H} \to \mathcal{H}$ is a compact operator which is diagonal in the trigonometric basis. The symbol $\xi \in [0,1]$ denotes the spatial variable. Further conditions on the spectrum of Q will be made precise below.

In a recent paper [EH01b], to which we also refer for further details about the model, it was shown that this equation possesses a unique invariant measure and satisfies the Strong Feller property. However, the question of the rate of convergence towards the invariant measure was left open. The aim of this paper is to show that this rate is exponential.

There is a fair amount of very recent literature about closely related questions, mainly concerning ergodic properties of the 2D Navier-Stokes equation. To the author's knowledge, the main results are exposed in the works of Kuksin and Shirikyan [KS00, KS01], Bricmont, Kupiainen and Lefevere [BKL00c, BKL00b], and E, Mattingly and Sinai [EMS01, Mat01], although the problem goes back to Flandoli and Maslowski [FM95]. The main differences between the model exposed here and the above papers is that we want to consider a situation where the *unstable* modes are *not* forced, whereas the forcing only acts onto the stable modes and is transmitted to the whole system through the nonlinearity. From this point of view, we are in a hypoelliptic situation where Hörmander-type conditions apply [Hör67, Hör85], as opposed to the essentially elliptic situation where the unstable modes are all forced and the (infinitely many) other modes are stabilized by the linear part of the equation.

Returning to the model (SGL), we denote by q_k the eigenvalue of Q corresponding to the kth trigonometric function (ordered in such a way that $k > 0$). We make the following assumption on the q_k:

Assumption 1.1 There exist constants $k_* > 0$, $C_1 > 0$, $C_2 > 0$, $\alpha \geq 2$ and $\beta \in (\alpha - 1/8, \alpha]$ such that

$$C_1 k^{-2\alpha} \leq q_k \leq C_2 k^{-2\beta} \;, \quad \text{for } k > k_*. \tag{1.1}$$

There are no assumptions on q_k for $k \leq k_*$, in particular one may have $q_k = 0$ in that region. Furthermore, k_* can be chosen arbitrarily large.

We denote by $\Phi_t(u)$ the solution of (SGL) at time t with initial condition $u \in \mathcal{H}$. If Φ_t exists and is sufficiently regular, one can define the semigroup $\mathcal{P}^t$ acting on bounded functions φ and the semigroup $\mathcal{P}^t_*$ acting on finite measures μ by

$$(\mathcal{P}^t\varphi)(u) = \mathbf{E}\Big((\varphi \circ \Phi_t)(u)\Big)\;, \qquad (\mathcal{P}^t_*\mu)(A) = \mathbf{E}\Big((\mu \circ \Phi_t^{-1})(A)\Big)\;.$$

In a recent paper [EH01b], to which we also refer for further details about the model, it was shown that the above model satisfies the following.

Theorem 1.2 *Under Assumption 1.1, the solution of (SGL) defines a unique stochastic flow Φ_t on $\mathcal{H}$, thus also defining a Markov semigroup $\mathcal{P}^t$. The semigroup $\mathcal{P}^t$ is Strong Feller and open set irreducible in arbitrarily short time. As a consequence, the semigroup $\mathcal{P}^t_*$ acting on measures possesses a unique invariant measure on $\mathcal{H}$.*

Recall that a semigroup is said "open set irreducible in arbitrarily short time" if the probability of reaching a given open set in a given time is always strictly positive.

We denote by μ_* the unique invariant probability measure of Theorem 1.2. We will show in this paper that for every probability measure μ, we have $\mathcal{P}^t_*\mu \to \mu_*$ and that this convergence takes place with an exponential rate (in time). More precisely, we introduce, for a given (possibly unbounded) Borel function $V : \mathcal{H} \to [1,\infty]$, the *weighted variational norm* defined on every signed Borel measure μ by

$$|\!|\!|\mu|\!|\!|_V \equiv \int_{\mathcal{H}} V(x)\,\mu_+(dx) + \int_{\mathcal{H}} V(x)\,\mu_-(dx)\;,$$

where $\mu_\pm$ denotes the positive (resp. negative) part of μ. When $V(x) = 1$, we recover the usual variational norm which we denote by $|\!|\!|\cdot|\!|\!|$. We also introduce the family of norms $\|\cdot\|_\gamma$ on $\mathcal{H}$ defined by

$$\|x\|_\gamma = \|L^\gamma x\|\;,$$

where L is the differential operator $1 - \partial_\xi^2$ and $\|\cdot\|$ is the usual norm on $\mathcal{H}$, *i.e.*

$$\|u\|^2 = \int_0^1 (|u|^2 + |\partial_\xi u|^2)\,d\xi\;.$$

The exact formulation of our convergence result is

Theorem 1.3 *There exists a constant $\lambda > 0$ such that for every $p \ge 1$, every $\gamma \le \alpha$, and every probability measure μ on $\mathcal{H}$, one has*

$$|\!|\!|\mathcal{P}^t_*\mu - \mu_*|\!|\!|_{V_{\gamma,p}} \le Ce^{-\lambda t}\;, \qquad \textit{with } V_{\gamma,p}(u) = \|u\|_\gamma^p + 1\;,$$

for every $t \ge 1$. The constant C is independent of the probability measure μ.

In the sequel, we will denote by Φ the Markov chain obtained by sampling the solution of (SGL) at integer times and by $\mathcal{P}(x,\,\cdot\,)$ the corresponding transition probabilities. Theorem 1.3 is a consequence of the following features of the model (SGL).

A. We construct a set K having the property that there exists a probability measure ν and a constant $\delta > 0$ such that $\mathcal{P}(x,\,\cdot\,) \geq \delta\,\nu(\,\cdot\,)$ for every $x \in K$. This means that K behaves "almost" like an atom for the Markov chain Φ. This is shown to be a consequence of the Strong Feller property and the irreducibility of the Markov semigroup associated to (SGL).

B. The dynamics has very strong contraction properties in the sense that it reaches some compact set very quickly. In particular, one can bound uniformly from below the transition probabilities to a set K satisfying property *A.*

These conditions yield some strong Doeblin condition and thus lead to exponential convergence results. The intuitive reason behind this is that, for any two initial measures, their image under $\mathcal{P}_*$ has a common part, the amount of which can be bounded uniformly from below and cancels out. This will be clarified in the proof of Proposition 2.1 below.

The remainder of the paper is organized as follows. In Section 2, we show how to obtain Theorem 1.3 from the above properties. The proof will be strongly reminiscent of the standard proof of the Perron-Frobenius theorem. In Section 3 we then show the contraction properties of the dynamics and in Section 4 we show that every compact set has the property *A.*

Acknowledgements

Several proofs in this paper are inspired by the recent work [RBT01] of Luc Rey-Bellet and Larry Thomas. The author is very grateful to them for having communicated their ideas before publication. The author also benefitted from several useful discussions with Jean-Pierre Eckmann and Jacques Rougemont. This work was partially supported by the Fonds National Suisse.

2 A Variant of the Perron-Frobenius Theorem

The following proposition shows, reformulated in a more rigorous way, why the properties *A.* and *B.* yield exponential convergence results towards the invariant measure.

Proposition 2.1 *Let Ψ be a Markov chain on a measurable space X and let Ψ satisfy the following properties:*

a. There exist a measurable set K, a positive constant δ and a probability measure ν_* such that for every measurable set A and every $x \in K$, one has $\mathcal{P}(x, A) \geq \delta\,\nu_*(A)$.

b. There exists a constant $\delta' > 0$ such that $\mathcal{P}(x, K) \geq \delta'$ for every $x \in \mathsf{X}$.

Then Ψ has a unique invariant measure μ_* and one has for every probability measure μ the estimate $|||\mathcal{P}_*^n\mu - \mu_*||| \leq 2(1-\delta\delta')^{-n/2}$.

Proof. The first observation we make is that for every probability measure μ one has by property *a.*,

$$(\mathcal{P}_*\mu)(K) = \int_{\mathsf{X}} \mathcal{P}(x, K)\,\mu(dx) \geq \delta' \;.$$

As a consequence of this and of property *b.*, one has for every measurable set A the bound

$$(\mathcal{P}_*^2\mu)(A) \geq \int_K \mathcal{P}(x, A)\,(\mathcal{P}_*\mu)(dx) \geq \delta\delta'\nu_*(A) \;. \tag{2.1}$$

Define the constant $\varepsilon = \delta\delta'$. An immediate consequence of (2.1) is that for any probability measure μ, one has

$$\|\mathcal{P}_*^2\mu - \varepsilon\nu_*\| = 1 - \varepsilon \;.$$

Now take any two probability measures μ and ν. Denote by $\eta_\pm$ the positive (resp. negative) part of $\mu - \nu$. Since μ and ν are probability measures, one has $\|\eta_+\| = \|\eta_-\| = \Delta$, say. Then, since $\mathcal{P}_*$ preserves probability, one has

$$\begin{aligned}\|\mathcal{P}_*^2\mu - \mathcal{P}_*^2\nu\| &= \|\mathcal{P}_*^2\eta_+ - \mathcal{P}_*^2\eta_-\| \le \|\mathcal{P}_*^2\eta_+ - \Delta\varepsilon\nu_*\| + \|\mathcal{P}_*^2\eta_- - \Delta\varepsilon\nu_*\| \\ &\le 2\Delta(1-\varepsilon) = (1-\varepsilon)\|\mu - \nu\| \;.\end{aligned}$$

This completes the proof of Proposition 2.1. □

Theorem 1.3 is then an easy consequence of the following lemmas.

Lemma 2.2 *For every $\gamma \le \alpha$, every $t > 0$, and every $p \ge 1$, there exists a constant $C_{\gamma,p,t}$ such that for every finite measure μ on $\mathcal{H}$ one has*

$$\|\mathcal{P}_*^t\mu\|_{V_{\gamma,p}} \le C_{\gamma,p,t}\|\mu\| \;, \tag{2.2}$$

with $\mathcal{P}_*^t$ the semigroup acting on measures solving (SGL).

Lemma 2.3 *For every compact set $K \subset \mathcal{H}$, there exists a probability measure ν_* and a constant $\delta > 0$ such that $\mathcal{P}(x,\,\cdot\,) \ge \delta\,\nu_*(\,\cdot\,)$ for every $x \in K$.*

Proof of Theorem 1.3. Fix once and for all $\gamma \le \alpha$ and $p \ge 1$. By Lemma 2.2, there exist constants C and δ such that the set $K = \{x \in \mathcal{H} \mid \|x\|_\gamma \le C\}$ satisfies $\mathcal{P}(x, K) \ge \delta$ for every $x \in \mathcal{H}$. By Lemma 2.3, we can apply Proposition 2.1 to find

$$\|\mathcal{P}_*^n\mu - \mu^*\| \le 2e^{-\lambda n} \;,$$

for some $\lambda > 0$ and for n any integer. Since $\mathcal{P}_*^t$ preserves positivity and probability, one immediately gets the same estimate for arbitrary real times. By Lemma 2.2 and the invariance of μ_*, this yields for some constant C,

$$\|\mathcal{P}_*^{t+1}\mu - \mu^*\|_{V_{\gamma,t}} \le Ce^{-\lambda t} \;.$$

The proof of Theorem 1.3 is complete. □

Remark 2.4 Writing V instead of $V_{\gamma,p}$, condition (2.2) is equivalent to the statement that $\mathbf{E}_x V(\Phi) \le C$ for all $x \in \mathcal{H}$. It is also possible to achieve exponential convergence results if this condition is replaced by the weaker condition that

$$\mathbf{E}_x V(\Phi) \le \begin{cases} cV(x) & \text{for } x \in \mathcal{H} \setminus K, \\ \Lambda & \text{for } x \in K, \end{cases} \tag{2.3}$$

with $c \in (0,1)$, $\Lambda > 0$ and K some compact set. The proof is somewhat lengthy and so we do not give it here. The interested reader is referred to [MT94, RBT01]. The difference in the results is that one gets an estimate of the type

$$\|\mathcal{P}_*^{t+1}\mu - \mu^*\|_V \le Ce^{-\lambda t}\|\mu\|_V \;.$$

So strong convergence towards the invariant measure holds for measures with finite $\|\cdot\|_V$-norm and not necessarily for every probability measure.

The remainder of the paper is devoted to the proof of Lemmas 2.2 and 2.3.

3 Contraction Properties of the Dynamics

This section is devoted to the proof of Lemma 2.2. We reformulate it in a more convenient way as

Proposition 3.1 *For every $p \geq 1$, every $\gamma \leq \alpha$, and every time $t > 0$, there is a constant $C_{p,t,\gamma} > 0$ such that, for every $x \in \mathcal{H}$, one has*

$$\mathbf{E}(\|\Phi_t(x)\|_\gamma^p) \leq C_{p,t,\gamma} \; . \tag{3.1}$$

Proof. We define the linear operator $L = 1 - \partial_\xi^2$ and the stochastic convolution

$$W_L(t) = \int_0^t e^{-L(t-s)}\, Q\, dW(s) \; .$$

With these notations, the solution of (SGL) reads

$$\Phi_t(x) = e^{-Lt}x - \int_0^t e^{-L(t-s)} \mathsf{P}(\Phi_s(x))\, ds + W_L(t) \; . \tag{3.2}$$

In a first step, we show that for every couple of times $0 < t_1 < t_2$, there exists a constant C_{p,t_1,t_2} independent of the initial condition x such that

$$\mathbf{E}\Big(\sup_{t_1<s<t_2} \|\Phi_s(x)\|_{\mathrm{L}^\infty}\Big) \leq C_{p,t_1,t_2} \; . \tag{3.3}$$

For this purpose, we introduce the auxiliary process $\Psi_t(x)$ defined by $\Psi_t(x) = \Phi_t(x) - W_L(t)$. We have for Ψ_t the equation

$$\Psi_t(x) = e^{-Lt}x - \int_0^t e^{-L(t-s)} \mathsf{P}(\Psi_s(x) + W_L(s))\, ds \; ,$$

i.e.$\Psi_t(x)$ can be interpreted pathwise as the solution of the PDE

$$\dot\Psi_t = -L\Psi_t - \mathsf{P}(\Psi_t + W_L(t)) \; , \quad \Psi_0 = x \; . \tag{3.4}$$

If we denote by q the degree of P (remember that $q \geq 3$), we have, thanks to the dissipativity of L, the inequality

$$\frac{D^-\|\Psi_t\|_{\mathrm{L}^\infty}}{Dt} \leq c_1 - c_2\|\Psi_t\|_{\mathrm{L}^\infty}^q + c_3\|W_L(t)\|_{\mathrm{L}^\infty}^q \; , \tag{3.5}$$

where the c_i are some strictly positive constants and D^-/Dt denotes the left lower Dini derivative. An elementary computation allows to verify that the solutions of the ordinary differential equation $\dot y = -cy^q + f(t)$ (with positive initial condition and $f(s) > 0$) satisfy the inequality

$$y(t) \leq (qct)^{-1/(q-1)} + \int_0^t f(s)\, ds \; , \tag{3.6}$$

independently of the initial condition. Standard estimates on Gaussian processes show furthermore that for every $t > 0$ and every $p \geq 1$, there exists a constant $C_{p,t}$ such that

$$\mathbf{E}\big(\sup_{s\in[0,t]} \|W_L(s)\|_{\mathrm{L}^\infty}^p\big) \leq C_{p,t} \; .$$

Combining this with (3.6), we get (3.3).

It remains to exploit the dissipativity of the linear operator L and the local boundedness of the nonlinearity to get the desired bound (3.1). We write for $s \in [t/2, t]$ the solution of (SGL) as

$$\Phi_s(x) = e^{-L(s-t/4)}\Phi_{t/4}(x) - \int_{t/4}^s e^{-L(s-r)}\mathsf{P}(\Phi_r(x))\,dr + \int_{t/4}^s e^{-L(s-r)}Q\,dW(r)\;.$$

Note that the last term of this equality has the same probability distribution as $W_L(s - t/4)$. Since $\|e^{-Lt}x\| \le t^{-1/2}\|x\|_{\mathrm{L}^\infty}$, we have (remember that $q = \deg \mathsf{P}$):

$$\begin{aligned}\mathbf{E}\Big(\sup_{t/2<s<t}\|\Phi_s(x)\|^p\Big) &\le C_{p,t} + C_{p,t}\mathbf{E}\big(\sup_{t/4<s<3t/4}\|W_L(s)\|^p\big)\\ &\quad + C\mathbf{E}\Big(\sup_{t/2<s<t}\Big(\int_{t/4}^s (s-r)^{-1/2}\|\mathsf{P}(\Phi_r(x))\|_{\mathrm{L}^\infty}\,dr\Big)^p\Big)\\ &\le C_{p,t} + C_{p,t}\mathbf{E}\Big(\sup_{t/4<s<t}\|\Phi_s(x)\|_{\mathrm{L}^\infty}^{pq}\Big) \le C_{p,t}\;.\end{aligned}$$

In these inequalities, we used (3.3) and the fact that $\mathbf{E}(\sup_{s\in[0,t]}\|W_L(s)\|_\gamma^p)$ is finite for every $\gamma \le \alpha$, every $t > 0$ and every $p \ge 1$. This technique can be iterated, using the fact that $\|e^{-Lt}x\|_{\gamma+1/2} \le t^{-1/2}\|x\|_\gamma$, until one obtains the desired estimate (3.1). The proof of Proposition 3.1 is complete. □

4 Strong Feller Chains and Small Sets

The aim of this section is to show that a sufficient condition for the existence of sets with the property *a.* of Proposition 2.1 is that the Markov chain is open set irreducible and has the Strong Feller property.

We follow closely [MT94] in our definitions. The main difference with their results is that we drop the assumption of local compactness of the topological base space and that our estimates hold globally with respect to the initial condition. We will adopt the following notations:

The symbol X stands for an arbitrary Polish space, *i.e.*a complete, separable metric space. The symbol Φ stands for a Markov chain on X. We denote by $\mathcal{P}(x, A)$ the transition probabilities of Φ. The m-step transition probabilities are denoted by $\mathcal{P}^m(x, A)$. The symbol $\mathscr{B}(\mathsf{X})$ stands for the Borel σ-field of X.

Definition 4.1 A set $K \in \mathscr{B}(\mathsf{X})$ is called *small* if there exists an integer $m > 0$, a probability measure ν on X, and a constant $\delta > 0$ such that $\mathcal{P}^m(x, A) \ge \delta\nu(A)$ for every $x \in K$ and every $A \in \mathscr{B}(\mathsf{X})$. If we want to emphasize the value of m, we call a set m*-small*.

With this definition, we reformulate Lemma 2.3 as

Theorem 4.2 *If Φ is irreducible and Strong Feller, every compact set is 2-small.*

The main step towards the proof of Theorem 4.2 is to show the existence of small sets which are sufficiently big to be "visible" by the dynamics. Recall that a set A is said to be *accessible* if $\mathcal{P}(x, A) > 0$ for every $x \in \mathsf{X}$. One has,

Proposition 4.3 *If Φ is irreducible and Strong Feller, there exist accessible small sets.*

Proof of Theorem 4.2. Recall that Doob's theorem guarantees the existence of a probability measure μ_0 such that the transition probabilities $\mathcal{P}(x,\,\cdot\,)$ are all equivalent to μ_0. This is a consequence of the Strong Feller property and the irreducibility of Φ.

By Proposition 4.3 there exists a small set A such that $\mu_0(A) > 0$. For every $x \in \mathsf{X}$ and every arbitrary $D \in \mathscr{B}(\mathsf{X})$, we then have

$$\mathcal{P}^{m+1}(x,D) \geq \int_A \mathcal{P}(y,D)\,\mathcal{P}^m(x,dy) \geq \mathcal{P}(x,A)\inf_{y\in A}\mathcal{P}(y,D) \geq \delta\mathcal{P}(x,A)\nu(D)\;,$$

for some $m > 0$, $\delta > 0$ and a probability measure ν. Since, by the Strong Feller property, the function $x \mapsto \mathcal{P}(x,A)$ is continuous and, by the accessibility of A, it is positive, there exists for every compact set $C \subset \mathsf{X}$ a constant $\delta' > 0$ such that

$$\inf_{x\in C}\mathcal{P}^{m+1}(x,D) \geq \delta'\nu(D)\;.$$

The proof of Theorem 4.2 is complete. □

The next subsection is devoted to the proof of Proposition 4.3.

4.1 Existence of accessible small sets

In this subsection, we will work with partitions of X. We introduce the following notation: if $\mathscr{P}$ is a partition of X, we denote by $\mathscr{P}(x)$ the (only) element of $\mathscr{P}$ that contains x. With this notation, one has the following theorem, a proof of which can be found *e.g.*in [Doo53, p. 344].

Theorem 4.4 (Basic Differentiation Theorem) *Let $(\mathsf{X},\mathscr{F},\mu)$ be a probability space and $\mathscr{P}_n$ be an increasing sequence of finite measurable partitions of X such that the σ-field generated by $\bigcup_n \mathscr{P}_n$ is equal to $\mathscr{F}$. Let ν be a probability measure on X which is absolutely continuous with respect to μ with density function h. Define the sequence of functions h_n by*

$$h_n(x) = \begin{cases} \dfrac{\nu(\mathscr{P}_n(x))}{\mu(\mathscr{P}_n(x))} & \text{if } \mu(\mathscr{P}_n(x)) > 0,\\ 0 & \text{if } \mu(\mathscr{P}_n(x)) = 0. \end{cases}$$

Then there exists a set N with $\mu(N) = 0$ such that $\lim_{n\to\infty} h_n(x) = h(x)$ *for every $x \in \mathsf{X}\setminus N$.*

This theorem is the main ingredient for the proof of Proposition 4.3.

The first point one notices is that if X is a Polish space, one can explicitly construct a sequence $\mathscr{P}_n$ of partitions that generate the Borel σ-field. Choose a sequence $\{x_i\}_{i=1}^\infty$ of elements which are dense in X (the existence of such a sequence is guaranteed by the separability of X) and a sequence $\{\varepsilon_j\}_{j=1}^\infty$ such that $\varepsilon_j > 0$ and $\lim_{j\to\infty}\varepsilon_j = 0$. Denote by $\mathcal{B}(x,r)$ the open ball of radius r and center x. We then define the sets M_i^j ($i \geq 1$ and $j \geq 0$) by

$$M_i^0 = \mathsf{X}\;,\quad M_i^j = \mathcal{B}(x_i,\varepsilon_j)\;.$$

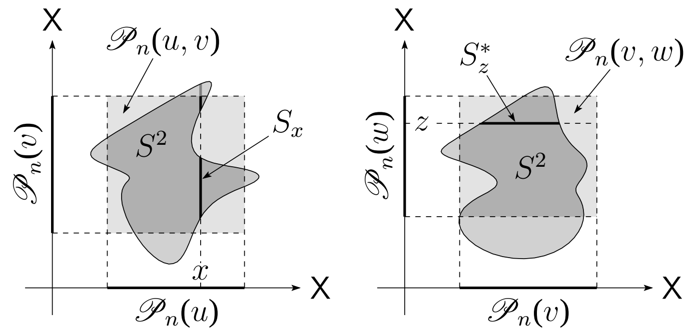

Figure 2: Construction of D and E.

This defines an increasing sequence of finite partitions $\mathscr{P}_n$ by $\mathscr{P}_n = \bigvee_{i,j\le n}\{M_i^j\}$ ($\vee$ denotes the refinement of partitions). We denote by $\mathscr{F}_\infty$ the σ-field generated by $\bigcup_n \mathscr{P}_n$. Since every open set $S \subset \mathsf{X}$ can be written as a countable union

$$S = \bigcup\{M_i^j \mid M_i^j \subset S\}\ ,$$

the open sets belong to $\mathscr{F}_\infty$ and so $\mathscr{F}_\infty = \mathscr{B}(\mathsf{X})$. This construction guarantees the applicability of the Basic Differentiation Theorem to our situation. We are now ready to give the

Proof of Proposition 4.3. Let us denote by $p(x,y)$ a jointly measurable version of the densities of $\mathcal{P}(x,\,\cdot\,)$ with respect to μ_0.

We define for every $x, y \in \mathsf{X}$ the sets $S_x \in \mathscr{B}(\mathsf{X})$ and $S_y^* \in \mathscr{B}(\mathsf{X})$ by

$$S_x = \{y \in \mathsf{X}\,|\,p(x,y) > \tfrac{1}{2}\}\ , \quad S_y^* = \{x \in \mathsf{X}\,|\,p(x,y) > \tfrac{1}{2}\}\ ,$$

and the set $S^2 \in \mathscr{B}(\mathsf{X}\times\mathsf{X})$ by

$$S^2 = \{(x,y) \in \mathsf{X}\times\mathsf{X}\,|\,p(x,y) > \tfrac{1}{2}\}\ .$$

Since $\mathcal{P}(x,\mathsf{X}) = 1$ for every $x \in \mathsf{X}$, one has $\mu_0(S_x) > 0$ for every x and therefore $\mu_0^2(S^2) = \int_{\mathsf{X}} \mu_0(S_x)\,d\mu_0(x) > 0$, where $\mu_0^2 = \mu_0 \times \mu_0$. Define the subset S^3 of X^3 by

$$S^3 = \{(x,y,z) \in \mathsf{X}^3 \mid (x,y) \in S^2 \text{ and } (y,z) \in S^2\}\ . \tag{4.1}$$

One has similarly $\mu_0^3(S^3) = \int_{S^2} \mu_0(S_y)\,d\mu_0^2(x,y) > 0$. Let us now define the sets $\mathscr{P}_n(x)$ as above and define $\mathscr{P}_n(x,y) = \mathscr{P}_n(x) \times \mathscr{P}_n(y)$.

By Theorem 4.4 with $\mu = \mu_0^2$ and $\nu = \mu_0^2|_{S^2}$, there exists a μ_0^2-null set N such that for $(x,y) \in S^2 \setminus N$ one has

$$\lim_{n\to\infty} \frac{\mu_0^2(S^2 \cap \mathscr{P}_n(x,y))}{\mu_0^2(\mathscr{P}_n(x,y))} = 1\ .$$

Since on the other hand $\mu_0^3(S^3) > 0$, there exist a triple (u,v,w) and an integer n such that $\mu_0^2(\mathscr{P}_n(u,v)) > 0$, $\mu_0^2(\mathscr{P}_n(v,w)) > 0$, and

$$\mu_0^2(S^2 \cap \mathscr{P}_n(u,v)) \ge \tfrac{7}{8}\mu_0^2(\mathscr{P}_n(u,v))\ , \tag{4.2a}$$

$$\mu_0^2(S^2 \cap \mathscr{P}_n(v,w)) \geq \tfrac{7}{8}\mu_0^2(\mathscr{P}_n(v,w)) \; . \tag{4.2b}$$

This means that S^2 covers simultaneously seven eights of the "surfaces" of both sets $\mathscr{P}_n(u,v)$ and $\mathscr{P}_n(v,w)$. (See Figure 2 for an illustration of this construction.) As a consequence of (4.2a), the set

$$D = \{x \in \mathscr{P}_n(u) \mid \mu_0(S_x \cap \mathscr{P}_n(v)) \geq \tfrac{3}{4}\mu_0(\mathscr{P}_n(v))\} \; ,$$

satisfies $\mu_0(D) \geq \frac{1}{2}\mu_0(\mathscr{P}_n(u))$. Similarly, the set

$$E = \{z \in \mathscr{P}_n(w) \mid \mu_0(S_z^* \cap \mathscr{P}_n(v)) \geq \tfrac{3}{4}\mu_0(\mathscr{P}_n(v))\} \; ,$$

satisfies $\mu_0(E) \geq \frac{1}{2}\mu_0(\mathscr{P}_n(w))$. On the other hand, one has by the definitions of E and D that for $x \in D$ and $z \in E$, $\mu_0(S_x \cap S_z^*) \geq \frac{1}{2}\mu_0(\mathscr{P}_n(v))$. Thus

$$p^2(x,z) \geq \int_{S_x \cap S_z^*} p(x,y)p(y,z)\,\mu_0(dy) \geq \frac{1}{4}\mu_0(S_x \cap S_z^*) \geq \frac{1}{8}\mu_0(\mathscr{P}_n(v)) \; , \tag{4.3}$$

for every $x \in D$ and every $y \in E$. Defining a probability measure ν by setting $\nu(\Gamma) = \mu_0(\Gamma \cap E)/\mu_0(E)$, there exists $\delta > 0$ such that for every $x \in D$, one has $\mathcal{P}(x,\Gamma) \geq \delta\nu(\Gamma)$ and thus D is small. Since $\mu_0(D) > 0$, the proof of Proposition 4.3 is complete. □

V. Exponential Mixing Properties of Stochastic PDEs Through Asymptotic Coupling

Abstract

We consider parabolic stochastic partial differential equations driven by white noise in time. We prove exponential convergence of the transition probabilities towards a unique invariant measure under suitable conditions. These conditions amount essentially to the fact that the equation transmits the noise to all its determining modes. Several examples are investigated, including some where the noise does *not* act on every determining mode directly.

1 Introduction

We are interested in the study of long-time asymptotics for parabolic stochastic partial differential equations. More precisely, the existence, uniqueness, and speed of convergence towards the invariant measure for such systems is investigated. The general setting is that of a stochastic PDE of the form

$$dx = Ax\,dt + F(x)\,dt + Q\,d\omega(t)\;, \qquad x(0) = x_0\;, \tag{1.1}$$

where x belongs to some Hilbert space $\mathcal{H}$, A is the generator of a C_0-semigroup on $\mathcal{H}$, $F:\mathcal{H} \to \mathcal{H}$ is some nonlinearity, ω is the cylindrical Wiener process on some other Hilbert space $\mathcal{W}$, and $Q:\mathcal{W} \to \mathcal{H}$ is a bounded operator. If the nonlinearity F is sufficiently "nice", there exists a unique solution $x(t)$ to (1.1) (see *e.g.*[DPZ92b]). In this paper, we investigate the asymptotic stability of (1.1). We say that the solutions of (1.1) are asymptotically stable if there exists a *unique* probability measure μ_* on $\mathcal{H}$ such that the laws of $x(t)$ converge to μ_*, independently of the initial condition x_0. We are interested in the situation where the asymptotic stability is a consequence of the noise (*i.e.*the deterministic equation $\dot{x} = Ax + F(x)$ is not asymptotically stable in the above sense), although the noise is weak, in the sense that the range of Q in $\mathcal{H}$ is "small".

The investigation of asymptotic behaviour for solutions of (1.1) goes back to the early eighties (see for example [MS99] for an excellent review article or the monograph [DPZ96] for a detailed exposition). Until recently, two approaches dominated the literature on this subject. For the first approach, sometimes called the "dissipativity method", one considers two solutions $x(t)$ and $y(t)$ of (1.1), corresponding to the same realization of the Wiener process ω, but with different initial conditions x_0 and y_0. If A and F are sufficiently dissipative, $\|x(t) - y(t)\|$ converges to 0 for large times in some suitable sense. If this convergence is sufficiently fast and uniform, it yields asymptotic stability results (see for example [DPZ92a]). Closely related to this approach are the Lyapunov function techniques, developed for semilinear equations in [Ich84]. The dissipativity method, as well as the Lyapunov function techniques, are limited by the requirement that the deterministic equation $\dot{x} = Ax + F(x)$ already shows stable behaviour.

The (linearly) unstable situations are covered by the second approach, to which we refer as the "overlap method". It consists in showing that the Markov transition semigroup associated to (1.1) has the strong Feller property and is topologically irreducible. Then, provided that the equation (1.1) shows some dissipativity, arguments as developed in the monograph [MT94], allow to bound the overlap between transition probabilities starting at two different initial points. This in turn yields strong asymptotic stability properties. The main technical difficulty of this approach is to show that the strong Feller property holds. This difficulty is usually mastered either by studying the infinite-dimensional backward Kolmogorov equation associated to (1.1) [DPZ91], or by showing that the Markov transition semigroup has good smoothing properties [DPEZ95, Cer99]. This technique is limited by the requirement that the noise be sufficiently non-degenerate. A typical requirement is that the range of Q *contains* the domain of some positive power of $-A$. To our knowledge, only one work [EH01b, Hai01] shows the strong Feller property for a stochastic PDE in a situation where the range of Q is not dense in $\mathcal{H}$ (but still of finite codimension).

Very recently, a third approach, to which we refer as the "coupling method", emerged in a series of papers on the 2D Navier-Stokes equation. (See [KS01, Mat01, MY01] and the references in Section 6.) The main idea of these papers is to make a splitting $\mathcal{H} = \mathcal{H}_L \oplus \mathcal{H}_H$ of the dynamics into a finite-dimensional, linearly unstable, low-frequency part $\mathcal{H}_L$ and a remaining infinite-dimensional stable part $\mathcal{H}_H$. An important assumption on Q is then that the range of Q contains $\mathcal{H}_L$. The space $\mathcal{H}_L$ is chosen in such a way that the long-time asymptotics of the dynamics is completely dominated by the behaviour of the low-frequency part. More precisely, for any given realization $x_L(t)$ of the low-frequency part, the dynamics of the high-frequency part $x_H(t)$ will loose memory of its initial condition exponentially fast. On the low-frequency part, in turn, the noise acts in a non-degenerate way. A clever coupling argument allows to combine these two facts in order to obtain asymptotic stability results. The argument consists in coupling two realizations of (1.1) in such a way that if the low-frequency parts meet at some time τ, they remain equal for all times $t > \tau$. (Of course, one has to show that τ is finite with probability 1.) In fact, this coupling method is very close to the Gibbsian approach developed in [KS00, BKL00b, EMS01], which consisted in transforming the infinite-dimensional Markovian system on $\mathcal{H}$ to a finite-dimensional non-Markovian system on $\mathcal{H}_L$. This finite-dimensional system was shown to have exponentially decaying memory and thus techniques from statistical mechanics can be applied.

Loosely speaking, the coupling method combines the arguments of both the dissipativity method (on $\mathcal{H}_H$) and the overlap method (on $\mathcal{H}_L$). The coupling method thus yields a very powerful approach to the problem of asymptotic stability of (1.1). The conditions of applicability of this coupling method have been successively weakened in the aforementioned papers, but the existing results always require, as we already mentioned, that the noise acts directly and independently on *every* determining mode of the equation. In this paper, we extend the coupling method to problems which do not satisfy this condition. Our overall approach is similar to the one exposed by Mattingly in [Mat01], and consequently some of our proofs are closely related to the arguments exposed there. Our main new idea is to construct a coupling for which the low-frequency parts of the dynamics do not actually meet at some finite time, but converge exponentially fast towards each other. This "asymptotic coupling" is achieved through a binding construction exposed in Section 2.3, which seems to be new and can in some cases be implemented even in very degenerate situations.

In the following section, we illustrate the method of asymptotic coupling for a simple finite dimensional problem.

1.1 A toy model

Consider the following system of stochastic differential equations in $\mathbf{R}^2$:

$$\begin{aligned} dx_1 &= (2x_1 + x_2 - x_1^3)\,dt + d\omega(t)\;, \\ dx_2 &= (2x_2 + x_1 - x_2^3)\,dt\;. \end{aligned} \tag{1.2}$$

This equation should be interpreted in the integral sense, with $\omega \in \Omega$ a Brownian motion. Applying Hörmander's condition [Hör85, Nor86], it is easy to see that the transition probabilities of (1.2) are smooth with respect to the Lebesgue measure on $\mathbf{R}^2$. Furthermore, an easy controllability argument shows that they have support everywhere and therefore are all mutually equivalent. Since (1.2) also exhibits a strong drift towards the center of the phase space at large amplitudes, it follows by standard arguments that (1.2) possesses a unique invariant measure μ_* and that every initial condition is exponentially (in variation norm) attracted by μ_*.

The problem with this argument is that it heavily relies on the existence of some reference measure (in this case the Lebesgue measure) which is equivalent to the transition probabilities. In the infinite-dimensional setting, such a reference measure will usually not exist when the noise is sufficiently degenerate. (For an account of some cases where such a reference measure does exist in the infinite-dimensional case, see [MS99, EH01b].) Furthermore, the fact that both directions in (1.2) are linearly unstable prevents one from applying the coupling method as it is presented in the previous section.

We will show that the invariant measure for (1.2) is unique, using a coupling construction which pushes solutions together at an exponential rate. This construction is asymptotic, compared to more conventional coupling constructions, which look for hitting times at which the coupled dynamics actually meets.

Before we proceed, let us explain briefly what we mean by "coupling". A coupling for (1.2) is a process $(x(t), y(t)) \in \mathbf{R}^2 \times \mathbf{R}^2$, whose marginals $x(t)$ and $y(t)$ taken separately are both solutions of (1.2) (but with different initial conditions). In general, one takes a measure P on $\Omega \times \Omega$, whose marginals are both equal to the Wiener measure W. Then a coupling for (1.2) can be constructed by drawing a pair $(\omega, \tilde{\omega}) \in \Omega \times \Omega$ distributed according to P and solving the equations

$$\begin{aligned} dx_1 &= (2x_1 + x_2 - x_1^3)\,dt + d\omega(t)\;, & dy_1 &= (2y_1 + y_2 - y_1^3)\,dt + d\tilde{\omega}(t)\;, \\ dx_2 &= (2x_2 + x_1 - x_2^3)\,dt\;, & dy_2 &= (2y_2 + y_1 - y_2^3)\,dt\;. \end{aligned} \tag{1.3}$$

We will carefully choose the measure P in such a way that the quantity $\|x - y\|$ converges exponentially to 0 for large times. This then yields the uniqueness of the invariant measure for (1.2).

Our main idea leading to the construction of P is to consider the following system in $\mathbf{R}^4$:

$$\begin{aligned} dx_1 &= (2x_1 + x_2 - x_1^3)\,dt + d\omega(t)\;, \\ dx_2 &= (2x_2 + x_1 - x_2^3)\,dt\;, \\ dy_1 &= (2y_1 + y_2 - y_1^3)\,dt + d\omega(t) + G(x_1, x_2, y_1, y_2)\,dt\;, \\ dy_2 &= (2y_2 + y_1 - y_2^3)\,dt\;, \end{aligned} \tag{1.4}$$

where $d\omega$ denotes twice the same realization of the Wiener process. We see that this equation is the same as (1.3) with $\tilde{\omega}$ defined by

$$\tilde{\omega}(t) = \omega(t) + \int_0^t G(x_1(s), x_2(s), y_1(s), y_2(s))\, ds\ . \tag{1.5}$$

The noise $\tilde{\omega} \in \Omega$ is distributed according to some measure $\tilde{\mathsf{W}}$ which is in general not equal to the Wiener measure W. Therefore, (1.4) does not yet define a coupling for (1.2). If G is small in the sense that the quantity

$$\int_0^\infty \|G(x_1(s), x_2(s), y_1(s), y_2(s))\|^2\, ds \tag{1.6}$$

is bounded with sufficiently high probability, then the measures $\tilde{\mathsf{W}}$ and W are equivalent. In this case, it is possible to construct a measure P on $\Omega \times \Omega$ whose marginals are W, with the important property that there exists a random time τ with $\mathbf{P}(\tau < \infty) = 1$ such that the solutions of the coupled system satisfy (1.4) for times $t \geq \tau$.

In view of the above, we have reduced the problem to finding a function G such that the solutions of (1.4) satisfy $\|y(t) - x(t)\| \to 0$ for $t \to \infty$ and (1.6) is bounded. We introduce the difference process $\varrho = y - x$, and we write

$$\dot{\varrho}_1 = 2\varrho_1 + \varrho_2 - \varrho_1(x_1^2 + x_1 y_1 + y_1^2) + G(x, y)\ , \tag{1.7a}$$
$$\dot{\varrho}_2 = 2\varrho_2 + \varrho_1 - \varrho_2(x_2^2 + x_2 y_2 + y_2^2)\ . \tag{1.7b}$$

It is easy to find a function G such that $\varrho_1 \to 0$, but this does not yet mean that ϱ_2 will go to zero. A closer look at (1.7b) shows that if we could force ϱ_1 to be very close to $-3\varrho_2$, (1.7b) could be written as

$$\dot{\varrho}_2 = -\varrho_2 + \varepsilon - \varrho_2(x_2^2 + x_2 y_2 + y_2^2)\ ,$$

which is asymptotically stable. Introduce the function $\zeta = \varrho_1 + 3\varrho_2$. We then have

$$\dot{\zeta} = (\ldots) + G(x_1, x_2, y_1, y_2)\ ,$$

with $(\ldots)$ an expression of the order $\|\varrho\|(1 + \|x\|^2 + \|y\|^2)$. Now we can of course choose $G = -(\ldots) - 2\zeta$. This way, the equation for ζ becomes $\dot{\zeta} = -2\zeta$ and we have the solution $\zeta(t) = \zeta(0)e^{-2t}$. Plugging this into (1.7b), we get

$$\dot{\varrho}_2 = -\varrho_2 + \zeta(0)e^{-2t} - \varrho_2(x_2^2 + x_2 y_2 + y_2^2)\ .$$

We thus have the estimate

$$|\varrho_2(t)| \leq |\varrho_2(0)|e^{-t} + |\zeta(0)|e^{-2t}\ .$$

Finally, ϱ_1 is estimated by using the definition of ζ and we get

$$|\varrho_1(t)| \leq |\varrho_2(0)|e^{-t} + 4|\zeta(0)|e^{-2t}\ .$$

This shows that, with G chosen this way, there exists a constant C such that

$$\|x(t) - y(t)\| \leq C\|x(0) - y(0)\|e^{-t}\ ,$$

for almost every realization of the noise. Since typical realizations of $x(t)$ do not grow faster than linearly, G is also of the order e^{-t}, with at most a polynomial factor in t multiplying the exponential. The main result of this paper, Theorem 4.1, shows that the above construction implies the existence and uniqueness of an invariant probability measure μ_* for the problem at hand. Furthermore, it shows that the transition probabilities converge exponentially fast towards μ_* in the Vaseršteĭn norm (the dual norm to the Lipschitz norm on functions).

This concludes our presentation of the toy model. For a more precise statement, the reader is encouraged to actually check that the above construction allows to verify the assumptions stated in Section 5.

The remainder of this paper is organized as follows. In Section 2, we give the precise definitions for the type of coupling we will consider. In Section 3, we state the properties of the coupling that are required for our purpose. In Section 4, we prove the abstract formulation of our main ergodic theorem. In Section 5, this abstract theorem is then specialized to the case of stochastic differential equations. In Section 6 finally, we present several examples where our construction applies, although the noise does not act directly on every determining mode of the equation.

Acknowledgements

The author thanks Jean-Pierre Eckmann, Guillaume van Baalen, Emmanuel Zabey, and Jacques Rougemont for several helpful discussions and comments. This work was partially supported by the Fonds National Suisse.

2 The Coupling Construction

In this section, we explain our coupling construction. Before we start with the actual definitions of the various objects appearing in the construction, we fix our notations.

2.1 Notations

If μ is a measure on a measurable space X (in the sequel, we will always consider Polish[8] spaces) and $f : \mathsf{X} \to \mathsf{Y}$ is a measurable map, we denote by $f^*\mu$ the measure on Y defined by $(f^*\mu)(A) \equiv \mu(f^{-1}(A))$. For example, if Π is a projection on one component of a product space, $\Pi^*\mu$ denotes the marginal of μ on this component. If a natural reference measure is given on the current space, we denote by $\mathcal{D}\mu$ the density of μ with respect to the reference measure.

We define for any two measures μ and ν the measures $\mu \wedge \nu$ and $\mu \setminus \nu$. If a common reference measure is given, these operations act on densities like

$$(\mathcal{D}(\mu \wedge \nu))(x) = \min\{\mathcal{D}\mu(x) \, , \mathcal{D}\nu(x)\} \; ,$$
$$(\mathcal{D}(\mu \setminus \nu))(x) = \max\{\mathcal{D}\mu(x) - \mathcal{D}\nu(x), 0\} \; .$$

It immediately follows that $\mu = (\mu \wedge \nu) + (\mu \setminus \nu)$ for any two measures μ and ν. We will use the equivalent notations $\mu \le \nu$ and $\nu \ge \mu$ to say that $\mu \wedge \nu = \mu$ holds. One can check the following relations:

$$f^*(\mu \wedge \nu) \le f^*\mu \wedge f^*\nu \; ,$$

[8] *i.e.*complete, separable, and metric

$$f^*(\mu \setminus \nu) \geq f^*\mu \setminus f^*\nu \;.$$

Equalities hold if f is injective.

For a given topological space X, we denote by $\mathscr{M}(\mathsf{X})$ the space of all finite signed Borel measures on X. We denote by $\mathscr{M}_1(\mathsf{X})$ the set of all probability measures on X. For $\mu \in \mathscr{M}(\mathsf{X})$, we denote by $\|\mu\|$ its total variation norm (which is simply its mass if μ has a sign).

2.2 Definition of coupling

In this section, and until the end of the paper, we will often consider families $\mathbf{Q}_y$ of measures indexed by elements $y \in \mathsf{Y}$, with Y some Polish space. One should think of y as the initial condition of a Markov chain on Y and of $\mathbf{Q}_y$ either as its transition probabilities, or as the measure on pathspace obtained by starting from y. We will always assume that the functions $y \mapsto \mathbf{Q}_y(A)$ are measurable for every Borel set A. If $\mathbf{Q}_y$ is a family of measures on Y^n and $\mathbf{R}_y$ is a family of measures on Y^m, a family of measures $(\mathbf{RQ})_y$ on $\mathsf{Y}^{n+m} = \mathsf{Y}^n \times \mathsf{Y}^m$ can be defined on cylindrical sets in a natural way by

$$(\mathbf{RQ})_y(A \times B) = \int_A \mathbf{R}_{z_n}(B)\, \mathbf{Q}_y(dz) \;, \tag{2.1}$$

where $A \subset \mathsf{Y}^n$, $B \subset \mathsf{Y}^m$, and z_n denotes the nth component of z.

We consider a discrete-time Markovian random dynamical system (RDS) Φ on a Polish space X with the following structure. There exists a “one-step” probability space $(\Omega, \mathscr{F}, \mathsf{P})$ and Φ is considered as a jointly measurable map $\Phi : (\mathsf{X}, \Omega) \to \mathsf{X}$. The iterated maps $\Phi^n : (\mathsf{X}, \Omega^n) \to \mathsf{X}$ with $n \in \mathbf{N}$ are constructed recursively by

$$\Phi^n(x, \omega_1, \ldots, \omega_n) = \Phi(\Phi^{n-1}(x, \omega_1, \ldots, \omega_{n-1}), \omega_n) \;,$$

This construction gives rise to a Markov chain on X (also denoted by Φ) with one-step transition probabilities

$$\mathsf{P}_x \equiv \Phi(x, \cdot)^* \mathsf{P} \;.$$

The n-step transition probabilities will be denoted by P_x^n. Our main object of study will be the family of measures on pathspace generated by Φ. Take a sequence $\{\omega_i\}_{i=0}^\infty$ and an initial condition $x \in \mathsf{X}$. We then define $x_0 = x$ and $x_{i+1} = \Phi(x_i, \omega_i)$. We will denote by $\mathbf{P}_x^n$ with $n \in \mathbf{N} \cup \{\infty\}$ the measure on X^n obtained by transporting P^n with the map $\{\omega_i\} \mapsto \{x_i\}$. It is also natural to view $\mathbf{P}_x^n$ as a measure on $\mathsf{X}^n \times \Omega^n$ by transporting P^n with the map $\{\omega_i\} \mapsto \{x_i, \omega_i\}$, so we will use both interpretations.

Remark 2.1 The above setup is designed for the study of stochastic differential equations driven by additive noise. In that case, Ω is some Wiener space and Φ maps an initial condition and a realization of the Wiener process on the solution after time 1. Nevertheless, our setup covers much more general cases.

The coupling needs two copies of the pathspace, *i.e.*we will consider elements $(x, y) \in \mathsf{X}^\infty \times \mathsf{X}^\infty$. It will be convenient to use several projectors from $\mathsf{X}^N \times \mathsf{X}^N$ to its components. We define therefore (for $n \leq N$):

$$\Pi_1 : (x, y) \mapsto x \;, \quad \Pi_2 : (x, y) \mapsto y \;, \quad \pi_n : (x, y) \mapsto (x_n, y_n) \;.$$

We also define $\pi_{i,n} \equiv \Pi_i \circ \pi_n$ for $i \in \{1, 2\}$.

Definition 2.2 Let Φ be a Markov chain on a Polish space X and let $\mathbf{P}_x^\infty$ be the associated family of measures on the pathspace X^∞. A *coupling* for Φ is a family $\mathbf{C}_{x,y}^\infty$ of probability measures on $\mathsf{X}^\infty \times \mathsf{X}^\infty$ satisfying

$$\Pi_1^*\mathbf{C}_{x,y}^\infty = \mathbf{P}_x^\infty \qquad \text{and} \qquad \Pi_2^*\mathbf{C}_{x,y}^\infty = \mathbf{P}_y^\infty \;,$$

where Π_1 and Π_2 are defined as above.

A trivial example of coupling is given by $\mathbf{C}_{x,y}^\infty = \mathbf{P}_x^\infty \times \mathbf{P}_y^\infty$. The interest of constructing a non-trivial coupling comes from the following observation. Take some suitable set of test functions $\mathcal{G}$ on X and define a norm on $\mathscr{M}(\mathsf{X})$ by

$$\|\mu\|_{\mathcal{G}} = \sup_{g\in\mathcal{G}} \langle g, \mu\rangle \;.$$

Once the existence of an invariant measure for the Markov chain Φ is established, one usually wishes to show its uniqueness by proving that Φ forgets about its past sufficiently fast, *i.e.*

$$\lim_{n\to\infty} \|\mathsf{P}_x^n - \mathsf{P}_y^n\|_{\mathcal{G}} = 0 \;, \quad \text{for all} \quad (x,y) \in \mathsf{X}^2 \;,$$

with suitable bounds on the convergence rate as a function of the initial conditions. Now take a coupling $\mathbf{C}_{x,y}^\infty$ for Φ. It is straightforward to see that by definition the equality

$$\langle \mathsf{P}_x^n, g\rangle = \int_{\mathsf{X}\times\mathsf{X}} g(z)\, (\pi_{1,n}^*\mathbf{C}_{x,y}^\infty)(dz)$$

holds, as well as the same equality where $\pi_{1,n}$ is replaced by $\pi_{2,n}$ and P_x^n is replaced by P_y^n. Therefore, one can write

$$\|\mathsf{P}_x^n - \mathsf{P}_y^n\|_{\mathcal{G}} = \sup_{g\in\mathcal{G}} \int_{\mathsf{X}\times\mathsf{X}} (g(\Pi_1 z) - g(\Pi_2 z))\, (\pi_n^*\mathbf{C}_{x,y}^\infty)(dz) \;. \tag{2.2}$$

This equation is interesting, because it is in many cases possible to construct a coupling $\mathbf{C}_{x,y}^\infty$ such that for n large, the measure $\pi_n^*\mathbf{C}_{x,y}^\infty$ is concentrated near the diagonal $\Pi_1 z = \Pi_2 z$, thus providing through (2.2) an estimate for the term $\|\mathsf{P}_x^n - \mathsf{P}_y^n\|_{\mathcal{G}}$. This is precisely what was shown in our toy model of Section 1.1, where we constructed f in such a way that $\|x(t) - y(t)\| \to 0$ for $t \to \infty$.

2.3 The binding construction

In this subsection, we describe a specific type of coupling for a given RDS Φ. Only couplings of that type will be taken under consideration in the sequel.

Let Φ and the associated probability space $(\Omega, \mathscr{F}, \mathsf{P})$ be as above. We consider a family $\psi_{x\to y} : \Omega \to \Omega$ (the pair (x,y) belongs to X^2) of measurable functions that also have measurable inverses. We will call these functions *binding functions* for Φ. The reason for this terminology is that, given a realization $\{\omega_n\}_{n=0}^\infty$ of the noise and a pair of initial conditions $(x_0, y_0) \in \mathsf{X}^2$, the binding functions allow us to construct two paths $\{x_n\}$ and $\{y_n\}$ by setting

$$\tilde\omega_n = \psi_{x_n\to y_n}(\omega_n) \;, \qquad x_{n+1} = \Phi(x_n, \omega_n) \;, \qquad y_{n+1} = \Phi(y_n, \tilde\omega_n) \;. \tag{2.3}$$

Our aim is to find a family $\psi_{x\to y}$ such that y_n converges towards x_n in a suitable sense for large values of n. Thus, the binding functions play the role of a spring between x and y. We will say that (2.3) is a *binding construction* for Φ. We denote the inverse of $\psi_{x\to y}$ by $\psi_{x\leftarrow y}$. The reason behind this notation should be clear from the diagram below.

$$\begin{array}{ccccccc} x_n & & & & & & y_n \\ \big\downarrow \Phi(\,\cdot\,,\omega_n) & \dashleftarrow & \omega_n & \underset{\psi_{x_n\leftarrow y_n}}{\overset{\psi_{x_n\to y_n}}{\rightleftarrows}} & \tilde\omega_n & \dashrightarrow & \Phi(\,\cdot\,,\tilde\omega_n)\big\downarrow \\ x_{n+1} & & & & & & y_{n+1} \end{array} \tag{2.4}$$

The solid arrows denote the various maps and the dashed arrows denote the influences of the appearing quantities on those maps. It shows that it is also possible to achieve the binding construction by first choosing a sequence $\{\tilde\omega_n\}_{n=0}^\infty$ and then using $\psi_{x_n\leftarrow y_n}$ to construct the ω_n, thus obtaining the same set of possible realizations for (x_n, y_n). This symmetry between $\psi_{x\to y}$ and $\psi_{x\leftarrow y}$ is also explicit in (2.6) below.

Guided by the above construction, we use the binding maps to construct a coupling Markov chain Ψ on $\mathsf{X}\times\mathsf{X}$ with transition probabilities $\mathbf{C}_{x,y}$ in the following way. Define the maps

$$\begin{aligned} \Psi_{x\to y} &: \Omega \to \Omega\times\Omega & \qquad \Psi_{x\leftarrow y} &: \Omega \to \Omega\times\Omega \\ &\omega \mapsto (\omega, \psi_{x\to y}(\omega))\;, & &\omega \mapsto (\psi_{x\leftarrow y}(\omega), \omega)\;. \end{aligned} \tag{2.5}$$

Notice that, up to some null set, the image of both maps is the set $\{(\omega,\tilde\omega) \mid \tilde\omega = \psi_{x\to y}(\omega)\}$. Then we define a family of measures $\mathsf{P}_{x,y}$ on $\Omega\times\Omega$ by

$$\mathsf{P}_{x,y} = (\Psi^*_{x\to y}\mathsf{P}) \wedge (\Psi^*_{x\leftarrow y}\mathsf{P}) = \Psi^*_{x\leftarrow y}(\mathsf{P}\wedge\psi^*_{x\to y}\mathsf{P})\;. \tag{2.6}$$

According to (2.4), the measure P_{x_n,y_n} is precisely the common part between the measure obtained for $(\omega_n,\tilde\omega_n)$ by distributing ω_n according to P and the one obtained by distributing $\tilde\omega_n$ according to P. Thus both marginals of the measure $\mathsf{P}_{x,y}$ are smaller (in the sense of Section 2.1) than P. In order to have a non-trivial construction, we impose that the measures P and $\psi^*_{x\to y}\mathsf{P}$ are equivalent. The density of $\psi^*_{x\to y}\mathsf{P}$ relative to P will be denoted by $\mathcal{D}_{x,y}(\omega)$.

Considering again (2.4), the family of measures $\mathsf{P}_{x,y}$ is transported on $\mathsf{X}\times\mathsf{X}$ by defining

$$\begin{aligned} \Phi_{x,y} &: \Omega\times\Omega \to \mathsf{X}\times\mathsf{X} \\ &(\omega,\tilde\omega) \mapsto (\Phi(x,\omega), \Phi(y,\tilde\omega))\;, \end{aligned}$$

and setting

$$\mathbf{Q}_{x,y} \equiv \Phi^*_{x,y}\mathsf{P}_{x,y}\;. \tag{2.7}$$

But this does not give a transition probability function yet, since the measures $\mathsf{P}_{x,y}$ are not normalized to 1. We therefore define the family of measures $\mathbf{P}_{x,y}$ by

$$\mathbf{P}_{x,y} = \mathsf{P}_{x,y} + c_{x,y}\,(\mathsf{P}\setminus\Pi_1^*\mathsf{P}_{x,y})\times(\mathsf{P}\setminus\Pi_2^*\mathsf{P}_{x,y})\;,$$

where the number $c_{x,y}$ is chosen in such a way that the resulting measure is a probability measure. By a slight abuse of notation, we used here the symbol Π_i to denote the projection on the

ith component of $\Omega \times \Omega$. As a matter of fact, $(\mathsf{P} \setminus \Pi_1^* \mathsf{P}_{x,y})$ and $(\mathsf{P} \setminus \Pi_2^* \mathsf{P}_{x,y})$ have the same mass, which is equal to $1 - \|\mathsf{P}_{x,y}\|$, so

$$c_{x,y} = \frac{1}{\|\mathsf{P} \setminus \Pi_2^* \mathsf{P}_{x,y}\|} \;,$$

for example. (Recall that the symbol $\|\cdot\|$ stands for the total variation norm, which is simply equal to its mass for a positive measure.) It is straightforward to show that the following holds:

Lemma 2.3 *The measures $\mathbf{P}_{x,y}$ satisfy $\Pi_i^* \mathbf{P}_{x,y} = \mathsf{P}$ for $i = 1, 2$.*

Proof. It is clear by (2.6) that $\Pi_i^* \mathsf{P}_{x,y} \le \mathsf{P}$. Thus

$$\begin{aligned} \Pi_1^* \mathbf{P}_{x,y} &= \Pi_1^* \mathsf{P}_{x,y} + c_{x,y} \|\mathsf{P} \setminus \Pi_2^* \mathsf{P}_{x,y}\| (\mathsf{P} \setminus \Pi_1^* \mathsf{P}_{x,y}) \\ &= (\mathsf{P} \wedge \Pi_1^* \mathsf{P}_{x,y}) + (\mathsf{P} \setminus \Pi_1^* \mathsf{P}_{x,y}) = \mathsf{P} \;, \end{aligned} \tag{2.8}$$

and similarly for $\Pi_2^* \mathbf{P}_{x,y}$. □

This finally allows us to define the transition probabilities for Ψ by

$$\mathbf{C}_{x,y} = \Phi_{x,y}^* \mathbf{P}_{x,y} \equiv \mathbf{Q}_{x,y} + \mathbf{R}_{x,y} \;. \tag{2.9}$$

In this expression, the only feature of $\mathbf{R}_{x,y}$ we will use is that it is a positive measure. We define $\mathbf{C}_{x,y}^\infty$ as the measure on the pathspace $\mathsf{X}^\infty \times \mathsf{X}^\infty$ obtained by iterating (2.1). Since $\Pi_1 \circ \Phi_{x,y} = \Phi(x, \cdot) \circ \Pi_1$ and similarly for Π_2, it is straightforward to verify, using Lemma 2.3, that the measure $\mathbf{C}_{x,y}^\infty$ constructed this way is indeed a coupling for Φ.

For a given step of Ψ, we say that the trajectories do couple if the step is drawn according to $\mathbf{Q}_{x,y}$ and that they don't couple otherwise.

Remark 2.4 Since $\mathbf{P}_{x,y}$ is a family of measures on $\Omega \times \Omega$, it is also possible to interpret $\mathbf{C}_{x,y}^n$ as a family of probability measures on $\mathsf{X}^n \times \mathsf{X}^n \times \Omega^n \times \Omega^n$. We will sometimes use this viewpoint in the following section. It is especially useful when the RDS Φ is obtained by sampling a continuous-time process.

Remark 2.5 It will sometimes be useful to have an explicit way of telling whether a step of Ψ is taken according to $\mathbf{Q}_{x,y}^\infty$ or according to $\mathbf{R}_{x,y}^\infty$ (*i.e.*whether the trajectories couple or not). To this end, we introduce a Markov chain $\hat{\Psi}$ on the augmented phase space $\mathsf{X} \times \mathsf{X} \times \{0, 1\}$ with transition probabilities

$$\mathbf{P}_{x,y} = \mathbf{Q}_{x,y} \times \delta_1 + \mathbf{R}_{x,y} \times \delta_0 \;.$$

The marginal of $\hat{\Psi}$ on $\mathsf{X} \times \mathsf{X}$ is of course equal to Ψ. By a slight abuse of notation, we will also write $\mathbf{C}_{x,y}^\infty$ for the probability measure on pathspace induced by $\hat{\Psi}$.

It will be useful in the sequel to have a map that "transports" the family of maps $\psi_{x \to y}$ on Ω^n via the RDS Φ. More precisely, fix a pair $(x, y) \in \mathsf{X} \times \mathsf{X}$ of starting points and a sequence $(\omega_0, \ldots, \omega_n)$ of realizations of the noise. We then define $x_0 = x$, $y_0 = y$, and, recursively for $i = 0, \ldots, n$

$$x_{i+1} = \Phi(x_i, \omega_i) \;, \qquad y_{i+1} = \Phi(y_i, \psi_{x_i \to y_i}(\omega_i)) \;.$$

This allows us to define the family of maps $\Xi^n_{x,y} : \Omega^n \to \Omega^n$ by

$$\Xi^{n+1}_{x,y}(\omega_0, \dots, \omega_n) \mapsto (\psi_{x_0 \to y_0}(\omega_0), \dots, \psi_{x_n \to y_n}(\omega_n)) \;. \tag{2.10}$$

Since $\psi^*_{x\to y}\mathsf{P}$ is equivalent to P, we see that $(\Xi^n_{x,y})^*\mathsf{P}^n$ is equivalent to P^n and we denote its density by $\mathcal{D}^n_{x,y}$. We also notice that the family of measures $\mathbf{Q}^n_{x,y}$ is obtained by transporting $(\Xi^n_{x,y})^*\mathsf{P}^n \wedge \mathsf{P}^n$ onto $\mathsf{X}^n \times \mathsf{X}^n$ with the maps $\Phi_{x_i,y_i} \circ \Psi_{x_i \to y_i}$. In particular, one has the equality

$$\|\mathbf{Q}^n_{x,y}\| = \|(\Xi^n_{x,y})^*\mathsf{P}^n \wedge \mathsf{P}^n\| = \int_{\Omega^n} \big(1 \wedge \mathcal{D}^n_{x,y}(\omega)\big)\, \mathsf{P}^n(d\omega) \;. \tag{2.11}$$

3 Assumptions on the Coupling

In this section, we investigate the properties of the coupling $\mathbf{C}^\infty_{x,y}$ constructed in the previous section. We give a set of assumptions on the binding functions $\psi_{x\to y}$ that ensure the existence and uniqueness of the invariant measure for Φ.

In order to achieve this, we want the map $\psi_{x\to y}$ to modify the noise in such a way that trajectories drawn according to $\mathbf{Q}_{x,y}$ tend to come closer together. This will be the content of Assumption **A3**. Furthermore, we want to know that this actually happens, so the noise should not be modified too much. This will be the content of assumptions **A4** and **A5**. All these nice properties usually hold only in a "good" region of the phase space. Assumptions **A1** and **A2** will ensure that visits to this good region happen sufficiently often.

3.1 Lyapunov structure

Since we are interested in obtaining exponential mixing, we need assumptions of exponential nature. Our first assumption concerns the global aspects of the dynamics. It postulates that Φ is attracted exponentially fast towards a "good" region of its state space. We achieve this by assuming the existence of a Lyapunov function for Φ.

Definition 3.1 Let Φ by a RDS with state space X as before. A *Lyapunov function* for Φ is a function $V : \mathsf{X} \to [0, \infty]$ for which there exist constants $a \in (0, 1)$ and $b > 0$, such that

$$\int_\Omega V(\Phi(x, \omega))\, \mathsf{P}(d\omega) \le aV(x) + b \;, \tag{3.1}$$

for every $x \in \mathsf{X}$ with $V(x) < \infty$.

Our first assumption then reads

A1 *There exist a Lyapunov function V for Φ. Furthermore, V is such that*

$$\mathsf{P}\{\omega \mid V(\Phi(x, \omega)) < \infty\} = 1 \;,$$

for every $x \in \mathsf{X}$.

For convenience, we also introduce the function $\tilde V : \mathsf{X} \times \mathsf{X} \to [0, \infty]$ defined by

$$\tilde V(x, y) = V(x) + V(y) \;.$$

Notice that $\tilde{V}$ is a Lyapunov function for Ψ by construction.

In some cases, when the control over the densities $\mathcal{D}_{x,y}$ is uniform enough or when the phase space is already compact (or bounded), a Lyapunov function is not needed. In such a situation, one can simply choose $V \equiv 1$.

In our case of interest, the RDS Φ is obtained by sampling a continuous-time process Φ_t at discrete times. In that setting, it is useful to have means to control excursions to large amplitudes that take place between two successive sampling times. To this end, we introduce a function $W : \mathsf{X} \times \Omega \to [0, \infty]$ given by

$$W(x,\omega) = \sup_{t\in[0,1]} V(\Phi_t(x,\omega))$$

in the continuous-time setting and by

$$W(x,\omega) = V(x)$$

in the discrete-time setting. In fact, any other choice of W is all right, as long as it satisfies the properties that are summarized in Assumption **A2** below.

Before stating these properties, we define two other functions that act on pairs of initial conditions that couple by

$$\begin{aligned} W_{x\to y}(\omega) &= W(x,\omega) + W(y, \psi_{x\to y}(\omega)) \;, \\ W_{x\leftarrow y}(\omega) &= W(x, \psi_{x\leftarrow y}(\omega)) + W(y,\omega) \;. \end{aligned} \tag{3.2}$$

We will assume that W and the binding functions are such that W, $W_{x\to y}$ and $W_{x\leftarrow y}$ do not behave much worse than V. More precisely, we will assume that:

A2 *There exists a function* $W : \mathsf{X} \times \Omega \to [0, \infty]$ *such that*

$$\operatorname*{ess\,inf}_{\omega\in\Omega} W(x,\omega) = V(x) \;, \tag{3.3a}$$

$$\int_\Omega W(x,\omega)\, \mathsf{P}(d\omega) \le c\, V(x) \;, \tag{3.3b}$$

for some constant $c > 0$*. Furthermore, there exist constants* $C > 0$ *and* $\delta \ge 1$ *such that the estimates*

$$\begin{aligned} W_{x\to y}(\omega) &\le C\big(1 + V(y) + W(x,\omega)\big)^\delta \;, \\ W_{x\leftarrow y}(\omega) &\le C\big(1 + V(x) + W(y,\omega)\big)^\delta \;, \end{aligned} \tag{3.4}$$

hold for the functions defined in (3.2).

The Lyapunov structure given by assumptions **A1** and **A2** ensures that W (and thus also V) does not increase too fast along a typical trajectory. In order to make this statement precise, we define for a given initial condition $x \in \mathsf{X}$ the sets $A_{x,k} \subset \Omega^\infty$ by

$$A_{x,k} = \{\omega \in \Omega^\infty \mid W(\Phi^n(x,\omega), \omega_n) \le kV(x) + kn^2 \quad \forall n > 0\} \;, \tag{3.5}$$

where k is some positive constant. The sets $A_{x,k}$ contain almost every typical realization of the noise:

Lemma 3.2 *Let Φ be a RDS satisfying assumptions* **A1** *and* **A2**. *Then, there exists a constant $C > 0$ such that*

$$\mathsf{P}^\infty(A_{x,k}) \geq 1 - \frac{C}{k} ,$$

for every $x \in \mathsf{X}$ and every $k > 0$.

Proof. For $\omega \in \Omega^\infty$, we define $x_n = \Phi^n(x,\omega)$. Notice that by (3.3b) and the Lyapunov structure, one has the estimate

$$\mathbf{E}(W(x_n, \omega_{n+1})) \leq c a^n V(x) + \frac{bc}{1-a} , \tag{3.6}$$

where $\mathbf{E}$ denotes expectations with respect to P^∞. We also notice that $A_{x,k} = \bigcap_{n>0} A_{x,k}^{(n)}$ with

$$A_{x,k}^{(n)} = \{\omega \mid W(x_n, \omega_{n+1}) \leq kV(x) + kn^2\} .$$

Combining this with (3.6), we see that

$$\mathsf{P}^\infty(A_{x,k}^{(n)}) \geq 1 - \frac{c}{k} \frac{a^n V(x) + b(1-a)^{-1}}{V(x) + n^2} .$$

Therefore, the worst possible estimate for $\mathsf{P}^\infty(A_{x,k})$ is

$$\mathsf{P}^\infty(A_{x,k}) \geq 1 - \frac{c}{k} \sum_{n=1}^\infty \frac{a^n V(x) + b(1-a)^{-1}}{V(x) + n^2} ,$$

which proves the claim. □

3.2 Binding property

The crucial property of the coupling is to bring trajectories closer together. In order to make this statement more precise, we introduce the Lipschitz norm $\|\cdot\|_{\mathrm{L}}$ defined on functions $g : \mathsf{X} \to \mathbf{R}$ by

$$\|g\|_{\mathrm{L}} = \sup_{x\in\mathsf{X}} |g(x)| + \sup_{x,y\in\mathsf{X}} \frac{|g(x)-g(y)|}{d(x,y)} ,$$

where $d(\cdot,\cdot)$ denotes the distance in X. The dual norm on $\mathscr{M}(\mathsf{X})$ is then given by

$$\|\mu\|_{\mathrm{L}} = \sup_{\|g\|_{\mathrm{L}}=1} \int_{\mathsf{X}} g(x)\,\mu(dx) .$$

With this definition at hand, we make the following assumption on the coupling part $\mathbf{Q}^\infty_{x,y}$.

A3 *There exist a positive constant γ_1 and a family of constants $K \mapsto C_K$ such that, for every $K > 0$,*

$$\|\pi^*_{1,n}\mathbf{Q}^\infty_{x,y} - \pi^*_{2,n}\mathbf{Q}^\infty_{x,y}\|_{\mathrm{L}} \leq C_K e^{-\gamma_1 n} , \tag{3.7}$$

holds when $\tilde V(x,y) \leq K$.

Remark 3.3 The sub-probability kernels $\mathbf{Q}_{x,y}$ are smaller than the transition probabilities for the binding construction (2.3). Thus, (3.7) is implied by an inequality of the type

$$\mathbf{E}(d(x_n, y_n)) \leq C\tilde{V}(x_0, y_0)e^{-\gamma_1 n} ,$$

where d denotes the distance in X and $\mathbf{E}$ denotes the expectation with respect to the construction (2.3).

Notice that this assumption is non-trivial only if our coupling is such that $\|\mathbf{Q}^\infty_{x,y}\| > 0$ for sufficiently many starting points. This will be ensured by the next assumption.

A4 *Let $\mathcal{D}^n_{x,y}$ be defined as in Section 2.3. We assume that for every $K > 0$, there exists a family of sets $\Gamma^K_{x,y} \subset \Omega^\infty$ and constants $c_1, c_2 > 0$ such that the estimates*

$$\mathsf{P}^\infty(\Gamma^K_{x,y}) > c_1 , \qquad \int_{\Gamma^K_{x,y}} (\mathcal{D}^n_{x,y}(\omega))^{-2}\, \mathsf{P}^n(d\omega) < c_2 , \tag{3.8}$$

hold for every $n \geq 0$, whenever $\tilde{V}(x,y) \leq K$. The integral over $\Gamma^K_{x,y}$ in (3.8) should be interpreted as the integral over the projection of $\Gamma^K_{x,y}$ onto its n first components.

A typical choice for $\Gamma^K_{x,y}$ is $\Gamma^K_{x,y} = A_{y,k}$ or $\Gamma^K_{x,y} = A_{x,k} \cap A_{y,k}$ with k sufficiently large as a function of K. In this case, Lemma 3.2 ensures that the conditions required on $\Gamma^K_{x,y}$ are satisfied. As a consequence of Assumption **A4**, we have

Proposition 3.4 *Let $\mathbf{Q}^\infty_{x,y}$ be defined as above and suppose that assumptions* **A1** *and* **A4** *hold. Then there exists for every K a constant C_K such that $\|\mathbf{Q}^\infty_{x,y}\| \geq C_K$, whenever $\tilde{V}(x,y) \leq K$.*

Proof. Notice first that if μ_1 and μ_2 are two equivalent probability measures with $\mu_2(dx) = \mathcal{D}(x)\,\mu_1(dx)$, then the condition

$$\int_A (\mathcal{D}(x))^{-2}\, \mu_1(dx) < c$$

implies that

$$(\mu_1 \wedge \mu_2)(A) \geq \frac{\mu_1(A)^2}{4c} ,$$

see, *e.g.*[Mat01]. Recalling (2.11), we use Lemma 3.2 and the above estimate with $\mu_1 = \mathsf{P}^n$, $\mathcal{D} = \mathcal{D}^n_{x,y}$, and $A = \Gamma^K_{x,y}$. Taking the limit $n \to \infty$ and using the assumption on $\Gamma^K_{x,y}$ proves the claim. □

Our last assumption will ensure that trajectories that have already coupled for some time have a very strong tendency to couple for all times.

In order to formulate our assumption, we introduce a family of sets $Q^n_K(x,y)$, which are the possible final states of a "coupled" trajectory of length n, starting from (x,y), and never leaving the set $\{(a,b) \mid V(a) + V(b) \leq K\}$. For a given pair of initial conditions $(x,y) \in \mathsf{X}^2$ with $\tilde{V}(x,y) \leq K$, we define the family of sets $Q^n_K(x,y) \subset \mathsf{X} \times \mathsf{X}$ recursively in the following way:

$$Q^0_K(x,y) = \{(x,y)\} ,$$

$$Q_K^{n+1}(x,y) = \bigcup_{(a,b)\in Q_K^n(x,y)} \{(\Phi_{a,b}\circ\Psi_{a\to b})(\omega) \mid \omega\in\Omega \text{ and } W_{a\to b}(\omega)\le K\} \;.$$

Notice that we would have obtained the same sets by reversing the directions of the arrows in the definition.

We also denote by $\mathcal{D}_{x,y}(\omega)$ the density of $\psi^*_{x\to y}\mathsf{P}$ relative to P.

A5 *There exist positive constants C_2, γ_2 and ζ, such that for every $K>0$, every $(x_0,y_0)\in\mathsf{X}^2$ with $\tilde V(x_0,y_0)\le K$, and every $(x,y)\in Q_K^n(x_0,y_0)$, the estimate*

$$\int_{W_{x\leftarrow y}(\omega)\le K}(1-\mathcal{D}_{x,y}(\omega))^2\,\mathsf{P}(d\omega)\le C_2e^{-\gamma_2 n}(1+K)^\zeta \;, \tag{3.9}$$

holds for $n>\zeta\ln(1+K)/\gamma_2$.

This assumption means that if the process couples for a time n, the density $\mathcal{D}_{x,y}$ is close to 1 on an increasingly large set, and therefore the probability of coupling for a longer time becomes increasingly large. This assumption is sufficient for the family of measures $(\mathbf{RQ}^n)_{x,y}$ to have an exponential tail at large values of n. More precisely, we have

Proposition 3.5 *Let assumptions* **A1**, **A2** *and* **A5** *hold. Then, there exists a positive constant γ_3 and, for every $K>0$, a constant C_K such that*

$$\|(\mathbf{RQ}^n)_{x,y}\|\le C_K e^{-\gamma_3 n} \;, \tag{3.10}$$

holds for every $n>0$, whenever $\tilde V(x,y)\le K$.

We first show the following elementary estimate (it is not optimal, but sufficient for our needs):

Lemma 3.6 *Let $\mu_1,\mu_2\in\mathscr{M}_1(\mathsf{X})$ be two equivalent probability measures with*

$$\mu_2(dx)=\mathcal{D}(x)\,\mu_1(dx) \;.$$

Then the conditions

$$\mu_1(A)\ge 1-\varepsilon_1 \quad \textit{and} \qquad \int_A (1-\mathcal{D}(x))^2\,\mu_1(dx)\le\varepsilon_2 \;,$$

for some measurable set A imply that

$$(\mu_1\wedge\mu_2)(A)\ge 1-\varepsilon_1-\varepsilon_2^{1/2} \;.$$

Proof. Define the set $E\subset\mathsf{X}$ by

$$E=A\cap\{x\in\mathsf{X}\mid\mathcal{D}(x)\ge 1\} \;.$$

We then have

$$(\mu_1\wedge\mu_2)(A)=\mu_1(E)+\int_{A\setminus E}\mathcal{D}(x)\,\mu_1(dx)$$

$$
\begin{aligned}
&= \mu_1(A) - \int_{A\setminus E} (1-\mathcal{D}(x))\,\mu_1(dx) \\
&\geq \mu_1(A) - \int_{A\setminus E} |1-\mathcal{D}(x)|\,\mu_1(dx) \\
&\geq 1-\varepsilon_1 - \sqrt{\int_{A\setminus E} (1-\mathcal{D}(x))^2\,\mu_1(dx)}\;.
\end{aligned}
$$

This shows the claim. □

Proof of Proposition 3.5. Fix the value n and the pair (x,y). For every $c_n \geq \tilde{V}(x,y)$ (we will fix it later), we have the estimate

$$
\begin{aligned}
\|(\mathbf{RQ}^n)_{x,y}\| &= \int_{\mathsf{X}^2} (1-\|\psi^*_{x_n\to y_n}\mathsf{P}\wedge\mathsf{P}\|)\,(\pi_n^*\mathbf{Q}^n_{x,y})(dx_n,dy_n) \\
&\leq (\pi_n^*\mathbf{Q}^n_{x,y})(\mathsf{X}^2\setminus Q^n_{c_n}(x,y)) \\
&\quad + \int_{Q^n_{c_n}(x,y)} (1-\|\psi^*_{x_n\to y_n}\mathsf{P}\wedge\mathsf{P}\|)\,(\pi_n^*\mathbf{Q}^n_{x,y})(dx_n,dy_n)\;.
\end{aligned}
\tag{3.11}
$$

Now choose another value w_n to be fixed later and consider for every (x_n,y_n) the set

$$
B_n = \{\omega\in\Omega \,|\, W_{x_n\leftarrow y_n}(\omega)\leq w_n\}\;.
$$

By the definition of $Q^n_{c_n}(x,y)$, its elements (x_n,y_n) satisfy in particular $\tilde{V}(x_n,y_n)\leq c_n$. By Assumption **A2** and the Lyapunov structure, we have for every $(x_n,y_n)\in Q^n_{c_n}(x,y)$ the estimate

$$
\mathsf{P}(B_n) \geq 1 - C\frac{c_n}{w_n^{1/\delta}}\;.
$$

Combining this and Assumption **A5** with Lemma 3.6 yields

$$
\|\psi^*_{x_n\to y_n}\mathsf{P}\wedge\mathsf{P}\| \geq 1 - C\frac{c_n}{w_n^{1/\delta}} - Ce^{-\gamma_2 n/2}(1+w_n)^{\zeta/2}\;,
$$

as long as w_n is such that

$$
w_n \geq c_n \quad\text{and}\quad n\geq \zeta\ln(1+w_n)/\gamma_2\;. \tag{3.12}
$$

It remains to give an upper bound for $(\pi_n^*\mathbf{Q}^n_{x,y})(\mathsf{X}^2\setminus Q^n_{c_n}(x,y))$ to complete our argument. Define the sets $A^n(K)\subset \mathsf{X}^n\times\mathsf{X}^n\times\Omega^n\times\Omega^n$ by

$$
A^n(K) = \{(x_i,y_i,\omega_i,\eta_i)_{i=1}^n \,|\, W(x_i,\omega_i)+W(y_i,\eta_i)\leq K\}\;.
$$

It is clear by the definition of $Q^n_{c_n}(x,y)$ that we have the equality

$$
(\pi_n^*\mathbf{Q}^n_{x,y})(\mathsf{X}^2\setminus Q^n_{c_n}(x,y)) = \mathbf{Q}^n_{x,y}(\mathsf{X}^n\times\mathsf{X}^n\times\Omega^n\times\Omega^n\setminus A^n(c_n))\;,
$$

where $\mathbf{Q}^n_{x,y}$ is considered as a measure on $\mathsf{X}^n\times\mathsf{X}^n\times\Omega^n\times\Omega^n$, following Remark 2.4. Since $\mathbf{Q}^n_{x,y}\leq\mathbf{C}^n_{x,y}$, we have

$$
(\pi_n^*\mathbf{Q}^n_{x,y})(\mathsf{X}^2\setminus Q^n_{c_n}(x,y)) \leq 1-\mathbf{C}^n_{x,y}(A^n(c_n)) \leq C\frac{n(\tilde{V}(x,y)+1)}{c_n}\;,
$$

for some constant C. This last estimate is obtained in a straightforward way, following the lines of the proof of Lemma 3.2. Plugging these estimates back into (3.11) yields

$$\|(\mathbf{RQ}^n)_{x,y}\| \le C\frac{n(\tilde V(x,y)+1)}{c_n} + C\frac{c_n}{w_n^{1/\delta}} + Ce^{-\gamma_2 n/2}(1+w_n)^{\zeta/2}\;.$$

At this point, we make use of our freedom to choose c_n and w_n. We set

$$c_n = \tilde V(x,y) + e^{\gamma_c n} \qquad \text{and} \qquad w_n = \tilde V(x,y) + e^{\gamma_w n}\;,$$

with γ_c and γ_w given by

$$\gamma_c = \frac{1}{2+2\delta\zeta}\gamma_2 \qquad \text{and} \qquad \gamma_w = \frac{\delta}{1+\delta\zeta}\gamma_2\;.$$

As a consequence, there exist for any $\gamma < \gamma_c$ some constants C and c such that

$$\|(\mathbf{RQ}^n)_{x,y}\| \le C(1+\tilde V(x,y))^c e^{-\gamma n}\;,$$

as long as $n \ge \zeta \ln(1+w_n)/\gamma_2$. (Such a value of n can always be found, because the exponent γ_w is always smaller than γ_2/ζ.) In order to complete the argument, we notice that (3.10) is trivially satisfied for small values of n because $\|(\mathbf{RQ}^n)_{x,y}\|$ is always smaller than 1 by definition: it suffices to choose C_K sufficiently big. The proof of Proposition 3.5 is complete. □

4 An Exponential Mixing Result

This section is devoted to the proof of the main theorem of this paper.

Theorem 4.1 *Let Φ be a RDS with state space X satisfying assumptions* **A1–A5***. Then, there exists a constant $\gamma > 0$ such that*

$$\|\mathsf{P}_x^n - \mathsf{P}_y^n\|_{\mathrm{L}} \le C(1+\tilde V(x,y))\,e^{-\gamma n}\;,$$

for every $(x,y) \in \mathsf{X}^2$ and every $n > 0$.

Remark 4.2 The proof of Theorem 4.1 does not rely on assumptions **A4** and **A5** directly, but on the conclusions of Propositions 3.4 and 3.5. Nevertheless, in the setting of stochastic differential equations, it seems to be easier to verify the assumptions rather than to show the conclusions of the propositions by other means.

Corollary 4.3 *If Φ satisfies assumptions* **A1–A5***, it possesses a unique invariant measure μ_* and*

$$\|\mathsf{P}_x^n - \mu_*\|_{\mathrm{L}} \le C(1+V(x))\,e^{-\gamma n}\;.$$

Proof of the corollary. To show the existence of the invariant measure μ_*, we show that for any given initial condition x with $V(x) < \infty$, the sequence of measures P_x^n is a Cauchy sequence in the norm $\|\cdot\|_{\mathrm{L}}$. We have indeed

$$\|\mathsf{P}_x^n - \mathsf{P}_x^{n+k}\|_{\mathrm{L}} = \sup_{\|g\|_{\mathrm{L}}\le 1}\int_{\mathsf{X}} g(z)(\mathsf{P}_x^n - \mathsf{P}_x^{n+k})(dz)$$

$$
\begin{aligned}
&= \sup_{\|g\|_{\mathrm{L}}\le 1} \int_{\mathsf{X}}\int_{\mathsf{X}} g(z)(\mathsf{P}_x^n - \mathsf{P}_y^n)(dz)\,\mathsf{P}_x^k(dy)\\
&\le \int_{\mathsf{X}} \|\mathsf{P}_x^n - \mathsf{P}_y^n\|_{\mathrm{L}}\,\mathsf{P}_x^k(dy) \le Ce^{-\gamma n}\int_{\mathsf{X}}(1+\tilde V(x,y))\,\mathsf{P}_x^k(dy)\\
&\le Ce^{-\gamma n}(1+V(x))\;,
\end{aligned}
$$

where we used the Lyapunov structure to get the last inequality.

The claim now follows immediately from the theorem, noticing that if μ_* is an invariant measure for Φ, then

$$
\int_{\mathsf{X}} V(x)\,\mu_*(dx) \le \frac{b}{1-a}\;,
$$

due to the Lyapunov structure and the fact that the dynamics immediately leaves the set $V^{-1}(\infty)$. □

Before we turn to the proof of Theorem 4.1, we introduce some notations and make a few remarks. By iterating (2.9), one sees that

$$
\mathbf{C}_{x,y}^{\infty} = \mathbf{Q}_{x,y}^{\infty} + \sum_{n=0}^{\infty}(\mathbf{C}^{\infty}\mathbf{R}\mathbf{Q}^n)_{x,y}\;, \tag{4.1}
$$

where the symbol $(\mathbf{C}^{\infty}\mathbf{R}\mathbf{Q}^n)_{x,y}$ is to be interpreted in the sense of (2.1). This expression is the equivalent, in our setting, of Lemma 2.1 in [Mat01]. Using (4.1), the Markov chain Ψ can be described by means of another Markov chain Υ on $\mathsf{Y} = (\mathsf{X}^2 \times \mathbf{N}) \cup \{\star\}$, where $\star$ corresponds to "coupling for all times" in the sense of Section 2.3. First, we define

$$
K_0 = \frac{4b}{1-a}\;, \qquad \tilde K_0 = \{(x,y)\mid \tilde V(x,y) \le K_0\}\;, \tag{4.2}
$$

where a and b are the constants appearing in the Lyapunov condition. This set is chosen in such a way that

$$
\int_{\mathsf{X}\times\mathsf{X}} \tilde V(x,y)\,\mathbf{C}_{x_0,y_0}(dx,dy) \le \frac{1+a}{2}\tilde V(x_0,y_0)\;, \qquad \forall\,(x_0,y_0)\notin \tilde K_0\;. \tag{4.3}
$$

At time 0, Υ is located at $(x,y,0)$. If it is located at (x,y,n) and $(x,y)\notin\tilde K_0$, then it makes one step according to $\mathbf{C}_{x,y}$ and n is incremented by one:

$$
\mathbf{P}_{(x,y,n)} = \mathbf{C}_{x,y}\times\delta_{n+1}\;.
$$

If Υ is located at (x,y,n) and $(x,y)\in\tilde K_0$, then it has a probability $\|\mathbf{Q}_{x,y}^{\infty}\|$ of jumping to $\star$ and a probability $\|(\mathbf{R}\mathbf{Q}^m)_{x,y}\|$ of making m steps according to $(\mathbf{R}\mathbf{Q}^m)_{x,y}$:

$$
\mathbf{P}_{(x,y,n)} = \|\mathbf{Q}_{x,y}^{\infty}\|\delta_\star + \sum_{m=0}^{\infty}\pi_m^*(\mathbf{R}\mathbf{Q}^m)_{x,y}\times\delta_{n+m+1}\;.
$$

If Υ is located at $\star$, it remains there:

$$
\mathbf{P}_\star = \delta_\star\;.
$$

The Markov chain Υ induces a family $\hat{\mathbf{P}}^\infty_{x,y}$ of probability measures on Y^∞. Let $\hat\tau : \mathsf{Y}^\infty \to \mathbf{N} \cup \{\infty\}$ be the function that associates to a sequence of elements in Y the largest value of n that is reached by the sequence ($\hat\tau = 0$ if the sequence is equal to $\star$ repeated). We also define $\hat\kappa : \mathsf{Y}^\infty \to \mathbf{N} \cup \{\infty\}$ as the value of n attained at the first non-vanishing time when the sequence hits the set $\tilde K_0 \times \mathbf{N}$ ($\hat\kappa = \infty$ if this set is never reached). The construction of Υ is very close to the coupling construction of [Mat01].

The crucial observation for the proof of Theorem 4.1 is

Lemma 4.4 *Let Φ be a RDS with state space X satisfying assumptions* **A1** *and* **A3**, *and let Υ be defined as above. Then, there exists a constant C such that*

$$\|\mathsf{P}^n_x - \mathsf{P}^n_y\|_{\mathrm{L}} \le \hat{\mathbf{P}}^\infty_{x,y}(\{\hat\tau \ge n/2\}) + Ce^{-\gamma_1 n/2} ,$$

for every $(x,y) \in \mathsf{X}^2$ and every $n > 0$.

Proof. Recall the Markov chain $\hat\Psi$ defined in Remark 2.5. We define a function τ_1 on its pathspace by

$$\begin{aligned}\tau_1 : \mathsf{X}^\infty \times \mathsf{X}^\infty \times \{0,1\}^\infty &\to \mathbf{N} \cup \{\infty\}\\ \{(x_i, y_i, b_i)\}_{i=1}^\infty &\mapsto \inf\{n \,|\, (x_n, y_n) \in \tilde K_0 \text{ and } b_i = 1 \ \forall i \ge n\} .\end{aligned}$$

Combining (2.2) with Assumption **A5** and the definition of τ_1, one sees that

$$\|\mathsf{P}^n_x - \mathsf{P}^n_y\|_{\mathrm{L}} \le \mathbf{C}^\infty_{x,y}(\{\tau_1 \ge n/2\}) + Ce^{-\gamma_1 n/2} .$$

From the construction of Υ and the definition of $\hat\tau$, we see furthermore that the probability distributions of τ_1 under $\mathbf{C}^\infty_{x,y}$ and of $\hat\tau$ under $\hat{\mathbf{P}}^\infty_{x,y}$ are the same. □

Proof of Theorem 4.1. It remains to show that $\hat{\mathbf{P}}^\infty_{x,y}(\{\hat\tau \ge n/2\})$ has an exponential tail. The key observation is the following. Let $x_n \in \mathbf{N} \cup \{-\infty\}$ with $n \ge 0$ be a Markov chain defined by

$$x_0 = 0 , \quad x_{n+1} = \begin{cases} -\infty & \text{with probability } p_\star, \\ x_n + m & \text{with probability } p_m, \end{cases}$$

where $m \ge 1$ and, of course, $p_\star + \sum_{m=1}^\infty p_m = 1$.

Lemma 4.5 *If the p_m have an exponential tail and we define $\tau = \max_n x_n$, then the probability distribution of τ also has an exponential tail.*

Proof. The claim is an easy consequence of Kendall's theorem, but for the sake of completeness, and because the proof is quite elegant, we outline it here. Define the analytic function $p(\zeta) = \sum_{m=1}^\infty p_m \zeta^m$ and define q_n as the probability of τ being equal to n. Notice that, because of the exponential tail, p is analytic in a disk of radius strictly bigger than 1 around the origin. A straightforward computation shows that $q_0 = p_\star$ and, for $n > 0$,

$$q_n = p_\star\Big(p_n + \sum_{k_1+k_2=n} p_{k_1}p_{k_2} + \sum_{k_1+k_2+k_3=n} p_{k_1}p_{k_2}p_{k_3} + \ldots\Big) ,$$

which is equal to the nth Taylor coefficient of the function

$$q(\zeta) = \frac{p_\star}{1 - p(\zeta)} \;.$$

Since $p(1) = 1 - p_\star < 1$, there exists an $\varepsilon > 0$ such that $p(1+\varepsilon) < 1$. Furthermore, since the p_n are all positive, one has the estimate $|p(\zeta)| \le p(|\zeta|)$. Using Cauchy's formula on a circle of radius $1+\varepsilon$, one gets

$$|q_n| \le \frac{p_\star}{1 - p(1+\varepsilon)} \frac{1}{(1+\varepsilon)^n} \;,$$

which shows the claim. □

Before we prove Theorem 4.1 in full generality, we restrict ourselves to the case when $(x, y) \in \tilde K_0$. It follows from the construction that $\hat\tau$ (seen as a random variable under the distribution induced by $\hat{\mathbf{P}}^\infty_{x,y}$) is dominated by the process x_n constructed above with the tail distribution of the p_m being equal to

$$\tilde p_m = \sup_{(x,y)\in\tilde K_0} \hat{\mathbf{P}}^\infty_{x,y}(\{\hat\kappa = m\}) \;.$$

This means that we define m_* as

$$m_* = \inf\Big\{m \;\Big|\; \sum_{n=m}^\infty \tilde p_n \le 1\Big\} \;,$$

and then set $p_m = \tilde p_m$ for $m \ge m_*$, $p_m = 0$ for $m < m_* - 1$, and p_{m_*-1} in such a way that the p_n sum up to 1.

Because of Lemma 4.5, it suffices to show that the tail distribution of the $\tilde p_m$ decays exponentially. We thus estimate the quantity $\hat{\mathbf{P}}^\infty_{x,y}(\{\hat\kappa \ge n\})$. To this end, we introduce the function $\tau_\Psi : \mathsf{X}^\infty \times \mathsf{X}^\infty \to \mathbf{N} \cup \{\infty\}$ defined by

$$\tau_\Psi(x, y) = \inf\{n > 0 \mid (x_n, y_n) \in \tilde K_0\} \;.$$

Notice that, in order to have $\hat\kappa \ge n$, there are two possibilities. Either the first step of Υ is taken according to $(\mathbf{RQ}^m)_{x,y}$ with some $m \ge n/2$, or the corresponding realization of Ψ stays outside of $\tilde K_0$ for a time longer than $n/2$. This yields the estimate

$$\hat{\mathbf{P}}^\infty_{x,y}(\{\hat\kappa \ge n\}) \le \sum_{m=n/2}^\infty \|(\mathbf{RQ}^m)_{x,y}\| + \frac{n}{2} \sup_{(x_0,y_0)\in\tilde K_0} \mathbf{C}^\infty_{x_0,y_0}(\{\tau_\Psi \ge n/2\}) \;,$$

holding for $(x, y) \in \tilde K_0$. The first term has an exponential tail by Proposition 3.5. The second term has also an exponential tail by (4.3) and standard Lyapunov techniques (see *e.g.*[MT94, Thm 15.2.5]). This concludes the proof of Theorem 4.1 for the case $(x, y) \in \tilde K_0$.

In order to conclude the proof for the case $(x, y) \notin \tilde K_0$, notice that

$$\hat{\mathbf{P}}^\infty_{x,y}(\{\hat\tau \ge n\}) \le \sum_{m=1}^\infty \mathbf{C}^\infty_{x,y}(\{\tau_\Psi = m\}) \sup_{(x_0,y_0)\in\tilde K_0} \hat{\mathbf{P}}^\infty_{x_0,y_0}(\{\hat\tau \ge n - m\})$$

$$\leq \frac{n}{2} \sup_{(x_0,y_0)\in\tilde K_0} \hat{\mathbf{P}}^{\infty}_{x_0,y_0}(\{\hat\tau \geq n/2\}) + \sum_{m=n/2}^{\infty} \mathbf{C}^{\infty}_{x,y}(\{\tau_\Psi = m\}) \;.$$

The first term is bounded by the construction above. The Lyapunov structure implies that there exists a constant $\gamma > 0$ such that the first hitting time τ_Ψ satisfies $\mathbf{E}_{(x,y)} e^{\gamma\tau_\Psi} = \mathcal{O}(\tilde V(x,y))$ for every $(x,y) \in \mathsf{X}^2$ (see again [MT94, Thm. 15.2.5]). This allows to bound the second term and concludes the proof of Theorem 4.1. □

5 Application to Stochastic Differential Equations

In this section, we will see how to apply Theorem 4.1 to the case when the RDS Φ is constructed by sampling the solution of a (possibly infinite-dimensional) stochastic differential equation. We will restrict ourselves to the case where the equation is driven by additive white noise. The case of multiplicative noise requires further estimates, but can also be described by the formalism exposed here.

Consider the equation described by

$$dx(t) = Ax\,dt + F(x)\,dt + Q\,d\omega(t)\;, \qquad x(0) = x_0\;, \tag{5.1}$$

where x belongs to some separable Hilbert space $\mathcal{H}$, ω is the cylindrical Wiener process on some separable Hilbert space $\mathcal{W}$, and A, F and Q satisfy the following assumptions:

B1 *a. The linear operator $A : \mathscr{D}(A) \to \mathcal{H}$ is the generator of a strongly continuous semigroup on $\mathcal{H}$.*

b. The operator $e^{At}Q : \mathcal{W} \to \mathcal{H}$ is Hilbert-Schmidt for every $t > 0$ and one has the estimate

$$\int_0^1 \|e^{At}Q\|^2_{\mathrm{HS}}\,dt < \infty\;. \tag{5.2}$$

c. The nonlinear operator $F : \mathscr{D}(F) \to \mathcal{H}$ is such that, for every $x_0 \in \mathcal{H}$, there exists a unique, continuous stochastic process $x(t)$ such that $x(s) \in \mathscr{D}(F)$ for $s > 0$ and

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-s)}F(x(s))\,ds + \int_0^t e^{A(t-s)}Q\,d\omega(s)\;, \tag{5.3}$$

for every $t > 0$.

Remark 5.1 This assumptions simply states that there exists a unique weak solution to (5.1). Notice that we do *not* make any assumptions on the tightness of the transition probabilities for (5.1). As a consequence, existence and uniqueness results for invariant measures can in principle be deduced from Theorem 5.5 below even in cases where the semigroup e^{At} is not compact.

In order to recover the formalism used in Section 2, we follow [DPZ92b] and introduce an auxiliary Hilbert space $\hat{\mathcal{W}}$ such that there exists a continuous embedding $\iota : \mathcal{W} \hookrightarrow \hat{\mathcal{W}}$, which is Hilbert-Schmidt. We can now set $\Omega = \mathcal{C}_0([0,1], \hat{\mathcal{W}})$, the space of continuous $\hat{\mathcal{W}}$-valued functions that vanish at 0, and define P as the Wiener measure on $\hat{\mathcal{W}}$ with covariance operator $\iota\iota^*$.

We define $\Phi : \mathcal{H} \times \Omega \to \mathcal{H}$ as the map that solves (5.3) up to time 1 given an initial condition and a realization of the noise. This map is defined P-almost everywhere on Ω. We also denote by $\Phi_t : \mathcal{H} \times \Omega^\infty \to \mathcal{H}$ the map that maps an initial condition and a realization of the noise onto the solution of (5.3) after a time t.

Our next assumption is the existence of an appropriate Lyapunov function V:

B2 *There exists a measurable function $V : \mathcal{H} \to [0, \infty]$ and constants $a < 1$ and $b, c, d > 0$ such that*

$$\begin{aligned} \mathbf{E}_\omega V(\Phi(x,\omega)) &\leq aV(x) + b\;, \\ \mathbf{E}_\omega\Big(\sup_{0\leq t\leq 1} V(\Phi_t(x,\omega))\Big) &\leq cV(x) + d\;, \\ \mathsf{P}(\{\omega \mid V(\Phi(x,\omega)) = \infty\}) &= 0\;, \end{aligned} \tag{5.4}$$

for every $x \in \mathcal{H}$. Furthermore, V dominates the norm in $\mathcal{H}$ in the sense that $\|x\| \leq C(1 + V(x))$ for some constant C.

As is Section 3, we define $\tilde V(x, y) = V(x) + V(y)$.

Remark 5.2 Take $\mathcal{H}$ equal to $\mathrm{L}^2(\mathcal{O})$ for some regular bounded domain $\mathcal{O} \subset \mathbf{R}^d$, A a second-order elliptic differential operator on $\mathcal{O}$ with sufficiently smooth coefficients, and F any polynomial non-linearity of odd degree having the correct sign. The assumptions **B1** and **B2** are satisfied with $V(x) = \|x\|_\star^p$ for every power $p \geq 1$ and every "reasonable" norm $\|\cdot\|_\star$, as long as Q is "small" enough. (One can for example take for $\|\cdot\|_\star$ the L^∞ norm or a Sobolev norm.)

We now turn to the binding construction for the problem (5.1). Take a function $G : \mathcal{H}^2 \to \mathcal{W}$ and consider the $\mathcal{H}^2$-valued process (x, y) solving

$$dx(t) = Ax\,dt + F(x)\,dt + Q\,d\omega(t)\;, \tag{5.5a}$$

$$dy(t) = Ay\,dt + F(y)\,dt + Q\,G(x,y)\,dt + Q\,d\omega(t)\;. \tag{5.5b}$$

Notice that the realization of ω is the same for both components. The process (5.5) yields our binding construction for (5.1). In order to give sense to (5.5b), we introduce the $\mathcal{H}$-valued process $\varrho(t) = y(t) - x(t)$ and we define it pathwise as the solution of the ordinary differential equation

$$\dot\varrho = A\varrho + F(x+\varrho) - F(x) + Q\,G(x, x+\varrho)\;. \tag{5.6}$$

We assume that G is sufficiently regular to ensure the existence and uniqueness of global weak solutions to (5.6) for almost every (with respect to the measure on pathspace induced by Φ_t) continuous function $x : [0,\infty) \to \mathcal{H}$. This allows us to *define* the stochastic process $y(t)$ by $y(t) = x(t) + \varrho(t)$. We will denote by $\vec\Phi_t : \mathsf{X} \times \mathsf{X} \times \Omega \to \mathsf{X}$ the map that solves (5.5b) up to time t, given an initial condition for x and y, and a realization of the noise.

The above construction is invertible in the following sense. Consider the $\mathcal{H}^2$-valued process

$$d\tilde x(t) = A\tilde x\,dt + F(\tilde x)\,dt - Q\,G(\tilde x, \tilde y)\,dt + Q\,d\tilde\omega(t)\;, \tag{5.7a}$$

$$d\tilde y(t) = A\tilde y\,dt + F(\tilde y)\,dt + Q\,d\tilde\omega(t)\;, \tag{5.7b}$$

where we give sense to the equation for $\tilde{x}$ as above by setting $\tilde{\varrho} = \tilde{y} - \tilde{x}$ and solving

$$\dot{\tilde{\varrho}} = A\tilde{\varrho} + F(\tilde{y}) - F(\tilde{y} - \tilde{\varrho}) + Q\,G(\tilde{y} - \tilde{\varrho}, \tilde{y}) .$$

We denote by $\overleftarrow{\Phi}_t : \mathsf{X} \times \mathsf{X} \times \Omega \to \mathsf{X}$ the map that solves (5.7a) up to time t, given an initial condition for $\tilde{x}$ and $\tilde{y}$, and a realization of the noise $\tilde{\omega} \in \Omega$. We see that (5.7) can be obtained from (5.5) by the substitution $d\tilde{\omega} = d\omega + G(x, y)\,dt$ and a renaming of the variables. This observation yields the invertibility of the maps $\psi_{x\to y}$ defined in Eq. (5.12) below.

We will state two more assumptions to make sure that the conclusions of Theorem 4.1 hold. First, we want G to become small as x and y become close.

B3 *There exists a constant $C > 0$ and exponents $\alpha, \beta > 0$ such that*

$$\|G(x, y)\|^2 \le C\|x - y\|^\alpha (1 + \tilde{V}(x, y))^\beta , \tag{5.8}$$

for every $x, y \in \mathcal{H}$.

The last assumption ensures that the process $y(t)$ converges towards $x(t)$ for large times.

B4 *There exist positive constants C and γ such that the solutions of (5.5) and (5.7) satisfy*

$$\|\Phi_t(x, \omega) - \vec{\Phi}_t(x, y, \omega)\| \le Ce^{-\gamma t}\Big(1 + V(y) + \sup_{s\le t} V(\Phi_s(x, \omega))\Big) , \tag{5.9a}$$

$$\|\overleftarrow{\Phi}_t(x, y, \omega) - \Phi_t(y, \omega)\| \le Ce^{-\gamma t}\Big(1 + V(x) + \sup_{s\le t} V(\Phi_s(y, \omega))\Big) , \tag{5.9b}$$

for P^∞-almost every $\omega \in \Omega^\infty$. Furthermore, there exists $\delta > 0$ such that one has the estimate

$$V(\vec{\Phi}_t(x, y, \omega)) \le C\Big(1 + V(y) + \sup_{s\le t} V(\Phi_s(x, \omega))\Big)^\delta , \tag{5.10a}$$

$$V(\overleftarrow{\Phi}_t(x, y, \omega)) \le C\Big(1 + V(x) + \sup_{s\le t} V(\Phi_s(y, \omega))\Big)^\delta , \tag{5.10b}$$

for P^∞-almost every $\omega \in \Omega^\infty$ and every $t \ge 0$.

Remark 5.3 One important particular case is the choice $V(x) = \|x\|^p$, where the power p is chosen in such a way that (5.9) is satisfied. Notice that in this case, the estimates (5.10) are a straightforward consequence of (5.9).

The function G is then only required to satisfy a bound of the type

$$\|G(x, y)\|^2 \le C\|x - y\|^\alpha (1 + \|x\| + \|y\|)^q ,$$

with α and q some arbitrary positive exponents.

It is also possible to choose $V(x) = \|x\|_\star^p$, with $\|\cdot\|_\star$ the norm of some Banach space $\mathcal{B} \subset \mathcal{H}$. In this case, (5.9) with the $\mathcal{B}$-norm replacing the $\mathcal{H}$-norm in the left-hand side implies (5.10).

Remark 5.4 An equivalent way of writing (5.9b) is

$$\|\Phi_t(x,\omega) - \vec{\Phi}_t(x,y,\omega)\| \le Ce^{-\gamma t}\Big(1 + V(x) + \sup_{s\le t} V(\vec{\Phi}_s(x,y,\omega))\Big)\;. \tag{5.11}$$

The equation (5.11) will be more natural in our examples, but (5.9) is more symmetric and more convenient for the proof of Theorem 5.5 below.

All these assumptions together ensure that exponential mixing takes place:

Theorem 5.5 *Let A, F and Q be such that assumptions* **B1** *and* **B2** *are satisfied. If there exists a function $G : \mathcal{H}^2 \to \mathcal{W}$ such that assumptions* **B3** *and* **B4** *hold, then the solution of (5.1) possesses a unique invariant measure μ_* and there exist constants $C, \gamma > 0$ such that*

$$\|\mathbf{P}_x^n - \mu_*\|_{\mathrm{L}} \le Ce^{-\gamma n}(1 + V(x))\;.$$

Proof. It suffices to show that assumptions **A1–A5** hold. Assumption **A1** follows immediately from Assumption **B2**. In order to check the other assumptions, we define the various objects appearing in the previous sections. We have already seen that $\mathsf{X} = \mathcal{H}$, $\Omega = \mathcal{C}_0([0,1], \hat{\mathcal{W}})$, and Φ is the solution of (5.1) at time 1.

We define the function $W : \mathcal{H} \times \Omega \to [1, \infty]$ by

$$W(x,\omega) = \sup_{t\in[0,1]} V(\Phi_t(x,\omega))\;.$$

The estimate (5.4) and the definition ensure that W satisfies (3.3a) and (3.3b). The bound (5.10) ensures that Assumption **A2** is also satisfied.

It remains to define the binding functions $\psi_{x\to y}$ and to compute the densities $\mathcal{D}_{x,y}^n$. According to the constructions (5.5) and (5.7), we define for $(x,y) \in \mathcal{H}^2$ the binding functions

$$(\psi_{x\to y}(\omega))(t) = \omega(t) + \int_0^t G(\Phi_s(x,\omega), \vec{\Phi}_s(x,y,\omega))\,ds\;, \tag{5.12a}$$

$$(\psi_{x\leftarrow y}(\omega))(t) = \omega(t) - \int_0^t G(\overleftarrow{\Phi}_s(x,y,\omega), \Phi_s(y,\omega))\,ds\;, \tag{5.12b}$$

with $t \in [0,1]$. It follows from the construction that these functions are each other's inverse. Furthermore, if we identify Ω^n with $\mathcal{C}_0([0,n], \hat{\mathcal{W}})$ in a natural way, we see that the maps $\Xi_{x,y}^n$ introduced in (2.10) are obtained from (5.12) by simply letting t take values in $[0,n]$. These observations allow us to compute the densities $\mathcal{D}_{x,y}^n(\omega)$ by Girsanov's theorem:

Lemma 5.6 *The family of densities $\mathcal{D}_{x,y}^n(\omega)$ is given by*

$$\mathcal{D}_{x,y}^n(\omega) = \exp\Big(\int_0^n G(\overleftarrow{\Phi}_t(x,y,\omega), \Phi_t(y,\omega))\,d\omega(t) - \frac{1}{2}\int_0^n \|G(\ldots)\|^2\,dt\Big)\;,$$

where the arguments of G in the second term are the same as in the first term.

Proof. If we can show that Girsanov's theorem applies to our situation, then it precisely states that

$$\psi^*_{x\leftarrow y}\hat{\mathsf{P}}^n = \mathsf{P}^n \;,$$

with $\hat{\mathsf{P}}^n(d\omega) = \mathcal{D}^n_{x,y}(\omega)\,\mathsf{P}^n(d\omega)$, and $\mathcal{D}^n_{x,y}(\omega)$ defined as above. Applying $\psi^*_{x\to y}$ to both sides of the equality shows the result.

We now show that Girsanov's theorem can indeed be applied. By [DPZ92b, Thm 10.14], it suffices to verify that

$$\int_{\Omega^n} \mathcal{D}^n_{x,y}(\omega)\,\mathsf{P}^n(d\omega) = 1 \;. \tag{5.13}$$

This can be achieved by a suitable cut-off procedure. Define for $N > 0$ the function

$$G_N(x,y) = \begin{cases} G(x,y) & \text{if } \|G(x,y)\| \le N, \\ 0 & \text{otherwise,} \end{cases}$$

and define

$$\mathcal{D}^{n,N}_{x,y}(\omega) = \exp\Bigl(\int_0^n G_N(\overleftarrow{\Phi}_t(x,y,\omega),\Phi_t(y,\omega))\,d\omega(t) - \frac{1}{2}\int_0^n \|G_N(\ldots)\|^2\,dt\Bigr) \;.$$

It is immediate that (5.13) holds for $\mathcal{D}^{n,N}_{x,y}$. Furthermore, it follows from Assumption **B4** that there exists a constant C_N such that $\mathcal{D}^{n,N}_{x,y}(\omega) = \mathcal{D}^n_{x,y}(\omega)$ on the set

$$\Gamma_N = \{\omega \in \mathcal{P}^n \mid \tilde{V}(\overleftarrow{\Phi}_s(x,y,\omega),\Phi_s(y,\omega)) < C_N \quad \forall s \in [0,n]\} \;.$$

The sets Γ_N satisfy $\lim_{N\to\infty} \mathsf{P}^n(\Gamma_N) = 1$ by (5.4) and (5.10b). This shows that (5.13) holds. Notice that the *a-priori* bounds of Assumption **B4** were crucial in this step in order to apply Girsanov's theorem. The bound (5.8) alone could lead to exploding solutions for which Girsanov's theorem does not apply. □

It is immediate that Assumption **A3** follows from Assumption **B4** and the definition of the norm $\|\cdot\|_{\mathrm{L}}$.

We now turn to the verification of Assumption **A4**. Recalling the definition (3.5), we see that in our case

$$A_{y,k} \subset B_{y,k} \equiv \{\omega \in \Omega^\infty \mid V(\Phi_s(y,\omega)) \le k(V(y) + s^2) \quad \forall s \ge 0\} \;.$$

As we see from the definition of $B_{y,k}$, a natural definition for a truncation $G_{y,k}$ of G (this time the truncation additionally depends on time) is

$$G_{y,k}(\tilde{x},\tilde{y},t) = \begin{cases} G(\tilde{x},\tilde{y}) & \text{if } V(\tilde{y}) \le k(V(y)+t^2), \\ 0 & \text{otherwise.} \end{cases}$$

As above, we define

$$\mathcal{D}^{n,k}_{x,y}(\omega) = \exp\Bigl(\int_0^n G_{y,k}(\overleftarrow{\Phi}_t(x,y,\omega),\Phi_t(y,\omega),t)\,d\omega(t) - \frac{1}{2}\int_0^n \|G_{y,k}(\ldots)\|^2\,dt\Bigr) \;.$$

By definition, $\mathcal{D}^{n,k}_{x,y}(\omega) = \mathcal{D}^n_{x,y}(\omega)$ for every $\omega \in B_{y,k}$. Setting $\xi = \delta(\alpha+\beta)$, we thus have the estimate

$$\begin{aligned}\int_{A_{y,k}} (\mathcal{D}^n_{x,y}(\omega))^{-2}\, \mathsf{P}^n(d\omega) &\le \int_{B_{y,k}} (\mathcal{D}^{n,k}_{x,y}(\omega))^{-2}\, \mathsf{P}^n(d\omega)\\ &\le \Big(\int_{B_{y,k}} \exp\Big(10\int_0^n \|G_{y,k}(\overleftarrow{\Phi}_t(x,y,\omega),\Phi_t(y,\omega),t)\|^2\,dt\Big)\,\mathsf{P}^n(d\omega)\Big)^{1/2}\\ &\le \Big(\int_{B_{y,k}} \exp\Big(10\int_0^n Ce^{-\gamma t}\Big(1+V(x)+\sup_{s\le t}V(\Phi_s(y,\omega))\Big)^{\xi}\,dt\Big)\,\mathsf{P}^n(d\omega)\Big)^{1/2}\\ &\le \exp\Big(C\int_0^n e^{-\gamma t}(1+k\tilde V(x,y)+kt^2)^{\xi}\,dt\Big)\;.\end{aligned}$$

In this expression, we used the Cauchy-Schwarz inequality to go from the first to the second line, and we used assumptions **B3** and **B4** to go from the second to the third line. Since the integral converges for $n\to\infty$, the bound is uniform in n and Assumption **A4** is verified.

The verification of Assumption **A5** is quite similar. Fix some positive constant $K > 0$ and use again the cutoff function

$$G_N(\tilde x,\tilde y) = \begin{cases} G(\tilde x,\tilde y) & \text{if } \|G(\tilde x,\tilde y)\|^2 \le N,\\ 0 & \text{otherwise.}\end{cases}$$

The precise value of N (as a function of K) will be fixed later. We also fix a pair $(x_0,y_0)\in\mathcal{H}^2$ with $\tilde V(x_0,y_0)\le K$, a value $n>0$, and initial conditions $(x,y)\in Q^n_K(x_0,y_0)$. By the definition of $Q^n_K(x_0,y_0)$, there exists an element $\tilde\omega\in\Omega^n$ such that

$$(x,y) = (\overleftarrow{\Phi}_n(x_0,y_0,\tilde\omega),\Phi_n(y_0,\tilde\omega))\;, \tag{5.14}$$

and such that

$$\sup_{s\in[0,n]} \tilde V(\overleftarrow{\Phi}_s(x_0,y_0,\tilde\omega),\Phi_s(y_0,\tilde\omega)) \le K\;. \tag{5.15}$$

Following the statement of Assumption **A5**, we define the set

$$B^K_{x,y} = \Big\{\omega\in\Omega\;\Big|\;\sup_{t\in[0,1]}\tilde V(\overleftarrow{\Phi}_t(x,y,\omega),\Phi_t(y,\omega))\le K\Big\}\;,$$

which is equal in our setup to the set over which integration is performed in (3.9). Being now accustomed to these truncation procedures, we define again

$$\mathcal{D}^{(K)}_{x,y}(\omega) = \exp\Big(\int_0^1 G_N(\overleftarrow{\Phi}_t(x,y,\omega),\Phi_t(y,\omega))\,d\omega(t) - \frac12\int_0^1 \|G_N(\ldots)\|^2\,dt\Big)\;.$$

By (5.14) and the cocycle property, we can write the integral in the above expression as

$$\int_0^1 G_N(\overleftarrow{\Phi}_{n+t}(x_0,y_0,\tilde\omega\omega),\Phi_{n+t}(y_0,\tilde\omega\omega))\,d\omega(t)\;,$$

where $\tilde\omega\omega$ is the realization of the noise which is equal to $\tilde\omega$ for a time n and then to ω for a time 1. Using the *a-priori* bound (5.15) as well as assumptions **B3** and **B4**, we thus see that there exists a constant C such that the choice

$$N = Ce^{-\alpha\gamma n}(1+K)^{\alpha+\beta}\;,$$

ensures that $\mathcal{D}_{x,y}^{(K)}(\omega)$ is equal to $\mathcal{D}_{x,y}(\omega)$ for $\omega \in B_{x,y}^K$.

We then have the estimate

$$\begin{aligned}\int_{B_{x,y}^K} (1 - \mathcal{D}_{x,y}(\omega))^2 \,\mathsf{P}(d\omega) &\le \int_\Omega (1 - \mathcal{D}_{x,y}^{(K)}(\omega))^2 \,\mathsf{P}(d\omega) \\ &= \int_\Omega (\mathcal{D}_{x,y}^{(K)}(\omega))^2 \,\mathsf{P}(d\omega) - 1 \\ &\le \Big(\int_\Omega \exp\Big(6 \int_0^1 \|G_N(\overleftarrow{\Phi}_t(x,y,\omega), \Phi_t(y,\omega))\|^2 \,dt\Big) \,\mathsf{P}(d\omega)\Big)^{1/2} - 1 \\ &\le \exp(Ce^{-\alpha\gamma n}(1+K)^{\alpha+\beta}) - 1 \;.\end{aligned}$$

If we take $n \ge \beta \ln(1+K)/\gamma$, the exponent is bounded by C and there exists a constant C' such that

$$\int_{B_{x,y}^K} (1 - \mathcal{D}_{x,y}(\omega))^2 \,\mathsf{P}(d\omega) \le C' e^{-\alpha\gamma n}(1+K)^{\alpha+\beta} \;,$$

thus validating Assumption **A5** with $\gamma_2 = \alpha\gamma$ and $\zeta = \alpha + \beta$.

The proof of Theorem 5.5 is complete. □

6 Examples

Numerous recent results show that the invariant measure for the 2D Navier-Stokes equation (and also for other dissipative PDEs) is unique if a sufficient number of low-frequency modes are forced by the noise [BKL00a, BKL00b, EMS01, Mat01, EL01, KS00, KS01, MY01]. These results are not covered directly by Theorem 5.5, but some more work is needed. The reason is that the sets A_x^k defined in (3.5) are not the natural sets that allow to control the influence of the low-frequency modes onto the high-frequency modes in the 2D Navier-Stokes equation.

On the other hand, our formulation of Theorem 5.5 makes it quite easy to verify that the n-dimensional Ginzburg-Landau equation (in a bounded domain) shows exponential mixing properties, if sufficiently many low-frequency modes are forced by the noise. We verify this in the following subsection.

6.1 The Ginzburg-Landau equation

We consider the SPDE given by

$$du = (\Delta u + u - u^3)\,dt + Q\,dw(t) \;, \qquad u(0) = u_0 \;, \tag{6.1}$$

where the function u belongs to the Hilbert space

$$\mathcal{H} = \mathrm{L}^2([-L,L]^n, \mathbf{R}) \;,$$

and Δ denotes the Laplacean with periodic boundary conditions. The symbol $Q\,d\omega(t)$ stands as a shorthand for

$$Q\,d\omega(t) \equiv \sum_{i=1}^N q_i e_i \,d\omega_i(t) \;,$$

where $\{q_i\}_{i=1}^N$ is a collection of strictly positive numbers, e_i denotes the ith eigenvector of the Laplacean, and the ω_i are N independent Brownian motions (for some finite integer N). We also denote by λ_i the eigenvalue of Δ corresponding to e_i and we assume that they are ordered by $\ldots \leq \lambda_2 \leq \lambda_1 \leq 0$. We will see that it is fairly easy to construct a binding function G for which the assumptions of the previous section hold with $V(u) = \|u\|$, where $\|\cdot\|$ denotes the norm of $\mathcal{H}$.

In [DPZ96], it is shown that (6.1) possesses a unique mild solution for initial conditions $u_0 \in \mathrm{L}^\infty([-L,L]^n)$. It is straightforward to extend this to every initial condition $u_0 \in \mathcal{H}$, by using the regularizing properties of the heat semigroup. Thus, Assumption **B1** holds and we denote by P_u^t the transition probabilities of the solution at time t starting from u. We have the following result:

Theorem 6.1 *There exist positive constants C and γ, and a unique measure $\mu_* \in \mathscr{M}_1(\mathcal{H})$ such that*

$$\|\mathsf{P}_u^t - \mu_*\|_{\mathrm{L}} \leq Ce^{-\gamma t}(1 + \|u\|)\;, \tag{6.2}$$

for every $u \in \mathcal{H}$ and every $t > 0$.

Proof. We verify that the assumptions of Theorem 5.5 hold. The bounds required for the verification of Assumption **B2** can be found in [Cer99, DPZ96], for example.

It remains to construct the forcing $G : \mathcal{H}^2 \to \mathbf{R}^N$ and to verify assumptions **B3** and **B4**. We consider two copies u_1 and u_2 of (6.1), with the noise $d\omega$ replaced by $d\omega + G\,dt$ in the second copy. We also denote by $\varrho = u_2 - u_1$ the difference process. It satisfies the differential equation

$$\dot\varrho = \Delta\varrho + \varrho - \varrho(u_1^2 + u_1u_2 + u_2^2)\; + Q\,G(u_1,u_2)\;. \tag{6.3}$$

We can project (6.3) onto the direction given by e_k. This yields

$$\dot\varrho_k = (\lambda_k + 1)\varrho_k - \Big(\varrho(u_1^2 + u_1u_2 + u_2^2)\Big)_k + q_kG_k(u_1,u_2)\;,$$

for $k = 1,\ldots,N$ and

$$\dot\varrho_k = (\lambda_k + 1)\varrho_k - \Big(\varrho(u_1^2 + u_1u_2 + u_2^2)\Big)_k\;,$$

for $k > N$. We choose G_k for $k = 1,\ldots,N$ as

$$G_k(u_1,u_2) = -\frac{2+\lambda_k}{q_k}\varrho_k\;.$$

Since G_k can only be defined this way if $q_k \neq 0$, we use at this point the fact that the noise acts directly and independently on *every unstable mode*. This requirement can be significantly weakened with the help of Theorem 5.5. We will focus next on more degenerate problems which illustrate the power of our technique.

This choice satisfies Assumption **B3**. With this choice, we can write down the evolution of the norm of ϱ as

$$\frac{d\|\varrho\|^2}{dt} = 2\langle\varrho, A\varrho\rangle - 2\langle\varrho,\varrho(u_1^2 + u_1u_2 + u_2^2)\rangle\;,$$

with A the linear operator given by adding up the contribution of $\Delta + 1$ and the contribution of G. By the condition we imposed on N, there exists a constant $a > 0$ such that $\langle\varrho, A\varrho\rangle \leq -a\|\varrho\|^2$. Furthermore, one has

$$\langle\varrho,\varrho(u_1^2 + u_1u_2 + u_2^2)\rangle \geq 0\;.$$

We thus have the differential inequality

$$\frac{d\|\varrho\|^2}{dt} \leq -2a\|\varrho\|^2 ,$$

which implies that

$$\|\varrho(t)\| \leq e^{-at}\|\varrho(0)\| .$$

This implies by Remark 5.3 that Assumption **B4** is also satisfied. The proof of Theorem 6.1 is complete. □

6.2 A reaction-diffusion system

Consider the following reaction-diffusion system:

$$\begin{aligned} du &= (\Delta u + 2u + v - u^3)\, dt + dw(t) , \\ dv &= (\Delta v + 2v + u - v^3)\, dt , \end{aligned} \tag{6.4}$$

where the pair (u, v) belongs to the Hilbert space

$$\mathcal{H} = \mathcal{H}_u \oplus \mathcal{H}_v = \mathrm{L}^2([-L, L], \mathbf{R}) \oplus \mathrm{L}^2([-L, L], \mathbf{R}) .$$

The symbol Δ again denotes the Laplacean with periodic boundary conditions, and $d\omega$ is the cylindrical Wiener process on $\mathcal{H}_u$ (meaning that it is space-time white noise).

Notice that, because of the presence of v, this system does not satisfy the assumptions stated in the papers mentioned at the beginning of this section. In other words, even though the forcing is infinite-dimensional, not all the determining modes for (6.4) are forced.

We take as our Lyapunov function

$$V(u, v) = \|u\|_\infty + \|v\|_\infty ,$$

with $\|\cdot\|_\infty$ the L^∞ norm. As in the previous subsection, one can show that with this choice of V, our problem satisfies assumptions **B1** and **B2**. We will now construct a binding function G which satisfies assumptions **B4** and **B3**. We consider, as in (5.5), two copies (u_1, v_1) and (u_2, v_2) of the system (6.4), but the noise is modified by G on the second copy. We also define $\varrho_u = u_2 - u_1$ and $\varrho_v = v_2 - v_1$. We then have

$$\begin{aligned} \dot\varrho_u &= \Delta\varrho_u + 2\varrho_u + \varrho_v - \varrho_u(u_1^2 + u_1u_2 + u_2^2) + G(u_1, u_2, v_1, v_2) , \\ \dot\varrho_v &= \Delta\varrho_v + 2\varrho_v + \varrho_u - \varrho_v(v_1^2 + v_1v_2 + v_2^2) . \end{aligned} \tag{6.5}$$

Our construction of G is inspired from the construction we presented in Section 1.1. We introduce the variable $\zeta = \varrho_u + 3\varrho_v$. Substituting this in (6.5), it defines the function G if we impose that the equation for ζ becomes

$$\dot\zeta = \Delta\zeta - \zeta , \tag{6.6}$$

so that $\|\zeta(t)\|^2 \leq \|\zeta(0)\|^2 e^{-t}$. Notice that the function G achieving this identity satisfies a bound of the type

$$\|G\| \leq C(\|\varrho_u\| + \|\varrho_v\|)(1 + \|u_1\|_\infty + \|u_2\|_\infty + \|v_1\|_\infty + \|v_2\|_\infty)^2 ,$$

thus satisfying Assumption **B3**. It remains to show that Assumption **B4** is satisfied. The equation for ϱ_v reads

$$\dot{\varrho}_v = \Delta\varrho_v - \varrho_v + \zeta - \varrho_v(v_1^2 + v_1 v_2 + v_2^2) \;.$$

Therefore, the norm of ϱ_v satisfies

$$\|\varrho_v(t)\|^2 \le \|\varrho_v(0)\|^2 e^{-t} + \frac{1 + \|\zeta(0)\|^2}{2} e^{-t} \;.$$

This in turn implies, through the definition of ζ and the bound on $\|\zeta(t)\|$, that a similar bound holds for $\|\varrho_u(t)\|$. This shows that the bound (5.9) is satisfied. Similar estimates hold with the L^∞ norm replacing the L^2 norm, and so Assumption **B4** is satisfied by Remark 5.3.

In fact, a straightforward computation, which can be found in [Cer99, Hai01, GM01] for example, shows that in this example, one can get a uniform estimate on the Lyapunov function V. More precisely, there exists a constant C such that for all initial conditions $x \in \mathcal{H}$,

$$\int_{\mathcal{H}} V(y)\, \mathsf{P}_x(dy) \le C \;. \tag{6.7}$$

Denoting by $\mathcal{P}_t^*$ the semigroup acting on measures generated by the solutions of (6.4), we thus have:

Theorem 6.2 *There exists a unique probability measure $\mu_* \in \mathscr{M}_1(\mathcal{H})$ such that $\mathcal{P}_t^*\mu_* = \mu_*$ for every $t \ge 0$. Furthermore, there exist constants C and γ such that*

$$\|\mathcal{P}_t^*\nu - \mu_*\|_{\mathrm{L}} \le C e^{-\gamma t} \;, \tag{6.8}$$

for every $\nu \in \mathscr{M}_1(\mathcal{H})$.

Proof. Combining (6.7) with the results of Theorem 4.1 and a computation similar to what was done in the proof of Corollary 4.3, we get (6.8) for integer times. The generalization to arbitrary times is straightforward, using the fact that the growth rate of the difference process (ϱ_u, ϱ_v) (with $G \equiv 0$) can easily be controlled. □

Remark 6.3 In fact, the dependence on u in the right-hand side of (6.2) can be removed similarly by checking that an estimate of the type (6.7) is verified for the solutions of the stochastic Ginzburg-Landau equation (6.1).

6.3 A chain with nearest-neighbour interactions

In the previous example, the noise acted on infinitely many degrees of freedom in a non-degenerate way. As a consequence, one step was sufficient to transmit the noise to the entire system. We will now look at a much more degenerate system, where the noise acts on only *one* degree of freedom, although an *arbitrary* number of modes are linearly unstable.

Our model is given by

$$\begin{aligned} dx_0 &= (a^2 x_0 + x_1 - x_0^3)\, dt + d\omega \;, \\ \dot{x}_k &= (a^2 - k^2) x_k + x_{k-1} + x_{k+1} - x_k^3 \;, \qquad k = 1, 2, \ldots \;, \end{aligned} \tag{6.9}$$

where $a \in \mathbf{R}$ is an arbitrary constant. One should think of the deterministic part of (6.9) as a very simple model for a dissipative PDE of the Ginzburg-Landau type. We will consider (6.9) in the (real) Hilbert space $\mathcal{H} = \ell^2$ endowed with its canonical orthonormal basis $\{e_k\}_{k=0}^\infty$. It is easy to verify that (6.9) possesses a unique solution. We denote again by $\mathcal{P}_t^*$ the semigroup acting on measures $\nu \in \mathscr{M}(\ell^2)$ generated by (6.9). We will show

Theorem 6.4 *For the problem (6.9), there exists a unique probability measure $\mu_* \in \mathscr{M}_1(\ell^2)$ such that $\mathcal{P}_t^* \mu_* = \mu_*$ for every $t \geq 0$. Furthermore, there exist constants C and γ such that*

$$\|\mathcal{P}_t^* \nu - \mu_*\|_{\mathrm{L}} \leq C e^{-\gamma t} ,$$

for every $\nu \in \mathscr{M}_1(\mathcal{H})$.

Proof. We will take as our Lyapunov function $V(x) = \|x\|^p$ for some power of p to be fixed later. It is a straightforward task to verify that the dynamics generated by (6.9) does indeed satisfy assumptions **B1** and **B2** for this choice of V.

We next show that a bound of the type (6.7) holds for the solutions of (6.9), thus yielding the uniformity in the convergence towards the invariant measure μ_*. Let us define the process $y(t) \in \ell^2$ by $y(t) = x(t) - \omega(t) e_0$. This process then satisfies the following system of differential equations:

$$\begin{aligned} \dot y_0 &= a^2(y_0 + \omega) + y_1 - (y_0 + \omega)^3 , \\ \dot y_1 &= (a^2 - 1) y_1 + y_0 + y_2 - y_1^3 + \omega , \\ \dot y_k &= (a^2 - k^2) y_k + y_{k-1} + y_{k+1} - y_k^3 , \qquad k = 2, 3, \ldots . \end{aligned} \tag{6.10}$$

We denote by $\|y\|_\infty$ the norm given by $\sup_k |y_k|$. It follows from [Lun95] that (6.10) possesses a strong solution for positive times. Furthermore, from (6.10) and the definition of the $\|\cdot\|_\infty$-norm, we see that there are constants $c_1, c_2 > 0$ such that

$$\frac{D_- \|y\|_\infty}{Dt} \leq -c_1 \|y\|_\infty^3 + c_2 (1 + |\omega(t)|^3) , \tag{6.11}$$

where D_-/Dt denotes the left-handed lower Dini derivative. A straightforward computation shows that (6.11) implies that there exists a constant C such such that

$$\|y(1/2)\|_\infty \leq C \sup_{t \in [0,1/2]} (1 + |\omega(t)|) ,$$

independently of the initial condition. In order to conclude the proof of the estimate (6.7), it suffices to show that there exists a constant C such that

$$\mathbf{E} \|y(1/2)\| \leq C (1 + \|y(0)\|_\infty) .$$

This follows easily from the dissipativity of the nonlinearity in $\mathcal{H}$ and the fact that the semigroup generated by the linear part of (6.10) is bounded from ℓ^∞ into ℓ^2.

It remains to verify that the assumptions **B1–B4** are indeed satisfied for some binding function G. This, together with the uniform bound obtained above, shows that the conclusions of

Theorem 6.4 hold. As for the toy model presented in Section 1.1, we consider a process $y \in \ell^2$ governed by the same equation as (6.9), but with $d\omega$ replaced by $d\omega + G(x,y)\,dt$. We then introduce the difference process $\varrho = y - x$, which is given by the solution of

$$\begin{aligned} \dot\varrho_0 &= a^2\varrho_0 + \varrho_1 - \varrho_0(x_0^2 + x_0y_0 + y_0^2) + G(x,y)\;, \\ \dot\varrho_k &= (a^2 - k^2)\varrho_k + \varrho_{k+1} + \varrho_{k-1} - \varrho_k(x_k^2 + x_ky_k + y_k^2)\;. \end{aligned} \tag{6.12}$$

The aim of the game is to find a function G for which $\varrho(t) \to 0$ as $t \to \infty$. We can split (6.12) into “low modes” and “high modes” by introducing

$$k_* = \inf\{k > 0 \mid k^2 - a^2 \ge 3\}\;.$$

At the level of the Hilbert space ℓ^2, we set $\ell^2 = \mathcal{H}_L \oplus \mathcal{H}_H$, where $\mathcal{H}_L \approx \mathbf{R}^{k_*}$ is generated by $e_0, \ldots, e_{k_*-1}$ and $\mathcal{H}_H$ is its orthogonal complement. We denote by ϱ_L and ϱ_H the components of ϱ and by A_H the restriction (as a symmetric quadratic form) of the linear part of (6.9) to $\mathcal{H}_H$. It is by construction easy to see that

$$\langle \varrho_H, A_H\varrho_H\rangle \ge \|\varrho_H\|^2\;.$$

As a consequence, we have for $\|\varrho_H\|^2$ the following estimate:

$$\|\varrho_H(t)\|^2 \le e^{-t}\|\varrho_H(0)\|^2 + \frac{1}{4}\int_0^t e^{t-s}|\zeta_1(s)|^2\,ds\;, \tag{6.13}$$

where we defined $\zeta_1 = \varrho_{k_*-1}$. (The reason for renaming ϱ_{k_*-1} this way will become clear immediately.) It remains to construct G in such a way to get good estimates on $\|\varrho_L(t)\|^2$. In order to achieve this, we use again the same method as for the first toy model. The variable ζ_1 obeys the equation

$$\dot\zeta_1 = c_1\varrho_{k_*-1} + \varrho_{k_*} + \varrho_{k_*-2} - \varrho_{k_*-1}(x_{k_*-1}^2 + x_{k_*-1}y_{k_*-1} + y_{k_*-1}^2)\;,$$

with some constant $c_1 \in \mathbf{R}$. We thus introduce a new variable ζ_2 defined by

$$\zeta_2 = (c_1+1)\varrho_{k_*-1} + \varrho_{k_*} + \varrho_{k_*-2} - \varrho_{k_*-1}(x_{k_*-1}^2 + x_{k_*-1}y_{k_*-1} + y_{k_*-1}^2)\;.$$

It is important to notice two facts about this definition. The first is that it yields for $|\zeta_1|^2$ the estimate

$$|\zeta_1(t)|^2 \le e^{-t}|\zeta_1(0)|^2 + \frac{1}{4}\int_0^t e^{t-s}|\zeta_2(s)|^2\,ds\;. \tag{6.14}$$

The second is that ζ_2 can be written in the form

$$\zeta_2 = \varrho_{k_*-2} + \mathcal{Q}_2(\varrho, x, y)\;,$$

where $\mathcal{Q}_2$ is a polynomial depending only on components ϱ_i, x_i and y_i with $i \ge k_* - 1$, and such that each of its terms contains at least one factor ϱ_i.

Now look at the equation for $\dot\zeta_2$. It is clear from the structure of ζ_2 and from the structure of the equations (6.9) and (6.12) that it can be written as

$$\dot\zeta_2 = -\zeta_2 + \zeta_3\;,$$

where

$$\zeta_3 = \varrho_{k_*-3} + \mathcal{Q}_3(\varrho, x, y) \; .$$

This time, the polynomial $\mathcal{Q}_3$ depends only on components with an index $i \geq k_* - 2$. This procedure can be iterated, yielding a whole family of variables

$$\zeta_l = \varrho_{k_*-l} + \mathcal{Q}_l(\varrho, x, y) \; , \tag{6.15}$$

where the $\mathcal{Q}_l$ are polynomials depending only on indices $i \geq k_* - l + 1$, and containing at least one factor ϱ_i in each term. Furthermore, one gets for every ζ_l the estimate

$$|\zeta_l(t)|^2 \leq e^{-t}|\zeta_l(0)|^2 + \frac{1}{4}\int_0^t e^{t-s}|\zeta_{l+1}(s)|^2\, ds \; . \tag{6.16}$$

Notice that (6.16) is valid for $l < k_*$. For $l = k_*$, we have

$$\dot{\zeta}_{k_*} = \mathcal{Q}_{k_*+1}(\varrho, x, y) + G(x, y) \; . \tag{6.17}$$

It thus suffices to choose G in such a way that (6.17) becomes

$$\dot{\zeta}_{k_*} = -\zeta_{k_*} \; . \tag{6.18}$$

Denoting by ζ the vector $\zeta_1, \ldots, \zeta_{k_*}$, we get from (6.16) and (6.18) the estimate

$$\|\zeta(t)\|^2 \leq Ce^{-\gamma t}\|\zeta(0)\|^2 \; , \tag{6.19}$$

for any $\gamma \in (0, 1)$. Plugging this into (6.13) yields for $\|\varrho_H\|$ the estimate

$$\begin{aligned}
\|\varrho_H(t)\|^2 &\leq Ce^{-\gamma t}(\|\varrho_H(0)\|^2 + \|\zeta(0)\|^2) \\
&\leq Ce^{-\gamma t}\|\varrho(0)\|^2(1 + \|x(0)\| + \|y(0)\|)^p \; ,
\end{aligned}$$

for some constants C, γ and p. It remains to get an estimate on $\|\varrho_L\|$. From (6.19) and the definition of ζ_1, we get immediately

$$|\varrho_{k_*-1}(t)|^2 \leq Ce^{-\gamma t}\|\varrho(0)\|^2(1 + \|x(0)\| + \|y(0)\|)^p \; .$$

From the definition of ζ_2, we get

$$|\varrho_{k_*-2}(t)|^2 \leq C\Big(|\zeta_2(t)|^2 + |\mathcal{Q}_2(\varrho(t), x(t), y(t))|^2\Big) \; .$$

But we know that $\mathcal{Q}_2$ only depends on components of ϱ, x, and y with an index $i \geq k_* - 1$. These are precisely the components of ϱ on which we already have an estimate. We thus get

$$|\varrho_{k_*-2}(t)|^2 \leq Ce^{-\gamma t}\|\varrho(0)\|^2(1 + \|x(0)\| + \|y(0)\| + \|x(t)\|)^p \; ,$$

for some other power p. Here we used the fact that $y(t) = x(t) + \varrho(t)$ to get rid of $\|y(t)\|$ in the estimate. The same reasoning can be applied to ϱ_{k_*-3}, and so forth down to ϱ_0. We finally get

$$\|\varrho_L(t)\|^2 \leq Ce^{-\gamma t}\|\varrho(0)\|^2(1 + \|x(0)\| + \|y(0)\| + \|x(t)\|)^p \; , \tag{6.20}$$

for some (large) power of p. We thus verified (5.9a). The bound (5.9b) is obtained in the same way, by noticing that we can as well get the estimate

$$\|\varrho_L(t)\|^2 \leq Ce^{-\gamma t}\|\varrho(0)\|^2(1 + \|x(0)\| + \|y(0)\| + \|y(t)\|)^p \; ,$$

instead of (6.20). The proof of Theorem 6.4 is complete. □

Remark 6.5 The whole construction is strongly reminiscent of what was done in [EPR99b] to control a finite Hamiltonian chain of non-linear oscillators with nearest-neighbour coupling driven by thermal noise at its boundaries.

Remark 6.6 The linearity of the nearest-neighbour coupling is not essential for our argument. We could as well have replaced (6.9) by

$$\begin{aligned} dx_0 &= (a^2 x_0 + V_2'(x_1 - x_0) - V_1'(x_0))\, dt + d\omega \;, \\ \dot{x}_k &= (a^2 - k^2)x_k + V_2'(x_{k-1} - x_k) + V_2'(x_{k+1} - x_k) - V_1'(x_k) \;, \end{aligned}$$

with V_1 and V_2 two polynomial-like functions, *i.e.*smooth functions such that

$$\frac{d^n V_i(x)}{dx^n} \approx x^{\alpha_i - n} \qquad \text{for} \qquad |x| \to \infty \;,$$

for some $\alpha_i \geq 2$. Imposing the condition $V_2''(x) \geq c$ for some $c > 0$ yields an effective coupling between neighbours at every point of the phase space. This is sufficient to apply our construction.

Références

[Bil68] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, 1968.

[Bis84] J.-M. Bismut, *Large Deviations and the Malliavin Calculus*, Birkhäuser Boston Inc., Boston, MA, 1984.

[BK37] N. N. Bogoluboff and N. M. Kriloff, *La Théorie Générale de la Mesure dans son Application à l'Etude des Systèmes Dynamiques de la Mécanique Non-Linéaire*, Ann. of Math. **38** (1937), 65–113.

[BKL00a] J. Bricmont, A. Kupiainen, and R. Lefevere, *Ergodicity of the 2D Navier-Stokes Equations with Random Forcing*, Preprint, 2000.

[BKL00b] J. Bricmont, A. Kupiainen, and R. Lefevere, *Exponential Mixing of the 2D Stochastic Navier-Stokes Dynamics*, Preprint, 2000.

[BKL00c] J. Bricmont, A. Kupiainen, and R. Lefevere, *Probabilistic Estimates for the Two Dimensional Stochastic Navier-Stokes Equations*, J. Stat. Phys. (2000), no. 3–4, 743–756.

[CDF97] H. Crauel, A. Debussche, and F. Flandoli, *Random Attractors*, J. Dynam. Differential Equations **9** (1997), no. 2, 307–341.

[Cer99] S. Cerrai, *Smoothing Properties of Transition Semigroups Relative to SDEs with Values in Banach Spaces*, Probab. Theory Relat. Fields **113** (1999), 85–114.

[Col94] P. Collet, *Non Linear Parabolic Evolutions in Unbounded Domains*, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci **437** (1994), 97–104.

[Col98] P. Collet, *Stochasticity and Chaos*, Probability Towards 2000, Lecture Notes in Statistics, vol. 128, Springer, New York, 1998, pp. 137–150.

[CP89] J. Carr and R. Pego, *Metastable Patterns in Solutions of* $u_t = \varepsilon^2 u_{xx} - f(u)$, Commun. Pure Appl. Math. **42** (1989), no. 5, 523–576.

[CP90] J. Carr and R. Pego, *Invariant Manifolds for Metastable Patterns in* $u_t = \varepsilon^2 u_{xx} - f(u)$, Proc. Roy. Soc. Edinburgh **116A** (1990), no. 1-2, 133–160.

[Die68] J. Dieudonné, *Foundations of Modern Analysis*, 7th ed., Academic Press, New York and London, 1968.

[Dob71] R. L. Dobrušin, *Markov Processes with a Large Number of Locally Interacting Components: Existence of a Limit Process and its Ergodicity*, Problemy Peredači Informacii **7** (1971), no. 2, 70–87.

[Doe38] W. Doeblin, *Exposé sur la Théorie des Chaînes Simples Constantes de Markoff à un Nombre Fini d'États*, Rev. Math. Union Interbalkanique **2** (1938), 77–105.

[Doo48] J. L. Doob, *Asymptotic Properties of Markoff Transition Probabilities*, Trans. Amer. Math. Soc. **63** (1948), 393–421.

[Doo53] J. L. Doob, *Stochastic Processes*, John Wiley & Sons, 1953.

[DPEZ95] G. Da Prato, K. D. Elworthy, and J. Zabczyk, *Strong Feller Property for Stochastic Semi-Linear Equations*, Stochastic Anal. Appl. **13** (1995), 35–45.

[DPZ91] G. Da Prato and J. Zabczyk, *Smoothing Properties of Transition Semigroups in Hilbert Spaces*, Stochastics Stochastics Rep. **35** (1991), 63–77.

[DPZ92a] G. Da Prato and J. Zabczyk, *Non-explosion, Boundedness and Ergodicity for Stochastic Semilinear Equations*, J. Diff. Equ. **98** (1992), 181–195.

[DPZ92b] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, University Press, Cambridge, 1992.

[DPZ96] G. Da Prato and J. Zabczyk, *Ergodicity for Infinite Dimensional Systems*, London Mathematical Society Lecture Note Series, vol. 229, University Press, Cambridge, 1996.

[EH00] J.-P. Eckmann and M. Hairer, *Non-Equilibrium Statistical Mechanics of Strongly Anharmonic Chains of Oscillators*, Commun. Math. Phys. **212** (2000), no. 1, 105–164.

[EH01a] J.-P. Eckmann and M. Hairer, *Invariant Measures for Stochastic PDE's in Unbounded Domains*, Nonlinearity **14** (2001), 133–151.

[EH01b] J.-P. Eckmann and M. Hairer, *Uniqueness of the Invariant Measure for a Stochastic PDE Driven by Degenerate Noise*, Commun. Math. Phys. **219** (2001), no. 3, 523–565.

[EL94] K. D. Elworthy and X.-M. Li, *Formulae for the Derivatives of Heat Semigroups*, J. Funct. Anal. **125** (1994), no. 1, 252–286.

[EL01] W. E and D. Liu, *Gibbsian Dynamics and Invariant Measures for Stochastic Dissipative PDEs*, Preprint, Princeton, 2001.

[EMS01] W. E, J. C. Mattingly, and Y. G. Sinai, *Gibbsian Dynamics and Ergodicity for the Stochastically Forced Navier-Stokes Equation*, to appear in Commun. Math. Phys., 2001.

[EPR99a] J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet, *Non-Equilibrium Statistical Mechanics of Anharmonic Chains Coupled to Two Heat Baths at Different Temperatures*, Commun. Math. Phys. **201** (1999), 657–697.

[EPR99b] J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet, *Entropy Production in Non-Linear, Thermally Driven Hamiltonian Systems*, J. Stat. Phys. **95** (1999), 305–331.

[ER98] J.-P. Eckmann and J. Rougemont, *Coarsening by Ginzburg-Landau Dynamics*, Commun. Math. Phys. **199** (1998), no. 2, 441–470.

[FLS96] E. Feireisl, P. Laurençot, and F. Simondon, *Global Attractors for Degenerate Parabolic Equations on Unbounded Domains*, J. Diff. Equ. **129** (1996), 239–261.

[FM95] F. Flandoli and B. Maslowski, *Ergodicity of the* 2*-D Navier-Stokes Equation Under Random Perturbations*, Commun. Math. Phys. **172** (1995), no. 1, 119–141.

[GLP99] G. Giacomin, J. L. Lebowitz, and E. Presutti, *Deterministic and Stochastic Hydrodynamic Equations Arising from Simple Microscopic Model Systems*, Stochastic partial differential equations: six perspectives, Amer. Math. Soc., Providence, RI, 1999, pp. 107–152.

[GM01] B. Goldys and B. Maslowski, *Uniform Exponential Ergodicity of Stochastic Dissipative Systems*, to appear in Czechoslovak Math. J., 2001.

[Hai01] M. Hairer, *Exponential Mixing for a Stochastic PDE Driven by Degenerate Noise*, Preprint, Geneva, 2001.

[Hen81] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, Lecture Notes in Mathematics, vol. 840, Springer, New York, 1981.

[Hör67] L. Hörmander, *Hypoelliptic Second Order Differential Equations*, Acta Math. **119** (1967), 147–171.

[Hör85] L. Hörmander, *The Analysis of Linear Partial Differential Operators I–IV*, Springer, New York, 1985.

[Ich84] A. Ichikawa, *Semilinear Stochastic Evolution Equations: Boundedness, Stability and Invariant Measures*, Stochastics **12** (1984), 1–39.

[JLM85] G. Jona-Lasinio and P. K. Mitter, *On the Stochastic Quantization of Field Theory*, Commun. Math. Phys. **101** (1985), 406–436.

[Köt83] G. Köthe, *Topological Vector Spaces I–II*, Springer, New York, 1983.

[KS00] S. B. Kuksin and A. Shirikyan, *Stochastic Dissipative PDE's and Gibbs Measures*, Commun. Math. Phys. **213** (2000), 291–230.

[KS01] S. B. Kuksin and A. Shirikyan, *A Coupling Approach to Randomly Forced Nonlinear PDE's. I*, Commun. Math. Phys. **221** (2001), 351–366.

[Lig77] T. M. Liggett, *The Stochastic Evolution of Infinite Systems of Interacting Particles*, École d'Été de Probabilités de Saint-Flour, VI, Lecture Notes in Mathematics, vol. 598, Springer-Verlag, Berlin, 1977, pp. 187–248.

[Lun95] A. Lunardi, *Analytic Semigroups and Optimal Regularity in Parabolic Problems*, Birkhäuser, Basel, 1995.

[Mal78] P. Malliavin, *Stochastic Calculus of Variations and Hypoelliptic Operators*, Proc. Intern. Symp. SDE (1978), 195–263.

[Mat01] J. C. Mattingly, *Exponential Convergence for the Stochastically Forced Navier-Stokes Equations and Other Partially Dissipative Dynamics*, Preprint, 2001.

[MS95a] A. Mielke and G. Schneider, *Attractors for Modulation Equations on Unbounded Domains – Existence and Comparison*, Nonlinearity **8** (1995), 743–768.

[MS95b] H. M. Möller and H. J. Stetter, *Multivariate Polynomial Equations with Multiple Zeros Solved by Matrix Eigenproblems*, Numerische Mathematik **70** (1995), 311–329.

[MS95c] C. Mueller and R. B. Sowers, *Random Travelling Waves for the KPP Equation with Noise*, J. Funct. Anal. **128** (1995), no. 2, 439–498.

[MS99] B. Maslowski and J. Seidler, *Invariant Measures for Nonlinear SPDE's: Uniqueness and Stability*, Archivum Math. **34** (1999), 153–172.

[MT94] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer, New York, 1994.

[Mue93] C. Mueller, *Coupling and Invariant Measures for the Heat Equation with Noise*, Ann. Probab. **21** (1993), no. 4, 2189–2199.

[MY01] N. Masmoudi and L.-S. Young, *Ergodic Theory of Infinite Dimensional Systems with Applications to Dissipative Parabolic PDEs*, Preprint, 2001.

[Nor86] J. Norris, *Simplified Malliavin Calculus*, Lecture Notes in Mathematics, vol. 1204, Springer, New York, 1986.

[RBT01] L. Rey-Bellet and L. Thomas, *Exponential Convergence to Non-Equilibrium Stationary States in Classical Statistical Mechanics*, Preprint, 2001.

[Rou99] J. Rougemont, *Dynamics of Kinks in the Ginzburg-Landau Equation: Approach to a Metastable Shape and Collapse of Embedded Pairs of Kinks*, Nonlinearity **12** (1999), no. 3, 539–554.

[RS80] M. Reed and B. Simon, *Methods of Modern Mathematical Physics I–IV*, Academic Press, San Diego, California, 1980.

[Str86] D. W. Stroock, *Some Applications of Stochastic Calculus to Partial Differential Equations*, Lecture Notes in Mathematics, vol. 976, Springer, New York, 1986.

[SV72] D. W. Stroock and S. R. S. Varadhan, *On the Support of Diffusion Processes with Applications to the Strong Maximum Principle*, Proc. 6th Berkeley Symp. Math. Stat. Prob. **III** (1972), 333–368.

[Vas69] L. N. Vaseršteĭn, *Markov Processes over Denumerable Products of Spaces Describing Large System of Automata*, Problems of Information Transmission **5** (1969), no. 3, 47–52.

[Wal64] W. Walter, *Differential- und Integralungleichungen*, Springer Tracts in Natural Philosophy, vol. 2, Springer, New York, 1964.