# THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE RENNES 1
COMUE UNIVERSITÉ BRETAGNE LOIRE

ÉCOLE DOCTORALE N° 601
*Mathématiques et Sciences et Technologies*
*de l'Information et de la Communication*
Spécialité : Mathématiques et leurs Interactions

Par

## Pierre PERRUCHAUD

# Homogénéisation pour le mouvement brownien cinétique

et quelques résultats sur son noyau

**Thèse présentée et soutenue à l'université de Rennes 1, le 21 octobre 2019**
**Unité de recherche : IRMAR (UMR 6625)**

**Rapporteurs avant soutenance :**

| | |
|---|---|
| Ana Bela CRUZEIRO | Professor catedrático, Universidade de Lisboa |
| Xue-Mei LI | Professor, Imperial College London |

**Composition du Jury :**

| | | |
|---|---|---|
| Président : | Thierry LÉVY | Professeur, Sorbonne Université |
| Rapporteur : | Ana Bela CRUZEIRO | Professor catedrático, Universidade de Lisboa |
| Examinateurs : | Shizan FANG | Professeur, Université de Bourgogne |
| | Jacques FRANCHI | Professeur, Université de Strasbourg |
| | Armen SHIRIKYAN | Professeur, Université de Cergy-Pontoise |
| Dir. de thèse : | Ismaël BAILLEUL | Maître de conférences, Université de Rennes 1 |
| Co-dir. de thèse : | Jürgen ANGST | Maître de conférences, Université de Rennes 1 |

# Table des matières

# Chapitre I

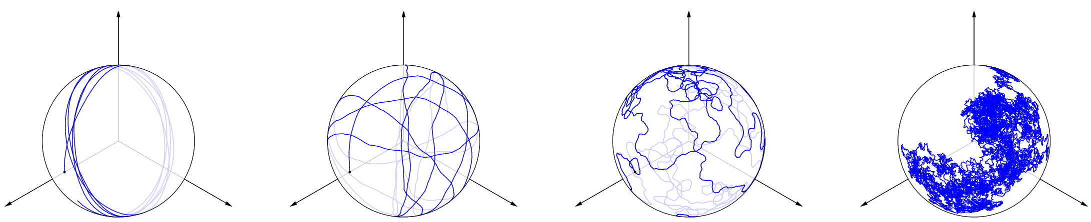
# Introduction

## 1 Le mouvement brownien cinétique

Le mouvement brownien cinétique est une famille de processus aléatoires décrivant une évolution intermédiaire entre la mécanique classique des objets massiques (un tasse, une planète), régie par des équations différentielles ordinaires, et le comportement probabiliste des corpuscules de masse très faible (un grain de pollen, une particule de poussière), modélisé par le mouvement brownien et ses analogues.

Un point matériel $x$ décrivant un mouvement brownien cinétique emprunte à la première de ces catégories les propriétés suivantes. Il a une vitesse $v$ bien définie, et une inertie qui impose à cette dernière de rester continue. On peut alors définir son énergie cinétique $\frac{1}{2}m|v|^2$, qui se trouve être conservée au cours du mouvement. Avec la seconde classe, il a en commun le caractère aléatoire de ses interactions avec son milieu. D'après la seconde loi de Newton, la dérivée de sa vitesse égale la somme des forces qu'exerce l'extérieur sur l'objet. En suivant ce principe, on suppose qu'à chaque instant, l'accélération induite par l'influence de son environnement est indépendante des autres instants, et ne privilégie aucune direction, sous la contrainte cependant de la préservation de l'énergie.

Cette famille de processus est indexée par un paramètre $\sigma$ positif, qui traduit la sensibilité de la particule au chaos qui l'entoure. Dans l'esprit des paragraphes précédents, on peut considérer que plus $\sigma$ est proche de zéro, plus l'objet est massique, et sa course difficile à infléchir. Au contraire, lorsque $\sigma$ est grand, il représente une particule de plus en plus légère, très vulnérable aux chocs et très mobile. Dans le cadre décrit ici, on dispose de nombreux outils mathématiques



$\sigma \in \{0{,}1\,;\,0{,}5\,;\,3\,;\,10\}$

FIGURE 1.1 – Le mouvement brownien cinétique sur la sphère.

qui expliquent comment s'effectue cette transition entre ces régimes très différents. Il existe différentes variantes de ce processus pour lesquelles ces résultats sont plus ou moins clairs ; le projet principal de ce mémoire est l'adaptation de ces méthodes pour montrer cette propriété d'interpolation dans un exemple de dimension infinie.

## 1.1 Premières propriétés

Traduisons mathématiquement les hypothèses physiques décrites plus haut. Pour un certain paramètre de bruit $\sigma \geq 0$, le mouvement brownien cinétique, disons à valeurs dans $\mathbb{R}^d$, est un processus $x^\sigma$ continûment dérivable. Sa dérivée $v^\sigma$ est de norme constante, fixée à 1 par convention. Puisque ses incréments sont indépendants et isotropes sous la contrainte $|v|^2 = 1$, il est naturel de modéliser $v^\sigma$ par un mouvement brownien sphérique, de diffusivité $\sigma^2$. En d'autres termes, pour $W$ un mouvement brownien standard sur la sphère, on définit le processus $(x^\sigma, v^\sigma)$ à valeurs dans $\mathbb{R}^d \times \mathbb{S}^{d-1}$ par

$$x^\sigma_t = x^\sigma_0 + \int_0^t v^\sigma_s \mathrm{d}s, \tag{1.1}$$

$$v^\sigma_t = W_{\sigma^2 t}. \tag{1.2}$$

Pour $\sigma = 0$, le processus $x^\sigma$ décrit le rayon issu de $x^\sigma_0$ et de vecteur directeur $v^\sigma_0$ (possiblement aléatoires) : c'est un cas particulier de mouvement géodésique. À l'inverse, lorsque $\sigma$ tend vers l'infini, la vitesse $v^\sigma$ a en un temps macroscopique le loisir de se déplacer et s'homogénéiser partout sur la sphère. Ainsi, $v^\sigma$ n'a pas de limite en tant que trajectoire, tandis que le comportement de $x^\sigma$ sur les intervalles compacts devient stationnaire en $x^\sigma_0$. On peut cependant observer une limite non triviale en effectuant un changement d'échelle en temps[1] : en posant $X^\sigma : t \mapsto x^\sigma_{\sigma^2 t}$, on constate que la loi de $X^\sigma$ converge vers celle d'un mouvement brownien lorsque $\sigma$ tend vers l'infini.

C'est ce que l'on appelle un résultat d'homogénéisation. Le processus $x^\sigma$ est un système dont la dépendance en $v^\sigma$ est déterministe, et le mouvement de $v^\sigma$ est extrêmement rapide relativement à celui de $x^\sigma$. Dans un système lent-rapide de ce type, on s'attend à ce que lorsque la vitesse du système rapide diverge, le processus lent converge vers la dynamique associée de manière déterministe à la moyenne du processus rapide. Ici, et après renormalisation, la vitesse $\mathrm{d}X^\sigma/\mathrm{d}t$, pour $\sigma$ très grand, est de l'ordre de $U_t/\sqrt{\mathrm{d}t}$ au temps $t$, où $U_t$ est uniformément distribué sur la sphère indépendamment de $U_s$, $s \neq t$. Autrement dit, les incréments de $X^\sigma$ sont essentiellement browniens.

Ce processus a été introduit par J. Angst, I. Bailleul et C. Tardif dans [ABT15], par analogie avec le cas relativiste où $v$ est à valeurs dans la sphère associée à la métrique hyperbolique $\mathrm{d}t^2 - \mathrm{d}x_1^2 - \cdots - \mathrm{d}x_n^2$. Plus précisément, ces auteurs définissent une version du mouvement brownien cinétique dans les variétés riemanniennes, que l'on survole dans la section 1.2. Le résultat élémentaire donné plus haut devient plus subtil dans ce cadre, et sa preuve constitue l'essentiel de la première moitié du travail [ABT15]. Notons que dans le cas des variétés compactes de dimension finie, X.-M. Li a donné la première preuve de l'homogénéisation dans [Li12], en utilisant des techniques très différentes. On trouvera dans la section 2 un traitement plus détaillé de l'historique du mouvement brownien dans la littérature.

Les méthodes introduites par [ABT15] constituent la base de l'article [ABP19] (chapitre III), résultat principal de ce travail de thèse, dans lequel mes directeurs et moi-même prouvons un

[1] De même qu'avec le mouvement brownien, pour lequel une dilatation en temps correspond à une contraction en espace, changer l'échelle de temps pour le mouvement brownien cinétique revient à modifier l'échelle en espace ainsi que le paramètre $\sigma > 0$, et plus généralement jouer sur l'un de ces paramètres est équivalent à adapter les deux autres. On choisit le changement d'échelle en temps pour son analogue plus direct sur les variétés.

résultat d'homogénéisation dans le cas d'une variété de dimension infinie dont les géodésiques représentent des solutions d'équations de la mécanique des fluides. Un ingrédient supplémentaire dans la résolution de ce dernier problème a été le traitement de phénomènes d'anisotropie, dont j'ai mené l'étude en dimension finie dans [Per18] (chapitre II). En marge de ce travail, je me suis intéressé à la description du noyau du mouvement brownien cinétique dans $\mathbb{R}^2$, dont je donne quelques pistes d'étude dans le chapitre IV.

## 1.2 Le cas des variétés

De même que pour les géodésiques et le mouvement brownien, il existe en dimension finie une extension naturelle du mouvement brownien cinétique aux variétés riemanniennes. Soit $(M, g)$ une variété riemannienne connexe, dont on note $d$ la dimension. On suppose pour cette exposition que $M$ est close (compacte sans bord).

**Flot géodésique.** On appelle géodésiques les courbes à valeurs dans $M$ qui, informellement, vont « tout droit. » Dans $\mathbb{R}^d$, ce sont les courbes affines $t \mapsto x_0 + tv_0$, de même que dans le tore plat $\mathbb{T}^d$. Sur la sphère $\mathbb{S}^2$, ce sont les grands cercles, parcourus à vitesse constante.

Mathématiquement, ces courbes $\gamma : [0; T] \to M$ sont par définition les minimiseurs locaux de l'énergie cinétique $\frac{1}{2}\int_0^T g(\dot\gamma_t, \dot\gamma_t)\mathrm{d}t$, à extrémités fixées. D'après le principe d'Euler-Lagrange, cette condition peut se reformuler en une équation différentielle sur le couple $(\gamma, \dot\gamma)$. La position $\gamma$ évolue selon la vitesse $\dot\gamma$, tandis que celle-ci est transportée parallèlement le long de la trajectoire. En particulier, la donnée d'une position et d'une vitesse initiales caractérise entièrement la courbe, et les courbes géodésiques sont précisément les intégrales d'un champ de vecteurs $\mathfrak{X} : TM \to T(TM)$, au sens où

$$\mathrm{d}(\gamma, \dot\gamma)_t = \mathfrak{X}(\gamma_t, \dot\gamma_t)\mathrm{d}t.$$

Dans un langage probabiliste, le champ $\mathfrak{X}$ est le générateur du flot géodésique. Puisque les géodésiques sont parcourues à vitesse constante, le champ $\mathfrak{X}$ est tangent aux hypersurfaces de $TM$ de norme constante, et en particulier la dynamique se restreint au fibré tangent unitaire $T^1M$.

**Mouvement brownien.** Une façon efficace de définir le mouvement brownien sur $M$ est de définir un laplacien $\Delta_M$, et de considérer le processus dont celui-ci est le générateur.[2] Parmi les nombreuses descriptions de cet opérateur, je trouve la suivante particulièrement adaptée à l'intuition probabiliste.

Imaginons qu'une famille de lois $(\mathbb{P}_x)_{x\in M}$ nous soit donnée, de même qu'un processus $X$ à valeurs dans $M$, qui correspondent aux critères attendus pour un hypothétique mouvement brownien issu de $x$. Pour un petit paramètre $\varepsilon > 0$, on peut imaginer que le processus partant de $x$ atteint la sphère $S(x, \varepsilon)$ en un point presque uniformément distribué par rapport au volume riemannien. Notons $T_\varepsilon$ le temps d'atteinte de la sphère. Alors pour toute fonction lisse $f : M \to \mathbb{R}$, on s'attend à ce que

$$\mathbb{E}_x[f(X_{T_\varepsilon}) - f(X_0)] = \mathbb{E}_x\left[\int_0^{T_\varepsilon} \frac{1}{2}\Delta_M f(X_t)\mathrm{d}t\right].$$

Or, le terme de gauche mesure la différence entre $f$ en $x$ et sa moyenne sur une sphère de petit rayon, tandis que celui de droite est proche du produit de $\Delta_M f(x)$ par l'espérance de $T_\varepsilon$. Dans

[2] On emploie ici la convention qui fait de $-\Delta_M$ un opérateur positif.

le cas euclidien, cette dernière vaut précisément $\varepsilon^2/2d$.[3] Ainsi, il paraît naturel de définir

$$\Delta_M f(x) := \lim_{\varepsilon\to 0} \frac{2d}{\varepsilon^2 Z_\varepsilon} \int_{S(x,\varepsilon)} \big(f(y)-f(x)\big)\,\mathrm{vol}_g(\mathrm{d}y),$$

où $Z_\varepsilon$ est le volume riemannien de la sphère $S(x,\varepsilon)$. Un petit bagage de géométrie riemannienne suffit à montrer que cette formule définit bien un opérateur différentiel à coefficients lisses sur $M$, ce qui complète la construction de $\Delta_M$, que l'on appelle opérateur de Laplace-Beltrami, et du processus associé à $\frac{1}{2}\Delta_M$, le mouvement brownien. On en esquisse une deuxième construction en termes d'équations différentielles stochastiques, plus géométrique, dans la section 3.2.

Un exemple important pour la construction du mouvement brownien cinétique est celui de la sphère unité $\mathbb{S}^{d-1}$, pour $d \geq 2$. On montre que pour une fonction lisse $f : \mathbb{R}^d \to \mathbb{R}$, le laplacien appliqué à $f_{|\mathbb{S}}$ s'écrit

$$\Delta_{\mathbb{S}} f_{|\mathbb{S}}(x) = \sum_i \big(1-(x^i)^2\big)\,\partial_i^2 f(x) - (d-1)\sum_i x^i \partial_i f(x). \tag{1.3}$$

On en déduit que le mouvement brownien $X$ sur la sphère vérifie l'équation différentielle stochastique

$$\mathrm{d}X_t = \sum_i \big(e_i - X_t^i X_t\big)\mathrm{d}W_t^i - \frac{d-1}{2} X_t \mathrm{d}t,$$

où $W$ est un mouvement brownien standard à valeurs dans $\mathbb{R}^d$. On note que $e_i - x^i x$ est la projection orthogonale de $e_i$ sur l'espace orthogonal à $x$ ; ainsi, en posant $P_{x^\perp}$ la projection en question, on peut aussi écrire

$$\mathrm{d}X_t = P_{X_t^\perp}\mathrm{d}W_t - \frac{d-1}{2} X_t \mathrm{d}t = P_{X_t^\perp} \circ \mathrm{d}W_t.$$

Ceci complète la définition du mouvement brownien cinétique euclidien décrit dans la première section.

**Mouvement brownien cinétique.** Avec le même niveau de rigueur que l'on a qualifié le flot géodésique de marche en avant, le mouvement brownien cinétique peut être décrit comme une promenade d'homme ivre. Celui-ci marche résolument à vitesse constante, mais tourne constamment et erratiquement de gauche et de droite.

En termes mathématiques, le mouvement brownien cinétique est une diffusion à valeurs dans $T^1M$, dépendant d'un paramètre $\sigma \geq 0$ (l'ébriété du personnage). Son générateur $L^\sigma$ est composé d'un terme de transport simple, le générateur $\mathfrak{X}$ du flot géodésique, et d'un terme brownien en vitesse, le laplacien $\Delta_v$ le long des fibres de $T^1M$. Ce dernier agit sur les fonctions lisses $f : T^1M \to \mathbb{R}$ par

$$\Delta_v f(x) = \Delta_{\pi^{-1}(x)} f_{|\pi^{-1}(x)}(x),$$

où $\pi : T^1M \to M$ est la projection canonique, c'est-à-dire que $\pi^{-1}(x)$ est la sphère unité de $T_xM$. Précisément, on définit

$$L^\sigma := \mathfrak{X} + \frac{\sigma^2}{2}\Delta_v.$$

Il est clair dans cette description que le mouvement brownien cinétique se réduit à une pure équation géodésique pour $\sigma = 0$. De même que pour la version euclidienne, on s'attend à ce

[3] Une preuve de ce fait consiste à prendre $f : x \mapsto |x^1|^2 + \cdots + |x^d|^2$ dans la formule ci-dessus.

que pour de grandes valeurs de $\sigma$, la vitesse s'homogénéise sur la sphère et que la position soit constante à la limite. Mieux encore, une bonne renormalisation du mouvement devrait faire apparaître un mouvement brownien en position, dans la limite $\sigma \to \infty$. Le théorème suivant, démontré par J. Angst, I. Bailleul et C. Tardif dans [ABT15], prouve précisément ce résultat.

**Théorème 1.1.** *Soit $(x^\sigma, v^\sigma)$ le mouvement brownien cinétique à valeurs dans $T^1M$ issu de $(x_0, v_0)$. Le processus renormalisé $t \mapsto x^\sigma_{\sigma^2 t}$ converge en loi vers un mouvement brownien de générateur*

$$\frac{2}{d(d-1)}\Delta_M.$$

Ce théorème est la principale inspiration de mon travail de thèse. Dans la partie 4, on décrit un mouvement géodésique qui apparaît naturellement sur une variété $M$ de dimension infinie : les équations d'Euler de la mécanique des fluides. Puisque le mouvement brownien cinétique est une perturbation naturelle et géométrique du flot géodésique, je me suis intéressé au traitement du mouvement brownien cinétique dans un cadre de dimension infinie. Cela m'a amené à considérer des phénomènes d'anisotropie, qui apparaîssent déjà en dimension finie et ont leur intérêt propre. Une exposition informelle de ces deux problèmes est donnée dans la partie 3, et le lecteur anglophone en trouvera un traitement détaillé dans les chapitres II et III.

## 1.3 Caractéristiques du noyau associé

On considère dans cette section le mouvement brownien cinétique sur $\mathbb{R}^d$ pour le paramètre $\sigma = 1$. Après avoir fixé la condition initiale $(x_0, v_0)$, on s'intéresse à la loi de $(x_t, v_t)$ pour un temps $t > 0$ petit. D'après la propriété d'interpolation entre le flot géodésique et le mouvement brownien, on s'attend à ce que celle-ci possède des propriétés partagées entre ces deux extrêmes. Or, les deux comportements en question sont radicalement différents : dans le premier cas, une condition de Dirac en $(x_0, v_0)$ est transportée en un Dirac en $(x_0 + tv_0, v_0)$, tandis que dans le second, la variable $x_t$ devient immédiatement à densité, elle-même analytique : en s'autorisant un acte de foi, on peut aller jusqu'à dire qu'en passant à la limite $\sigma \to \infty$ avec une renormalisation convenable, le couple $(x_t, v_t)$ est distribué selon la loi

$$\mathcal{N}_d(0, tI) \otimes \mathrm{Unif}(\mathbb{S}^{d-1}).$$

On peut donc imaginer que pour le cas $\sigma = 1$, le comportement de $(x_t, v_t)$ soit subtil à étudier.

Notons déjà deux propriétés de cette loi, empruntées à l'une et l'autre de ces limites. De même que le mouvement brownien, et comme le lecteur familier avec les équations aux dérivées partielles l'aura remarqué, le couple $(x_t, v_t)$ admet une densité lisse pour tout $t > 0$. En effet, en notant $u_t$ cette densité et $\Delta_v f$ le laplacien sur la sphère appliqué à la seconde composante de $f : \mathbb{R}^d \times \mathbb{S}^{d-1} \to \mathbb{R}$ et $\partial_x f$ son gradient selon la première coordonnée, $u_t$ est solution de

$$\partial_t u_t = -v \cdot \partial_x u_t + \frac{1}{2}\Delta_v u_t,$$

où $(x, v)$ est l'élément générique de $\mathbb{R}^d \times \mathbb{S}^{d-1}$. On verra par la suite que cette équation est dite hypoelliptique, au sens introduit par L. Hörmander, et que cette condition est suffisante pour assurer l'existence d'une densité lisse pour $(x_t, v_t)$.

Cependant, de même que le flot géodésique, la densité $u_t$ se déplace à vitesse bornée. Au vu de l'équation différentielle stochastique définissant le mouvement brownien cinétique, il est clair que $|x_t - x_0| < t$ presque sûrement. En utilisant un résultat de D. Stroock et S. R. S. Varadhan, on montrera plus tard que le support de $u_t$ est en fait précisément $\{(x, v), |x - x_0| \leq t\}$. Ainsi, la densité possède un défaut d'analyticité, au moins le long du produit $t\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$. De plus, puisque $v_t$ s'éloigne de $v_0$ en $\sqrt{t}$, la position de $x_t$ est très proche de $x_0 + tv_0$, *a priori* à $t\sqrt{t}$ près. Or, la densité est précisément nulle en ce point. Il faudra donc conduire une analyse fine pour comprendre les principaux aspects de $u_t$.

Dans la partie 5, on présente la preuve des deux énoncés élémentaires donnés plus haut, concernant le support et la régularité du noyau $u_t$. Cela nous donne l'occasion de discuter quelques propriétés connues et attendues des diffusions elliptiques et sous-elliptiques. Le cas hypoelliptique est beaucoup moins connu, et est le sujet d'une part de mon travail de thèse. Les grandes lignes de ce projet sont présentées dans cette section 5 ; elles sont développées dans la partie IV.

# 2 Repères bibliographiques

## 2.1 Homogénéisation

Le mouvement brownien relativiste, analogue du mouvement brownien cinétique dans les variétés lorentzienne, a été introduit par J. Franchi et Y. Le Jan dans l'article [FLJ07], en se basant sur des idées de Dudley que l'on pourra trouver dans [Dud66]. Cette diffusion a notamment été étudiée explicitement par J. Angst, I. Bailleul et C. Tardif. Ces derniers ont introduit le mouvement brownien cinétique dans [ABT15], article dans lequel ils prouvent le résultat d'homogénéisation dans le cas isotrope de dimension finie, et décrivent la frontière de Poisson dans les variétés symétriques. L'approche choisie est celle des chemins rugueux. Une approche différente a été introduite plus tôt et indépendamment par X.-M. Li dans [Li12, Theorem 4.3]. Elle se base sur des outils d'analyse fonctionnelle et de calcul stochastique, et impose des conditions sur la géométrie de la variété sous-jacente, en l'occurence la compacité. Ces résultats sont prolongés dans [Li16b] au variétés complètes dont le rayon d'injectivité est borné inférieurement.

Ce type d'interpolation entre le flot géodésique et le mouvement brownien rappelle le travail de J.-M. Bismut sur le laplacien hypoelliptique, qui sert d'inspiration plutôt que de base de travail. On pourra par exemple consulter [Bis15]. Les motivations du laplacien hypoelliptique sont d'ordre géométrique : en effectuant un lien entre flot géodésique et mouvement brownien, et en adaptant ce processus aux formes différentielles, J.-M. Bismut éclaire quelques propriétés du théorème d'indice pour le laplacien.

L'extension aux mouvements anisotropes, via des méthodes ergodiques, s'inspire de l'article [BFH09] de E. Breuillard, P. Friz et M. Huesmann. Mon travail est à mon avis très différent du leur, mais ils se basent tous deux sur la structure de groupe de Lie qui apparaît dans la définition des chemins rugueux.

Pour le cas de dimension infinie, l'article fondateur est dû à V. Arnol'd. Dans [Arn66], il exprime les équations d'Euler de mécanique des fluides incompressibles comme une équation des géodésiques dans un espace de difféomorphismes. Cette discussion informelle est rendue rigoureuse par D. Ebin et J. Marsden dans [EM69], qui en donnent aussi une version visqueuse via les équations de Navier-Stokes. De manière plus anecdotique, cette formulation lagrangienne a été redécouverte récemment pour la modélisation numérique, voir par exemple l'article [PMT+11].

Leurs méthodes viennent de la formulation de l'analyse globale non linéaire en termes de variétés de dimension infinie, qui se développe dans les années 60 en lien avec la théorie de l'indice. Je me suis basé dans ma thèse sur le texte [Pal68] de R. Palais, à la lecture duquel il est clair que les idées originales y côtoient les résultats bien connus de l'époque qui n'avaient pas encore été compilés dans un manuscrit cohérent. Pour des références alternatives, ce mémoire paraît en parallèle de celui de J. Eells, [Eel66], dont les motivations sont similaires quoique plus appliquées et moins fonctorielles, et un traitement plus moderne peut être trouvé dans le livre [Gli11] de Y. Gliklikh, qui traite aussi d'analyse stochastique.

Récemment, une classe de perturbations aléatoires des équations d'Euler a été introduites par D. Holm dans [Hol15]. Dans la formulation eulérienne de ces équations, cette approche consiste à ajouter un bruit additif ; dans la description lagrangienne, les champs de vecteurs le long desquels le bruit se diffuse sont invariants à droite. D. Holm et ses coauteurs ont déjà produit une riche littérature sur le sujet, voir par exemple [CFH18] ou [CHR18].

Une autre école stochastique de la mécanique des fluides a été fondée par M. Arnaudon et A. B. Cruzeiro, inspirés par K. Yasue. Contrairement à l'approche développée dans ce manuscrit, il s'agit d'une théorie probabiliste d'équations déterministes : l'objectif est d'exprimer la solution (déterministe) des équations de Navier-Stokes comme la solution d'un problème variationnel probabiliste. On pourra par exemple consulter les articles [AC12] et [CFM07].

### 2.2 Asymptotique en temps petit

L'article fondateur pour les noyaux de la chaleurs dans le cas sous-elliptique, moins singulier que celui qui nous occupe, est sans doute [BA88], par G. Ben Arous. On y trouve un développement en série du noyau de la chaleur hors du cut locus pour une variété sous-riemannienne. Ce travail est complété par le traitement dans [BA89] de l'asymptotique sur la diagonale. L'approche est probabiliste : elle se base sur le principe de grandes déviations de D. Stroock et S.R.S. Varadhan, le calcul stochastique de P. Malliavin, et des idées géométriques de J.-M. Bismut.

Pour des travaux plus récent, on peut citer par exemple les travaux de D. Barilari, U. Boscain et R. Neel [BBN12], pour le cas du cut locus, et [BBCN16], avec G. Charlot, pour le cas de variétés génériques de petite dimension.

Au contraire, les deux approches que l'on se propose de mettre en œuvre se basent sur les travaux de V. Kolokoltsov, tels qu'on les trouve dans son livre [Kol00]. Les résultats qui y sont décrits s'appliquent à des cas algébriquement contraints ; cependant, la stratégie se veut plus générale. Elle se base sur une reformulation hamiltonienne, inspirée par l'analyse semiclassique, puis sur la méthode de la paramétrix.

Pour ce qui est du problème hamiltonien, les techniques employées sont élémentaires, et ne dépassent pas les textes de référence. Je me suis basé pour ma part sur le chapitre 2 du livre [Kol00] déjà évoqué. La partie concernant la méthode de la paramétrix est inspirée ce que R. Melrose appelle le *heat calculus,* décrit en détail dans le chapitre 7 de son livre [Mel93], que l'on pourrait dans notre cas appeler *l'hypoheat calculus.* Une version abrégée et très accessible est consultable dans les notes [Gri04] de D. Grieser. L'ambition de ces deux textes est d'obtenir des formules asymptotiques pour la trace du noyau de la chaleur, afin de parvenir à des estimations de valeurs propres ou des résultats de théorie de l'indice.

Dans le travail [Fra19], J. Franchi décrit le noyau associé à une diffusion dont la singularité est comparable à celle qui nous occupe. En termes techniques, il s'agit d'une sous-diffusion d'un certain processus introduit dans la partie IV.1, en quelque sorte une approximation du mouvement brownien cinétique à l'ordre 2. Les méthodes sont analytiques, basées sur une transformée de Fourier et une méthode du point col. Le travail que nous avons mené avec V. Kolokoltsov est de nature très différente en termes d'outils, et en un sens complémentaire : la construction proposée dans ce manuscrit exprime le noyau exact comme un développement en série, et la diffusion approchée en question constitue à la fois son premier terme et la brique de base pour décrire les suivants.

# 3 Homogénéisation

## 3.1 Le cas isotrope de dimension finie

Avant de discuter les résultats d'homogénéisation obtenus au cours de cette thèse, il est édifiant de se pencher sur la preuve du cas initialement étudié par J. Angst, I. Bailleul et C. Tardif dans [ABT15]. Soit $(M, g)$ une variété riemannienne connexe, complète, stochastiquement complète, et de dimension $d \geq 2$. Pour tout $\sigma \geq 0$, on définit sur le fibré tangent unitaire $\pi : T^1M \to M$ le mouvement brownien cinétique $(x^\sigma, v^\sigma)$ comme le processus de Feller de générateur

$$L^\sigma := \mathfrak{X} + \frac{1}{2}\widehat{\Delta}_v.$$

On rappelle que $\mathfrak{X}$ est le générateur du flot géodésique sur $T^1M$, et $\Delta_v$ est le laplacien le long des fibres, tels qu'ils sont décrits dans la section 1.2.

On rappelle le théorème 1.1 : la partie position renormalisée $t \mapsto x^\sigma_{\sigma^2 t}$ du mouvement brownien cinétique $(x^\sigma, v^\sigma)$ issu de $(x_0, v_0)$, converge en loi dans $\mathcal{C}^0([0; 1], M)$ vers un mouvement brownien de générateur

$$\frac{2}{d(d-1)}\Delta_M.$$

Comme on l'a vu en introduction, le résultat est clair dans la variété modèle $(\mathbb{R}^d, g_{\text{eucl}})$. L'idée de la preuve de [ABT15] est de réduire le cas général au cas euclidien. Il existe une procédure, que l'on appelle le développement de Cartan, qui exprime le mouvement brownien cinétique sur $M$ en fonction de son analogue euclidien au sens suivant. On peut définir un certain fibré $\pi_{OM} : OM \to M$ sur $M$, le fibré orthonormal, et une équation différentielle (stochastique)

$$\mathrm{d}z^\sigma_t = f(z^\sigma_t) \circ \mathrm{d}x^\sigma_t$$

sur $OM$, dirigée par la partie position $x^\sigma$ du mouvement brownien cinétique euclidien $(x^\sigma, v^\sigma)$ et de solution $z^\sigma$, telle que la projection $q^\sigma := \pi_{OM}(z^\sigma)$ décrit un mouvement brownien cinétique sur $M$. La figure 3.2 en donne une représentation schématique.

$$\begin{tikzcd}
 & z^\sigma_t := \int_0^t f(z^\sigma_s) \circ \mathrm{d}x^\sigma_s \in OM \arrow[dr, mapsto, "\pi_{OM}"] & \\
x^\sigma_t \in \mathbb{R}^d \arrow[ur, mapsto, "I"] \arrow[rr, mapsto, "\text{Développement de Cartan}"'] & & q^\sigma_t := \pi_{OM}(z^\sigma_t) \in M
\end{tikzcd}$$

*Les applications ne dépendent pas seulement de la valeur à un temps $t > 0$ fixé ni de la trajectoire, mais aussi de la structure de semimartingale.*

FIGURE 3.2 – Construction du développement de Cartan.

On appelle $I(x) = I_f(x)$, $I$ pour application d'Itô, la solution de l'équation différentielle de Stratonovich

$$\mathrm{d}I(x)_t = f(I(x)_t) \circ \mathrm{d}x_t.$$

L'emploi de l'intégrale de Stratonovich peut paraître superflu ici. Son intérêt vient du fait qu'en ce sens, le développement de Cartan d'un mouvement brownien est un mouvement brownien ; en fait, cette procédure, connue sous le nom de construction d'Eells-Elworthy-Malliavin, est devenue la méthode standard pour écrire intrinsèquement en termes d'équation différentielle stochastique le mouvement brownien sur une variété riemannienne. Une stratégie émerge alors clairement de cette formulation : en deux étapes, on montre d'abord que le mouvement brownien cinétique euclidien converge vers un mouvement brownien, ce qui est, comme on l'a vu, aisé, puis que le développement de Cartan est continu, et garantit ainsi la convergence de $q^\sigma$ par celle de $x^\sigma$.

Malheureusement, il est classique que l'application d'Itô n'est que mesurable, disons de $\mathcal{C}^\gamma([0;1],\mathbb{R}^d)$ dans $\mathcal{C}^0([0;1],OM)$ pour $0 \leq \gamma < 1/2$. En fait, le développement de Cartan lui-même n'est pas continu en ce sens, on en verra un exemple précis dans la partie 3.2 qui suit. Au contraire, il gagnerait à être connu qu'il existe un cadre robuste dans lequel l'équivalent de $I$ est effectivement continu. La théorie des chemins rugueux, développée par T. Lyons, traite isolément les parties probabiliste et géométrique du problème de la façon suivante. À une semimartingale $X$, on associe un objet $\mathbf{X}$ par des méthodes probabilistes, que l'on appelle un chemin rugueux. On peut ensuite définir une application continue $\mathbf{I}$ de l'ensemble $\mathrm{RP}([0;1],\mathbb{R}^d)$ des chemins rugueux[4] sur $\mathbb{R}^d$ à valeurs dans l'ensemble des courbes $\mathcal{C}^0([0;1],OM)$, que l'on appelle l'application d'Itô-Lyons. L'avantage de cette approche est que le chemin rugueux ne dépend pas du problème $f$ considéré, et que cet objet est en un certain sens de nature euclidienne, et ainsi facilement manipulable.

Le plan de preuve est donc le suivant : on montre que le chemin rugueux $\mathbf{X}^\sigma$ associé au processus renormalisé $X^\sigma : t \mapsto x^\sigma_{\sigma^2 t}$ converge en loi vers le chemin rugueux $\mathbf{X}$ associé à un mouvement brownien $X$ ; puis on en déduit que le mouvement brownien cinétique sur $M$, qui s'écrit $\pi_{OM} \circ \mathbf{I}(\mathbf{X}^\sigma)$, converge vers le mouvement brownien sur $M$, décrit par $\pi_{OM} \circ \mathbf{I}(\mathbf{X})$, par continuité de $\pi \circ \mathbf{I}$.

## 3.2 Développement de Cartan et chemins rugueux

Cette approche exploite deux outils majeurs avec lesquels on aura besoin d'une certaine familiarité : le développement de Cartan et les chemins rugueux. On se propose dans les paragraphes qui suivent d'en esquisser les grands principes dans le cadre de la dimension finie.

**Développement de Cartan.** Soit $M$ une surface compacte plongée dans $\mathbb{R}^3$. On cherche à décrire un procédé qui transforme le mouvement brownien cinétique euclidien, disons dans l'espace $T_{x_0}M$ pour un certain $x_0$ fixé, en un mouvement brownien cinétique sur $M$. Si sa description mathématique ne saute pas aux yeux du lecteur novice, la question au contraire ne se pose pas pour un habitant microscopique de $M$ ; appelons-le Mr Hyde. Soit un certain Dr Jeckyll, résidant au contraire dans l'espace euclidien $\mathbb{R}^2$. En supposant qu'ils décrivent tous deux un mouvement brownien cinétique tel que décrit dans la section 1, une manière de faire correspondre leurs déambulations est de synchroniser leurs changements de vitesse : tous deux marchent droit devant eux à vitesse constante, et tournent, bien qu'aléatoirement, l'un en même temps que l'autre et de la même façon. De même que je ne prête pas attention à la courbure de la terre lorsque je prends mon vélo pour mes trajets quotidiens, cette correspondance est parfaitement claire pour nos protagonistes dès qu'ils sont de taille suffisamment petite. Plus généralement, on définit donc informellement le développement de Cartan d'une courbe lisse de la façon suivante.

**Définition 3.1.** Le développement de Cartan dans $M$ d'une courbe $\gamma$ tracée dans $\mathbb{R}^2$ est celle que décrit Mr Hyde lorsque le docteur Jekyll marche le long de $\gamma$. △

[4]En fait, cet espace dépend d'un exposant de régularité $0 < \gamma < 1$, que l'on doit la plupart du temps choisir strictement entre $1/3$ et $1/2$. On note donc $\mathrm{RP}^\gamma([0;1],\mathbb{R}^d)$ cet espace.

En particulier, le développement de Cartan d'une ligne droite est une géodésique.

Tournons notre attention vers la figure 3.3, qui illustre ce procédé. À gauche, on trouve la trajectoire décrite par le Dr Jeckyll dans le plan euclidien. À droite, on a développé ce trajet sur la sphère, qui représente donc le déplacement de Mr Hyde. Au centre, l'identification des plans tangents représente la correspondance entre les incréments infinitésimaux euclidiens et sphériques. Il faut imaginer que cette correspondance est revue en permanence, en planquant en tout temps le plan euclidien sur la sphère, sans glisser autour du point de contact.

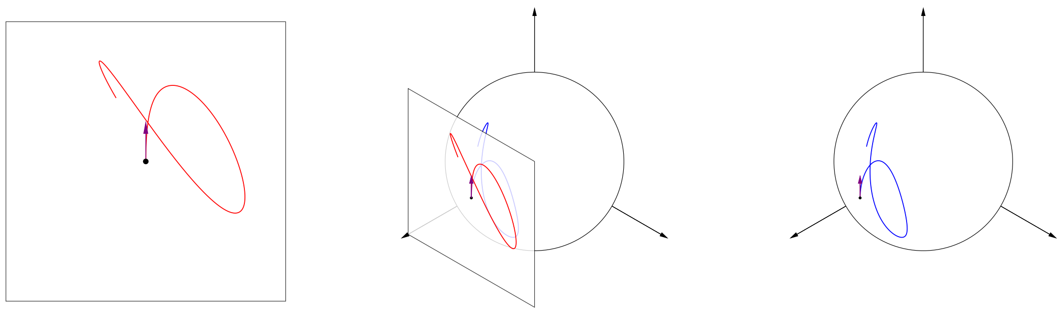


FIGURE 3.3 – Développement de Cartan sur la sphère.

Notons que Mr Hyde peut revenir sur ses pas avec une orientation différente, auquel cas un même déplacement de Dr Jeckyll n'aura pas le même effet que lors de son premier passage. En particulier, cela impose que l'on ne peut pas décrire la position $h$ du premier en fonction de $j$, celle du second, par une simple équation différentielle contrôlée de la forme

$$\mathrm{d}h_t = f(h_t)\mathrm{d}j_t.$$

Il faut donc connaître l'orientation de Mr Hyde en tout temps pour pouvoir décrire l'évolution de sa trajectoire : c'est ici qu'intervient le fibré $OM$ introduit plus haut. En pratique, pour ne pas dépendre de l'orientation du docteur Jekyll, qui dans le cas du mouvement brownien est mal définie, on décrit plutôt la différence des orientations entre ces deux personnages : pour toute direction dans $\mathbb{R}^2$, on observe la façon dont elle est perçue par Dr Jekyll (devant lui, à sa gauche...), et on lui associe la direction $v \in T_hM$ que Mr Hyde perçoit de la même façon. L'orientation $u$ du couple $(h, u) \in OM$ est donc une isométrie de $\mathbb{R}^2$ vers $T_hM$. Malheureusement, il est malaisé de donner ici plus de détails sur cette construction. On se contente de reprendre de façon semi-formelle la discussion ci-dessus, étendue en dimension $\mathbb{R}^d$. Le lecteur intéressé pourra consulter l'ouvrage [Hsu02] de E. P. Hsu.

- Étant donnée une courbe lisse $\gamma$ dans $\mathbb{R}^d$ tracée par le docteur Jekyll et son vaisseau multi-dimensionnel, son développement de Cartan est le chemin suivi par Mr Hyde, lorsque celui-ci maintient la même vitesse et les mêmes indicateurs de virage multi-rotationnels.

- Cette procédure est décrite par une équation différentielle contrôlée de la forme

$$\mathrm{d}(h, u)_t = f(h_t, u_t)\mathrm{d}j_t,$$

où $u$ décrit l'orientation de Mr Hyde en fonction de celle de Dr Jeckyll, et l'ensemble des positions orientées $(j, u)$ forme un fibré $OM$ au dessus de l'ensemble $M$ des positions de Mr Hyde.

- Si le Dr Jekyll suit un mouvement brownien standard, et que l'on interprète l'équation différentielle contrôlée ci-dessus au sens de Stratonovich, alors Mr Hyde décrit lui aussi un mouvement brownien standard sur $M$.

**Défaut de continuité.** Le défaut de continuité du développement de Cartan s'observe facilement. On discute à présente l'exemple d'une certaine courbe sur la sphère, illustrée sur la figure 3.4. On a représenté sur le plan tangent une courbe rouge rectiligne, et une courbe bleue à boucles, qui la suit grossièrement. Elles ont respectivement pour équation, dans un système de coordonnées bien choisi,

$$t \mapsto (0, t) \qquad \text{et} \qquad t \mapsto \big(1 - \varepsilon \cos(2t/\varepsilon^2), t + \varepsilon \sin(2t/\varepsilon^2)\big).$$

Sur la sphère, on a tracé leurs développements de Cartan.

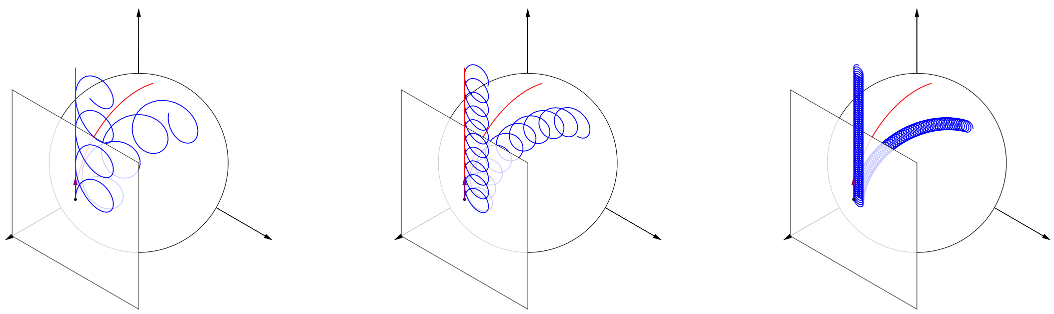


FIGURE 3.4 – Phénomène de dérive dans le développement de Cartan.

Lorsque $\varepsilon$ tend vers 0, la courbe bleue converge uniformément[5] vers la courbe rouge ; cependant, on observe que son développement converge vers une autre courbe, qui a tendance à se déporter sur la droite. Il existe plusieurs manières de comprendre ce comportement, une première étant de noter que les cercles de courbure constante dans le plan euclidien sont plus longs que les cercles de même courbure sur la sphère, ce qui permet au point sur la sphère de tourner pendant un peu plus d'un cercle quand le point sur le plan parcourt exactement une circonférence, et ainsi d'accumuler, boucle par boucle, une petite dérive qui reste présente à la limite. Une vision plus quantitative se base sur le théorème de Gauss-Bonnet. Une conséquence de ce théorème est la suivante. Soit $\Omega$ un domaine de la sphère unité $\mathbb{S}^2$ bordé par une courbe lisse $\partial\Omega$. On suppose que $\partial\Omega$ est le développement de Cartan d'une courbe euclidienne.[6] Alors Mr Hyde, au bout d'un tour, est revenu sur ses pas avec même orientation. Au contraire, Dr. Jekyll, dans le plan euclidien, n'a pas de raison d'être revenu à son point de départ ; cependant, il est possible dans son monde de comparer son orientation initiale à son orientation finale, et d'en déduire qu'il a tourné d'un certain angle $\theta$.

**Théorème 3.2.** *Lorsque Mr Hyde décrit une courbe fermée simple lisse $\partial\Omega$ sur la sphère unité bordant un domaine $\Omega$ d'aire $A$, Dr Jekyll tourne d'un angle $\theta$ dans le plan euclidien. Ces quantités satisfont alors la relation*

$$2\pi + \theta = A.$$

[5] On montre en fait facilement qu'elle converge vers la courbe rouge en norme hölderienne d'exposant $1/2 + \eta$ si et seulement si $\eta < 0$.

[6] C'est en fait toujours le cas : l'antidéveloppement de Cartan, qui produit le résultat que l'on attend, se définit essentiellement de la même façon.

Il y a bien sûr des subtilités d'orientation à prendre en compte. On dit que $\theta$ augmente lorsque le docteur Jekyll tourne vers la gauche, et que le domaine $\Omega$ se trouve sur la droite de Mr Hyde. On suppose bien sûr qu'initialement, ils ont tous les deux les pieds au point de contact entre la sphère et le plan, et la tête à l'extérieur de la sphère. Ayant ce théorème en tête, on peut expliquer intuitivement que, puisque chaque boucle décrite par la courbe bleue est d'aire $\varepsilon^2$ et que l'on en trouve $1/\varepsilon^2$ par unité de temps, le Dr Jekyll entoure une aire qui devient rapidement invisible à l'œil nu, mais qui croît linéairement en temps, et qui se traduit par un dérive constante de Mr Hyde vers la droite.[7]

**Chemins rugueux.** L'idée de la théorie des chemins rugueux, introduite par T. Lyons, est de considérer cette aire cachée comme une donnée du problème. Un chemin rugueux à valeurs dans $\mathbb{R}^d$, dont on note $\epsilon_1, \cdots, \epsilon_d$ la base canonique, est un couple $\mathbf{X} = (X, \mathbb{X})$ vérifiant certaines propriétés analytiques et algébriques, où $X$ décrit le comportement macroscopique du chemin, et $\mathbb{X}$ ses propriétés microscopiques. En pratique, $X : [0;1] \to \mathbb{R}^d$ est une courbe classique, et pour tout $0 \leq s \leq t \leq 1$, la quantité $\mathbb{X}^{ij}_{ts} - \mathbb{X}^{ji}_{ts}$ est un réel décrivant l'aire entourée par la courbe hypothétique $\mathbf{X}$ dans le plan orienté engendré par $(\epsilon_i, \epsilon_j)$. La partie symétrique de $\mathbb{X}$, dans les cas que l'on considère, se déduit de $X$ et est donc de peu d'intérêt pour nous.

L'exemple fondamental est celui des courbes lisses (disons $\mathcal{C}^1$, même si l'on pourrait descendre à $\mathcal{C}^{1/2+\eta}$ pour tout $\eta > 0$). Pour toute courbe $X : [0;1] \to \mathbb{R}^d$, il existe un chemin rugueux canonique $\mathbf{X} = (X, \mathbb{X})$ défini par

$$\mathbb{X}^{ij}_{ts} := \int_s^t (X^i_u - X^i_s) \mathrm{d}X^j_u,$$

au sens de Stieltjes. Dans notre utilisation des chemins rugueux, en toute généralité, la quantité de droite n'est pas bien définie ; typiquement, $X$ est de classe $\mathcal{C}^\gamma$ pour $1/3 < \gamma < 1/2$. Au contraire, la méthode des chemins rugueux utilise le membre de gauche, donné *a priori,* pour jouer le rôle de celui de droite. Une définition indépendante des coordonnées est donnée par

$$\mathbb{X}_{ts} := \int_s^t (X_u - X_s) \otimes \mathrm{d}X_u \in \mathbb{R}^d \otimes \mathbb{R}^d. \tag{3.4}$$

Le lecteur non familier avec les produits tensoriels peut considérer $a \otimes b$ comme n'étant qu'une notation pour la collection de produits $(a^i b^j)_{ij}$, et se convaincre que le produit tensoriel a toutes les propriétés attendues d'un produit — ce qui n'est en un sens pas loin d'être une définition.

En se basant sur cet exemple, on donne les deux propriétés attendues d'un chemin rugueux. Un chemin rugueux $\mathbf{X} = (X, \mathbb{X}) \in \mathrm{RP}^\gamma([0;1], \mathbb{R}^d)$ de régularité $1/3 < \gamma \leq 1/2$ se compose d'un chemin $X : [0;1] \to \mathbb{R}^d$ de régularité $\mathcal{C}^\gamma$, et d'une collection d'accroissements $\mathbb{X} : \{0 \leq s < t \leq 1\} \to \mathbb{R}^d \otimes \mathbb{R}^d$ de régularité hölderienne $2\gamma$, au sens où

$$|\mathbb{X}_{ts}| \leq C|t-s|^{2\gamma}$$

uniformément en $s < t$. On impose de plus la condition algébrique suivante pour tout $s < u < t$, appelée relation de Chen.

$$\mathbb{X}_{ts} = \mathbb{X}_{tu} + X_{us} \otimes X_{tu} + \mathbb{X}_{us}$$

Cette relation rend compte du fait que, dans le cas où $X$ est une courbe lisse,

$$\int_s^t (X_a - X_s) \otimes \mathrm{d}X_a = \int_u^t (X_a - X_u) \otimes \mathrm{d}X_a + \int_u^t (X_u - X_s) \otimes \mathrm{d}X_a + \int_s^u (X_a - X_s) \otimes \mathrm{d}X_a.$$

[7] Il est difficile de rendre cette heuristique plus précise puisqu'il est difficile de relier l'aire $A'$ entourée par le Dr Jekyll à l'angle $\theta'$ décrit par Mr Hyde, cette dernière quantité étant elle-même mal définie.

L'ensemble ainsi défini n'est pas un espace vectoriel, mais c'est un espace métrique complet muni d'une distance appropriée, se comportant comme une norme $\mathcal{C}^\gamma$ sur $X$ et $\mathcal{C}^{2\gamma}$ sur $\mathbb{X}$.

Ce cadre des chemins rugueux permet de restaurer la continuité du développement de Cartan au sens suivant. Soit $M$ une variété riemannienne close de dimension $d$, et $\mathrm{RP}^\infty([0;1],\mathbb{R}^d)$ l'ensemble des chemins rugueux lisses, c'est-à-dire l'ensemble des couples $(X,\mathbb{X})$ où $X$ est une courbe lisse et $\mathbb{X}$ est défini par la relation (3.4). Alors l'application

$$\mathrm{RP}^\infty([0;1],\mathbb{R}^d) \xrightarrow{\text{Développement de Cartan}} \mathcal{C}^0([0;1],M)$$

est continue par rapport au topologies rugueuses et uniformes, respectivement à la source et au but. Elle se prolonge ainsi à l'adhérence du membre de gauche, vu dans l'ensemble $\mathrm{RP}^\gamma([0;1],\mathbb{R}^d)$, que l'on appelle l'ensemble des chemins rugueux géométriques. Autrement dit, à condition de se limiter en un sens aux limites de courbes lisses, le défaut de continuité est entièrement décrit par les boucles infinitésimales, et si l'on comprend la façon dont les boucles du mouvement brownien cinétique euclidien se comportent dans la limite $\sigma \to \infty$, on comprendra son analogue sur $M$.

Plus généralement, il existe une unique notion de solution aux équations différentielles contrôlées à valeurs dans une variété $N$ de dimension $d$, de la forme

$$\mathrm{d}z_t = f(z_t)\mathrm{d}\mathbf{X}_t,$$

pour $\mathbf{X}$ un chemin rugueux géométrique à valeurs dans $\mathbb{R}^n$, qui en dépende de manière continue et qui coïncide avec la solution classique dirigée par $X$ dans le cas où $\mathbf{X} = (X,\mathbb{X})$ est un chemin rugueux lisse au sens décrit ci-dessus.[8] La construction donnée dans la plupart des ouvrages de référence permet en prolonger cette correspondance en une application continue

$$\mathbf{I}_f : \mathrm{RP}^\gamma([0;1],\mathbb{R}^n) \to \mathcal{C}^0([0;1],N),$$

l'application d'Itô-Lyons, définie sur l'ensemble des chemins rugueux de régularité $\gamma$.

La dernière propriété des chemins rugueux que l'on a utilisée dans la section précédente est le fait que lorsque $X$ est un mouvement brownien, on peut canoniquement lui associer un chemin rugueux (aléatoire) $\mathbf{X} = (X,\mathbb{X})$ en définissant $\mathbb{X}$ par la relation (3.4), au sens de Stratonovich. On appelle ce chemin rugueux le chemin rugueux brownien. Alors la solution de

$$\mathrm{d}z_t = f(z_t)\mathrm{d}\mathbf{X}_t,$$

au sens des chemins rugueux, coïncide presque sûrement avec la solution de

$$\mathrm{d}z_t = f(z_t) \circ \mathrm{d}X_t,$$

au sens de Stratonovich. En particulier, si $\mathbf{X}^\sigma$ est un chemin rugueux aléatoire tel que $\mathbf{X}^\sigma \to \mathbf{X}$ en loi pour $\sigma \to \infty$, avec $\mathbf{X}$ un mouvement brownien isotrope, le développement de Cartan de $X^\sigma$ converge en loi vers le développement de Cartan de $X$, qui est en fait le mouvement brownien isotrope sur $M$, comme annoncé dans la section précédente.

## 3.3 Plan de preuve

Il est temps de se tourner vers la preuve des résultats présentés dans [Per18] et [ABP19]. Ces deux articles, dont on trouvera des versions équivalentes dans les parties II et III, décrivent des résultats d'homogénéisation pour des variantes du mouvement brownien cinétique : le premier concerne

[8] Possiblement jusqu'à un temps d'explosion fini.

des processus dont on relâche certaines conditions concernant la partie vitesse, en particulier l'isotropie ; le second se place dans le cadre de dimension infinie des flots de difféomorphismes d'une variété. S'ils sont de difficulté variable selon les parties, leur structure générale est identique, et suit de près celle du premier travail [ABT15].

Soit $\mathscr{M}$ une variété riemannienne, possiblement de dimension infinie, dans un sens à définir. On commence par définir le mouvement brownien cinétique sur $\mathscr{M}$, issu disons de $(x_0, v_0) \in T^1\mathscr{M}$. On peut décomposer cette construction en trois étapes. La première est la construction d'un processus vitesse $v^\sigma$ dans $T_{x_0}\mathscr{M}$ ; la seconde consiste à intégrer la vitesse pour obtenir un analogue euclidien $x^\sigma$ du mouvement brownien cinétique ; enfin, on définit un procédé géométrique pour le traduire en un processus $q^\sigma$ sur $\mathscr{M}$. On retrouvera ces étapes dans la preuve de l'homogénéisation : on montre d'abord un résultat de mélange pour la vitesse, d'où l'on déduit un théorème central limite fonctionnel pour le mouvement brownien cinétique euclidien. Ce théorème doit être ensuite renforcé en un théorème de convergence pour le chemin rugueux correspondant, qui induit le résultat de convergence pour le mouvement brownien cinétique sur $\mathscr{M}$ si la troisième étape de la construction s'exprime en termes d'équation différentielle contrôlée par $x^\sigma$, par exemple une équation différentielle stochastique.

**Variétés.** Dans le premier travail [Per18], on se place dans le cadre d'une variété riemannienne $\mathscr{M}$ connexe de dimension $d$. On ne fait pas d'hypothèse de complétude, ce qui implique que la convergence sera en un sens locale en espace. Étant donné un point initial $x_0$, le plan tangent $T_{x_0}\mathscr{M}$ est isomorphe à $\mathbb{R}^d$ ; quitte à fixer un isomorphisme, on travaille avec un processus vitesse $v^\sigma$ à valeurs dans $\mathbb{R}^d$.

Dans le second travail [ABP19], la variété $\mathscr{M}$ consiste en l'ensemble des difféomorphismes de $M$ de régularité fixée, pour une certaine variété riemannienne $(M, g)$ de dimension $d$. On considère aussi la sous-variété des transformations préservant le volume riemannien. On ne discute ici que ce second cas, que l'on dit incompressible. La régularité en question est celle de Sobolev $\mathrm{H}^s$, d'indice $s > 1+d/2$, et on suppose $M$ close (compacte sans bord). Il est naturel de choisir comme point initial $x_0$ l'identité de $M$, ce que l'on fait. Le plan tangent $T_{x_0}\mathscr{M}$ est donc l'ensemble des perturbations infinitésimales de l'identité ; conformément à l'intuition, ses éléments s'identifient à des champs de vecteurs à divergence nulle sur $M$.

**Vitesse et mélange.** Discutons d'abord le cas des difféomorphismes, qui motive le second. Par analogie avec le mouvement brownien cinétique standard sur une variété riemannienne close, et puisque $T_{\mathrm{id}}\mathscr{M}$ est doté du produit scalaire $\mathrm{H}^s$, on aimerait définir la vitesse $v^\sigma$ comme la solution de l'équation différentielle stochastique

$$\mathrm{d}v_t^\sigma = \sigma P_{(v^\sigma)^\perp} \circ \mathrm{d}W_t,$$

où $P_{w^\perp}$ est la projection sur l'orthogonal de $w$, c'est-à-dire sur le plan tangent à la sphère unité de $\mathrm{H}^s$, et $W$ est un mouvement brownien standard sur $T_{\mathrm{id}}\mathscr{M}$. On s'abstient ici de donner une définition de mouvement brownien standard, mais l'on peut dire en peu de mots que la covariance d'un mouvement brownien à valeurs dans un certain espace vectoriel peut être vue comme un produit scalaire sur ce même espace. On voudrait donc que la covariance de $W$ corresponde précisément au produit scalaire de $T_{\mathrm{id}}\mathscr{M}$ ; cependant, il est bien connu qu'il n'existe pas de mouvement brownien standard sur cet espace, ou plus précisément, qu'un tel processus doit être à valeurs dans un espace plus grand de fonctions moins régulières. En effet, si par exemple $(\epsilon_1, \epsilon_2, \cdots)$ est une base hilbertienne de $T_{\mathrm{id}}\mathscr{M}$, la somme $\sum_{k\geq 1} W^k \epsilon_k$, où $(W^1, W^2, \cdots)$ est une suite de mouvements browniens standards indépendants à valeurs dans $\mathbb{R}$, ne converge pas dans $T_{\mathrm{id}}\mathscr{M}$. Il est donc nécessaire d'associer à chaque mouvement brownien $W^k$ un coefficient $\alpha_k^2$

de façon à ce que la somme $\sum_k \alpha_k^2$ soit convergente. Dans ce cas, on peut poser formellement $A = \mathrm{diag}(\alpha_1, \alpha_2, \cdots)$, et définir la vitesse comme la solution de l'équation différentielle

$$\mathrm{d}v_t^\sigma = \sigma P_{(v^\sigma)^\perp} \circ A\mathrm{d}W_t.$$

En pratique, on définit un mouvement brownien $B$ de covariance $A^2$, au sens où

$$\mathbb{E}[\phi(B)^2] = \sum_k \alpha_k^2 \phi(\epsilon_k)^2$$

pour toute forme linéaire $\phi \in (T_{\mathrm{id}}\mathscr{M})^*$, qui joue le rôle de $AW$ dans l'équation ci-dessus. Il existe des méthodes pour choisir des matrices $A$ naturelles ; voir la section III.4 du travail en question pour un traitement plus détaillé.

Le mouvement ainsi décrit est donc anisotrope, pour des raisons d'existence importantes. Les problèmes posés par cette anisotropie peuvent être étudiés dès le cadre de la dimension finie, en choisissant la même définition pour $v^\sigma$, pour $B$ un mouvement brownien anisotrope de dimension finie ; on dit que $v^\sigma$ décrit un mouvement brownien sphérique anisotrope parcouru à vitesse $\sigma^2$, d'anisotropie $A$. Les calculs explicites utilisés pour montrer la convergence dans le cas isotrope n'ont pas d'équivalent direct dans ce cas, mais on s'attend au même comportement qualitatif. En pratique, on constate que les hypothèses utiles dans le cas fini-dimensionnel sont les suivantes. Le processus $v^\sigma$ s'exprime comme le changement de temps $t \mapsto \widehat{v}_{\sigma^2 t}$ d'un processus $\widehat{v}$ à valeurs dans $T_{x_0}\mathscr{M}$, et vérifie quatre hypothèses, décrites ci-dessous de manière informelle. Un exemple standard est celui du mouvement brownien standard sur la sphère.

1. Il vérifie une certaine propriété d'invariance : c'est un processus de Feller, ou markovien, ou une chaîne de Markov cachée...

2. Il est borné, ou de taille contrôlée en termes de moments.

3. Il admet une mesure invariante $\mu$, par rapport à laquelle il est suffisamment mélangeant, au sens où la loi de $\widehat{v}_t$ sachant $\widehat{v}_0$ est en un certain sens proche de $\mu$. De plus, on suppose que $\widehat{v}_0$ est distribué selon $\mu$.

4. Il admet suffisamment de symétrie, une condition suffisante étant qu'il soit invariant en loi par rapport aux réflexions
$$\widehat{v} = (\widehat{v}^1, \cdots, \widehat{v}^d) \mapsto (\widehat{v}^1, \cdots, -\widehat{v}^i, \cdots, \widehat{v}^d).$$

Ces hypothèses sont discutées dans la partie II.4.2, à travers différents exemples.

En dimension finie, il existe de nombreux outils pour montrer ces propriétés. En particulier, hormis la troisième, ces conditions sont toutes évidentes dans le cas où $\widehat{v}$ est un mouvement brownien sphérique anisotrope. L'hypothèse manquante, bien que non évidente, est bien connue : $\widehat{v}$ étant la solution d'une équation différentielle stochastique elliptique sur une variété compacte, il existe des méthodes qui montrent directement la convergence en variation totale de $\widehat{v}_t$ vers une unique mesure invariante $\mu$, à vitesse exponentielle et uniformément en le point initial ; voir par exemple le couplage explicite introduit par W. Kendall dans [Ken86], et son développement [Cra91] par M. Cranston. Ce fait est exprimé dans l'équation (II.2.6) du premier travail, et conclut la première étape du résultat de convergence décrit plus haut.

Au contraire, en dimension infinie, la loi de $\widehat{v}_t$ est en général singulière par rapport à $\widehat{v}_s$ pour $s \neq t$, et il est illusoire de chercher une formulation du point 3. en termes de variation totale. Dans le cas très particulier du mouvement brownien sphérique anisotrope, on montre un résultat de convergence exponentielle en distance de Wasserstein 2 dans la proposition III.2.2. Notre preuve impose l'hypothèse suivante sur la matrice d'anisotropie $A$ décrite plus haut.

$$3\alpha_1^2 < \sum_{k\geq 1} \alpha_k^2$$

Cependant, nous nous accordons, peut-être avec différents degrés de conviction, à penser que cette condition devrait pouvoir être levée.

**Cadre euclidien et homogénéisation.** La position $x^\sigma$ du processus euclidien est défini simplement comme l'intégrale de la vitesse $v^\sigma$. Comme discuté précédemment, on cherche à montrer que le chemin rugueux $\mathbf{X}^\sigma$ canoniquement associé à $X^\sigma : t \mapsto x^\sigma_{\sigma^2 t}$ converge en loi vers le chemin rugueux brownien $\mathbf{X}$. Notons que d'après la description de $v$ en fonction de $\widehat{v}$, on peut écrire

$$X_t^\sigma = \frac{1}{\sigma^2} \int_0^{\sigma^4 t} \widehat{v}_s \mathrm{d}s,$$

forme sous laquelle on pressent un théorème central limite. Le problème de convergence est traité de manière classique : on montre que l'ensemble des lois de $\mathbf{X}^\sigma$ est tendu, puis que tout point d'accumulation de cet ensemble est nécessairement la loi de $\mathbf{X}$.

On montre la tension de cette famille *via* des majorations de type Kolmogorov-Lamperti : uniformément en $\sigma > 0$ et $0 \leq s < t \leq 1$, on montre que pour tout $p \geq 2$, il existe une constante $C_p > 0$ telle que

$$\mathbb{E}[|X_t - X_s|^p] \leq C_p |t-s|^{p/2}, \qquad \mathbb{E}[|\mathbb{X}_{ts}|^p] \leq C_p |t-s|^p.$$

Classiquement, la majoration de gauche implique que l'ensemble de lois $\{\mathcal{L}(X^\sigma)\}_{\sigma>0}$ est tendu dans $\mathcal{C}^{1/2-\eta}([0;1], E)$ pour tout $\eta > 0$. Il n'est alors pas étonnant que les deux inégalités montrent la tension de $\{\mathcal{L}(\mathbf{X}^\sigma)\}_{\sigma>0}$ dans $\mathrm{RP}^{1/2-\eta}([0;1], E)$ pour tout $0 < \eta < 1/2$.[9] En pratique, celles-ci sont la conséquence des hypothèses de bornitude et de mélange de $\widehat{v}$ ; on utilise en dimension finie des arguments combinatoires élémentaires, et en dimension infinie des techniques classiques de décomposition en martingales, de façon à montrer

$$\mathbb{E}\left[\left|\int_0^T \widehat{v}_t \mathrm{d}t\right|^p\right] \leq C_p T^{p/2}$$

et son équivalent rugueux. Ces majorations sont les lemmes II.3.1 et II.3.2, et les résultats III.2.6 et III.2.13, respectivement dans les travaux en dimension finie et infinie.

La seconde partie du résultat de convergence euclidien est la caractérisation des limites éventuelles. Supposons que le chemin rugueux $\mathbf{Y} = (Y, \mathbb{Y})$ ait pour loi un point d'accumulation de l'ensemble $\{\mathcal{L}(\mathbf{X}^\sigma)\}_{\sigma>0}$. Alors $Y$ est à accroissements stationnaires, puisque $X^\sigma$ l'est aussi. De plus, si $(Y_{t_i} - Y_{s_i})_{1\leq i\leq I}$ est une famille d'accroissements correspondant à des intervalles compacts disjoints, elle est en fait constituée de variables indépendantes. En effet, les accroissements équivalents de $X^\sigma$ font intervenir les valeurs de $\widehat{v}$ sur des plages de temps distantes de $\sigma^4\varepsilon$, pour un certain $\varepsilon > 0$. D'après l'hypothèse (3) de la partie précédente, ces accroissements sont asymptotiquement indépendants, et $Y$ étant continu presque sûrement, on en conclut que $Y$ est à

[9] On n'a pas défini ici ce qu'est un chemin rugueux de régularité $0 < \gamma \leq 1/3$. Cette discussion nous emmènerait trop loin, mais pour ce qui nous intéresse, il existe une injection continue canonique de $\mathrm{RP}^{\gamma_+}$ dans $\mathrm{RP}^{\gamma_-}$ pour tout $\gamma_+ > \gamma_-$, et c'est la famille ces lois images qui est tendue, à proprement parler.

accroissements indépendants stationnaires. Il s'agit donc d'un mouvement brownien avec dérive. Cependant, l'hypothèse de symétrie (4) de la partie précédente montre que la dérive doit être symétrique par rapport aux réflexions préservant les axes, donc nulle. Il reste donc à caractériser la covariance de $Y$, au sens disons de la donnée de

$$\phi \in (T_{\mathrm{id}}\mathscr{M}) \mapsto \mathbb{E}[\phi(Y_1)^2].$$

Or, on montre facilement que celle-ci ne peut être que

$$\phi \mapsto 2\int_0^\infty \mathbb{E}[\phi(\widehat{v}_0)\phi(\widehat{v}_t)]\mathrm{d}t$$

dans le cas où l'intégrande est effectivement intégrable en temps. L'hypothèse de décorrélation (3), lorsque la convergence est suffisamment rapide, garantit cette intégrabilité, et $Y$ est entièrement caractérisé. Ainsi, on a déjà montré que $X^\sigma$ converge faiblement vers le mouvement brownien dont la covariance est décrite ci-dessus. Ce résultat est le sujet des propositions II.3.5 et III.2.5.

Comme on l'a dit dans la section 3.2, la partie symétrique de $\mathbb{Y}$ est donc elle aussi caractérisée. Le même type d'arguments que précédemment montre que la partie antisymétrique $\mathbb{A}^{\mathbf{Y}}$ ressemble à un mouvement brownien avec dérive. Or, sans pouvoir rentrer ici dans les détails, on a vu plus haut que $\mathbb{A}^{\mathbf{Y}}_{t0}$ est d'ordre $t$, et ainsi ne peut pas admettre de partie brownienne. De plus, l'existence d'une partie de dérive se traduirait informellement par l'accumulation de boucles parcourues dans le sens direct, le long d'un certain plan orienté. Mais d'après l'hypothèse de symétrie (4), on devrait trouver, avec probabilité égale, une accumulation de boucles parcourues dans le sens indirect. C'est absurde, et la dérive est nulle elle aussi, ce qui montre que $\mathbb{A}^{\mathbf{Y}}$ doit être nul.[10] Il est bien connu que ces propriétés sont partagées par le chemin rugueux brownien, et ainsi $\mathbf{Y}$ ne peut être que le chemin rugueux canoniquement associé au mouvement brownien de covariance

$$\phi \mapsto 2\int_0^\infty \mathbb{E}[\phi(\widehat{v}_0)\phi(\widehat{v}_t)]\mathrm{d}t.$$

Ceci conclut la preuve du cas euclidien ; on trouvera ce résultat et sa preuve rigoureuse sous les noms de théorème II.3.6 dans la partie concernant la dimension finie, et de théorème III.2.14 dans celle concernant les variétés de difféomorphismes.

**Développement de Cartan.** Dans le cas de la dimension finie, il suffit de se reporter à la section 3.2 pour conclure la construction et la démonstration. Le mouvement cinétique sur une variété $\mathscr{M}$ de dimension $d$ correspondant au processus de vitesse choisi est défini comme le développement de Cartan de son analogue euclidien. Notons que même si la vitesse est markovienne dans le cas euclidien, le processus défini sur $T\mathscr{M}$ ne l'est plus dès que celle-ci est anisotrope : si l'on suit une courbe dans $\mathscr{M}$, il n'est pas possible en connaissant uniquement la direction avant de différentier l'axe haut-bas de l'axe gauche-droite, le long desquels la vitesse peut se comporter de manière très différente. Il s'agit donc d'un processus markovien si l'on considère le point en mouvement comme étant muni non seulement d'une vitesse mais aussi d'une orientation.

Ceci fait, le développement de Cartan étant une fonction continue du chemin rugueux à la source, le résultat d'homogénéisation rugueuse euclidien est converti en un résultat d'homogénéisation pour le mouvement à valeurs dans $\mathscr{M}$, anisotrope au même sens que celui décrit ci-dessus : étant donnée une orientation, l'intensité du bruit gaussien dans la direction haut-bas a le loisir d'être plus importante que celle présente dans la direction gauche-droite, à condition de garder

[10] Il est en fait, en un sens, aussi nul qu'il puisse être, c'est-à-dire non nul malgré tout.

ce choix le long du mouvement. Ce résultat correspond au théorème II.1.3 pour le cas du mouvement brownien cinétique anisotrope, et au théorème II.4.1 et ses extensions dans la partie II.4.2 pour le cas où la vitesse suit un processus plus général.

Le cas de la dimension infinie ajoute une difficulté supplémentaire : la construction du développement de Cartan. Celui-ci n'a à ma connaissance pas été défini dans ce cadre, et combler ce manque fait partie du travail que j'ai effectué pour traiter ce cas. Le lecteur devra attendre la partie 4.4 pour en trouver une description rapide. La construction effective est donnée dans la section III.5, à partir des considérations des sections III.3.3 et III.3.4.

Il sera important de noter que cette construction s'exprime sous la forme d'une équation différentielle contrôlée. Ainsi, on peut conclure de la même manière que pour le cas de dimension finie, en invoquant les théorèmes de continuité évoqués, dans le cadre des chemins rugueux à valeurs dans un espace de Banach. On obtient alors le théorème III.4.4. Son équivalent dans le cadre où l'on n'impose pas la conservation du volume est le théorème III.4.3.

# 4 Mécanique des fluides

## 4.1 Mécanique classique et géodésiques

Comme on l'a vu plus haut, il existe des phénomènes physiques concrets dont la modélisation fait intervenir l'équation des géodésiques dans un cadre de dimension infinie. Celui qui est étudié dans cette thèse est l'exemple de la mécanique des fluides.

Pour comprendre comment ces équations apparaissent, le plus simple est de s'inspirer de la dimension finie. En l'absence de force extérieure, comme par exemple à bord de la station spatiale internationale, le mouvement d'un objet ponctuel de masse $m$ est rectiligne, parcouru à vitesse constante $v$. En tant que courbe allant d'un point $A$ à un point $B$, elle minimise localement l'intégrale son l'énergie cinétique, qui dans ce cas est simplement $\frac{1}{2}m|v|^2$ ; autrement dit, elle réalise une géodésique dans l'espace ambiant, muni de l'énergie en question.

Si l'objet n'est plus ponctuel, mais par exemple une balle, on constate qu'en plus de se déplacer en ligne droite, elle tourne sur elle-même à vitesse constante. Quel est l'espace dans lequel vit cette vitesse ? On peut voir l'ensemble des couples (position, orientation) comme le groupe $G$ des isométries de $\mathbb{R}^3$, en fixant une configuration de référence. Il s'agit d'un groupe de Lie de dimension 6, pouvant être représenté par une partie de translation et une partie de rotation : $G = \mathbb{R}^3 \rtimes SO_3(\mathbb{R})$. Une variation de configuration pour la balle peut se décrire soit par rapport à l'espace ambiant par l'action à gauche par un élément de $G$, soit dans le référentiel de la balle elle-même, par l'action à droite. La vitesse $\dot{g}$ de la balle dans une certaine configuration $g \in G$ est alors un élément de $T_gG$, et on peut la voir par rapport à l'observateur comme un vecteur $\dot{g}g^{-1}$ de l'algèbre de Lie $\mathfrak{g} := T_eG$, ou par rapport à l'objet comme le vecteur $g^{-1}\dot{g}$ de la même algèbre. L'énergie cinétique naturelle associée serait alors $\frac{1}{2}m|\dot{g}|^2$, pour une certaine métrique riemannienne sur $G$.

Bien sûr, le choix de la métrique n'est pas anodin. Deux notions d'énergie différentes vont mener à deux ensembles de trajectoires différentes. La propriété principale qu'elle doit vérifier est l'invariance par changement de repère : la vitesse mesurée par un astronaute situé juste à côté de l'objet est la même que celle qu'enregistre sa collègue canadienne plus loin dans le module. En termes mathématiques, cela impose que la métrique est invariante par multiplication à gauche. L'espace[11] des métriques riemanniennes sur $G$ devient alors fini-dimensionnel, et un tel produit scalaire est entièrement caractérisé par ses valeurs sur $\mathfrak{g}$. D'autres considérations physiques lui donnent la forme suivante. En fixant un point de $\mathbb{R}^3$, on identifie dans $G$ une copie de $SO_3(\mathbb{R})$

[11] À proprement parler, il s'agit d'un cône convexe.

et un élément de $\mathfrak{g}$ peut être vu comme une partie $v \in \mathbb{R}^3$ de translation, et une partie $r \in \mathbb{R}^3$ de rotation, colinéaire à l'axe de rotation, et de magnitude proportionnelle à sa vitesse angulaire. Alors l'énergie cinétique prend la forme $\frac{1}{2}m|v|^2 + I(r)$, pour une certaine forme quadratique $I$ appelée matrice d'inertie. Pour l'exemple de la balle, l'isotropie de l'objet impose à $I$ d'être elle-même isotrope. On constate alors que les courbes géodésiques sont effectivement les trajectoires décrites plus haut : rotation et translation uniformes.

Pour un corps non isotrope, il peut être plus difficile de le mettre en rotation selon certaines directions. Un stylo, par exemple, a un axe de rotation privilégié selon lequel il tourne plus facilement. On trouvera sur internet des vidéos qui permettent de se convaincre que malgré l'apparente simplicité du problème, le mouvement d'un solide anisotrope hors gravité peut être complexe, et il est frappant qu'une équation aussi intuitive que celle des géodésiques soit effectivement adaptée pour décrire de tels comportements. On pourra par exemple consulter la page Wikipédia de l'effet Djanibekov, pour les corps dont la matrice d'inertie $I$ a trois valeurs propres distinctes, que je trouve particulièrement édifiante.

On vient de voir sur un exemple que la trajectoire d'un corps libre est régie par une équation géodésique, dont la métrique associée correspond à la notion physique d'énergie cinétique. Un autre cas nous aidera à interpréter le cas de la mécanique des fluides, celui des mouvements contraints. Étant données deux masses $x$ et $y$ reliées par une tige sans masse de longueur fixée, la discussion précédente ne permet pas de modéliser leur évolution. En effet, l'inertie de cette objet est la même que celle de deux masses identiques sans interaction, qui devraient se déplacer en ligne droite, indépendamment l'une de l'autre.[12] En termes mathématiques, la contrainte se modélise par la restriction de l'espace des configurations possibles à une sous-variété. Ici, il s'agit de l'ensemble des couples $(x, y)$ tels que la distance $|y - x|$ est égale à la longueur de la tige.

Dans un tel système, la restriction de l'énergie à ce sous-ensemble de configurations, c'est-à-dire la structure riemannienne induite sur la sous-variété par la métrique ambiante, permet encore une fois d'exprimer les trajectoires comme la solution des équations géodésiques. Celle-ci peut être informellement décrite de la manière suivante, disons dans le cas d'une sous-variété $M$ de l'espace euclidien $\mathbb{R}^d$. Partant d'un point initial $x$, par exemple un $n$-uplet $(x_1, \cdots, x_n) \in M \subset (\mathbb{R}^3)^N$, et d'une vitesse initiale $v$, à l'instant $t + \mathrm{d}t$, le point est à une position $x + v\mathrm{d}t$. La vitesse devrait rester constante ; cependant, elle ne serait alors plus tangente à la sous-variété $M$ à l'ordre 1. Il faut donc plutôt considérer $\Pi_{x+\mathrm{d}x}(V)$, où $\Pi_y$ la projection orthogonale sur l'espace tangent $T_yM$ à $M$ en $y$. Cette heuristique ne peut pas être rigoureuse, puisque, *a priori*, $x + v\mathrm{d}t$ n'appartient pas à $M$. Cependant, en remplaçant $\Pi_y$ par la projection orthogonale sur l'espace $\ker \mathrm{d}\Phi_y$, où $\Phi$ est une submersion définissant $M = \Phi^{-1}\{0\}$ localement, on peut montrer que l'on décrit un schéma d'Euler qui converge effectivement vers l'équation des géodésiques sur $M$ en un sens naturel lorsque $\mathrm{d}t$ tend vers zéro.[13] Il n'est pas difficile d'adapter cette démarche au cas d'une sous-variété $M \subset \mathscr{M}$, en supposant que l'on dispose déjà d'une procédure pour simuler les accroissements géodésiques dans la variété ambiante $\mathscr{M}$.

## 4.2 Milieux continus

Considérons un ensemble de particules de fluide dans un volume fini sans bord. Disons par exemple que l'on dispose d'un (grand) nombre fini $N$ de points dans un domaine cubique périodisé (mathématiquement, un tore $\mathbb{T}^3$), et que l'on considère les cellules de Voronoï associées. Chacune de ces particules subit la pression de ses voisines, et on suppose donc que la contrainte de

[12] Ceci n'est que partiellement vrai ; on pourrait considérer une métrique infinie sauf sur un sous-espace vectoriel tangent à un certain feuilletage, à la manière d'une variété sous-riemannienne.

[13] Si ce schéma est efficace pour l'exposition, il l'est beaucoup moins pour des considérations numériques sérieuses. Pour un traitement de qualité, demander l'aide d'un professionnel.

préservation du volume est respectée : au cours du mouvement du fluide, le volume de chaque particule est fixé. En termes mathématiques, on dispose sur la variété totale $(\mathbb{T}^3)^N$ de $N$ fonctions $\mathcal{A}_i$ correspondant au volume de chaque cellule, et on considère un ensemble de la forme $M = \{\mathcal{A}_i = A_i^0 \text{ pour tout } i\}$. Les fonctions $\mathcal{A}_i$ sont lisses localement autour d'un point générique, donc $M$ ressemble à une sous-variété. On suppose que le fluide est non-visqueux, de manière à ce que le système soit purement conservatif. L'énergie totale du système suffit donc à le caractériser ; il s'agit de la somme des énergies cinétiques de chaque cellule, que l'on choisit de la forme $\frac{1}{2}\mu\mathcal{A}_i|v_i|^2$, où $v_i$ est la vitesse du $i^{\text{e}}$ germe, et $\mu$ la masse volumique du fluide.

Loin des singularités (lorsque deux points se rencontrent, par exemple), les équations du mouvement sont bien définies, et le raisonnement précédent permet de décrire la dynamique du fluide combinatoire. L'idée de l'approche de V. Arnol'd pour la mécanique des fluides est d'imaginer que cette discrétisation ne modifie pas substantiellement la nature de l'évolution : les équations des géodésiques permettent aussi de décrire le fluide continu, dans un espace de configurations de dimension infinie. Notons cependant que cette approche des particules de fluides est trop naïve ; pour un traitement similaire par des chercheurs en physique, voir les articles [SE01, SEZ05] de M. Serrano, P. Español et I. Zúñiga. Arnol'd, d'ailleurs, se garde bien de discrétiser l'espace. La plupart de ce qui se trouve dans cette section est introduit dans son article [Arn66] de 1966, y compris les considérations de dimension finie. Cette interprétation est fondatrice, pour ne pas dire révolutionnaire, mais pas entièrement rigoureuse. Le travail de formalisation est entrepris par D. Ebin et J. Marsden, et s'achève par l'article [EM69], publié trois ans plus tard.

De même que précédemment, on décrit un espace de configurations muni d'une notion d'énergie et d'un ensemble de contraintes. Ces objets devront rendre compte des équations d'Euler pour la mécanique des fluides, qui modélisent un fluide parfait incompressible et non visqueux, disons contenu dans le volume d'une variété riemannienne close $M$. On discute le lien entre notre description et elle en termes d'équations aux dérivées partielles dans la section suivante.

Un point donné de l'espace ambiant est caractérisé par la position de toutes les particules de fluide. On les indexe de manière naturelle par leur position initiale $x$, de façon à ce qu'une configuration soit donnée par une fonction $\varphi : M \to M$, initialement représentée par l'identité. La vitesse du fluide, dans cet espace de dimension infinie, consiste en la donnée de la vitesse de chaque particule. Il s'agit donc d'une collection $V$ de vecteurs tangents à $M$, tel que le vecteur $V(x)$ indexé par $x$ soit enraciné en $\varphi(x)$. Chaque élément de volume $\mathrm{d}x$ a donc une énergie $\frac{1}{2}\mu(x)|V(x)|^2\mathrm{d}x$, où $\mu$ est la masse volumique du fluide. On considère le fluide homogène, de façon à ce que $\mu$ soit une constante, et par extensivité de l'énergie, on obtient la formule suivante pour l'énergie du fluide.

$$\frac{1}{2}\mu\int_M |V(x)|^2\mathrm{d}x = \frac{1}{2}\mu|V|^2_{\mathrm{L}^2}.$$

Il reste à comprendre l'influence des contraintes imposées. Dans notre cas, nous avons pour l'instant ignoré la préservation du volume. On pourrait en implémenter une version faible en restreignant les fonctions $\varphi$ admissibles à celles qui préservent la mesure. Cette approche n'est pas adaptée aux outils de géométrie différentielle auxquels on voudrait avoir accès ; on fait donc l'hypothèse beaucoup plus restrictive que $\varphi$ est de classe $\mathcal{C}^1$, et que la forme volume $\omega$ associée à la structure riemannienne est préservée : $\varphi^*\omega = \omega$.[14]

Une propriété notable de cette construction est que l'énergie ainsi définie est invariante à droite. En effet, remplacer une évolution $t \mapsto \varphi_t$ par $t \mapsto \varphi_t \circ \psi$ consiste simplement à réindexer les particules. À condition que $\psi$ préserve le volume, l'énergie d'une trajectoire est donc rigoureusement égale à celle de l'autre ; mathématiquement, cela s'observe dans la définition de $|V|_{\mathrm{L}^2}$,

[14] On a supposé ici $M$ orientable. Si elle ne l'est pas, on considère une densité plutôt qu'une forme volume.

par le constat que le jacobien de $\varphi$ vaut 1 dans le changement de variable.

En substance, le point de vue d'Arnol'd consiste à dire que l'on a effectivement construit ci-dessus une sous-variété dans une variété riemannienne ambiante de dimension infinie, et que les géodésiques associées sont les solutions des équations d'Euler.

## 4.3 Les équations d'Euler

Dans la littérature, la description la plus courante des équations d'Euler se fait en termes d'équations aux dérivées partielles. Dans le cas qui nous intéresse (incompressible homogène), elles s'écrivent ainsi.

$$\frac{\partial u}{\partial t} + \nabla_u u = -\nabla p, \qquad \operatorname{div} u = 0.$$

Un certain nombre de notations est à éclaircir. Tout d'abord, les variables $u_t(x)$ et $p_t(x)$ représentent respectivement la vitesse et la pression de la particule située au point $x$ à l'instant $t$ ; dans le formalisme de la partie précédente, $u = V \circ \phi^{-1}$. On peut déjà noter que l'équation de droite rend compte du fait que le fluide est incompressible : puisque $\phi$ préserve le volume, $V$ est à divergence nulle si et seulement si $u$ l'est.

Ensuite, $\nabla_u u$ représente la dérivée (covariante) de $u$ dans la direction $u$. La quantité

$$\frac{\partial u}{\partial t} + \nabla_u u$$

est ce que l'on appelle la dérivée particulaire de $u$. Elle est égale dans notre cas à l'accélération $\dot{V} \circ \phi^{-1}$.[15] D'après la seconde loi de Newton, cette accélération est égale à la somme des forces subies par la particule. Ici, ces forces dérivent d'un potentiel, la pression, dont on a noté $\nabla p$ le gradient. Dans notre introduction en dimension finie, on a décrit les accroissements de la vitesse des géodésiques restreintes à une sous-variété, notre actuel $V$, en les exprimant au premier ordre comme la projection des incréments dans la variété ambiante sur le plan tangent en $\phi + \mathrm{d}\phi$ à la sous-variété. Or, l'espace tangent à l'ensemble des difféomorphismes préservant le volume est celui des vecteurs de divergence nulle. Sous la condition que la variété $M$ contenant le fluide soit simplement connexe, son orthogonal est exactement l'ensemble des gradients. Autrement dit, puisque la vitesse $V$ définie par l'équation des géodésique doit avoir une dérivée orthogonale aux champs à divergence nulle, elle doit s'exprimer comme un certain gradient, ce qui nous mène à l'équation de gauche.

Inversement, si $u$ satisfait les équations d'Euler pour une certaine pression $p$, on définit $\phi$ en intégrant la vitesse, via l'équation différentielle

$$\phi_t(x) = x + \int_0^t u_s \circ \phi_s(x) \mathrm{d}s,$$

puis $V := \dot{\phi}$. Puisque $u$ est à divergence nulle, $V$ l'est aussi. De plus, sa dérivée (l'accélération $\dot{V}$) n'a de composante non nulle que sur l'espace des gradients, donc sa composante de divergence nulle est en un sens constante. La vitesse $V$ est donc aussi constante qu'elle peut l'être sous la contrainte de préservation du volume, elle satisfait donc l'équation des géodésiques.

Les équations d'Euler sous la forme donnée ci-dessus sont donc formellement équivalentes à celles des géodésiques décrites plus haut.

[15] Cette équation n'a de sens que dans une variété $M$ dont tous les espaces tangents sont comparables, par exemple un groupe de Lie. Dans le cas général, il faut user de notions plus avancées de géométrie riemannienne et parler de $K(\dot{V}) \circ \phi^{-1}$, où $K : TTM \to TM$ est le connecteur, c'est-à-dire la projection orthogonal sur l'espace des vecteurs verticaux.

## 4.4 Analyse globale non-linéaire

Le raisonnement décrit plus haut est purement formel. Pour donner un sens précis à ces objets, Ebin et Marsden utilisent les outils de l'analyse globale non-linéaire. On en donne ici une exposition proche de celle donnée par R. Palais dans [Pal68].

**Variétés d'applications.** L'analyse globale est le nom donné à l'étude des variétés selon l'approche de l'analyse (analyse fonctionnelle, équations aux dérivées partielles...). Elle se base couramment sur l'introduction d'espaces de fonctions bien adaptés. Par exemple, étant donné un fibré $\pi : V \to M$ sur une variété close $M$ munie d'un volume $\mathrm{d}x$,[16] on peut définir l'ensemble des distributions à valeurs dans $V$ comme un élément du dual de l'espace $\mathcal{C}^\infty(V^*)$ vérifiant certaines conditions de continuité, dont des exemples réguliers sont donnés par les sections localement intégrable $f : M \to V$, via l'identification

$$\rho \mapsto \langle f, \rho \rangle := \int_M \rho_x(f_x)\mathrm{d}x.$$

Étant donné un fibré vectoriel $V$ sur une variété close $M$, respectivement de rang $k$ et de dimension $d$, il est bien connu que l'on peut définir l'espace $\mathrm{H}^s(V)$ des sections de $V$ de régularité de Sobolev $s \in \mathbb{R}$. Il existe des dizaines de définitions équivalentes de ces espaces, s'appliquant parfois à des exposants $s$ particuliers ; on en donne quelques unes ci-dessous, pas toujours indépendantes.

- L'espace $\mathrm{H}^0(V)$ est égal à l'ensemble $\mathrm{L}^2(V)$, et leurs topologies coïncident.
- Une section $f : M \to V$ est de régularité $\mathrm{H}^s$ pour $s \geq 0$ si en tout point $x$, il existe une carte locale $\varphi : U \subset M \to \varphi(U) \subset \mathbb{R}^d$ autour de $x$ et une trivialisation $(\tau, \pi) : F_{|U} \to \mathbb{R}^k \times U$ telles que $\tau f \varphi^{-1}$ appartienne à $\mathrm{H}^s(\varphi(U), \mathbb{R}^k)$. Une norme sur cet espace est donnée par la somme des $\left|\tau_i f \varphi_i^{-1}\right|_{\mathrm{H}^s}$, où l'ensemble (fini) des $U_i$ choisis recouvre $M$.
- Une distribution $f$ à valeurs dans $V$ est de régularité $\mathrm{H}^s$ si pour une certaine constante $C > 0$,
$$|\langle f, \rho \rangle| \leq C|\rho|_{\mathrm{H}^{-s}}$$
pour tout $\rho \in \mathcal{C}^\infty(V^*)$. Une norme sur cet espace est donnée par la plus petite constante $C$ satisfaisant cette propriété.
- Une distribution $f$ à valeurs dans $V$ est de régularité $\mathrm{H}^{s+k}$ si, pour tout opérateur différentiel $L : \mathcal{D}'(V) \to \mathcal{D}'(M \times \mathbb{R})$ à coefficients lisses d'ordre $k \in \mathbb{N}$, la distribution $Lf$ est à valeurs dans $\mathrm{H}^s(M \times \mathbb{R})$.
- Pour toute connexion $\nabla$ sur $V$, on note $\Delta$ le laplacien de Bochner associé, de domaine $D(\Delta) \subset \mathrm{L}^2(V)$. Son spectre $\sigma(\Delta)$ est positif, et est constitué de valeurs propres de multiplicité finie. L'espace $\mathrm{H}^s(V)$ est alors le complété de la somme directe algébrique
$$\bigoplus_{\lambda \in \sigma(\Delta)} \ker(\Delta - \lambda\, \mathrm{id}),$$
relativement à la norme
$$|f|^2_{\mathrm{H}^s} := \sum_\lambda \lambda^{2s} |f_\lambda|^2_{\mathrm{L}^2},$$
les valeurs propres étant comptées avec multiplicité.

[16] La donnée d'un volume n'est qu'une commodité. En son absence, les distributions sont des éléments du dual de $\mathcal{C}^\infty(V^* \otimes |\wedge^d|)$, où $|\wedge^d|$ est le fibré des densités sur $M$.

Pour montrer l'équivalence de la dernière définition avec les autres, on utilise les expressions en Fourier de la régularité de Sobolev et du laplacien.

On note qu'aucune norme sur ces espaces n'est canonique. Cependant, leur topologie l'est, et on se contente souvent de choisir un produit scalaire arbitraire qui l'engendre.

On parle d'analyse globale non-linéaire quand le fibré considéré n'est plus vectoriel, mais seulement différentiel. Il peut par exemple s'agir du fibré orthonormal $OM$, constitué des couples $(x, e)$, où $e : \mathbb{R}^d \to T_xM$ est un isométrie, du fibré trivial $(x, y) \in M \times N \mapsto x$ sur $M$ lorsque $N$ est une variété, etc. On note $\pi : F \to M$ un tel fibré, et on dit qu'une application lisse $\Phi : F \to F'$ est un morphisme de fibré si $\pi_{F'} \circ \Phi = \pi_F$. Puisqu'une variété n'est définie qu'à difféomorphisme près, toute notion de régularité $(\mathcal{R})$ pour une section $f : M \to F$ doit être invariante par composition au but par un difféomorphisme de fibré.

Cette contrainte est en fait écrasante. Par exemple, il n'est pas très difficile de montrer que pour une section $f : M \to V$ d'un fibré vectoriel, $\Phi \circ f : M \to V$ est dans $\mathrm{L}^2$ pour tout difféomorphisme de fibré $\Phi$ si et seulement si $f$ est en fait essentiellement bornée. Plus généralement, puisque les variétés ne ressemblent que localement à des espaces vectoriels, les espaces de fonctions non continues ont tendance à se comporter pathologiquement. Au contraire, le résultat suivant donne une caractérisation très satisfaisante des espaces de Sobolev $\mathrm{H}^s$ pour $s > \frac{d}{2}$, c'est-à-dire lorsqu'ils s'injectent dans l'ensemble des fonctions continues.

**Théorème 4.1.** *Soit $s > \frac{d}{2}$.*

*Il existe un unique foncteur $\mathrm{H}^s$ de la catégorie des fibrés différentiels (muni des morphismes de fibrés) dans celle des variétés de Hilbert (munie des applications lisses) tel que*

- *lorsque $V$ est un fibré vectoriel, $\mathrm{H}^s(V)$ coïncide avec l'espace hilbertisable*[17] *défini précédemment ;*
- *$\mathrm{H}^s(F)$ est constitué de sections continues ;*
- *l'application associée à un morphisme $\Phi : F \to F'$ est la composition au but :*

$$\begin{aligned}\mathrm{H}^s(F) &\to \mathrm{H}^s(F')\\ f &\mapsto \Phi \circ f.\end{aligned}$$

On appelle $\omega_\Phi$ l'application décrite par le dernier point. Dans la littérature, on trouve sous le nom de lemme oméga le fait que les applications de cette forme sont lisses. C'est l'ingrédient principal de la preuve du théorème, dans le cas où $\Phi$ est un morphisme de fibrés différentiels entre fibrés vectoriels. C'est aussi un outil fondamental : montrer qu'une fonction est lisse dans des espaces de dimension infinie est souvent ardu ; ce résultat en main, la difficulté se réduit à l'élaboration, en dimension finie, de fibrés et de morphismes de fibrés bien choisis. Parfois, l'on a aussi besoin de considérer des opérateurs différentiels ; une propriété importante de ces espaces est que les opérateurs différentiels à coefficients compacts, dans un sens raisonnable, sont bien définis de $\mathrm{H}^{s+k}(F)$ dans $\mathrm{H}^s(F')$ dès que l'ordre de l'opérateur est au plus $k$.

Tentons de décrire la différentielle d'une application $\omega_\Phi$, pour un morphisme $\Phi : F \to F'$. Intuitivement, étant donnée une section $f : M \to F$, une variation infinitésimale $\delta f$ de $f$ s'exprime comme une collection $\delta f(x)$ de variations infinitésimales de $f(x)$. Chacune de ces variations est un élément de l'espace tangent $T(F_{f(x)})$ dans la fibre. On appelle espace tangent vertical $VF$ la collection de ces fibres tangentes. Elle est muni d'une structure de fibré sur $M$ naturelle. Cette

[17] Espace vectoriel topologique dont la topologie est issue d'une norme de hilbertienne.

discussion informelle peut être rendue rigoureuse, et se traduit en un isomorphisme $T\mathrm{H}^s(F) \simeq \mathrm{H}^s(VF)$.

Sous l'image de $\Phi$, une variation $\delta f(x)$ de $f(x)$ se traduit en une variation $(T_{f(x)}\Phi)(\delta f(x))$, et ainsi, la restriction $V\Phi : VF \to VF'$ de $T\Phi$ permet de définir $\omega_{V\Phi}$, qui s'identifie à $T\omega_\Phi$ :

$$\begin{tikzcd} T\mathrm{H}^s(F) \arrow[r, "T\omega_\Phi"] \arrow[d, "\simeq"'] & T\mathrm{H}^s(F') \arrow[d, "\simeq"] \\ \mathrm{H}^s(VF) \arrow[r, "\omega_{V\Phi}"] & \mathrm{H}^s(VF') \end{tikzcd}$$

La preuve usuelle du lemme oméga se base sur cet isomorphisme : il suffit de montrer que $\omega_\Phi$ est $\mathcal{C}^1$ et que $T\omega_\Phi = \omega_{V\Phi}$, puisqu'alors le raisonnement s'applique ensuite aussi à $\omega_{V\Phi}$, etc.

Ceci complète la boîte à outils élémentaire de l'analyste global et non-linéaire. Dans la suite, on trouvera une description de la structure riemannienne de la variété des difféomorphismes d'une variété compacte, avec comme objectif la discussion du développement de Cartan dans ce cadre.

**Variété de difféomorphismes.** Ces quelques paragraphes se basent sur le travail [EM69] d'Ebin et Marsden. Soit $M$ une variété riemannienne close de dimension $d$. D'après l'exposition qui précède, l'ensemble $\mathscr{M} := \mathrm{H}^s(M, M)$ des applications de $M$ dans $M$ de régularité $\mathrm{H}^s$ est bien défini pour $s > d/2$. On peut le voir comme un ensemble de sections du fibré trivial $M \times M \to M$. Comme on l'a vu plus haut, l'espace tangent $T_\phi\mathscr{M}$ en $\phi$ s'identifie au sous-ensemble de $\mathrm{H}^s(M, TM)$ constitué des champs de vecteurs $V : M \to TM$ enracinés en $\phi$, c'est-à-dire tels que $\pi \circ V = \phi$. Alors on peut définir sur $T_\phi\mathscr{M}$ l'application

$$\int_M V_x \cdot W_x \mathrm{d}x,$$

une application bilinéaire dont la forme quadratique associée est l'énergie évoquée dans la section 4.2. On peut montrer que c'est effectivement une métrique riemannienne sur $\mathscr{M}$, faible au sens où la topologie qu'elle induit (essentiellement la topologie $\mathrm{L}^2$) est plus faible que celle de la variété. Si de plus $s > 1 + \frac{d}{2}$, il n'y a pas de problème à considérer le sous-ensemble $\mathscr{M}_0$ des difféomorphismes préservant la forme volume, tels que décrits dans la partie 4.2. On peut en fait montrer qu'il s'agit d'une sous-variété, et on a donc défini rigoureusement tous les objets évoqués dans cette dernière partie.

La nature locale de cette métrique permet de construire un grand nombre d'opérations point par point. Par exemple, alors qu'une métrique faible n'admet pas forcément de connexion de Levi-Civita, on peut construire explicitement une telle dérivée covariante point par point. En notant $K : TTM \to TM$ le connecteur de $M$, c'est-à-dire l'opérateur de projection orthogonale sur les champs de vecteurs verticaux, la connexion $\overline{\nabla}$ est donnée par

$$\overline{\nabla}_X Y = \omega_{\mathrm{id},K} \circ TY \circ X.$$

Décomposons cette formule. $X$ et $Y$ sont des champs de vecteurs sur $\mathscr{M}$, donc des fonctions $\mathscr{M} \to T\mathscr{M}$. La différentielle $TY$ est donc une application $T\mathscr{M} \to TT\mathscr{M}$, et $TY \circ X : \mathscr{M} \to TT\mathscr{M}$. D'après les isomorphismes décrit dans la partie précédente, $TT\mathscr{M}$ s'identifie à $\mathrm{H}^s(M, TTM)$. Ainsi, puisque $(\mathrm{id}, K)$ envoie $M \times TTM$ sur $M \times TM$, le membre de droite est bien une application de $\mathscr{M}$ vers $T\mathscr{M}$, et il n'est pas difficile de se convaincre qu'il s'agit en effet d'un champ de vecteurs.

Muni de cette connexion, on constate que $t \mapsto \varphi_t \in \mathscr{M}$ est une géodésique dans cet espace si et seulement si $t \mapsto \varphi_t(x) \in M$ est une géodésique pour tout $x \in M$ ; de même, $t \mapsto V_t \in T\mathscr{M}$

est transporté parallèlement si et seulement si $t \mapsto V_t(x) \in TM$ est transporté parallèlement pour tout $x \in M$ ; etc. Or, on a vu dans la partie 4.1 qu'il existe une stratégie pour construire les géodésiques sur les sous-variétés : il suffit de projeter la vitesse sur le plan tangent à la sous-variété à chaque incrément infinitésimal. C'est effectivement ce qui est fait dans [EM69], et de cette façon, les auteurs prouvent des résultats très forts d'existence, de régularité et de continuité par rapport aux conditions initiales pour les équations d'Euler sur les variétés closes.

**Développement de Cartan.** Le développement de Cartan, dans le cadre de la dimension finie, a été décrit dans la partie 3.2. Un élément fondamental de mon travail sur le mouvement brownien cinétique en dimension infinie est sa construction, qui repose entièrement sur l'adaptation de cette procédure à la dimension infinie. Les détails en sont donnés dans le chapitre III, dans les sections III.3 et III.5. Malheureusement, il me faut supposer que le lecteur est familier avec la construction fini-dimensionnelle du développement de Cartan ; en particulier, je ne décrirai ici que l'idée de la construction du transport parallèle.

Le développement de Cartan d'une courbe $v$ à valeurs dans $T_{\mathrm{id}}\mathscr{M}$ est une courbe $z$ à valeurs dans un espace $\mathscr{O}$, qui joue en dimension infinie le rôle d'un hypothétique fibré orthonormal $O\mathscr{M}$. L'introduction de cet espace $\mathscr{O}$ doit sa raison d'être au fait que $O\mathscr{M}$ est très grand, et en particulier sa structure topologique et géométrique est complexe. En effet, sa définition ne fait appel qu'à la structure de variété riemannienne de $\mathscr{M}$, en particulier elle oublie en quelque sorte le fait que $\mathscr{M}$ est une variété d'applications ; le lien entre $O\mathscr{M}$ et $M$ est difficile à éclaircir, et un grand nombre d'éléments de $O\mathscr{M}$ n'ont pas de sens physique (opérations non-locales, discontinuités, etc.). On se restreint donc à un sous-fibré $\mathscr{O}$, dont la structure est plus claire et plus aisément manipulable, et qui se trouve contenir tous les éléments de $O\mathscr{M}$ qui nous occupent. À peu de choses près, il s'agit d'un espace de sections.

La courbe $v$ dont on veut construire le développement, à valeurs dans $T_{\mathrm{id}}M$, est un champ de vecteurs dépendant du temps, que l'on peut voir comme une collection de mouvements $v(x)$ indexés par $x \in M$ et à valeurs dans $T_xM$. Chacune de ces courbes admet un développement de Cartan dans $OM$, ce qui définit une collection de mouvements $z(x)$ à valeurs dans $OM$, qui n'est autre que $z$. En un sens, cette construction est faite point-par-point, ce qui est rendu possible par l'emploi de la métrique $\mathrm{L}^2$, qui ne fait pas intervenir de corrélation entres points proches comme pourrait le faire, par exemple, une norme de type Sobolev.

Quant à $\mathscr{M}_0$, l'idée est de représenter une orientation dans l'espace $\mathscr{O}$ décrit plus haut, et d'introduire la correction due à la contrainte de la préservation du volume dans un facteur indépendant. Cette stratégie n'a rien de spécifique à la dimension infinie : on illustre dans la figure 4.5 le cas d'une sous-variété de dimension 1 dans $\mathbb{R}^2$. D'un côté, on transporte à l'aide de $\mathscr{O}$ un repère le long d'un chemin dans $\mathscr{M}_0$, parallèlement au sens de la métrique de $\mathscr{M}$. Cette procédure correspond au déplacement du repère bleu. On fait alors une erreur, mesurée par une rotation dans l'espace tangent d'arrivé. Mais le repère de $\mathscr{O}$ permet de ramener cette erreur dans l'espace tangent de départ, qui, lui, est fixe. On peut alors corriger cette erreur dans un espace euclidien, dans lequel les manipulations sont plus directes. La correction est représentée en rose sur le graphique. La superposition de ces deux mouvements permet d'exprimer le transport parallèle, et donc le développement de Cartan, comme la solution d'une équation différentielle.

## 5 Asymptotique en temps petit

À côté du projet principal discuté plus haut, j'ai étudié en temps petit de la densité du mouvement brownien cinétique, dans le cas euclidien de dimension finie. Dans cette optique, j'ai passé cinq mois entre février et juin 2018 à l'université de Warwick en Angleterre, pour collaborer avec

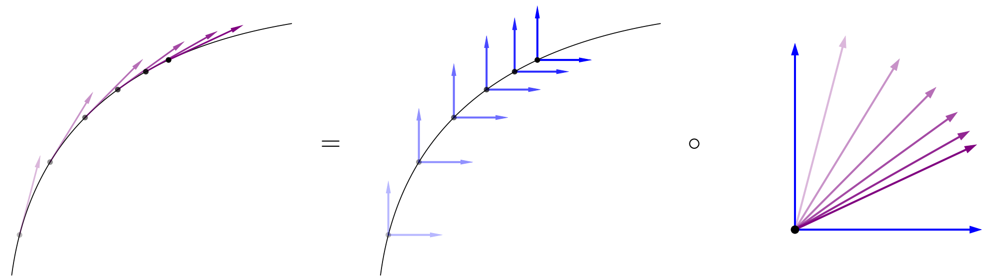


FIGURE 4.5 – Transport parallèle pour une sous-variété.

Vassili Kolokoltsov. Je suis profondément reconnaissant à l'université de Warwick et à V. Kolokltsov, où et avec qui j'ai beaucoup appris, et leur adresse mes sincères remerciements. Le fruit de ce travail, toujours en cours, est développé dans le chapitre IV, et la partie que le lecteur a sous les yeux lui sert d'introduction.

On considère le mouvement brownien cinétique dans $\mathbb{R}^d$ pour $d \geq 2$, de paramètre $\sigma = 1$ et de condition initiale $(x_0, v_0)$ fixés, dont on note $L$ le générateur. On appelle $u_t$ la densité du couple $(x_t, v_t)$. Elle est solution, au sens des distributions, de l'équation

$$\partial_t u_t = -v\partial_x u_t + \frac{1}{2}\Delta_v u_t = L^* u_t, \tag{5.5}$$

où $\partial_x f$ représente le gradient de $f : \mathbb{R}^d \times \mathbb{S}^{d-1}$ selon sa première composante $x \in \mathbb{R}^d$, et $\Delta_v f$ le laplacien selon sa seconde composante $v \in \mathbb{S}^{d-1}$, au sens de Laplace-Beltrami.

## 5.1 Support

Comme promis, on montre que le support de $u_t$ est précisément $\{|x - x_0| \leq t\}$ pour tout $t > 0$. Puisque $x$ se déplace à vitesse 1, il est clair que $\mathbb{P}(|x_t| > t) = 0$, et ainsi le support de $u_t$ est inclus dans $\{|x| \leq t\}$. Pour ce qui est de l'inclusion contraire, on se base sur le théorème de Stroock-Varadhan, dont une version très faible et adaptée à nos besoins est la suivante. Soit $h : [0; t] \to \mathbb{R}^d$ une courbe $\mathcal{C}^1$ telle que $h_0 = 0$. On note $(x^h, v^h)$ la solution de l'équation contrôlée

$$\mathrm{d}x_s^h = v^h \mathrm{d}s, \qquad \mathrm{d}v_s^h = P_{(v_s^h)^\perp} \mathrm{d}h_s,$$

où $P_{v^\perp}$ est la projection sur l'espace orthogonal à $v$. Alors pour un tel $h$ quelconque, le point $(x_t^h, v_t^h)$ appartient au support de $u_t$.

Soit donc un point $(x, v)$ dont on veut montrer qu'il est dans le support de $u_t$. On suppose que $|x - x_0| < t$. Ainsi, le segment reliant $x_0$ à $x$ est de longueur strictement plus petite que $t$, et on peut modifier très légèrement cette courbe au voisinage de ses extrémités pour disposer d'une courbe lisse $\gamma : [0; \tau] \to \mathbb{R}^d$ paramétrée par longueur d'arc, telle que $\tau < t$, $(\gamma_0, \dot\gamma_0) = (x_0, v_0)$ et $(\gamma_\tau, \dot\gamma_\tau) = (x, v)$. Quitte à ajouter une partie oscillante (disons $s \mapsto \varepsilon\chi(s)\sin(s/\varepsilon^2)\epsilon_1$ pour un certain $\chi$ à support dans $(0; 1)$), on peut augmenter la longueur de $\gamma$ et trouver une nouvelle courbe $c$ vérifiant les mêmes hypothèses avec $\tau = t$. Puisque $\dot c$ est unitaire, on constate que

$$\mathrm{d}c_s = \dot c_s \mathrm{d}s, \qquad \mathrm{d}\dot c_s = P_{\dot c_s^\perp} \mathrm{d}\dot c_s.$$

Ainsi, en posant $h = \dot{c}$ et par unicité, on trouve $(x^h, v^h) = (c, \dot{c})$ et ainsi $(x, v) = (c_t, \dot{c}_t)$ appartient au support. Puisque celui-ci est fermé, $\{|x| \le t\} \subset \mathrm{Supp} u_t$, ce qui conclut.

## 5.2 Régularité

En notant $(\epsilon_1, \cdots, \epsilon_d)$ la base canonique de $\mathbb{R}^d$, on définit les champs de vecteurs $V_i : \mathbb{S}^{d-1} \to T\mathbb{S}^{d-1}$ par

$$V_i(v) = \epsilon_i - v^i x.$$

En d'autres termes, $V_i(v)$ est la projection orthogonale de $\epsilon_i$ sur le plan tangent à la sphère en $v$. Ces champs de vecteurs définissent canoniquement des champs de vecteurs sur $\mathbb{R}^d \times \mathbb{S}^{d-1}$ constants par rapport au premier facteur, que l'on appelle aussi $V_i$. On pose aussi le champ de vecteurs

$$V_0 : (x, v) \in \mathbb{R}^d \times \mathbb{S}^{d-1} \mapsto (-v, 0) \in T(\mathbb{R}^d \times \mathbb{S}^{d-1}).$$

Alors, en se basant par exemple sur l'expression (1.3) pour le laplacien sphérique, l'équation (5.5) se récrit

$$\partial_t u_t = \frac{1}{2} \sum_i V_i \cdot (V_i \cdot u_t) + V_0 \cdot u_t = \Big( \frac{1}{2} \sum_i V_i^2 + V_0 \Big) u_t,$$

où l'on note $A \cdot f := \mathrm{d}f(A)$ la dérivée de $f$ dans la direction $A$ dans la première forme, et on identifie $A$ à l'opérateur $f \mapsto A \cdot f$ dans la seconde.

On a décomposé l'opérateur $L^*$ en deux termes, une somme de carrés de champs de vecteurs et un champ de vecteurs. On dit que $L^*$ est mis sous la forme d'Hörmander — on pourrait aussi y ajouter un terme d'ordre 0. Sous cette forme, L. Hörmander a développé une théorie pour décrire la régularité des solutions d'équations de ce type, à partir de la géométrie des champs de vecteurs $V_0, \cdots, V_d$. On n'en utilisera qu'une fraction, plus particulièrement les concepts importants pour nous de diffusions elliptique, sous-elliptique et hypoelliptique, ordonnés par stricte inclusion. Les deux premiers serviront de modèle, ou du moins d'inspiration, le dernier s'applique au mouvement brownien cinétique.

On considère donc, sur une variété $M$ de dimension $n$, un opérateur $H$ donné sous la forme d'Hörmander

$$H = \frac{1}{2} \sum_{1 \le i \le d} A_i^2 + A_0 + a, \tag{5.6}$$

pour des champs de vecteurs lisses $A_0, \cdots, A_m : M \to TM$ et une fonction lisse $a : M \to \mathbb{R}$. On appelle $\rho_t$, lorsqu'elle existe,[18] une solution de l'équation

$$\partial_t \rho_t = H \rho_t \tag{5.7}$$

au sens où $\rho$ est une distribution sur $\mathbb{R} \times M$ à support dans $\mathbb{R}_+ \times M$, et $\partial_t \rho - H_x \rho$ est à support dans $\{0\} \times M$. On s'intéresse au cas où la condition initiale est, en un sens, une distribution de Dirac, mais les résultats de régularité s'appliquent pour n'importe quelle condition initiale. Le cas du mouvement brownien cinétique consiste à choisir $n = 2d - 1$, $M = \mathbb{R}^d \times \mathbb{S}^{d-1}$, $A_i = V_i$ et $a = 0$. On énonce une version simple du théorème de régularité d'Hörmander. Celui-ci en a par exemple donné une preuve dans [Hör85, Theorem 22.2.1]

**Théorème 5.1.** *Soient $H$ un opérateur donné sous la forme d'Hörmander (5.6), et $\rho_t$ satisfaisant l'équation de diffusion (5.7) associée à $H$. Alors si $H$ est hypoelliptique (par exemple s'il est elliptique ou sous-elliptique), $\rho_t$ coïncide avec une fonction lisse pour tout $t > 0$, et plus généralement $\rho$ coïncide avec une fonction lisse sur $\mathbb{R}_+^* \times M$.*

[18] Le cas qui nous intéresse est bien sûr celui où $\rho_t$ est la loi au temps $t$ d'une diffusion ; alors l'existence d'une solution à valeurs mesure est donnée, avec éventuellement une perte de masse s'il y a explosion en temps fini avec probabilité strictement positive.

Dans le reste de cette section, on définit l'hypoellipticité au sens du théorème ci-dessus à travers différents exemples de complexité croissante.

**Ellipticité.** L'opérateur $H$ est dit elliptique si pour tout $x \in M$, l'espace vectoriel engendré par les vecteurs $A_1(x), \cdots, A_d(x)$ est l'espace tangent $T_xM$ tout entier. En particulier, on doit avoir $n \leq d$, ce qui n'est pas le cas du mouvement brownien cinétique. L'exemple fondamental d'opérateur elliptique est celui du laplacien dans l'espace euclidien $M = \mathbb{R}^n = \mathbb{R}^d$, associé au mouvement brownien. En effet, il est clair que pour le laplacien $H = \Delta/2$ des probabilistes,

$$\frac{1}{2}\Delta = \frac{1}{2}\sum_i \epsilon_i^2,$$

pour $\epsilon_1, \cdots, \epsilon_d$ les champs de vecteurs constants correspondants à la base canonique de $\mathbb{R}^d$. La solution de (5.7) avec condition initiale $x_0 \in \mathbb{R}^d$ est bien sûr explicite :

$$\rho_t(x) = \frac{1}{\sqrt{2\pi t}^d} \exp\left(-\frac{|x - x_0|^2}{2t}\right).$$

On note qu'en un sens, la régularisation de la condition initiale autour de $x_0$ se fait sur une échelle d'ordre $\sqrt{t}$.

Un autre exemple important, et qui nous servira de référence, est celui du laplacien sur une variété riemannienne, décrivant donc le mouvement brownien associé à la métrique. Pour les besoins de l'exposition, on se place sur $M = \mathbb{R}^n$, muni d'une métrique lisse $g$ uniformément bornée inférieurement et à dérivées premières bornées.[19] Alors le mouvement brownien $(B_t)_{t\geq 0}$ issu de $x_0$ est bien défini pour tout $x_0 \in \mathbb{R}^n$. Soit $(A_1, \cdots, A_d)$ un famille de $d = n$ champs de vecteurs tels que $A_1(x), \cdots, A_d(x)$ est une base $g$-orthonormale de $\mathbb{R}^n$ en tout point, construits par exemple point par point par l'algorithme de Gram-Schmidt à partir de la base canonique $(\epsilon_1, \cdots, \epsilon_d)$. Alors le générateur du brownien $B$ peut être mis sous la forme d'Hörmander de façon à ce que les carrés qui y apparaissent soient les champs $A_i^2$, et son adjoint est de la forme d'Hörmander avec les mêmes carrés. Ce dernier est donc automatiquement elliptique, et la densité $\rho_t$ du mouvement brownien issu de $x_0$ est lisse d'après le théorème de Hörmander. On peut montrer qu'elle satisfait, en un sens à définir, l'approximation

$$\rho_t(x) \simeq \frac{1}{\sqrt{\det g_{x_0}} \cdot \sqrt{2\pi t}^d} \exp\left(-\frac{d_g(x, x_0)^2}{2t}\right)$$

où $d_g$ est la distance riemannienne associée à $g$. Une fois encore, on constate que la régularisation s'effectue sur une échelle spatiale de l'ordre de $\sqrt{t}$.

**Sous-ellipticité.** L'hypothèse d'ellipticité peut être reformulée en ces termes. L'application $M \times \mathbb{R}^d$ qui associe à $\big(x_0, (s_1, \cdots, s_d)\big)$ l'image $\phi^d_{s_d} \circ \cdots \circ \phi^1_{s_1}(x_0)$ de $x_0$ sous l'action des flots $\phi^i$ engendrés par les champs de vecteurs $A_i$ est une submersion aux points $(x_0, 0)$. Autrement dit, informellement, on peut atteindre tout point infiniment proche de $x_0$ en suivant chaque champ de vecteurs $A_i$ pendant un temps infinitésimal bien choisi. Notons que le résultat ne dépend pas de l'ordre dans lequel on applique les flots, puisqu'à l'ordre 1, la composition de flots correspond à l'addition de champs de vecteurs : dans une carte et à $x_0$ fixé,

$$\phi^d_{s_d} \circ \cdots \circ \phi^1_{s_1}(x_0) = x_0 + s_1 A_1(x_0) + \cdots + s_d A_d(x) + O(|s|). \tag{5.8}$$

[19] Il existe $\varepsilon > 0$ tel que $g_x(v, v) \geq \varepsilon |v|^2_{\text{eucl}}$ pour tout $x, v \in \mathbb{R}^d$.

Il est alors naturel de considérer les points $x$ que l'on peut atteindre à partir de $x_0$ sous l'action de ces champs de vecteurs. On aboutit au problème de contrôle suivant : trouver, pour un tel $x_0$ fixé, une courbe raisonnablement lisse $h : [0; t] \to \mathbb{R}^d$ telle que la solution $X^h$ de l'équation différentielle contrôlée

$$\mathrm{d}X^h_s = \sum_i A_i(X^h_s)\mathrm{d}h^i_s, \qquad X^h_0 = x_0$$

vérifie $X^h_t = x$. D'après, par exemple, le théorème d'inversion locale, il est clair que l'hypothèse d'ellipticité implique que les lignes droites permettent d'atteindre tout point $x$ suffisamment proches de $x_0$.

Dans l'expression (5.8), on n'a considéré que le développement à l'ordre 1 des flots en question, ce qui suffit pour montrer un résultat de surjectivité dans le cas elliptique. En termes de contrôle, il a suffit de considérer un sous-espace de dimension $d$ pour $h$. Dans le cas où ce n'est pas possible, il est naturel de tenter d'exploiter les termes d'ordre 2. Or, il est bien connu que le défaut de commutativité des flots $\phi$ et $\psi$ associés aux champs de vecteurs $A$ et $B$, qui comme on l'a dit est nul à l'ordre 1, s'exprime à l'ordre 2 sous la forme

$$\psi_u^{-1} \circ \phi_s^{-1} \circ \psi_u \circ \phi_s(x_0) = su[A, B](x_0) + O\big((|s| + |u|)^3\big)$$

dans une carte autour de $x_0$, où $[A, B] := (A^i\partial_i B^j - B^i\partial_i A^j)_{1\le j\le n}$ est le crochet de Lie. On s'attend donc à ce que, quitte à choisir un contrôle $h$ de dérivée plus grande pour renforcer les contributions d'ordre 2, on puisse atteindre un voisinage de $x_0$ à la condition que les champs $A_i$ et $[A_i, A_j]$ engendrent tout l'espace tangent. Il n'y a d'ailleurs pas de raison de s'arrêter au premier crochet de Lie.

Soit $E$ le plus petit sous-espace vectoriel de l'ensemble des champs de vecteurs lisses sur $M$ tel que

- $E$ contient les champs $A_1, \cdots, A_d$ ;
- $E$ est stable par multiplication par une fonction lisse ;
- $E$ est stable par crochet de Lie.

On dit que $E$ est la $\mathcal{C}^\infty(M)$-algèbre de Lie engendrée par $(A_1, \cdots, A_d)$. Alors l'ensemble $E_x := \{A(x), A \in E\}$ est un sous-espace vectoriel de $T_xM$ pour tout $x \in M$. On dit que $A_1, \ldots, A_d$ satisfont la condition d'Hörmander si $E_x = T_xM$ pour tout $x \in M$, et dans ce cas $H$ est dit sous-elliptique. On peut montrer que cette condition est suffisante pour que l'application $h \mapsto X^h_t$ définie plus haut, restreinte aux courbes lisses, soit surjective sur un voisinage de $x$.

On peut déjà remarquer que le mouvement brownien cinétique est un exemple archétypique de diffusion qui ne peut pas être sous-elliptique. Si $x \in \mathbb{R}^d$, alors la sous-variété $\{x\} \times \mathbb{S}^{d-1}$ est une sous-variété intégrale pour les champs de vecteurs $V_1, \cdots, V_d$, c'est-à-dire que tous ces champs de vecteurs sont tangents à cette sphère. En particulier, tous les crochets de Lie de $V_1, \cdots, V_d$, aussi itérés qu'ils soient, seront eux aussi tangents à la sous-variété, et ainsi $E_{(x,v)}$ ne contient que $\{0\} \times T\mathbb{S}^{d-1}$.

Notre guide pour les équations sous-elliptiques est l'équation de la chaleur sur une variété sous-riemannienne $M$ munie d'un volume lisse. Informellement, une variété sous-riemannienne est une variété riemannienne dans laquelle il est interdit de se déplacer dans certaines directions. Par exemple, l'ensemble des configuration pour une voiture dans le plan, données par une position, une orientation et un angle au volant, est une variété de dimension 4, mais on ne peut modifier une telle configuration que selon deux dimensions : la position, selon le cercle déterminé par l'orientation et l'angle au volant, et cet angle lui-même. Pourtant, on peut atteindre n'importe quelle configuration proche d'une configuration initiale donnée, si l'on sait manœuvrer avec suffisamment d'habileté. C'est une condition que l'on impose aussi aux variétés sous-riemannienne : on suppose par définition l'ensemble des champs de vecteurs admissibles vérifie la condition d'Hörmander. Si l'on associe à chaque action infinitésimale un coût quadratique, on obtient une structure de variété sous-riemannienne.[20]

Il existe sur une variété sous-riemannienne une notion de distance naturelle : la distance $d(x,y)$ entre $x$ et $y$ est l'infimum des longueurs de courbes admissibles joignant $x$ à $y$. On voit que l'existence d'une telle courbe pour $y$ proche de $x$ est équivalente à l'existence d'une solution pour le problème de contrôle décrit plus haut ; ainsi, la distance $d(x,y)$ est toujours finie lorsque $M$ est connexe. De même que pour le cas des variétés riemanniennes, il existe un laplacien naturel $H$ sur une variété sous-riemannienne munie d'un volume, et on peut considérer l'équation de diffusion associée (5.7) de condition initiale $x_0$. Les champs de vecteurs apparaissant dans la partie d'ordre 2 de $H$ génèrent l'espace des vecteurs admissibles, et satisfont ainsi la condition d'Hörmander, ce qui garantit que $\rho_t$ est lisse pour tout temps $t > 0$. D'après un résultat de G. Ben Arous, on a pour un $x$ fixé générique[21] la représentation

$$\rho_t(x) = \frac{C_{x,x_0}}{\sqrt{t}^d} \exp\left(-\frac{d(x,x_0)^2}{2t}\right)\left(1 + O(t)\right)$$

pour une certaine constante $C_{x,x_0} > 0$ indépendante de $t$.

Cette fois-ci, l'échelle de la régularisation est moins claire. Comme on l'a vu heuristiquement, on peut atteindre les points proches de $x_0$ dans la direction engendrée par les champs de vecteurs admissibles en développant leur action au premier ordre. Disons que dans une carte, les points distants de $\varepsilon > 0$ dans cette direction sont atteint en suivant certain flots pendant une distance $\varepsilon$. En ce qui concerne les points atteints par un crochet de Lie, il faut pour les atteindre aller à l'ordre deux, donc suivre le flot pendant une distance $s$ telle que $s^2 \simeq \varepsilon$ : c'est une longueur beaucoup plus importante, de l'ordre de $\sqrt{\varepsilon}$. De même, un point situé dans la direction d'un crochet de Lie de la forme $[A_i,[A_j,A_k]]$ est atteint en un temps $\varepsilon^{1/3}$, et ainsi de suite. La longueur caractéristique pour la régularisation dépend donc des espaces selon lesquels on l'étudie, et est de l'ordre de $\sqrt{t}^\ell$, où $\ell$ est la longueur minimale du crochet requis pour atteindre le point $x$.

**Hypoellipticité.** La condition d'hypoellipticité que l'on discute à présent dépend cette fois-ci de la partie d'ordre 1 de l'opérateur, et est plus faible que les deux précédentes notions. Soit $F \supset E$ le plus petit sous-espace vectoriel de l'ensemble des champs de vecteurs lisses sur $M$ tel que

- $F$ contient les champs $A_1, \cdots, A_d$ ;
- $F$ est stable par multiplication par une fonction lisse ;
- $F$ est stable par crochet de Lie.
- $F$ est stable par le crochet de Lie $[\cdot, A_0] : A \mapsto [A, A_0]$.

En termes algébriques, $F$ est l'idéal engendré par $(A_1, \cdots, A_d)$ dans la $\mathcal{C}^\infty(M)$-algèbre de Lie engendrée par $(A_0, \cdots, A_d)$. On dit que $(A_0, \cdots, A_d)$ vérifie la condition parabolique d'Hörmander si $F_x = T_xM$ pour tout $x \in M$, et dans ce cas, $H$ est dit hypoelliptique.

[20] Une dernière condition importante est que l'espace des directions admissibles à partir du point $x$ est de dimension indépendante de $x$.

[21] Le point $x$ ne doit pas appartenir au cut-locus de $x_0$.

La loi du mouvement brownien cinétique satisfait une équation hypoelliptique. En effet, on montre que le sous-espace de $T_xM$ engendré par la famille $\big([V_0, V_i](x,v)\big)_i$ contient les vecteurs $(w,0)$ pour tout $w$ orthogonal à $v$ ; en fait,

$$[V_0, V_i] = \partial_{x_i} - v^i \times (v,0).$$

Puisque l'on sait déjà que $V_1(x,v), \cdots, V_d(x,v)$ engendre $\{0\} \times T\mathbb{S}^{d-1}$, l'espace $F_{(x,v)}$ est de dimension au moins $2d-2$, et il suffit de montrer que $(v,0)$ est aussi dans cet espace pour conclure. Or, si $u \in \mathbb{R}^d$ est orthogonal à $v$,

$$\sum_{ij} u^i u^j \big[V_i, [V_0, V_j]\big](x,v) = |u|^2 \times (v,0),$$

et on vérifie donc la condition parabolique d'Hörmander.

À ma connaissance, il n'existe pas pour les diffusions hypoelliptiques de cas prototypique comme dans les cas elliptiques et sous-elliptiques. L'existence d'une distance $d(x,x_0)$ telle que l'on pourrait dire, par exemple,

$$\ln\big(u_t(x,x_0)\big) \simeq -\frac{d(x,x_0)^2}{2t}$$

en un certain sens, est loin d'être claire dans le cas général. D'ailleurs, on trouve plutôt dans la littérature l'opinion qu'une telle distance n'existe pas. Une possibilité pour restaurer ce genre d'équivalent serait d'autoriser la distance à dépendre du temps, comme le propose par exemple J. Franchi dans [Fra19], mais elle reste pour l'instant à l'état de conjecture. On pourra consulter [DM10] pour un cas où une expression satisfaisante peut être prouvée. On admet que les crochets $[V_i, V_j]$, pour $i,j > 0$, jouent un rôle différent dans l'analyse de celui des crochets $[V_0, V_i]$ pour $i > 0$ ; cependant, il est actuellement hors de portée de donner un cadre géométrique général, par exemple *via* un problème de contrôle, dans lequel les propriétés analytiques du noyau de la chaleur $\rho_t$ associé soient facilement comprises.

Dans la suite de cette section, je détaille les méthodes que j'ai mises en œuvre pour décrire certains aspects du noyau $u$ dans le cadre de la dimension $d = 2$. Il s'agit de deux approximations indépendantes du problème : l'une montre la convergence d'une série d'approximations vers le noyau $u$, tandis que la seconde considère le problème variationnel associé, dans le même esprit que la discussion autour du support donnée plus haut.

## 5.3 Méthode de la paramétrix

Dans la première partie du chapitre IV, on applique la méthode dite de la paramétrix à pour trouver une expression pour $u$ sous la forme d'un développement en série. Le cas du mouvement brownien cinétique étant relativement subtil, on illustre plutôt cette méthode par l'exemple de l'équation de la chaleur dans $\mathbb{R}^d$ muni d'une métrique $g$ lisse. Les premières considérations sont cependant valides pour l'une comme pour l'autre.

On fait sur $g$ l'hypothèse que ses dérivées d'ordre 1 sont uniformément bornées, et que $g_x(v,v) \geq \varepsilon |v|^2$ pour un certain $\varepsilon > 0$, uniformément en $x$ et $v$. L'objectif est de trouver une expression asymptotique de $\rho$, où $\rho_t(x_0, \cdot)$ est la distribution au temps $t$ du mouvement brownien $B$ associé à $g$ et issu de $x_0$. Les hypothèses sur $g$ assurent que $B$ n'explose pas en temps fini, et ainsi $\rho_t$ est une mesure de probabilité pour tout $t > 0$.

On interprète $\rho$ comme une distribution sur $\mathbb{R} \times \mathbb{R}^d$ à support dans $\mathbb{R}_+ \times \mathbb{R}^d$ ; elle agit sur les fonctions test, disons $\phi \in \mathcal{C}_0^\infty(M, \mathbb{R})$, par

$$\langle \rho(x_0, \cdot), \phi \rangle = \int_0^\infty \mathbb{E}_{x_0}[\phi(B_s)]\mathrm{d}s.$$

C'est ce que l'on appelle une solution fondamentale de l'équation de la chaleur : à $x_0 \in \mathbb{R}^d$ fixé, $\rho(x_0, \cdot)$ est une distribution sur $\mathbb{R} \times \mathbb{R}^d$ à support dans $\mathbb{R}_+ \times \mathbb{R}^d$, satisfaisant l'équation

$$(\partial_t - L)\rho(x_0, \cdot) := \Big(\partial_t - \frac{1}{2}\Delta_g^*\Big)\rho(x_0, \cdot) = \delta_{0,x_0}$$

au sens des distributions, où $\Delta_g$ est le laplacien associé à $g$ agissant selon la seconde variable, $\Delta_g^*$ son adjoint et $\delta_{0,x_0}$ la distribution de Dirac en $(t, x) = (0, x_0)$. D'après la partie précédente, et le théorème 5.1 en particulier, $(t, x) \mapsto \rho_t(x_0, x)$ est en fait une fonction lisse sur $\mathbb{R}_+^* \times M$. En effet, $\Delta_g^*$ est elliptique comme l'a vu dans la partie précédente, et comme on s'y attend pour un opérateur de type laplacien.

La solution fondamentale permet de résoudre l'équation correspondant à une condition initiale $f \in \mathrm{L}^1$

$$(\partial_t - L)\rho_t^f(x) = 0 \text{ pour tout } t > 0, x \in \mathbb{R}^d, \qquad \rho_0^f = f.$$

Une solution explicite est donnée par la convolution

$$\rho_t^f(x) = \int_{\mathbb{R}^d} f(x_0)\rho_t(x_0, x)\mathrm{d}x_0\,;$$

intuitivement, la solution $\rho^f$ est la somme des solutions $f(x_0)\rho(x_0, \cdot)$ associées aux conditions initiales $f(x_0)\delta_{x_0}$. Dans la suite, on note $P_t f$ la solution $\rho^f$, en pensant à $P_t$ comme un opérateur de convolution.

**Principe de Duhamel.** Une paramétrix $\widetilde{\rho}$ est par définition une approximation de $\rho$. On trouve dans la littérature plusieurs manières de quantifier ce que l'on appelle une approximation ; plus précisément, en écrivant

$$(\partial_t - L)\widetilde{\rho}(x_0, \cdot) = \delta_{0,x_0} - E(x_0, \cdot),$$

on peut faire différentes hypothèses sur le terme d'erreur $E(x_0, \cdot)$. Plus bas, on en donnera la forme dans les cas qui nous intéressent. La propriété qui nous intéresse est que $E$ soit suffisamment régulière pour pouvoir appliquer la démarche qui suit. Pour les équations aux dérivées partielles de ce type, qui sont associées à un semi-groupe $P_t$, on peut représenter $\widetilde{\rho}$ à partir de $E$. Sans se préoccuper du fait que les objets considérés sont bien définis, on peut écrire

$$\begin{aligned}
\widetilde{\rho}_t(x_0, x) &= P_t\widetilde{\rho}_0(x_0, x) + \int_0^t \frac{\mathrm{d}}{\mathrm{d}s}(P_{t-s}\widetilde{\rho}_s)(x_0, x)\mathrm{d}s \\
&= (P_t\delta_{0,x_0})(x_0, x) + \int_0^t \big(P_{t-s}(\partial_t - L)\widetilde{\rho}\big)_s(x_0, x)\mathrm{d}s \\
&= \rho_t(x_0, x) + \int_0^t (P_{t-s}E)_s(x_0, x)\mathrm{d}s \\
&= \rho_t(x_0, x) - \int_0^t \int_{\mathbb{R}^d} E_s(x_0, y)\rho_{t-s}(y, x)\mathrm{d}y\mathrm{d}s.
\end{aligned}$$

Cette représentation est appelée principe de Duhamel. Elle est par exemple valide en faisant le choix de paramétrix

$$\widetilde{\rho}_t(x_0, x) := \frac{1}{\sqrt{\det g_{x_0}} \cdot \sqrt{2\pi t}^d} \exp\left(-\frac{g_x(x - x_0, x - x_0)}{2t}\right),$$

c'est-à-dire la solution de l'équation à coefficients constants, gelés en $x_0$.

En posant $\mathcal{E} : v \mapsto \mathcal{E}v$ l'opérateur de convolution par $E$ défini par

$$(\mathcal{E}v)_t(x_0, x) = \int_0^t \int_{\mathbb{R}^d} E_s(x_0, y) v_{t-s}(y, x) \mathrm{d}y \mathrm{d}s,$$

le principe de Duhamel s'écrit $\widetilde{\rho} = (\mathrm{id} - \mathcal{E})\rho$. Formellement, on exprime $\rho$ en fonction de $\widetilde{\rho}$ par

$$\rho = \widetilde{\rho} + \mathcal{E}\widetilde{\rho} + \mathcal{E}^2\widetilde{\rho} + \cdots$$

La difficulté de la méthode de la paramétrix consiste à montrer la convergence de la série ci-dessus. Notons au passage qu'une interprétation probabiliste de la méthode la paramétrix est donnée dans [BKH15] par V. Bally et A. Kohatsu-Higa, en reformulant les convolutions par des compositions d'accroissements de chaînes de Markov.

Dans la section 5.2, on a donné une expression approchée de $\rho$. En fait, la méthode de la paramétrix donne une expression de la forme

$$\rho_t(x_0, x) = \widetilde{\rho}_t(x_0, x)\big(1 + O(t)\big) + O\big(\exp(-1/Ct)\big).$$

La stratégie que l'on peut mettre en œuvre dans ce cas est la suivante. On constate que $\widetilde{\rho}_t$ se met sous la forme $t^{-d/2}\check{\rho}\big(t, x_0, (x - x_0)/\sqrt{t}\big)$, où $\check{\rho}$ est une fonction lisse sur $\mathbb{R} \times (\mathbb{R}^d)^2$, exponentiellement décroissante par rapport à sa troisième variable. Un calcul direct montre que l'erreur $E$ commise est de la forme $t^{-(d-1)/2}\check{E}$, où $\check{E}$ est aussi une fonction lisse des même paramètres, et décroît aussi exponentiellement. Ceci vient de la forme de la paramétrix : *a priori*, les opérateurs $\partial_t$ et $L$ devraient créer une singularité en $1/t$, or $E$ ne voit apparaître qu'une singularité en $1/\sqrt{t}$. En fait, en termes de développement en puissances de $\sqrt{t}$, le terme principal est nul, et le terme suivant est plus régulier. On peut montrer que ce gain d'un facteur $\sqrt{t}$ se propage lors des convolution : une fonction de la forme $t^{-d/2+a}\check{A}$ avec $a > 0$, une fois convolée par $E$, diminue sa singularité à une puissance $t^{-(d-1)/2+a}$. La singularité en vient à s'effacer, et les convolutions d'ordre élevé s'annulent de plus en plus fortement au voisinage de $t = 0$. On peut alors montrer la convergence de la série conjecturée plus haut, et en déduire des estimées du type donné ci-dessus.

La difficulté du cas hypoelliptique strict est alors double. Tout d'abord, il convient de trouver une paramétrix adaptée, de façon à ce que, en un sens, les termes principaux dans l'expression de l'erreur s'annulent. Cela se fait en choisissant un certain $\widetilde{u}$ solution de l'équation dont les coefficients sont approchés à l'ordre 2 — à comparer, donc, avec l'approximation à l'ordre 0 faite dans le cas elliptique. Dans la partie IV.1, on définit $\widetilde{u}$ par l'équation (IV.1.2), et on étudie ses propriétés dans la section IV.1.5. Cela permet de montrer effectivement que l'erreur $E$ est en un sens petite.

Ensuite, il faut comprendre si, et comment, la régularité se propage lorsque l'on effectue des convolutions. Le fait de pouvoir isoler les termes $(x - x_0)/\sqrt{t}$ et absorber ainsi une partie de la singularité dans cette renormalisation est intrinsèquement lié à l'ellipticité de l'équation. En effet, on ne voit pas apparaître la stratification construite par l'application successive des crochets de Lie, et une longueur caractéristique de l'ordre de $\sqrt{t}$ dans toutes les directions n'est pas ce que l'on attend. Trouver la bonne description des échelles, et en déduire une description adaptée des quantités en jeu, est au centre de notre approche. De la même façon que la régularité des fonctions apparaissant dans le problème de la chaleur était contrôlée par la puissance de $t$, la définition IV.1.4 donne une collection d'espaces de fonctions indexée par un paramètre de régularité $a$. Le fait que la convolution propage la régularité est alors une conséquence du théorème IV.1.6.

## 5.4 Problème variationnel

Dans la partie IV.2, on trouvera une approche variationnelle pour l'étude du noyau $u$, inspirée de la méthode WKB et de l'analyse semiclassique. Bien qu'elle ait été abordée avec l'ambition d'appliquer la méthode de la paramétrix, cette tentative échoue, mais met en lumière les difficultés de la tâche, et la géométrie du problème sous-jacent.

Le problème variationnel est précisément celui décrit dans l'étude du support faite ci-dessus dans la section 5.1. En dimension 2, on peut le reformuler, et chercher les minimiseurs $h$ d'une certaine énergie $\frac{1}{2}|h|^2$ parmi les courbes telles que la solution $(x^h, y^h, \phi^h)$ de l'équation

$$\mathrm{d}x^h_s = \cos(\phi^h_s)\mathrm{d}s, \qquad \mathrm{d}y^h_s = \sin(\phi^h_s)\mathrm{d}s, \qquad \mathrm{d}\phi^h_s = \mathrm{d}h_s$$

atteigne un certain point $(x, y, \phi)$ fixé en un temps $t$, partant d'un certain point $(x_0, y_0, \phi_0)$ fixé. Cette énergie n'est autre que la norme $\mathrm{H}^1$ :

$$|h|^2 := \int_0^t |\dot{h}_s|^2 \mathrm{d}s.$$

Comme on l'a vu lors de l'étude du support, ces contrôles $h$ et les solutions $(x^h, y^h, \phi^h)$ qu'ils engendrent sont liés à celle du noyau $u$. Dans notre cas, ce lien est, à ma connaissance, encore heuristique. On ne discute pas ici des connections conjecturées, mais on trouvera dans l'introduction de la partie IV.2 un peu plus de contexte, et en particulier le lien avec l'analyse classique et la méthode WKB introduite pour l'étude de l'équation de Schrödinger. Une des difficultés que l'on rencontre pour rendre plus concrètes ces stratégies est que le problème hamiltonien associé est singulier dans le sens suivant.

Dans le cas de ce problème de minimisation, on peut trouver un lagrangien $L$ tel que

$$\frac{1}{2}|h|^2 = \int_0^t L(X^h_s, \dot{X}^h_s)\mathrm{d}s,$$

où l'on a posé $X = (x, y, \phi)$. En effet, il suffit de définir $L(X, \dot{X}) = \infty$ pour $\dot{x} \neq \cos\phi$ ou $\dot{y} \neq \sin\phi$, et $\frac{1}{2}|\dot{\phi}|^2$ dans le cas contraire. Ce lagrangien admet un hamiltonien associé $H$ défini par

$$H(X, P) := \sup_V (P \cdot V - L(X, V)).$$

Les équations de Hamilton qui en découlent, c'est-à-dire l'équation différentielle ordinaire

$$\dot{X}_t = \frac{\partial H}{\partial P}(X_t, P_t), \qquad \dot{P}_t = \frac{\partial H}{\partial X}(X_t, P_t)$$

sur $(\mathbb{R}^3)^2$, sont dans ce cas

$$\begin{aligned} \dot{x}_t &= \cos(\phi_t) & \dot{p}_t &= 0 \\ \dot{y}_t &= \sin(\phi_t) & \dot{q}_t &= 0 \\ \dot{\phi}_t &= \psi_t & \dot{\psi}_t &= \sin(\phi_t)p_t - \cos(\phi_t)q_t, \end{aligned}$$

où l'on a noté $P = (p, q, \psi)$.

Notons $(X_t(X_0, P_0), P_t(X_0, P_0))$ les solutions des équations de Hamilton de condition initiale $(X_0, P_0)$. Pour certains hamiltoniens $H$, les minimiseurs de l'action sont donnés, pour $X$ proche de $X_t(X_0, 0)$, par les courbes $s \mapsto X_s(X_0, P_0)$, que l'on appelle les caractéristiques. Pour que cela soit utilisable en pratique, il faut pouvoir atteindre les points proches de $X_t(X_0, 0)$ par des images $X_t(X_0, P_0)$. Une façon automatique de vérifier cette condition est de montrer que $P_0 \mapsto X_t(X_0, P_0)$ est un difféomorphisme pour $P_0$ proche de 0. Ce cas n'est pas rare, par exemple le problème de contrôle associé à une diffusion elliptique vérifie cette condition. Alors des arguments classiques montrent que les minimiseurs de l'action de régularité suffisante sont en fait les caractéristiques du hamiltonien $H$.

Cependant, on peut constater que pour notre hamiltonien et pour tout $p_0 \in \mathbb{R}$, les courbes

$$s \mapsto (x_0 + s \cos \phi_0, y_0 + s \sin \phi_0, \phi_0; p_0 \cos \phi_0, p_0 \sin \phi_0, 0)$$

sont solutions des équations de Hamilton. Or, toutes ces courbes atteignent le même point en temps $t$ ; $X_t$ n'a donc aucune chance d'être un difféomorphisme local, ni d'ailleurs localement surjective, puisque $X_t$ est lisse.

Le résultat le plus central de mon travail sur ce problème est la description de l'image d'un voisinage de 0 par l'application $P_0 \mapsto X_t(X_0, P_0)$. Entre autres, on peut montrer en un sens que l'obstruction à l'injectivité décrite ci-dessus est la seule, au moins localement ; ceci est fait dans les propositions IV.2.3 et IV.2.5. À partir de cette première avancée, je montre que les minimiseurs de l'action, sous la contrainte d'une régularité suffisante, sont en fait effectivement les caractéristiques associées à $H$, ce qui correspond au théorème IV.2.6.

# Chapter II

# Homogenisation for anisotropic kinetic random motions

This chapter is based on the article [Per18] of the author, which was sent to arXiv.org in November 2018. This work is currently in revision for the Electronic Journal of Probability.

## 1 Introduction

We consider a class of anisotropic and kinetic random motions on the unit tangent space of a general Riemannian manifold $(\mathcal{M}, g)$ of dimension $d \geq 2$. In the simplest case when the base manifold is the Euclidean space $\mathbb{R}^d$, the typical process we have in mind can be described as follows: let $\sigma > 0$ be a positive parameter and let $(B_t)_{t\geq 0}$ be a Brownian motion in $\mathbb{R}^d$ with (non identity) covariance matrix $\Sigma = A^*A$. We construct an anisotropic diffusion process $(v_t)_{t\geq 0} = (v_t^\sigma)_{t\geq 0}$ on the Euclidean sphere $\mathbb{S}^{d-1} \subset \mathbb{R}^d$ by solving the Stratonovich differential equation

$$\mathrm{d}v_t = \sigma \Pi_{v_t^\perp} \circ \mathrm{d}B_t, \tag{1.1}$$

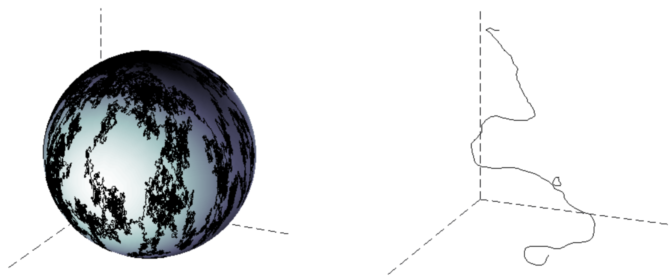
where $\Pi_{v_t^\perp}$ denotes the projection on the orthogonal of $v_t$. We then integrate the velocity process $(v_t)_{t\geq 0}$ to obtain a process $(x_t)_{t\geq 0} = (x_t^\sigma)_{t\geq 0}$ with values in $\mathbb{R}^d$

$$x_t := x_0 + \int_0^t v_s \mathrm{d}s. \tag{1.2}$$

The process $(x_t, v_t)_{t\geq 0}$ is thus a diffusion process with values in the unit tangent space $T^1\mathbb{R}^d = \mathbb{R}^d \times \mathbb{S}^{d-1}$. The first projection $(x_t)_{t\geq 0}$ is a $\mathcal{C}^1$ curve in $\mathbb{R}^d$, which inherits the anisotropy of the velocity process $(v_t)_{t\geq 0}$, and the positive parameter $\sigma$ allows one to slow or speed up the clock of the latter. The next figure shows an approximation of a sample path of the resulting process.

On a general Riemannian manifold $(\mathcal{M}, g)$, an analogue process $(x_t, v_t)_{t\geq 0}$ with values in the unit tangent bundle $T^1\mathcal{M}$ can be constructed starting from the above Euclidean process and using the classical stochastic development/parallel transport machinery. Namely, the process $(x_t, v_t)_{t\geq 0}$ in $T^1\mathcal{M}$ is characterised by the fact that the image of $v_t \in T^1_{x_t}\mathcal{M}$ in the fixed unit tangent space $T^1_{x_0}\mathcal{M} \simeq \mathbb{S}^{d-1}$ by the inverse stochastic parallel transport along $(x_s)_{0\leq s\leq t}$ solves equation (1.1) above.

The isotropic analogue of the process, i.e. the process associated with $\Sigma = \mathrm{Id}$, was introduced in [ABT15] under the name kinetic Brownian motion, where its was shown that as the parameter



*On the left, the velocity* $(v_t)_{0\leq t\leq 10}$ *with values on* $\mathbb{S}^2$*; on the right, the corresponding position* $(x_t)_{0\leq t\leq 10}$ *with values in* $\mathbb{R}^3$*. The chosen covariance matrix is* $\Sigma = \mathrm{diag}(1, 1.1, 1.2)$.

FIGURE 1.1 – Sample path of the velocity (left) and position (right) processes.

$\sigma$ goes from zero to infinity, then the sample paths of the process $(x_{\sigma^2 t})_{t\geq 0}$ interpolates in a precise sense between geodesics and Brownian paths on the based manifold $\mathcal{M}$. For a fixed intensity parameter $\sigma$, the Poisson boundary of the process was also fully determined if the base manifold is rotationally invariant.

The motivation to introduce anisotropy in this context is twofold. From an applied point of view, the kinetic Brownian motion is a simple, yet very reasonable model for the dynamics of a mesoscopic spherical particle with bounded velocity in an isotropic heat bath. Compared to the standard Langevin dynamics where the velocities are Gaussian, the fact that the velocities are here of unit norm is perfectly consistent with special relativity theory. The homogenisation phenomena shown in [ABT15] illustrates the fact that the scaling limit of the process, i.e. the macroscopic behaviour of the particle is nevertheless diffusive, as anticipated. Now, if the geometry of the mesoscopic particle under consideration is not spherical, or if the heat bath is anisotropic, the dynamics of the velocity process has to be anisotropic, see e.g. [HBR13, CP03, Kam88] and the references therein. In that context, the velocity evolution given by the stochastic differential equation (1.1) with $\Sigma \neq \mathrm{Id}$ is very natural. As we will see below and with this applied point of view, the main result of this article guarantees that the macroscopic behaviour of the particle is still diffusive, with an explicit anisotropy matrix.

From a more theoretical point of view, the introduction of anisotropy is also unavoidable if one wants to generalise the results of [ABT15] to an infinite dimensional setting, say to an infinite dimensional Hilbert space. Indeed, doing so, one quickly faces the problem of defining spherical Brownian motion in this context. Looking at equation (1.1), the orthogonal projection makes perfect sense in a Hilbert setting but we have to give meaning to the driving Brownian motion $B$. This can naturally be done using the notion of abstract Wiener space, see e.g. [Gro67, Gro70] or [Str93, Chapter 8]. Roughly speaking, in that framework the driving process in (1.1) has to belong to the image of a radonifying injection, hence introducing a Hilbert-Schmidt covariance operator. In a finite dimensional setting, the action of this Hilbert-Schmidt operator amounts to replacing the standard Brownian motion $B$ by a Brownian motion with covariance $\Sigma \neq \mathrm{Id}$, i.e. to replace the isotropic noise driving kinetic Brownian motion by an anisotropic noise; this justifies our choice of dynamics for the velocity process.

Our goal in this paper is to exhibit the asymptotics of the time rescaled process $(x^\sigma_{\sigma^2 t}, v^\sigma_{\sigma^2 t})_{t\geq 0}$ as the intensity parameter $\sigma$ goes to infinity. More precisely, we show that in both Euclidean and Riemannian contexts, its first projection converges to an anisotropic Brownian motion.

The presence of anisotropy drastically complexifies the approach and computations compared to the isotropic framework. Namely, in the isotropic Euclidean setting considered in Section 2.2 of [ABT15] and which is the core of the proof when associated with rough paths techniques, the homogenisation of kinetic Brownian motion was proved using Itô calculus and standard martingale techniques. As it will be clear in Section 2 below, the Doob–Meyer decomposition of the velocity process given by equation (1.1) gets more involved here, its invariant measure is not likely to be easy to describe, and martingale techniques need explicit solutions of the Poisson equation which seems hopeless in this context. In fact, guessing a formula for the invariant measure of the $v^\sigma$ on the sphere before reading the statement of Proposition 1.1 does not seem obvious.

For this reason, we adopt a different approach and point of view here. Our proof of homogenisation for the time rescaled version of the process $(x_t^\sigma, v_t^\sigma)_{t\geq 0}$ is indeed essentially based on quantitative mixing properties of the velocity process. We show in particular that

**Proposition 1.1** (Lemma 2.1 and Proposition 2.4 below)**.** *The process $v_t^\sigma$ solution of (1.1) is ergodic in $\mathbb{S}^{d-1}$ with an explicit invariant measure $\mu$ whose density with respect to the uniform measure $\mathrm{d}\theta$ on the sphere is given by*

$$\frac{\mathrm{d}\mu}{\mathrm{d}\theta}(\theta) = \frac{\|A^{-1}\theta\|^{1-d}}{\int_{\mathbb{S}^{d-1}} \|A^{-1}\theta\|^{1-d}\mathrm{d}\theta}.$$

*In particular, the invariant measure $\mu$ as well as the trajectories are invariant under all the coordinate reflections*

$$(\theta^1, \cdots, \theta^i, \cdots, \theta^d) \mapsto (\theta^1, \cdots, -\theta^i, \cdots, \theta^d), \quad 1 \leq i \leq d. \tag{1.3}$$

*Moreover, there exists a positive constant $\tau$ such that, if $\mathcal{F}_{[a,b]}$ denotes the $\sigma$-algebra generated by the unit speed ($\sigma = 1$) velocity process $v_t$, for $a \leq t < b$, then for any $0 \leq s < t$ and any bounded measurable real-valued random variables $P$ and $F$ that are $\mathcal{F}_{[0,s]}$ and $\mathcal{F}_{[t,\infty]}$−measurable, respectively, we have*

$$\big|\mathbb{E}_\mu[PF] - \mathbb{E}_\mu[P]\,\mathbb{E}_\mu[F]\big| \lesssim |P|_\infty |F|_\infty\, \mathsf{e}^{-(t-s)/\tau}. \tag{1.4}$$

The above strong mixing and symmetry properties of the velocity process are the key ingredients to establish the homogenisation of the anisotropic version of kinetic Brownian motion in the Euclidean setting. Indeed, we have the following result.

**Theorem 1.2** (Theorem 3.6 below)**.** *Let $(x_t^\sigma, v_t^\sigma)_{t\geq 0}$ with values in $T^1\mathbb{R}^d$ be the solution of equation (1.1) and (1.2), starting from $(x_0, v_0)$ where $x_0$ is fixed and $v_0$ chosen at random according to $\mu$. Then as $\sigma$ goes to infinity, the time rescaled process $(x_{\sigma^2 t}^\sigma)_{t\in[0,1]}$ converges in law to a Brownian motion in the Euclidean space $\mathbb{R}^d$, with covariance matrix $\mathrm{diag}(\gamma_1, \cdots, \gamma_d)$ where*

$$\gamma_i := 2\int_0^{+\infty} \mathbb{E}_\mu[v_0^i v_t^i]\mathrm{d}t, \quad 1 \leq i \leq d.$$

Our strategy of proof consists in establishing that the rough path lift of $(x_{\sigma^2 t}^\sigma)_{t\geq 0}$ converges to the Stratonovich rough path lift of a Brownian motion with the above covariance. To do so, we use again the strong mixing properties of the velocity process, associated with a Lamperti-type criterion to ensure the tightness of the lift in rough path topology — see Lemmas 3.1 and 3.2 below. We then identify the limit process by showing that it has to be a stationary process with independent Gaussian increments on the nilpotent group associated with the rough path structure, see Theorem 3.6.

Using the fact that the notion of stochastic development amounts to solving a stochastic differential equation and that the Itô map is continuous with respect to the rough paths topology, one can conclude that the previous Euclidean statement actually holds on a general Riemannian manifold. Anisotropic Brownian motion on $\mathcal{M}$ is defined as the stochastic development of an anisotropic Brownian motion in $T_{x_0}\mathcal{M}$.

**Theorem 1.3.** *Let $(\mathcal{M}, g)$ be a complete and stochastically complete Riemannian manifold and let $(x_t^\sigma, v_t^\sigma)_{t\geq 0}$ be the process with values in $T^1\mathcal{M}$ characterised by the fact that the image of $v_t \in T^1_{x_t}\mathcal{M}$ in the fixed unit tangent space $T^1_{x_0}\mathcal{M} \simeq \mathbb{S}^{d-1}$ by the inverse stochastic parallel transport along $(x_s)_{0\leq s\leq t}$ solves equation (1.1) in $T^1_{x_0}\mathcal{M}$. Then as $\sigma$ goes to infinity, the time rescaled process $(x^\sigma_{\sigma^2 t})_{t\in[0,1]}$ converges in law to an anisotropic Brownian motion on the base manifold $\mathcal{M}$.*

As it will be clear from the proof of Theorem 1.2, the homogenisation phenomenon holds as soon as the mixing properties of the velocity process and the symmetry of the trajectories described in Proposition 1.1 hold. In other words, the conclusion of Theorem 1.3 is valid as soon as the process $(x_t^\sigma, v_t^\sigma)_{t\geq 0}$ we consider is the stochastic development of a velocity process satisfying the conclusions of Proposition 1.1. In particular, our proof actually applies even if $(v_t)_{t\geq 0} = (v_t^\sigma)_{t\geq 0}$ is an ergodic Markov process with jumps on $\mathbb{S}^{d-1}$ as soon as the conditions (1.3) and (1.4) are fulfilled.

**Theorem 1.4** (Theorem 4.1 below)**.** *Let $(\mathcal{M}, g)$ be a complete and stochastically complete Riemannian manifold and let $(x_t^\sigma, v_t^\sigma)_{t\geq 0}$ be the process with values in $T\mathcal{M}$ characterised by the fact that the image of $v_t \in T_{x_t}\mathcal{M}$ in the fixed tangent space $T_{x_0}\mathcal{M} \simeq \mathbb{R}^d$ by the inverse stochastic parallel transport along $(x_s)_{0\leq s\leq t}$ satisfies the conditions (1.3) and (1.4). Then as $\sigma$ goes to infinity, the time rescaled process $(x^\sigma_{\sigma^2 t})_{t\in[0,1]}$ converges in law to an anisotropic Brownian motion on the base manifold $\mathcal{M}$.*

See Theorem 4.1 for a precise statement. In this level of generality, in Section 4.2 we recover classical results, amongst which Pinsky's so-called random flight [Pin76] and time-dependent variations of it; the anisotropic Langevin diffusion, where $v$ is an anisotropic Ornstein-Uhlenbeck process; and linear interpolation of symmetric random walks as in [BFH09]. It is unclear whether or not the methods of X.M. Li [Li16a, Li16b] or Herzog, Hottovy and Volpe [HHV16] can get back such a result. In a somewhat independent direction, the interesting work [CFK+17] of Chevyrev and coauthors studies this kind of convergence in deterministic systems.

The outline of the article is the following. In the next Section 2, we study the velocity process solution of equation (1.1). We characterise its invariant measure and establish the mixing properties which are the key ingredients in our approach of the homogenisation phenomenon. Section 3 is then devoted to the proofs of our main Theorem 1.2 and 1.3. More precisely, in Section 3.1, we show the tightness of the rough path lift of the process in the Euclidean setting. In Section 3.2, we then identify the limit as a Brownian motion on the underlying two-step nilpotent Lie group. This completes the proof of Theorem 1.2 in the Euclidean setting. Finally, in Section 3.3, we use the continuity of the Itô map to extend the proof of homogenisation to an arbitrary complete stochastically complete Riemannian manifold. The last section consists in developments, including Theorem 1.4 and comments in Section 4.1, and various examples in Section 4.2.

## 2 Mixing properties of the velocity process

Let $(B_t)_{t\geq 0}$ be a Euclidean Brownian motion in $\mathbb{R}^d$ with non degenerate covariance matrix $\Sigma$. Without loss of generality, up to an appropriate choice of coordinate system, we can assume that the matrix $\Sigma$ is diagonal, with square root $A$, namely

$$\Sigma = \mathrm{diag}\left(\alpha_1^2, \cdots, \alpha_d^2\right), \quad A = \mathrm{diag}\left(\alpha_1, \cdots, \alpha_d\right).$$

Let us recall that, by definition, the anisotropic velocity process $(v_t) = (v_t^1, \cdots, v_t^d)$ with values in $\mathbb{S}^{d-1} \subset \mathbb{R}^d$ and with intensity $\sigma > 0$ is the solution of the Stratonovich stochastic differential equation

$$\mathrm{d}v_t = \sigma \Pi_{v_t^\perp} \circ \mathrm{d}B_t,$$

where $\Pi_{v_t^\perp}$ denotes the projection on the orthogonal of $v_t$. Equivalently, there exist a standard Euclidean Brownian motion $(W_t)_{t\geq 0}$ such that $v_t$ satisfies the Itô stochastic differential equation

$$\mathrm{d}v_t = \sigma \Pi_{v_t^\perp} A \mathrm{d}W_t - \frac{\sigma^2}{2}\big(\Sigma + \mathrm{tr}(\Sigma)\mathrm{Id} - 2\langle v_t, \Sigma v_t\rangle \mathrm{Id}\big) v_t \mathrm{d}t,$$

or even more explicitly in Euclidean coordinates, for $1 \leq i \leq d$

$$\mathrm{d}v_t^i \;=\; -\frac{\sigma^2}{2} v_t^i \left[\alpha_i^2 + \sum_{j=1}^d \alpha_j^2 - 2\sum_{j=1}^d \alpha_j^2 |v_t^j|^2\right] \mathrm{d}t + \sigma \left(\alpha_i \mathrm{d}W_t^i - v_t^i \sum_{j=1}^d \alpha_j v_t^j \mathrm{d}W_t^j\right). \tag{2.5}$$

In the following, $d$ and $\Sigma$ are fixed, and we write $f \lesssim g$ for some quantities $f$ and $g$ whenever $f \leq Cg$ for a constant $C > 0$ independent of any other parameter.

### 2.1 Invariant measure

The object of this section is to establish that the velocity process $(v_t)_{t\geq 0}$ is ergodic in $\mathbb{S}^{d-1}$ and to write down its invariant measure explicitly. From equation (2.5), it is not difficult to express the infinitesimal generator $L$ of the process and try to solve the equation $L^*\mu = 0$. Nevertheless, since we are working on the sphere, integrations by parts and computations are quite unpleasant, and we prefer to introduce a natural Euclidean lift of the velocity process. Namely, if $\|\cdot\|$ denotes the standard Euclidean norm, consider the $\mathbb{R}^d$-valued process $(u_t)_{t\geq 0}$ starting from $u_0 \neq 0$ such that $v_0 = u_0/\|u_0\|$, and solution of the stochastic differential equation system

$$\mathrm{d}u_t^i = \frac{\sigma^2}{2}\left(-u_t^i \|u_t\|^2 + \alpha_i^2 u_t^i\right) \mathrm{d}t + \sigma \alpha_i \|u_t\| \mathrm{d}W_t^i, \quad 1 \leq i \leq d.$$

Equivalently, it is the solution to the Stratonovich stochastic differential equation

$$\mathrm{d}u_t = -\frac{\sigma^2}{2}\|u_t\|^2 u_t \,\mathrm{d}t + \sigma \|u_t\| \circ \mathrm{d}B_t.$$

Then, a direct application of Itô's formula shows that the projection $u_t/\|u_t\|$ on $\mathbb{S}^{d-1}$ satisfies equation (2.5). To show that $u_t$ is ergodic and find an explicit expression for its invariant measure, let us now perform the simple linear change of variable $y_t := A^{-1}u_t = \Sigma^{-1/2}u_t$. By Itô's formula we get

$$\mathrm{d}y_t^i = \frac{\sigma^2}{2}\left(-\|Ay_t\|^2 y_t^i + \alpha_i^2 y_t^i\right) \mathrm{d}t + \sigma \|Ay_t\| \mathrm{d}W_t.$$

Setting $V_A(y) := -\log\|Ay\| + \frac{1}{2}\|y\|^2$, the infinitesimal generator $L_y$ of $y_t$ is given by

$$L_y = \frac{\sigma^2}{2}\|Ay\|^2 L_0, \quad \text{where} \quad L_0 := (-\nabla V_A \cdot \nabla + \Delta).$$

The diffusion process with generator $L_0$ is naturally ergodic with invariant measure proportional to $\mathsf{e}^{-V_A}$ so that $(y_t)_{t\geq 0}$ is also ergodic with invariant measure

$$\nu(\mathrm{d}y) := C_A\,\|Ay\|^{-1}\mathsf{e}^{-\frac{1}{2}\|y\|^2}\,\mathrm{d}y,$$

where $C_A$ is a normalizing constant. In other words, the Euclidean lift $(u_t)_{t\geq 0}$ of $(v_t)_{t\geq 0}$ is ergodic in $\mathbb{R}^d$ and its invariant measure is proportional to $\|\cdot\|^{-1}$ times the centred Gaussian measure with covariance $\Sigma$. One can then compute the invariant measure of the velocity process as the image measure of the latter with respect to the projection on the sphere.

**Lemma 2.1.** *The velocity process $(v_t)_{t\geq 0}$ is ergodic in $\mathbb{S}^{d-1}$ and its invariant measure $\mu$ is absolutely continuous with respect to the uniform measure $\mathrm{d}\theta$ on the sphere, with a density given by*

$$\frac{\mathrm{d}\mu}{\mathrm{d}\theta}(\theta) = \frac{\|A^{-1}\theta\|^{1-d}}{\int_{\mathbb{S}^{d-1}}\|A^{-1}\theta\|^{1-d}\mathrm{d}\theta}.$$

*In particular, the invariant measure $\mu$ of the velocity process is invariant under all the coordinate reflections $(\theta_1,\cdots,\theta_i,\cdots,\theta_d)\mapsto(\theta_1,\cdots,-\theta_i,\cdots,\theta_d)$, for $1\leq i\leq d$.*

*Proof.* For any bounded measurable test function $f$ on $\mathbb{S}^{d-1}$, we have

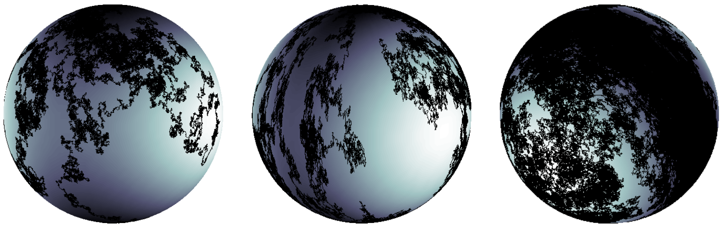
$$\begin{aligned}
\int_{\mathbb{S}^{d-1}} f(v)\mu(\mathrm{d}v) &= C_A\int_{\mathbb{R}^d} f\left(\frac{Ay}{\|Ay\|}\right)\frac{\mathsf{e}^{-\frac{1}{2}\|y\|^2}}{\|Ay\|}\mathrm{d}y \\
&= C_A\int_{\mathbb{R}^d} f\left(\frac{u}{\|u\|}\right)\|u\|^{-1}\mathsf{e}^{-\frac{1}{2}\|A^{-1}u\|^2}\frac{\mathrm{d}u}{\det A} \\
&= C'_A\int_0^{+\infty}\int_{\mathbb{S}^{d-1}} f(\theta)\,r^{-1}\mathsf{e}^{-\frac{1}{2}r^2\|A^{-1}\theta\|^2}r^{d-1}\mathrm{d}r\mathrm{d}\theta \\
&= \frac{\int_{\mathbb{S}^{d-1}} f(\theta)\|A^{-1}\theta\|^{1-d}\mathrm{d}\theta}{\int_{\mathbb{S}^{d-1}}\|A^{-1}\theta\|^{1-d}\mathrm{d}\theta}.
\end{aligned}$$

□

The next figures illustrate the relation between the covariance matrix $\Sigma$, the sample paths of the velocity process $(v_t)$ and its invariant measure $\mu$. The colour map on the sphere is chosen according to the value of the density of the invariant measure: small values of $\|A^{-1}\theta\|^{1-d}$ are represented in light grey whereas large values are represented in dark grey.

**Remark 2.2.** Let us emphasise here that the invariant measure $\mu$ of the velocity process actually differs from the projected Gaussian measure with covariance $\Sigma$, also known as angular Gaussian distribution, which, at first sight, could seem like a natural candidate for the velocity's equilibrium measure. Namely, if $f$ is a bounded measurable test function on the sphere, and if $X$ is a Gaussian variable in $\mathbb{R}^d$ with law $\mathcal{N}(0,\Sigma)$, we have indeed

$$\mathbb{E}\left[f\left(\tfrac{X}{\|X\|}\right)\right] = \frac{\int_{\mathbb{S}^{d-1}} f(\theta)\|A^{-1}\theta\|^{-d}\mathrm{d}\theta}{\int_{\mathbb{S}^{d-1}}\|A^{-1}\theta\|^{-d}\mathrm{d}\theta}.$$



*The colormap on the spheres correspond to the density of the invariant probability measure: the darker the likelier. From left to right, the chosen covariance matrices are* $\Sigma = \mathrm{diag}(1, 1.1, 1.2)$*,* $\Sigma = \mathrm{diag}(1, 4, 9)$*, and* $\Sigma = \mathrm{diag}(1, 100, 100)$*.*

FIGURE 2.2 – Sample paths of the velocity process and invariant probability measures.

In other words, the invariant measure $\mu$ admits a density proportional to $\|A^{-1}\theta\|$ with respect to the standard projected Gaussian measure of covariance $\Sigma$. $\triangle$

**Remark 2.3.** Going back to the modelisation point of view mentioned in the introduction, where $(v_t)_{t\geq 0}$ is thought as the velocity of a mesoscopic particle in an anisotropic heat bath, the invariant measure $\mu$ also differs from the standard choices for equilibrium measure in directional statistics, such as the Von Mises–Fisher distribution, Fisher–Bingham distribution or wrapped Brownian distributions, see Sections 9.3 and 9.4 of [MJ00] and the references therein. We emphasise here the fact that the dynamics governed by equation (1.1) is fully intrinsic so that the measure $\mu$ is a simple and natural candidate to model anisotropic data; it also has natural interpretation in terms of projection of the invariant measure of the Euclidean lift $(u_t)_{t\geq 0}$. $\triangle$

## 2.2 Mixing properties

Let us now establish the strong mixing properties of the velocity process that will be our main tool in the proof of the homogenisation result, Theorem 1.2. To avoid changes in the time scale, we fix $\sigma = 1$, from here to the end of the section. We also introduce a few additional notations. If $\lambda$ is a probability distribution on $\mathbb{S}^{d-1}$, let $\mathbb{P}_\lambda$ be a probability measure under which the velocity $(v_t)_{t\geq 0}$ solves equation (1.1) with initial condition $v_0 \sim \lambda$, and $\mathbb{E}_\lambda$ its associated expectation. We denote by $(P_t)_{t\geq 0}$ the semigroup associated to $v$, acting on continuous functions $f : \mathbb{S}^{d-1} \to \mathbb{R}$, and by $(P_t^*)_{t\geq 0}$ its dual, acting on probability measures on $\mathbb{S}^{d-1}$. In other words,

$$P_t f(x) := \mathbb{E}_{\delta_x}[f(v_t)] \qquad \text{and} \qquad P_t^* \lambda := \mathcal{L}(v_t | v_0 \sim \lambda),$$

for any such $f$ and $\lambda$.

To get to the second part of Proposition 1.1, we use the well-known fact that since the velocity process $(v_t)_{t\geq 0}$ is an elliptic diffusion in a compact Riemannian manifold, here the unit sphere, with invariant probability measure $\mu$, we have the estimate

$$\|P_t^* \lambda - \mu\|_{\mathrm{TV}} \lesssim \exp(-t/\tau) \tag{2.6}$$

for any probability $\lambda$ on $\mathbb{S}^{d-1}$, for some positive constant $\tau$. Given an interval $[a, b)$ of $[0, \infty)$, define $\mathcal{F}_{[a,b)}$ as the $\sigma$-algebra generated by the unit speed velocity process $v_t$, for $a \leq t < b$. We write $A \in \mathcal{F}_{[a,b)}$ to say that a random variable is $\mathcal{F}_{[a,b)}$-measurable.

**Proposition 2.4.** *For any $0 \leq s < t$ and any bounded measurable real-valued random variables $P \in \mathcal{F}_{[0,s)}$ and $F \in \mathcal{F}_{[t,\infty)}$, we have*

$$\big|\mathbb{E}_\mu[PF] - \mathbb{E}_\mu[P]\,\mathbb{E}_\mu[F]\big| \lesssim |P|_\infty |F|_\infty \, \mathsf{e}^{-(t-s)/\tau}.$$

*Proof.* Since

$$\big|\mathbb{E}_\mu[PF] - \mathbb{E}_\mu[P]\mathbb{E}_\mu[F]\big| \leq |P|_\infty \mathbb{E}_\mu\Big[\big|\mathbb{E}_\mu[F|\mathcal{F}_{[0,s]}] - \mathbb{E}_\mu[F]\big|\Big],$$

by the Markov property, it suffices to prove that one has

$$\big|\mathbb{E}_{P_u^*\lambda}[G] - \mathbb{E}_\mu[G]\big| \lesssim |G|_\infty \mathsf{e}^{-u/\tau}, \tag{2.7}$$

for any probability measure $\lambda$ on $\mathbb{S}^{d-1}$ and any real-valued measurable functional $G$. By a monotone class argument, it suffices to prove estimate (2.7) for elementary functionals of the form $G = g(v_{t_1}, \ldots, v_{t_k})$, for some bounded continuous real-valued function $g$ on $(\mathbb{R}^d)^k$ and times $t_1 \leq \cdots \leq t_k$. But since the diffusion has the Feller property, the function $\overline{g}(v_0) := \mathbb{E}_{v_0}\big[g(v_{t_1}, \ldots, v_{t_k})\big]$ is continuous on the sphere, so we get (2.7) in that case by applying (2.6) to $\overline{g}$. ☐

The remainder of the section is devoted to the proof of the technical Lemma 2.6, that states an estimate about iterated integrals involving the covariances between the coordinates of the unit speed velocity process. Given a collection of positive times $s_1, \ldots, s_n$, set $\Delta := \max_{1\leq k<n}(s_k \wedge s_{k+1})$. We denote by $k_0 \in [\![1, n-1]\!]$ an index where this maximum is attained.

**Proposition 2.5.** *Under $\mathbb{P} = \mathbb{P}_\mu$, and for any indices $1 \leq j_1, \ldots, j_n \leq d$ and times $s_1, \cdots, s_n \geq 0$,*

$$\Big|\mathbb{E}\big[v_{s_1}^{j_1} \cdots v_{s_1+\cdots+s_n}^{j_n}\big]\Big| \lesssim \mathsf{e}^{-\Delta/\tau}.$$

*Proof.* For $1 \leq i \leq n$, set $t_i := s_1 + \cdots + s_i$, and define the bounded quantities

$$V_- := v_{t_1}^{j_1} \cdots v_{t_{k_0-1}}^{j_{k_0-1}}, \qquad V_0 := v_{t_{k_0}}^{j_{k_0}}, \qquad V_+ := v_{t_{k_0+1}}^{j_{k_0+1}} \cdots v_{t_n}^{j_n}.$$

Note that $V_0$ is centred. Applying Proposition 2.4 twice, this decomposition gives

$$\begin{aligned}
\Big|\mathbb{E}\big[v_{s_1}^{j_1} \cdots v_{s_1+\cdots+s_n}^{j_n}\big]\Big| &= \big|\mathbb{E}[V_-V_0V_+] - \mathbb{E}[V_-]\mathbb{E}[V_0]\mathbb{E}[V_+]\big| \\
&\leq \big|\mathbb{E}[V_-V_0V_+] - \mathbb{E}[V_-]\mathbb{E}[V_0V_+]\big| + |V_-|_\infty \,|\mathbb{E}[V_0V_+] - \mathbb{E}[V_0]\mathbb{E}[V_+]| \\
&\lesssim |V_-|_\infty |V_0V_+|_\infty \mathsf{e}^{-s_{k_0}/\tau} + |V_-|_\infty |V_0|_\infty |V_+|_\infty \mathsf{e}^{-s_{k_0+1}/\tau} \\
&\lesssim \mathsf{e}^{-\Delta/\tau}.
\end{aligned}$$
☐

**Lemma 2.6.** *Suppose $\mathbb{P} = \mathbb{P}_\mu$. Given a positive integer $n$, we have*

$$\int_{0\leq t_1\leq\cdots\leq t_{2n}\leq T} \big|\mathbb{E}_\mu\big[v_{t_1}^{i_1} \cdots v_{t_{2n}}^{i_{2n}}\big]\big| \, \mathrm{d}t_1 \ldots \mathrm{d}t_{2n} \lesssim_n T^n$$

*for any indices $1 \leq i_1, \cdots, i_{2n} \leq d$.*

*Proof.* The idea is to apply Proposition 2.5 with the largest $\Delta$ possible for each tuple $(t_1, \cdots, t_{2n})$. In practice, we write first

$$\int_{0\le t_1\le\cdots\le t_{2n-1}} \left|\mathbb{E}\left[v^{i_1}_{t_1}\cdots v^{i_{2n}}_{t_{2n}}\right]\right| \,\mathrm{d}t_1\ldots\mathrm{d}t_{2n} \le \int_{[0,T]^{2n}} \left|\mathbb{E}\left[v^{i_1}_{s_1}\cdots v^{i_{2n}}_{s_1+\cdots+s_{2n}}\right]\right| \,\mathrm{d}s_1\ldots\mathrm{d}s_{2n}.$$

Fix now the tuple $(s_1, \cdots, s_{2n})$, and set

$$\Delta(s) := \max_{1\le k<2n} (s_k \wedge s_{k+1}),$$

so the integrand in the right hand side above is bounded above by a constant multiple of $\mathsf{e}^{-\Delta(s)/\tau}$, from Proposition 2.5.

The rest is combinatorics. We first sort the indices $k$ of the gaps $s_k$ according to the value of $s_k$ with respect to $\Delta = \Delta(s)$. Set $a := \min\{k \in [\![1, 2n]\!] : s_k = \Delta\}$. Then, note that there are at most $n$ gaps $s_k$ of size larger than $\Delta$: otherwise, two of them would be consecutive, and $\Delta$ would not be optimal. This is the same as saying that there are at least $n$ small gaps $s_k \le \Delta$, including $s_a$. Define $1 \le b_1 < \cdots < b_{n-1} \le 2n$ as the first $(n-1)$ indices different from $a$ corresponding to gaps of size at most $\Delta$. In other words, if $s_k \le \Delta$, then either $k = b_i$ for some $1 \le i < n$, $k = a$, or $k > a, b_{n-1}$. Finally, denote by $1 \le c_1 < \cdots < c_n \le 2n$ the other indices, so that we have a partition of $\{1, \cdots, 2n\}$ in three sets $A(s) := \{a\}$, $B(s) := \{b_1, \cdots, b_{n-1}\}$ and $C(s) := \{c_1, \cdots, c_n\}$ of fixed sizes. Now, given a fixed partition $(\alpha, \beta, \gamma)$ of $[\![1, 2n]\!]$ with $\alpha = \{\alpha_0\}$ of size 1, and the set $\beta = \{\beta_1, \ldots, \beta_{n-1}\}$ of size $n-1$, we have

$$\begin{aligned}\left|\mathbb{E}\left[v^{i_1}_{s_1}\cdots v^{i_{2n}}_{s_1+\cdots+s_{2n}}\right]\right| \mathbf{1}_{(A(s),B(s),C(s))=(\alpha,\beta,\gamma)} &\lesssim \mathsf{e}^{-\Delta(s)/\tau}\mathbf{1}_{s_{\beta_1},\ldots,s_{\beta_{n-1}}\le\Delta(s)} \\ &\lesssim \mathsf{e}^{-s_{\alpha_0}/\tau}\mathbf{1}_{s_{\beta_1},\ldots,s_{\beta_{n-1}}\le s_{\alpha_0}},\end{aligned}$$

from which we get

$$\begin{aligned}\int_{[0,T]^{2n}} &\left|\mathbb{E}\left[v^{i_1}_{s_1}\cdots v^{i_{2n}}_{s_1+\cdots+s_{2n}}\right]\right| \mathbf{1}_{(A(s),B(s),C(s))=(\alpha,\beta,\gamma)} \,\mathrm{d}s_1\ldots\mathrm{d}s_{2n} \\ &\lesssim T^n \int_0^T \mathsf{e}^{-s/\tau} s^{n-1}\mathrm{d}s \lesssim_n T^n\end{aligned}$$

and the result of the lemma, by summing over the set of all partitions $(\alpha, \beta, \gamma)$ of $[\![1, 2n]\!]$ with the above size. $\square$

# 3 Proof of the main result

Let us now describe how the mixing properties of the velocity process derived in Section 2.2 imply the homogenisation for the time rescaled position process $(x^\sigma_{\sigma^2 t})_{t\ge0}$, as $\sigma$ goes to infinity, in both Euclidean and Riemannian framework. As mentioned in the introduction, we will actually work with a rough path lift of the kinetic process. We refer the reader to [FH14, Bai15b] for gentle introductions to rough paths theory, and given $\gamma \in (0, 1)$, we denote by $\mathsf{RP}(\gamma) = \mathsf{RP}^\gamma([0, 1], \mathbb{R}^d)$ the set of weak geometric $\gamma$-Hölder rough paths.

**Notations.** We are interested in the stationary case $\mathbb{P} := \mathbb{P}_\mu$, where $\mu$ is the invariant measure of the velocity, as described in Lemma 2.1. Define $X^\sigma : t \mapsto x^\sigma_{\sigma^2 t}$, so that we are interested in the limiting behaviour of $(X^\sigma_t)_{t\geq 0}$. To make good use of the mixing properties of $v$ such as Proposition 2.4 without having to change the time scale, from now on $(v_t)_{t\geq 0}$ *will always stand for* $(v^\sigma_t)_{t\geq 0}$ *with* $\sigma = 1$. With this convention, we can express the increments of $X^\sigma$ as

$$X_t - X_s = \frac{1}{\sigma^2}\int_{\sigma^4 s}^{\sigma^4 t} v_u \mathrm{d}u.$$

The process $X_\sigma$ being $\mathcal{C}^1$, it admits a canonical rough path lift $\mathbf{X}^\sigma = (X^\sigma, \mathbb{X}^\sigma)$, where $\mathbb{X}^\sigma$ is defined by

$$\mathbb{X}^\sigma_{ts} := \int_s^t (X^\sigma_u - X^\sigma_s)\otimes \mathrm{d}X^\sigma_u = \frac{1}{\sigma^4}\int_{\sigma^4 s}^{\sigma^4 t}\int_{\sigma^4 s}^{u} v_z \otimes v_u \mathrm{d}z\, \mathrm{d}u.$$

Our proof relies on the algebraic properties of rough paths. Namely, that in the 2-step nilpotent group $G \subset \mathbb{R}\oplus\mathbb{R}^d\oplus(\mathbb{R}^d)^{\otimes 2}$, the process $\mathbf{x}^\sigma : t \mapsto (1, X^\sigma_t, \mathbb{X}^\sigma_{t0})$ has increments

$$(\mathbf{x}^\sigma_s)^{-1}\mathbf{x}^\sigma_t = (1, X^\sigma_t - X^\sigma_s, \mathbb{X}^\sigma_{ts}),$$

which, using the above expressions, are measurable with respect to $\sigma\big((v_u)_{\sigma^4 s\leq u<\sigma^4 t}\big)$.

Recall that we write $f \lesssim g$ for some quantities $f$ and $g$ when there exists a positive constant $C > 0$ depending on $\Sigma$ alone such that $f \leq Cg$. If $C$ is allowed to depend on a parameter, say $p$, we write $f \lesssim_p g$.

## 3.1 Tightness in rough paths space

We first establish that the family of processes $(X^\sigma_t)$ and their rough paths lifts are tight for the corresponding topology. To do so, we use a standard Lamperti criterion, namely we have the following lemma.

**Lemma 3.1.** *For every $a \geq 1$,*

$$\sup_{\sigma>0} \mathbb{E}\big[|X^\sigma_t - X^\sigma_s|^a\big] \lesssim_a |t-s|^{a/2}.$$

*Proof.* Given any positive time $T$ and any positive integer $n$, we show that one has

$$\mathbb{E}\left[\left|\int_0^T v_t\, \mathrm{d}t\right|^{2n}\right] \leq C_n T^n \tag{3.8}$$

for some positive constant $C_n$ depending only on $n$. The inequality of the lemma follows as a consequence since for any positive integer $n$ such that $2n \geq a$, we have

$$\mathbb{E}\big[|X^\sigma_t - X^\sigma_s|^a\big] = \mathbb{E}\big[|X^\sigma_{t-s}|^a\big] \leq \frac{1}{\sigma^{2a}}\,\mathbb{E}\left[\left|\int_0^{\sigma^4(t-s)} v_u\, \mathrm{d}u\right|^{2n}\right]^{a/2n} \leq C_n^{a/2n}(t-s)^{a/2}.$$

Given $T > 0$ and $n \in \mathbb{N}^*$, we have

$$\begin{aligned}\mathbb{E}\left[\left|\int_0^T v_t\, \mathrm{d}t\right|^{2n}\right] &= \mathbb{E}\left[\left(\sum_{1\leq i\leq d}\left(\int_0^T v^i_t\, \mathrm{d}t\right)^2\right)^n\right]\\ &= \sum_{1\leq i_1,\cdots,i_n\leq d}\int_{[0,T]^{2n}} \mathbb{E}\left[v^{i_1}_{t_1} v^{i_1}_{t_2}\cdots v^{i_n}_{t_{2n-1}} v^{i_n}_{t_{2n}}\right] \mathrm{d}t_1\cdots \mathrm{d}t_{2n},\end{aligned}$$

with the following estimate for each individual term on the right hand side. Fix $1 \leq j_k \leq d$, for $1 \leq k \leq 2n$. For any permutation $\phi \in \mathfrak{S}_{2n}$, we have from Lemma 2.6

$$\int_{[0,T]^{2n}} \mathbb{E}\Big[v_{t_1}^{j_1} \cdots v_{t_{2n}}^{j_{2n}}\Big] \mathbf{1}_{t_{\phi(1)}<\cdots<t_{\phi(2n)}} \,\mathrm{d}t = \int_{0\leq t_1\leq\cdots\leq t_{2n}\leq T} \mathbb{E}\Big[v_{t_1}^{j_{\phi(1)}} \cdots v_{t_{2n}}^{j_{\phi(2n)}}\Big] \,\mathrm{d}t_1 \cdots \mathrm{d}t_{2n}$$
$$\lesssim_n T^n,$$

from which the result of the Lemma follows by summation over $\phi$ and $j$. □

We use the Hilbert-Schmidt norm $|\cdot|$ on $\mathbb{R}^d \otimes \mathbb{R}^d \simeq L(\mathbb{R}^d) \simeq \mathbb{R}^{d^2}$; it coincides with the Euclidean norm on $\mathbb{R}^{d^2}$.

**Lemma 3.2.** *For every $a > 0$,*

$$\sup_{\sigma>0} \mathbb{E}\big[|\mathbb{X}_{ts}^\sigma|^a\big] \lesssim_a |t-s|^a.$$

*Proof.* As above, the inequality of the statement follows from an inequality of the form

$$\mathbb{E}\left[\left|\int_{0\leq s\leq t\leq T} v_s \otimes v_t \,\mathrm{d}s\mathrm{d}t\right|^{2n}\right] \leq C_n T^{2n},$$

for some positive constant $C_n$ depending only on $n$. Fix $T > 0$ and $n \in \mathbb{N}^*$, and set for $\ell \in [\![1,d]\!]^{4n}$

$$I_\ell := \int_{0\leq s_1\leq t_1\leq T} \cdots \int_{0\leq s_{2n}\leq t_{2n}\leq T} \mathbb{E}\Big[v_{t_1}^{\ell_1} v_{s_1}^{\ell_2} \cdots v_{t_{2n}}^{\ell_{4n-1}} v_{s_{2n}}^{\ell_{4n}}\Big] \,\mathrm{d}s_1\mathrm{d}t_1 \cdots \mathrm{d}s_{2n}\mathrm{d}t_{2n},$$

so we have

$$\mathbb{E}\left[\left|\int_0^T \int_0^t v_s \otimes v_t \,\mathrm{d}s\mathrm{d}t\right|^{2n}\right] = \mathbb{E}\left[\left(\sum_{1\leq i,j\leq d} \left(\int_0^T \int_0^t v_s^i v_t^j \mathrm{d}s\mathrm{d}t\right)^2\right)^n\right]$$
$$= \sum_{i,j\in[\![1,d]\!]^n} I_{i*j}$$

with $i * j = (i_1, j_1, i_1, j_1, \cdots, i_k, j_k, i_k, j_k)$. As in Lemma 3.1, estimating each $I_{i*j}$ using Lemma 2.6 does the job. □

One can then apply the Kolmogorov-Lamperti tightness criterion for rough paths stated in Corollary A.12 of [FV10] to get the following result from Lemma 3.1 and Lemma 3.2.

**Corollary 3.3.** *Pick $1/3 < \gamma < 1/2$. The family $\{\mathcal{L}(\mathbf{X}^\sigma)\}_{\sigma>0}$ of distributions on $\mathsf{RP}(\gamma)$ is tight.*

**Remark 3.4.** Although the criterion does ensure the tightness of both $X^\sigma$ and $\mathbb{X}^\sigma$ in suitable path spaces, the usual rough path space $\mathsf{RP}(\gamma)$ actually involves more iterated integrals as $\gamma \to 0$. For instance, one has to consider the quantity

$$\int_s^t \int_s^u (X_v - X_s) \otimes \mathrm{d}X_v \otimes \mathrm{d}X_u$$

for rough differential equations to be well-posed whenever $1/4 < \gamma \leq 1/3$. This is merely technical, since it is known that the topology of $\mathsf{RP}(\gamma')$ for $0 < \gamma' < \gamma$ is weaker than that induced by $\mathsf{RP}(\gamma)$ on the first few levels, due to algebraic constraints. However, we choose not to bother with these details, and to state Proposition 3.5 with any $\gamma > 0$, while assuming $\gamma > 1/3$ in Theorem 3.6. △

## 3.2 Brownian limit

The family of processes $(X_t^\sigma)$ and their lifts being tight for the rough paths topology, in order to establish its convergence, we are left to identify the possible limit process. Our strategy here is to prove that the latter is necessarily a stationary process with independent Gaussian increments on the underlying nilpotent group, and therefore is a Brownian motion. Let us set

$$\gamma_i := 2\int_0^\infty \mathbb{E}[v_0^i v_t^i]\,\mathrm{d}t.$$

**Proposition 3.5.** *For every $\gamma < 1/2$, the processes $X^\sigma$ converge in law in $\mathcal{C}^\gamma([0;1],\mathbb{R}^d)$ to the Brownian motion on $\mathbb{R}^d$ with covariance matrix $\mathrm{diag}(\gamma_1,\cdots,\gamma_d)$, as $\sigma$ goes to $\infty$.*

*Proof. Stationarity and independence.* We first show that any $\mathbb{R}^d$-valued process $X$ whose law $\widehat{\mathbb{P}}$ is a limit point of $(\mathcal{L}(X^\sigma))_{\sigma>0}$ in $\mathcal{C}^\gamma([0;1],\mathbb{R}^d)$ as $\sigma$ tends to $\infty$ has stationary independent increments.

Indeed, since $v_0$ has distribution the invariant measure of the diffusion $v$, the increments of $X^\sigma$ are stationary for every $\sigma$, so the increments of $X$ are stationary as well. Fix now $0 \le s_1 < t_1 \le \cdots \le s_n < t_n \le 1$, and bounded continuous functions $F_i : \mathbb{R}^d \to \mathbb{R}$, for $1 \le i \le n$. Fix $\varepsilon > 0$ small enough. From a repetitive use of Proposition 2.4, as used in Proposition 2.5, we have

$$\left|\mathbb{E}\left[\prod_{1\le i\le n} F_i\big(X^\sigma_{t_i-\varepsilon} - X^\sigma_{s_i}\big)\right] - \prod_{1\le i\le n}\mathbb{E}\left[F_i\big(X^\sigma_{t_i-\varepsilon} - X^\sigma_{s_i}\big)\right]\right| \lesssim_n |F_1|_{L^\infty}\cdots|F_n|_{L^\infty}\,\mathrm{e}^{-\sigma^4\varepsilon/\tau}$$

for some positive constant $\tau$, and we see that

$$\widehat{\mathbb{E}}\left[\prod_{1\le i\le n} F_i\big(X_{t_i-\varepsilon} - X_{s_i}\big)\right] = \prod_{1\le i\le n}\widehat{\mathbb{E}}\big[F_i\big(X_{t_i-\varepsilon} - X_{s_i}\big)\big],$$

sending $\sigma$ to $\infty$ along a proper subsequence. Using the boundedness and continuity of the functions $F_i$ and the continuity of the process $X$, we can send $\varepsilon$ to 0 and see that $X$ has independent increments. So $X$ is a Brownian motion; it has null mean since every $X_1^\sigma$ has null mean, and its covariance is given by the limit of the covariances of the $X_1^\sigma$.

*Covariance formula.* First, it follows from the identity

$$\mathcal{L}(v^1,\cdots,v^i,\cdots,v^n) = \mathcal{L}(v^1,\cdots,-v^i,\cdots,v^n)$$

that different components of $X_1$ have null covariance since this is the case for different components of $X_1^\sigma$. Now, for $1 \le i \le d$, we have

$$\begin{aligned}
\mathbb{E}\big[\big((X_1^\sigma)^i\big)^2\big] &= \frac{1}{\sigma^4}\int_0^{\sigma^4}\int_0^{\sigma^4}\mathbb{E}[v_s^i v_t^i]\,\mathrm{d}s\,\mathrm{d}t = \frac{2}{\sigma^4}\int_0^{\sigma^4}\int_t^{\sigma^4}\mathbb{E}[v_s^i v_t^i]\,\mathrm{d}s\,\mathrm{d}t \\
&= \frac{2}{\sigma^4}\int_0^\infty\int_0^\infty \mathbf{1}_{t+u\le\sigma^4}\mathbb{E}[v_{t+u}^i v_t^i]\mathrm{d}u\,\mathrm{d}t = 2\int_0^\infty\left(1-\frac{u}{\sigma^4}\right)_+\mathbb{E}[v_u^i v_0^i]\mathrm{d}u
\end{aligned}$$

with $(\cdot)_+$ the positive part. According to Proposition 2.4, the integrand is smaller than a constant multiple of $\exp(-u/\tau)$, uniformly on $\sigma$. It is integrable, so we see from Lebesgue dominated convergence theorem that the above variance tends to $\gamma_i$. $\square$

**Theorem 3.6** (Main Theorem — Euclidean version)**.** *Pick* $1/3 < \gamma < 1/2$*. The processes* $\mathbf{X}^\sigma$ *converge in law in* $\mathsf{RP}(\gamma)$*, as* $\sigma$ *goes to* $\infty$*, to the Brownian rough path on* $\mathbb{R}^d$ *with covariance matrix* $\mathrm{diag}(\gamma_1, \cdots, \gamma_d)$*.*

*Proof. G-valued Lévy process.* As above, we first notice that any limit measure of the laws of $(\mathbf{X}^\sigma)_{\sigma>0}$ turns the canonical process on $\mathsf{RP}(\gamma)$ into a random process with stationary independent increments, in the free nilpotent Lie group of step 2, as a consequence of the corresponding property for $\mathbf{X}^\sigma$. The canonical process on the free nilpotent Lie group of step 2 is thus a continuous Lévy process under any limit law, so, according to Hunt's theorem, we can identify the former from its generator. More specifically, such a process $Y$ is characterised by the action

$$f \mapsto \lim_{t\to 0} \frac{1}{t}\mathbb{E}_e[f(Y_t) - f(e)] \in \mathbb{R}$$

of its generator on smooth functions $f : G \to \mathbb{R}$ with compact support, where $e$ is the unit of $G$; see [App14, Theorems 5.3.3] or [Lia04, Theorem 1.1].

*Generator.* Let $\widehat{\mathbb{P}}$ be any limit point of the laws of $\mathbf{X}^\sigma$ on $\mathsf{RP}(\gamma)$, and denote by $\mathbf{X} = (X, \mathbb{X})$ its canonical variable. We know from Proposition 3.5 that $X$ is a Brownian motion $W$; denote by $\mathbf{W} = (W, \mathbb{W})$ its canonical Stratonovich rough path lift, also defined on the space $(\mathsf{RP}(\gamma), \widehat{\mathbb{P}})$. Since the velocity process $v = (v^1, \dots, v^d)$ and $(v^1, \dots, v^{i-1}, -v^i, v^{i+1}, \dots, v^d)$ have the same law for every $1 \le i \le d$, for $v_0$ distributed according to the invariant measure $\mu$, the antisymmetric part $\mathbb{A}^{\mathbf{X}}_{ts} := \frac{1}{2}(\mathbb{X}_{ts} - {}^t\mathbb{X}_{ts})$ is centred for any $0 \le s \le t \le 1$. We also know from the uniform estimates proved in Lemmas 3.1 and 3.2 that

$$\widehat{\mathbb{E}}\big[|X_t|^2\big] \lesssim t, \quad \widehat{\mathbb{E}}\big[|\mathbb{A}^{\mathbf{X}}_{t0}|^2\big] \lesssim \widehat{\mathbb{E}}\big[|\mathbb{X}_{t0}|^2\big] \lesssim t^2, \tag{3.9}$$

uniformly in $t \in [0, 1]$.

A last piece of notation. Since the set of antisymmetric matrices lies in the tangent space to the free nilpotent Lie group $G$ of step 2, at any point $\mathbf{z} \in G$, any smooth real-valued function $f$ defined on $G$, with compact support, has a well-defined partial differential $\partial_{\mathbb{A}} f(\mathbf{z})$ in the direction of antisymmetric matrices, defined by the identity

$$\partial_{\mathbb{A}} f(\mathbf{z})(\mathbb{A}) = \frac{\mathrm{d}}{\mathrm{d}t}_{|t=0} f\big(\mathbf{z} + t(0, 0, \mathbb{A})\big),$$

for any $\mathbf{z} = (1, Z, \mathbb{Z}) \in G$ and any antisymmetric matrix $\mathbb{A}$. Setting $\overline{\mathbf{z}} := \big(1, Z, \frac{1}{2}(\mathbb{Z} + {}^t\mathbb{Z})\big)$, we further have

$$\big|f(\mathbf{z}) - f(\overline{\mathbf{z}}) - (\partial_{\mathbb{A}} f)(\overline{\mathbf{z}})(\mathbb{A}^{\mathbf{z}})\big| \lesssim_f |\mathbb{A}^{\mathbf{z}}|^2,$$

since $f$ has compact support. Denote by $e$ the unit of the group $G$. Denote by $\mathbb{A}^{\mathbf{W}}$ the antisymmetric part of $\mathbb{W}$ and set $\overline{\mathbf{X}}_t := \big(1, X_t, \frac{1}{2}X_t^{\otimes 2}\big) \in G$, so that $\mathbf{X}_t = \overline{\mathbf{X}}_t + (0, 0, \mathbb{A}^{\mathbf{X}})$ and $\mathbf{W}_t = \overline{\mathbf{X}}_t + (0, 0, \mathbb{A}^{\mathbf{W}})$. We have, for some fixed $f$ smooth with compact support,

$$\begin{aligned}
&\left|\frac{1}{t}\widehat{\mathbb{E}}\big[f(\mathbf{X}_t) - f(e)\big] - \frac{1}{t}\widehat{\mathbb{E}}\big[f(\mathbf{W}_t) - f(e)\big]\right| \\
&\qquad = \frac{1}{t}\left|\widehat{\mathbb{E}}\Big[f\big(\overline{\mathbf{X}}_t + (0, 0, \mathbb{A}^{\mathbf{X}}_t)\big) - f\big(\overline{\mathbf{X}}_t + (0, 0, \mathbb{A}^{\mathbf{W}}_t)\big)\Big]\right| \\
&\qquad \lesssim_f \frac{1}{t}\left|\widehat{\mathbb{E}}\Big[\Big((\partial_{\mathbb{A}} f)(\overline{\mathbf{X}}_t) - (\partial_{\mathbb{A}} f)(e)\Big)(\mathbb{A}^{\mathbf{X}}_t - \mathbb{A}^{\mathbf{W}}_t)\Big]\right| + \frac{1}{t}\left|\widehat{\mathbb{E}}\big[(\partial_{\mathbb{A}} f)(e)(\mathbb{A}^{\mathbf{X}}_t - \mathbb{A}^{\mathbf{W}}_t)\big]\right| \\
&\qquad\quad + \frac{1}{t}\Big(\widehat{\mathbb{E}}\big[|\mathbb{A}^{\mathbf{X}}_t|^2\big] + \widehat{\mathbb{E}}\big[|\mathbb{A}^{\mathbf{W}}_t|^2\big]\Big) \\
&\qquad \lesssim_f (1) + (2) + (3).
\end{aligned}$$

We show that each term vanishes as $t$ goes to 0, which implies that the two Markov processes $\mathbf{X}$ and $\mathbf{W}$ have the same generator, hence the same distribution. We have first from estimates (3.9) the upper bound

$$\begin{aligned}(1) &\leq \frac{1}{2t}\widehat{\mathbb{E}}\Big[\sqrt{t}\,\big\|(\partial_{\mathbb{A}}f)(\overline{\mathbf{X}}_t)-(\partial_{\mathbb{A}}f)(e)\big\|^2\Big]+\frac{1}{2t}\widehat{\mathbb{E}}\Big[\frac{1}{\sqrt{t}}\big|\mathbb{A}_t^{\mathbf{X}}-\mathbb{A}_t^{\mathbf{W}}\big|^2\Big]\\ &\lesssim_f \frac{1}{\sqrt{t}}\widehat{\mathbb{E}}\big[|\overline{\mathbf{X}}_t-e|^2\big]+\frac{1}{t\sqrt{t}}\widehat{\mathbb{E}}\Big[\big|\mathbb{A}_t^{\mathbf{X}}\big|^2\Big]+\frac{1}{t\sqrt{t}}\widehat{\mathbb{E}}\Big[\big|\mathbb{A}_t^{\mathbf{W}}\big|^2\Big]\\ &\lesssim_f \sqrt{t}.\end{aligned}$$

We also have $(2)=0$, since $\mathbb{A}_t^{\mathbf{X}}$ and $\mathbb{A}_t^{\mathbf{W}}$ are centred and $\partial_{\mathbb{A}}f(e)$ is linear. Finally, we have $(3)\lesssim t$ from the upper bounds (3.9). We thus have the upper bound

$$\left|\frac{1}{t}\widehat{\mathbb{E}}\big[f(\mathbf{X}_t)-f(e)\big]-\frac{1}{t}\widehat{\mathbb{E}}\big[f(\mathbf{W}_t)-f(e)\big]\right|\lesssim_f \sqrt{t},$$

from which the result follows. □

## 3.3 From Euclidean space to Riemannian manifolds

Let $(\mathcal{M},g)$ be a Riemannian manifold of dimension $d$, without boundary. We emphasised in the introduction that anisotropic Brownian motion describes the random motion of a non-point-like object, with its own notion of local orientation. Such an object is represented by a point in the orthonormal frame bundle $O\mathcal{M}$ of $\mathcal{M}$, where its dynamics is described by a stochastic differential equation. We refer to Hsu's book [Hsu02] for a reference textbook on stochastic differential geometry.

In this subsection, we use Einstein summation convention: indices appearing twice are implicitly summed.

### The orthonormal frame bundle

Denote by $z=(q,e)$ a generic point of the orthonormal frame bundle $O\mathcal{M}$ of $\mathcal{M}$, with $q\in\mathcal{M}$ and $e:\mathbb{R}^d\to T_q\mathcal{M}$, an orthonormal frame of $T_q\mathcal{M}$; we write $\pi:O\mathcal{M}\to\mathcal{M}$ for the canonical projection map. The Levi-Civita connection on $T\mathcal{M}$ induces a notion of horizontal vectors on $T\mathcal{M}$ or $O\mathcal{M}$. Let $\mathrm{H}$ stand for the horizontal lift operator, meaning the map $O\mathcal{M}\times\mathbb{R}^d\to TO\mathcal{M}$ uniquely characterised by the property that $\mathrm{H}_z(u)\in T_zO\mathcal{M}$ is horizontal and

$$\mathrm{d}\pi_z\big(\mathrm{H}_z(u)\big)=e(u),$$

for any $u\in\mathbb{R}^d$ and $z=(q,e)\in O\mathcal{M}$. Letting $(\epsilon_1,\dots,\epsilon_d)$ be the canonical basis of $\mathbb{R}^d$, local coordinates $q^i$ on $\mathcal{M}$ induce canonical coordinates on $O\mathcal{M}$ by writing

$$e_i:=e(\epsilon_i)=e_i^j\frac{\partial}{\partial q^j}.$$

Denoting by $\Gamma_{ij}^k$ the Christoffel symbols of the Levi-Civita connection associated with the above coordinates, the vector fields $\mathrm{H}(u)$ have the following expression.

$$\mathrm{H}_z(\epsilon_\alpha)=e_\alpha^i\frac{\partial}{\partial q^i}-\Gamma_{ij}^k(q)e_\alpha^ie_l^j\frac{\partial}{\partial e_l^k}.$$

**Cartan's development map and anisotropic kinetic Brownian motion**

Roughly speaking, Cartan's development map associates in its simplest form a $\mathcal{C}^1$ path in $\mathcal{M}$, started from $q_0 \in \mathcal{M}$, to any $\mathcal{C}^1$ path in the Euclidean space $\mathbb{R}^d$. Technically, given a $\mathcal{C}^1$ path $(x_t)_{t\geq 0}$ in $\mathbb{R}^d$, and $z_0 = (q_0, e_0) \in O\mathcal{M}$, the Cartan development of $(x_t)_{t\geq 0}$ on $\mathcal{M}$ is defined as the projection $(q_t)_{0\leq t<T}$ on $\mathcal{M}$ of the horizontal $O\mathcal{M}$-valued path $(z_t) =: (q_t, e_t)_{0\leq t<T}$ solution of the ordinary differential equation

$$\mathrm{d}z_t = \mathrm{H}_{z_t}(\mathrm{d}x_t), \qquad \text{i.e.} \qquad \dot{z}_t = \mathrm{H}_z(\dot{x}_t) \tag{3.10}$$

started from $q_0$, possibly up to some explosion time $T$. Note that the choice of $x : t \mapsto tu$ for some $u \in \mathbb{R}^d$ leads to $q$ being a geodesic with initial condition $\dot{q}_0 = e_0(u)$; in particular, the development of $X^\sigma$ tends to a geodesic with random initial condition as $\sigma \to 0$.

Classical stochastic analysis (in the Stratonovich sense) can be used to make sense of the preceding equation for $x$ a semimartingale, defining Cartan's stochastic development — refer to Hsu's book [Hsu02] for a pedagogical account of the theory. For example, one of the many equivalent constructions of Brownian motion on $\mathcal{M}$ started at $q_0$ consists in developing a standard Euclidean Brownian motion. Accordingly, we define anisotropic Brownian motion on $\mathcal{M}$ as the development of the Euclidean Brownian motion with covariance $\mathrm{diag}(\gamma_1, \cdots, \gamma_d)$.

Anisotropic kinetic Brownian motion $(q_t^\sigma)_{0\leq t<T}$ on $\mathcal{M}$ is the stochastic development of the anisotropic kinetic Brownian motion $(X_t^\sigma)_{t\geq 0}$ on $\mathbb{R}^d$; it is indexed by the speed parameter $\sigma$ of its flat counterpart. This is a $\mathcal{C}^1$ random path which depends on the entire frame $e_0$ — its isotropic counterpart only depends *in law* on $e_0$, from symmetry properties of Wiener measure on $\mathbb{R}^d$. Although $X^\sigma$ converges in law to an anisotropic Brownian motion $B$ on $\mathbb{R}^d$, the poor regularity properties of the Itô solution map does not allow to conclude that anisotropic Brownian motion $x^\sigma$ on $\mathcal{M}$ converges to projection on $\mathcal{M}$ of the solution of the equation

$$\mathrm{d}z_t = \mathrm{H}(z_t) \circ \mathrm{d}B_t.$$

This is exactly the kind of conclusion that rough paths theory provides.

**Rough paths and rough differential equations with values in manifolds**

We discuss a few results of rough paths theory with values in manifolds. These results are all classical, and their Euclidean counterparts can be found e.g. in [FH14] or [FV10]. Let $\mathcal{N}$ be a manifold, and, for a collection $A = (A_1, \cdots, A_n)$ of smooth vector fields on $\mathcal{N}$ and an initial condition $p \in \mathcal{N}$, consider the (deterministic) controlled differential equation

$$\mathrm{d}z_t = A(z_t)\mathrm{d}x_t, \quad z_0 = p$$

on $\mathcal{N}$, where $x$ is a driving curve with values in $\mathbb{R}^n$. The equation makes sense whenever $x$ is of class $\mathcal{C}^1$ (dividing each side by $\mathrm{d}t$, one might say), and if moreover $x$ is of class $\mathcal{C}^2$, its solution is characterised by the fact that for any fixed $t \geq 0$ and $f : \mathcal{N} \to \mathbb{R}$ smooth with compact support,

$$f(z_t) = f(z_s) + (A_i f)(z_t)(x_t^i - x_s^i) + O(|t-s|^2)$$

as $s \to t$. Now if $\mathbf{X} = (X, \mathbb{X})$ is a rough path of Hölder regularity $1/3 < \gamma \leq 1/2$, we consider the following notion of solution: a continuous path $z : [0, T) \to \mathcal{N}$ is a solution of the rough differential equation

$$\mathrm{d}z_t = A(z_t)\mathbf{X}_{\mathrm{d}t}\,, \quad z_0 = p \tag{3.11}$$

if one can find some $a > 1$ such that any choice of $t \geq 0$ and $f : \mathcal{N} \to \mathbb{R}$ smooth with compact support yields

$$f(z_t) = f(z_s) + (A_i f)(z_t)(X_t^i - X_s^i) + (A_i A_j f)(z_t)\mathbb{X}_{ts}^{ij} + O(|t-s|^a)$$

as $s \to t$. This point of view is taken from [Bai10, Bai15b], in the mindset of [Dav08]. We say that $z$ explodes as $t \to T$ if $z$ leaves any compact set.

In a probabilistic mindset, the fundamental remark is that, for $\mathbf{X}$ the Stratonovich rough path lift of some standard Brownian motion $W$, such a solution coincides almost surely with the solution of the Stratonovich equation

$$\mathrm{d}z_t = A(z_t) \circ \mathrm{d}W_t, \quad z_0 = p.$$

It is a striking feature of rough paths theory that not only does (3.11) admit a unique solution $z$ for any (deterministic) rough path $\mathbf{X}$, in the above sense and up to some explosion time $T > 0$, but also the Itô-Lyons map $\mathbf{X} \mapsto z$ is continuous in the following sense. Fix $d$ a Riemannian distance on $\mathcal{N}$. If $T' < T$ and $\varepsilon > 0$, there exists some $\delta > 0$ such that for any $\mathbf{X}'$ at rough path distance at most $\delta$ from $\mathbf{X}$, the solution $z'$ of

$$\mathrm{d}z'_t = A(z'_t)\mathbf{X}'_{\mathrm{d}t}\,, \quad z'_0 = p$$

is defined on $[0, T']$ and satisfies $d(z_t, z'_t) < \varepsilon$ for all $0 \leq t \leq T'$.

This kind of continuity in enough to ensure convergence in law: namely, if $(\mathbf{X}^n)_{n\geq 0}$ is a family of random rough paths converging in law to $\mathbf{X}$ with respect to the rough path topology, then in a sense, the (random) solution $z^n$ of (3.11) driven by $\mathbf{X}^n$ converges to the solution of that driven by $\mathbf{X}$. Let us make that point precise. Denote by $\widehat{\mathcal{N}}$ the one point compactification of $\mathcal{N}$ ($\widehat{\mathcal{N}} = \mathcal{N}$ if $\mathcal{N}$ is compact) and set $C_p$ the space of continuous paths $z : [0,1] \to \widehat{\mathcal{N}}$ starting at $p$ such that $z_{t+\cdot} \equiv \infty$ whenever $z_t = \infty$. Fix $d$ a Riemannian metric on $\mathcal{N}$ such that $d(p, p') \to \infty$ as $p' \to \infty$, and define on $C_p$ the smallest topology containing, for any $\gamma \in C_p$ and $R, \varepsilon > 0$, the set of paths $z \in C_p$ satisfying

$$\max_{\substack{t\geq 0\\ d(p,\gamma_t)\leq R}} d(z_t, \gamma_t) < \varepsilon.$$

The topology does not depend on $d$, and a sequence $z^n$ of curves in $C_p$ converges to $z^\infty$ if and only if for all $R$, the curves $z^n_{\cdot\wedge\tau_R}$ stopped when they get at distance $R$ of $p$ converge uniformly to $z^\infty_{\cdot\wedge\tau_R}$. We can now state what one might call a theorem of continuity in law, in the following form.

**Theorem 3.7.** *For some fixed $1/3 < \gamma \leq 1/2$, let $(\mathbf{X}^n)_{n\geq 0}$ be a sequence of random $\gamma$-rough paths with values in $\mathbb{R}^d$, whose distributions converge weakly to that of $\mathbf{X}^\infty$. These processes might be defined on different probability spaces.*

*Then, for any $0 \leq n \leq \infty$, there exists a unique random variable $z^n$ with values in $C_p$ such that it solves the rough differential equation*

$$\mathrm{d}z^n_t = A(z^n_t)\mathbf{X}^n_{\mathrm{d}t}\,, \quad z^n_0 = p$$

*almost surely up to explosion, and the distributions of $z^n$ converge to that of $z^\infty$ with respect to the topology of $C_p$ described above.*

**The interpolation result**

The proof of Theorem 1.3 then follows from the rough path convergence of the rough path lift $\mathbf{X}^\sigma$ of anisotropic kinetic Brownian motion $X^\sigma$ in $\mathbb{R}^d$, Theorem 3.6, and the continuity properties of the Itô-Lyons solution map to rough differential equations. As in [ABT15], one needs to use the stochastic and geodesic completeness of $(\mathcal{M}, g)$ to conclude that the convergence of the $O\mathcal{M}$-valued development of anisotropic kinetic Brownian in $\mathbb{R}^d$ in not only local, but that weak convergence holds true; see Proposition 2.4.3 and Lemma 2.4.4 in [ABT15]. Stochastic completeness refers here to the isotropic Brownian motion on $\mathcal{M}$. We implicitly use here the fact that for a complete and stochastically complete Riemannian manifold, the anisotropic Brownian motion on $\mathcal{M}$ is also stochastically complete.

# 4 A wider class of kinetic ergodic motions

In this section, we take a step back, and see what remains of Theorem 1.3 in a higher level of generality. Suppose that $(v_t^\sigma)_{t\geq 0}$ is of the form $v_t^\sigma = I(\overline{v}_{\sigma^2 t})$, with $(\overline{v}_t)_{t\geq 0}$ a càdlàg Markov process with values in some manifold $\mathcal{W}$ and $I : \mathcal{W} \to \mathbb{R}^d$ bounded continuous — Theorem 1.3 deals with the case $I : \mathcal{W} = \mathbb{S}^{d-1} \hookrightarrow \mathbb{R}^d$ and $\overline{v}$ the anisotropic Brownian motion with time scale 1. Because the path $(x_t^\sigma)_{t\geq 0}$ integrating the velocity is Lipschitz, its development on a Riemannian manifold is well-defined, and the objects described in Theorem 1.4 make sense. We first restate and prove it, in the form of Theorem 4.1, then discuss some examples in Section 4.2.

## 4.1 From kinetic Brownian motion to kinetic ergodic processes

This subsection is devoted to the proof of the following rewriting of Theorem 1.4. We discuss its hypotheses in the following Subsection 4.2.

**Theorem 4.1** (Main theorem — Abstract version)**.** *Let $(\mathcal{M}, g)$ be a Riemannian manifold of dimension $d$, and $(q_t^\sigma)_{t\geq 0}$ a process on $\mathcal{M}$ whose velocity $\dot{q}_t^\sigma \in T_{q_t}\mathcal{M}$ has image $v_t^\sigma \in T_{q_0}\mathcal{M} \simeq \mathbb{R}^d$ under the inverse stochastic parallel transport along $q$. Suppose that, for some càdlàg Markov process $\overline{v}$ on a manifold $\mathcal{W}$, $(v_t^\sigma)_{t\geq 0}$ is the continuous image of $(\overline{v}_{\sigma^2 t})_{t\geq 0}$, i.e. $v_t^\sigma = I(\overline{v}_{\sigma^2 t})$ with $I : \mathcal{W} \to T_{q_0}\mathcal{M}$ bounded continuous. Suppose that $\overline{v}$ admits an invariant measure $\mu$ such that under $\mathbb{P} = \mathbb{P}_\mu$,*

1. *equation (1.4) holds with $\mathcal{F}_{[a,b]}$ the $\sigma$-algebra generated by $\{I(\overline{v}_t)\}_{a\leq t<b}$;*
2. *for all $1 \leq i \leq d$, the flippings $(v^1, \cdots, v^{i-1}, -v^i, v^{i+1}, \cdots, v^d)$ have the same distribution as $v = v^\sigma = (v^1, \cdots, v^d)$ for some, hence all, $\sigma > 0$.*

*Then as $\sigma \to \infty$, the time rescaled process $(q_{\sigma^2 t}^\sigma)_{t\in[0,1]}$ converges in law to an anisotropic Brownian motion on $\mathcal{M}$ with covariance $\mathrm{diag}(\gamma_1, \cdots, \gamma_d)$,*

$$\gamma_i := \int_0^\infty \mathbb{E}\left[I(\overline{v}_0)^i I(\overline{v}_t)^i\right] \mathrm{d}t.$$

**Remark 4.2.** Condition (2) above is indeed necessary. Assuming only that $I(\overline{v})$ is centred, the tightness result stated in Corollary 3.3 still holds, as well as the Brownian behaviour of the Euclidean path as shown in Proposition 3.5; however, the limit rough path need not be Brownian — see example 4.2 below. In particular, there is no reason for the manifold-valued result to hold.

In the common 'rolling without slipping' analogy used to described stochastic development of Brownian motion, one might think of the resulting non-Brownian effect as a force rotating the paper around the contact point, so that the path on the manifold may have a tendency to lean to one side. △

**Remark 4.3.** Throughout our study, we have worked at equilibrium, with $\mathbb{P} = \mathbb{P}_\mu$. Although it simplifies the proofs, it is merely a cosmetic concern in the case of kinetic Brownian motion. In fact, under the assumption (2.6), Theorem 4.1 holds for any $\mathbb{P}_\lambda$: see Proposition 4.4 below. For instance, it will be the case in examples 4.2 and, to some extent, 4.2 below. It is not clear whether the result should hold without this additional property. △

To establish Theorem 4.1, let us review the ingredients of the proof of Theorem 1.3. The tightness results, more specifically Corollary 3.3, are essentially a consequence of Lemma 2.6. It holds whenever (1.4) is satisfied (condition (1)), $I(\overline{v}_0)$ is centred (condition (2)) and $I$ is bounded. On the other hand, the convergence towards Brownian motion relies, in addition, on the symmetry property (condition (2)) and independence of the increments. Equation (1.4) ensures the latter, so that the proof of Theorem 4.1 is essentially that of Theorem 1.3.

**Proposition 4.4.** *Replace condition* (1) *in Theorem 4.1 by the following variant of* (2.6)*. There exists some mixing time* $\tau > 0$ *such that for all* $x \in \mathcal{W}$ *and* $t > 0$*,*

$$\|P_t^* \delta_x - \mu\|_{\mathrm{TV}} \leq f(x) \exp(-t/\tau) \tag{4.12}$$

*for some function* $f : \mathcal{W} \to \mathbb{R}_+$ *integrable with respect to* $\mu$*.*

*Then the conclusion also holds under* $\mathbb{P}_\lambda$*, for any probability measure* $\lambda$ *on* $\mathcal{W}$ *such that* $\lambda(f) := \int f \mathrm{d}\lambda < \infty$*.*

*Proof.* It is enough to show the convergence of the Euclidean rough paths $(\mathbf{X}^\sigma)_{\sigma>0}$.

*Tightness.* We claim that Proposition 2.4 holds for $\mathbb{E}_\lambda$. Indeed, by the same arguments, we see that

$$|\mathbb{E}_{P_u^* \delta_x}[G] - \mathbb{E}_\mu[G]| \leq |G|_\infty \, f(x) \, \mathsf{e}^{-u/\tau} \tag{4.13}$$

holds in lieu of (2.7). From this we deduce

$$\begin{aligned} |\mathbb{E}_{P_{t-s}^* \delta_x}[G] - \mathbb{E}_{P_t^* \lambda}[G]| &\leq |\mathbb{E}_{P_{t-s}^* \delta_x}[G] - \mathbb{E}_\mu[G]| + |\mathbb{E}_\mu[G] - \mathbb{E}_{P_t^* \lambda}[G]| \\ &\leq (f(x) + \lambda(f) \, \mathsf{e}^{-s/\tau}) |G|_\infty \, \mathsf{e}^{-(t-s)/\tau}, \end{aligned}$$

which is enough for rest of the proof to hold. It is then an easy exercise to adapt the proof of Corollary 2.5, and from this point every idea leading to tightness is the same, even if some care must be given to non-stationarity in the actual computations, e.g. regarding equation (3.8).

*Brownian limit.* Let $\widehat{\mathbb{P}}_\mu$ be the law of the Brownian rough path on $\mathsf{RP}(\gamma)$, and $\widehat{\mathbb{P}}_\lambda$ a limit point of the laws of $\mathbf{X}^\sigma$ under $\mathbb{P}_\lambda$. We only need to show that $\widehat{\mathbb{P}}_\lambda = \widehat{\mathbb{P}}_\mu$.

Define the translation operator $T_h$ on $\mathsf{RP}(\gamma)$ as

$$T_h(Y, \mathbb{Y}) := (Y_{h+\cdot} - Y_h, \mathbb{Y}_{h+\cdot, h+\cdot}).$$

Now, for any continuous bounded map $F : \mathcal{C}([0,1], G) \to \mathbb{R}$ and $\varepsilon > 0$, equation (4.13) gives

$$|\mathbb{E}_\lambda[F(T_\varepsilon \mathbf{X}^\sigma)] - \mathbb{E}_\mu[F(T_\varepsilon \mathbf{X}^\sigma)]| \leq \lambda(f) |F|_\infty \, \mathsf{e}^{-\sigma^4 \varepsilon/\tau},$$

which, taking limits along a proper subsequence, implies that

$$\widehat{\mathbb{E}}_\lambda[F(T_\varepsilon \mathbf{X})] = \widehat{\mathbb{E}}_\mu[F(T_\varepsilon \mathbf{X})].$$

But $T_\varepsilon \mathbf{X} \to \mathbf{X}$ in $\mathcal{C}^0([0,1], G)$, so the above equation holds for $\varepsilon = 0$, and $\widehat{\mathbb{P}}_\lambda$ is the law of the announced anisotropic Brownian motion. Note that $T_\varepsilon \mathbf{X}$ has no reason to converge to $\mathbf{X}$ in the rough path topology, so tightness had to be proved beforehand. $\square$

## 4.2 Discussion of the hypotheses

Because we introduced Theorem 4.1 merely as a reformulation of Theorem 1.3 with highlight on the key assumptions, there is much room for improvement. However, it is tedious to state a general theorem without hiding the ideas under a layer of abstract context, unnecessary in most applications. Instead, we chose in this last section to discuss the hypotheses of 4.1 through a variety of examples, which we believe conveys the versatility of the method.

### Spinning motion

The first example illustrates what happens when the motion does not satisfy the symmetry condition (2) in Theorem 4.1 above. Set $I : \mathcal{W} = \mathbb{R}/2\pi\mathbb{Z} \to \mathbb{C} \simeq \mathbb{R}^2$ the exponential $v \mapsto \mathsf{e}^{\mathsf{i}v}$, and define $\overline{v}$ as the spinning motion

$$\mathrm{d}\overline{v}_t = \mathrm{d}t + \mathrm{d}W_t, \quad \text{i.e.} \quad \overline{v}_t = \overline{v}_0 + t + W_t \pmod{2\pi},$$

where $W$ is a standard Brownian motion on $\mathbb{R}$. Its dynamics is of course very simple: it admits a unique invariant measure $\mu(\mathrm{d}v) = \frac{1}{2\pi}\mathrm{d}v$ and satisfies equation (2.6), so all hypotheses but condition (2) in Theorem 4.1 above are satisfied.

As mentioned in Remark 4.2 above, the laws of $(X^\sigma)_{\sigma>0}$ do converge to that of a Brownian process. As of those of the lifts $(\mathbf{X}^\sigma)_{\sigma>0}$, however, some drift appears in the limit. Indeed, setting $\mathbb{A}^\sigma$ the antisymmetric part of $\mathbb{X}^\sigma$,

$$(\mathbb{A}^\sigma_{t0})^{12} = \frac{1}{2\sigma^4}\int_0^{\sigma^4 t}\int_0^s \sin(\overline{v}_s - \overline{v}_u)\mathrm{d}s\mathrm{d}u = \int_0^\infty\int_0^\infty \frac{1}{2\sigma^4}\mathbf{1}_{u+\tau\le\sigma^4 t}\sin(\overline{v}_{u+\tau} - \overline{v}_u)\mathrm{d}u\mathrm{d}\tau,$$

so we get

$$\mathbb{E}\big[(\mathbb{A}^\sigma_{t0})^{12}\big] = \int_0^\infty\int_0^\infty \frac{1}{2\sigma^4}\mathbf{1}_{u+\tau\le\sigma^4 t}\sin(\tau)\,\mathsf{e}^{-\tau/2}\mathrm{d}u\mathrm{d}\tau = \frac{1}{2}\int_0^\infty \Big(t - \frac{\tau}{\sigma^4}\Big)_+ \sin(\tau)\,\mathsf{e}^{-\tau/2}\mathrm{d}\tau$$

with $(\cdot)_+$ the positive part. The limit is a non-zero linear function of $t$, so the limit of the lifts cannot be Brownian.

Such drift phenomena in the Lévy area have arisen and been studied in different works recently, particularly in the context of random walks. See e.g. the articles [LS18, LS17] of Lopusanschi and Simon, and those of Ishiwata, Kawabi and Namba, [IKN18a, IKN18b].

### Random flight

The so-called random flight is perhaps the most elementary situation where the velocity has jumps. Studied by Pinsky under the name of isotropic transport process, it is shown in [Pin76] to exhibit a limiting Brownian behaviour in manifolds. It can be described in terms of our $I$ and $\overline{v}$ by setting $I : \mathcal{W} = \mathbb{S}^{d-1} \hookrightarrow \mathbb{R}^d$ and $\overline{v}$ a pure jump process, with rate 1 and uniform measure. In this case, the mixing property (1.4) is a consequence of the stronger statement (2.6) that the dynamics converges exponentially fast to equilibrium in total variation, in the same way we treated anisotropic Brownian motion, so that our Theorem 4.1 applies readily. There are no complications in dealing with jumps.

Because the velocity is isotropic, the limit covariance $\mathrm{diag}(\gamma_1, \cdots, \gamma_d)$ is proportional to Id. Setting $T$ the first jump time,

$$\gamma_i = \frac{2}{d} \int_0^\infty \mathbb{E}[\overline{v}_0 \cdot \overline{v}_t] \mathrm{d}t = \frac{2}{d} \int_0^\infty \mathbb{P}(T \leq t) \mathrm{d}t = \frac{2}{d},$$

and we recover the result of [Pin76].

**Donsker invariance principle for random walks**

In certain situations, the ergodic properties of the system are most easily described in discrete time. An example, studied in [BFH09] by Breuillard, Friz and Huesmann, is that of random walks. If $(Y_k)_{k \geq 0}$ is a sequence of independent *bounded* random variables with values in $\mathbb{R}^d$, *symmetric* in the sense that their common law is invariant with respect to the flippings as described in condition (2) in Theorem 4.1 above, we consider the piecewise linear process $W$ that, on each interval $[k, k+1]$, is affine and increases by $Y_k$. In other words, we express it as

$$W : t \mapsto \sum_{k < \lfloor t \rfloor} Y_k + (t - \lfloor t \rfloor) Y_n.$$

The classical invariance principle of Donsker states that the rescaled processes $W^\sigma : t \mapsto W_{\sigma^4 t}/\sigma^2$ converge in law to an anisotropic Brownian motion, with respect to the uniform convergence on compact sets; the result of [BFH09] strengthen it to rough path convergence.

Let us translate this dynamics in our framework. Set $\mathcal{W} = \mathbb{R}/\mathbb{Z} \times \mathbb{R}^d$, $I : (\alpha, y) \mapsto y$, and define the dynamics of $(\overline{v}_t)_{t \geq 0} = (\alpha_t, y_t)_{t \geq 0}$ as follows. Given initial conditions $(\alpha_0, y_0) \in [0, 1) \times \mathbb{R}^d$, $\alpha$ grows continuously with rate 1, i.e. $\alpha_t = \alpha_0 + t \pmod 1$, whereas $y$ stays constant on time intervals of length 1, then jumps independently of the past according to the law of $Y_k$, i.e. $y_t = Y_{\lfloor t - \alpha_0 \rfloor}$ with the convention $Y_{-1} = y_0$. Under the initial condition $\delta_0 \otimes \mathcal{L}(Y_0)$, we see that the law of $X^\sigma$ is exactly that of $W^\sigma$.

The process $\overline{v}$ is Markovian, although not Feller, and admits an invariant measure $\mathrm{Unif}(\mathbb{R}/\mathbb{Z}) \otimes \mathcal{L}(Y_0)$. Because it is not ergodic, there is no hope for equation (2.6) to hold. Maybe surprisingly, even if $I$ kills the non-mixing coordinate, it is also false that condition (1) of Theorem 4.1 holds: in the case where $Y$ has no atoms, take $P$ to be the first jump time in $[0, 1]$, and $F$ the first jump time in $[n, n+1]$. However, it is true for any $\mu_\alpha := \delta_\alpha \otimes \mathcal{L}(Y_0)$, with constants independent of $\alpha$: indeed, it is obvious that for $t > 1$ and any probability law $\lambda$ on $\mathbb{R}^d$,

$$P_t^*(\delta_\alpha \otimes \lambda) = \mu_{\alpha + t} = P_t^* \mu_\alpha$$

holds in lieu of (2.6). Remarkably, nothing more than this is needed throughout the proof. It should be clear that Proposition 2.5 holds for any $\mu_\alpha$, and that tightness follows in the same fashion. Independence of increments, as stated in Propositions 3.5 and Theorem 3.6, hides no difficulty either. It is true that one has to be careful about the limit variance in Proposition 3.5, because the Markov property is used in a crucial way. In our case, for any $\alpha \in [0, 1)$ and $\sigma > 1$, we end up with

$$\begin{aligned}
\mathbb{E}_{\mu_\alpha}\big[\big((X_1^\sigma)^i\big)^2\big] &= \frac{1}{\sigma^4}\int_0^{\sigma^4}\int_0^{\sigma^4}\mathbb{E}_{\mu_\alpha}[y_s^i y_t^i]\,\mathrm{d}s\,\mathrm{d}t \\
&= \sum_{n\geq -1}\frac{1}{\sigma^4}\int_0^{\sigma^4}\int_0^{\sigma^4}\mathbf{1}_{n+\alpha\leq s,t<n+1+\alpha}\mathbb{E}_{\mu_\alpha}[y_s^i y_t^i]\,\mathrm{d}s\,\mathrm{d}t \\
&= \mathbb{E}[|Y_0^i|^2]\cdot\frac{(1-\alpha)^2+\lfloor\sigma^4+\alpha-1\rfloor+\{\sigma^4+\alpha-1\}^2}{\sigma^4}
\end{aligned}$$

with $\{\cdot\}$ the fractional part. In the limit, the variance converges to $\mathbb{E}[|Y_0^i|^2]$, and the result of Theorem 1.3 holds with covariance $\mathbb{E}[Y_0Y_0^*]$, in accordance with [BFH09].

Surprisingly enough, the symmetry condition (2) is not mandatory here: see [BFH09]. In particular, a drift in the antisymmetric part of $\mathbb{X}^\sigma$ that does not vanish in the limit, as mentioned in example 4.2 above, must come from additional structure: in [LS18], the hidden Markov chain; in [LS17], the underlying directed graph; etc.

Note that in the case of random walks, as a consequence of the work of Chevyrev, see [Che18, Example 5.8], convergence of $X^\sigma$ as stated in Proposition 3.5 in enough to ensure convergence of $\mathbf{X}^\sigma$ to some random rough path. It is not clear from this approach, however, that this limit is indeed Brownian.

### Donsker invariance principle for Markov chains

The reader may have noticed that in the above example 4.2, independence of the variables $(Y_k)_{k\geq 0}$ is a bit much, and one could work with covariances vanishing exponentially fast. Suppose for instance that $(Y_k)_{k\geq 0}$ is a time-homogeneous Markov chain with invariant measure $\mu$ with compact support, whose correlations decrease as $\mathsf{e}^{-k/\tau}$, $\tau>0$; namely, letting $Q$ be the transition kernel of $Y$,

$$\|\delta_y\, Q^k-\mu\|_{\mathrm{TV}}\lesssim \mathsf{e}^{-k/\tau}$$

for all $y$ in the support of $\mu$. Then, setting $\mu_0:=\delta_0\otimes\mu$, we get, for any probability measure $\lambda$ with Supp $\lambda\subset$ Supp $\mu$,

$$\|P_t^*(\delta_0\otimes\lambda)-P_t^*\mu_\alpha\|_{\mathrm{TV}}=\left\|\int(\delta_y\,Q^{\lfloor t\rfloor}-\mu)\lambda(\mathrm{d}y)\right\|_{\mathrm{TV}}\lesssim \mathsf{e}^{-t/\tau}.$$

Again, this inequality can be substituted for equation (2.6) in the proof of Proposition 2.4, and under the same symmetry condition as above, the convergence result still holds true.

Examples of such Markov chains are any aperiodic irreducible finite state Markov chain; or any Markov chain with transition kernel $Q(y,\mathrm{d}y')$ absolutely continuous with respect to some measure $\nu$, and such that $\frac{\mathrm{d}Q(y,\cdot)}{\mathrm{d}\nu}$ is bounded below by a positive constant $m>0$, uniformly in $y,y'$. Note however that the symmetry condition (2) of Theorem 4.1 is a bit stronger than in the independent case, since we need the flippings to leave the law of the whole sequence invariant.

### Time-dependent Brownian motion

The way we wrote our convergence theorems is ill-suited to treat time-dependent randomness, i.e. when the motion of the velocity is not homogeneous in time. However, there are cases where randomness can be somewhat dissociated from the time dependence, and our methods do in fact yield interesting convergence results. In the present example, we set to recover, in the limit, the Brownian motion on a manifold $\mathcal{M}$ endowed with a time-dependent metric $g_t$, as introduced in [ACT08] by Arnaudon, Coulibaly and Thalmaier.

Such an approach has already been set up in [Kuw12], in a similar fashion as the random flight described in example 4.2 above. The idea is to freeze the metric in small time intervals $[t_i, t_{t+1}]$, say of size $1/\sigma^4$, over which the movement $q$ is purely geodesic with respect to the metric $g_{t_i}$, the initial condition being chosen uniformly at $t_i$ on the unit $g_{t_i}$-sphere of the tangent space of $\mathcal{M}$ at $q_{t_i}$. Suitably renormalised, this process converges to the time-dependent Brownian motion described above. We introduce a similar random flight which lets the metric vary continuously, and may be considered more natural in this respect, then prove its convergence to time-dependent Brownian motion.

We begin by describing time-dependent Brownian motion and its surroundings. Suppose $g_t$ is smooth, as a function on $\mathbb{R}_+ \times T\mathcal{M} \otimes T\mathcal{M}$. Let $F\mathcal{M}$ be the frame bundle over $\mathcal{M}$, and choose a point $q_0 \in \mathcal{M}$ together with a $g_0$-orthonormal frame $e_0$ of $T_{g_0}\mathcal{M}$. For a $\mathcal{C}^1$ path $(x_t)_{t\geq 0}$ in $\mathbb{R}^d$, we define the time-dependent development of $x$ as the solution $(z_t)_{t\geq 0} = (q_t, e_t)_{t\geq 0}$ of the following equation, whose terms we describe below.

$$\mathrm{d}z_t = \mathrm{H}_{t,z_t}(\mathrm{d}x_t) - \frac{1}{2}\frac{\partial g_t}{\partial t}(e_t\epsilon_i, e_t\epsilon_j)\mathrm{V}^{ij}_{z_t}\mathrm{d}t, \quad z_0 = (q_0, e_0). \tag{4.14}$$

We use Einstein notation. As in Section 3.3, $(\epsilon_1, \cdots, \epsilon_d)$ is the canonical basis of $\mathbb{R}^d$, and the $\mathrm{H}_{t,z}\epsilon_i$, resp. $\mathrm{V}^{ij}_z$, are the canonical horizontal vector fields, resp. vertical vector fields. Note that because the metric $g$ is time-dependent, the associated horizontal vector fields H must depend on $t$ as well. In coordinates,

$$\mathrm{H}_{t,z}(\epsilon_\alpha) = e^i_\alpha \frac{\partial}{\partial q^i} - \big(\Gamma_t(q)\big)^k_{ij}\, e^i_\alpha e^j_l \frac{\partial}{\partial e^k_l}, \qquad \mathrm{V}^{ij}_z = e^k_j \frac{\partial}{\partial e^k_i}.$$

If we compare (4.14) to (3.10), the added vertical fields are there to ensure that $e_t$ is at all times orthonormal for $g_t$. We refer to [CP11] for an insight about why this definition is a sensible choice.

In particular, the time-dependent geodesics are the solutions of the equation associated to $x_t = tu$ for some fixed $u \in \mathbb{R}^d$, and the time-dependent Brownian motion is the solution driven by some standard Brownian motion $W$ in the Stratonovich sense, or, equivalently, by the standard Stratonovich rough path $\mathbf{W}$ in the rough sense.

Note that we did not discuss time-dependent rough differential equations in Section 3.3. In the case of an equation driven by a $\mathcal{C}^1$ control $x$, the standard technique is of course to consider $t \mapsto (t, x_t)$ as the control. The same trick works with rough paths: associated to any rough path $\mathbf{Y} = (Y, \mathbb{Y})$ is a canonical lift $\widehat{\mathbf{Y}}$ of $t \mapsto (t, Y_t)$ compatible with $\mathbf{Y}$. The solution of time-dependent rough differential equations is then well-defined. In what follows, we will also use the fact that $\mathbf{Y} \mapsto \widehat{\mathbf{Y}}$ is continuous in the rough path topology, so $\widehat{\mathbf{X}}^\sigma \to \widehat{\mathbf{X}}$ in law whenever $\mathbf{X}^\sigma \to \mathbf{X}$ in law.

We define a kind of interpolated random walk on $\mathcal{M}$ whose limit will be the Brownian motion described above. Fix $\sigma > 0$, and define $W^\sigma$ successively on each interval $[s,t] = [\frac{n}{\sigma^4}, \frac{n+1}{\sigma^4}]$ as follows: $\xi^\sigma_n$ is chosen independently of all the rest according to the uniform measure on the unit $g_s$-sphere of $T_{W^\sigma_s}\mathcal{M}$, and $W^\sigma$ on $[s,t]$ is a time-dependent geodesic in the above sense, with initial condition $\dot{W}^\sigma_s = \sqrt{d}\,\xi^\sigma_n$.

As in the previous example, there is a direct equivalent of this dynamics in our framework. Set $\mathcal{W} = \mathbb{R}/\mathbb{Z} \times \mathbb{S}^{d-1}$ and $I : (\alpha, y) \hookrightarrow \sqrt{d}y$, following the same dynamics as in 4.2, with $Y_0$ uniformly distributed on $\mathbb{S}^{d-1}$. We choose the initial condition to be $\delta_0 \otimes \mathrm{Unif}(\mathbb{S}^{d-1})$; for the same reasons as in example 4.2 above, $(\mathbf{X}^\sigma)_{\sigma>0}$ converges to the Brownian rough path $\mathbf{X}$ with covariance $d\,\mathbb{E}[Y_0Y_0^*] = \mathrm{Id}$.

Everything described so far is essentially time-invariant — the time-dependence appears when we use this family of rough paths to describe a motion on $\mathcal{M}$. Fix $q_0 \in \mathcal{M}$, and $e_0$ a $g_0$-orthonormal frame of $T_{q_0}\mathcal{M}$. Define the solution $(z_t)_{t\geq 0} = (q_t, e_t)_{t\geq 0}$ (up to explosion) on the frame bundle $F\mathcal{M}$ of equation (4.14) driven by $\mathbf{X}$, in the rough sense.

By definition, $q_t$ defined as above is the Brownian motion associated to the time-dependent metric $g_t$, as described in [ACT08]. If we set $z^\sigma = (q^\sigma, u^\sigma)$ the solution of the equation driven by $\mathbf{X}^\sigma$, we get instead $q^\sigma = W^\sigma$ in law. The convergence of $\mathbf{X}^\sigma$, together with the general theory of rough paths (see Theorem 3.7), ensures that $q^\sigma$, hence $W^\sigma$, converges in law to the time-dependent Brownian motion $q$.

### Langevin Process

We conclude with an example where the velocity $v$ has unbounded support, which goes against the implicit assumption that $I$ be bounded. We consider the process with anisotropic Ornstein-Uhlenbeck velocity, i.e. satisfying

$$\mathrm{d}\overline{v}_t = -\overline{v}_t \mathrm{d}t + \mathrm{d}B_t$$

for $B$ an anisotropic Brownian motion of covariance $\Sigma$. In the isotropic case, it is a simple scalar example of the hypoelliptic Laplacian of Bismut; see [Bis15]. The anisotropic case is also treated in [BHVW17].

Here, $I : \mathcal{W} = \mathbb{R}^d \to \mathbb{R}^d$ is simply the identity, and hence does not quite fit the hypotheses of Theorem 1.4. However, it is well known that $\overline{v}$ admits as an invariant measure the Gaussian distribution $\mu = \mathcal{N}(0, \frac{1}{2}\Sigma)$ with covariance $\frac{1}{2}\Sigma$. Using the coupling $B'_t = -B_t$, it is known, and not difficult to see, that

$$\|P_t^*\delta_x - \mu\|_{\mathrm{TV}} \lesssim (1 \vee |x|)\, \mathsf{e}^{-t},$$

from whence, because $1 \vee |x|$ is in $L^1(\mu)$, we derive Proposition 2.4; see Proposition 4.4.

In our proof, boundedness of the velocity is essentially used twice: for proving the decorrelation of coordinates in Proposition 2.5, and to show that the variance of the limit must be the limit of the variances in 3.5. Because $\mu$ has moments of all order, the latter will add no difficulty — in fact, any moment of order $> 2$ would suffice. As for the former, it is a bit trickier. We use the following variation of Proposition 2.5.

**Proposition 4.5.** *Fix some $\varepsilon > 0$ and some positive integer $n \in \mathbb{N}^*$. There exists $\tau' = \tau'(\tau, n, \varepsilon) > 0$ such that under $\mathbb{P} = \mathbb{P}_\mu$, and for any indices $1 \leq j_1, \cdots, j_n \leq d$ and times $s_1, \cdots, s_n \geq 0$,*

$$\left|\mathbb{E}\left[v_{s_1}^{j_1} \cdots v_{s_1+\cdots+s_n}^{j_n}\right]\right| \lesssim |v_0^{j_1}|_{L^{n+\varepsilon}} \cdots |v_0^{j_n}|_{L^{n+\varepsilon}}\, \mathsf{e}^{-\Delta/\tau'}.$$

We give only hints of the proof. In the spirit of the proof of Proposition 2.5, set

$$V_- := \prod_{1\leq k<k_0} \left(v_{t_k}^{j_k} / |v_0^{j_k}|_{L^{n+\varepsilon}}\right)$$

and similarly for $V_0$ and $V_+$. Write $V_* = W_* + R_*$ with $W_* := V_* \mathbf{1}_{|V_*|\geq M}$; for $M = \exp(\eta\Delta)$ with $\eta > 0$ small enough, the proof of Proposition 2.5 applied to $W_*$, together with a careful handling of the remainder $R_*$, are enough to get to the above result. It automatically implies Lemma 2.6, since $\mu$ has moments of all order, hence the conclusion of Theorem 1.3.

Note that the treatment of unboundedness is not specifically designed for the Langevin process, so it can be applied to the study of the random walk as well. Moreover, it is not necessary for all moments to exist: moments of order $\alpha > 2/(1-2\gamma)$ are enough to ensure tightness in $\mathsf{RP}(\gamma)$. Indeed, our proof, enhanced by the above corollary, will hold with moments of order $2n > 2/(1-2\gamma)$ for any positive integer $n$; but adding an easy truncation argument at the beginning of the proofs of Lemma 3.1 and 3.2 will strengthen the result to non even integral moments. In this respect, our moment assumption is a bit weaker than that of [BFH09] in the symmetrical case.

# Chapter III

# Homogenisation for configuration spaces of manifolds

This chapter is based on the article [ABP19] written with my advisors. J. Angst, I. Bailleul and myself sent this work to arXiv.org in May 2019, and it is currently under review.

## 1 Introduction

Kinetic Brownian motion is a purely geometric random perturbation of geodesic motion. In its simplest form, in $\mathbb{R}^d$, the sample paths of kinetic Brownian motion are $C^1$ random paths run at unit speed, with velocity a Brownian motion on the unit sphere, run at speed $\sigma^2$, for a positive speed parameter $\sigma$. More formally, it is a hypoelliptic diffusion with state space $\mathbb{R}^d \times \mathbb{S}^{d-1}$, solution to the stochastic differential equation

$$\begin{aligned} dx_t^\sigma &= v_t^\sigma\, dt, \\ dv_t^\sigma &= \sigma\, P_{v_t^\sigma}(\circ dW_t), \end{aligned}$$

for $P_a : \mathbb{R}^d \to \langle a \rangle^\perp$, the orthogonal projection on the orthogonal of $\langle a \rangle$, for $a \neq 0$ in $\mathbb{R}^d$, and $W$ a standard $\mathbb{R}^d$-valued Brownian motion. If $\sigma = 0$, we have a straight line motion with constant velocity. For a fixed $0 < \sigma < +\infty$, we have a $C^1$ random path, whose typical behavior is illustrated in Figure 1.1 below.

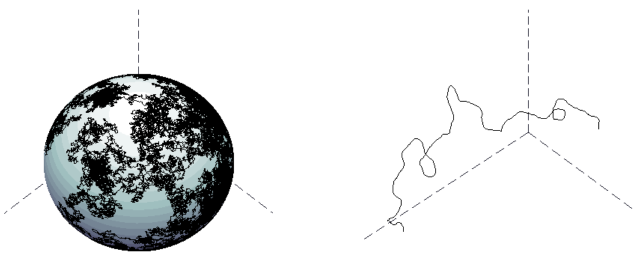


FIGURE 1.1 – Brownian motion on the sphere and its integral path in $\mathbb{R}^3$.

For $\sigma$ increasing to infinity, the exponentially fast decorrelation of the velocity process $v^\sigma$ on the sphere implies that the position process $x^\sigma$ converges to the constant path $x_0$, if the latter is fixed independently of $\sigma$. One has to rescale time and look at the evolution at the time scale $\sigma^2$ to see a non-trivial limit. It is indeed elementary to prove that the time rescaled position process $(x^\sigma_{\sigma^2 t})_{0\leq t\leq 1}$ of kinetic Brownian motion converges weakly in $C\big([0,1],\mathbb{R}^d\big)$ to a Brownian motion with generator $\kappa_d\,\Delta_{\mathbb{R}^d}$, with $\kappa_d := \frac{4}{d(d-1)}$. This homogenization result is in fact valid on a general finite dimensional Riemannian manifold $M$, under very mild geometric assumptions.

Kinetic Brownian motion on a $d$-dimensional Riemannian manifold $M$ is defined as Cartan development $\big(m^\sigma_t, \dot m^\sigma_t\big)$ in the unit tangent bundle $T^1M$ of $M$ of kinetic Brownian motion in $\mathbb{R}^d$. It is a geodesic for $\sigma = 0$, and a $C^1$ random path for a finite positive value of $\sigma$. It was first proved by X.-M. Li in [Li12] that the time-rescaled position process $(m^\sigma_{\sigma^2 t})_{0\leq t\leq 1}$ converges weakly to Brownian motion with generator $\kappa_d\,\Delta_M$. See Figure 2 below for an illustration in the setting of the flat 2-dimensional torus. The manifold $M$ was assumed to be compact and martingale methods were used to prove that homogenization result. X.-M. Li then extended this result in [Li16a] to non-compact manifolds subject to a growth condition on their curvature tensor. In [ABT15], Angst, Bailleul and Tardif gave the most general result, assuming only geodesic and stochastic completeness, using rough paths theory as a working horse to transport a rough path convergence result about kinetic Brownian motion in $\mathbb{R}^d$ to the manifold setting.

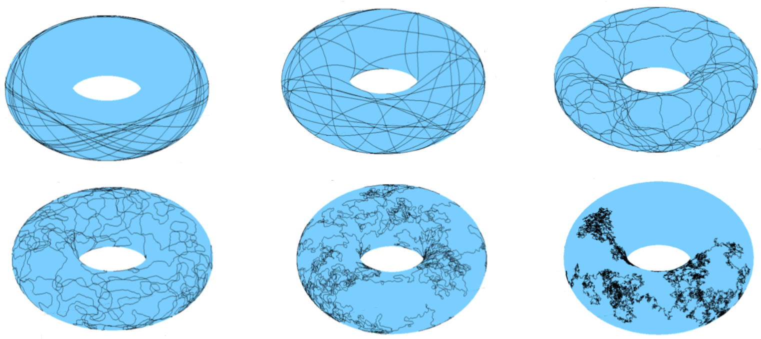


FIGURE 1.2 – On the two dimensional flat torus, typical examples of sample paths of kinetic Brownian motion $\big(x^\sigma_{\sigma^2 t}\big)_{0\leq t\leq 1}$ as $\sigma$ increases.

See also [Li18] for further results in homogeneous spaces, and [Per18] for a generalization of the homogenization result of [ABT15] to anisotropic kinetic Brownian motion, or more general Markov processes on $T^1M$. Note that the dynamically obvious convergence of the unrescaled kinetic Brownian motion to the geodesic motion has been studied from the spectral point of view in [Dro17], for compact manifolds with negative curvature, showing that the $L^2$ spectrum of the generator of the unrescaled kinetic Brownian motion converges to the Pollicott-Ruelle resonances of $M$. Other examples of homogenization results for Langevin-type processes include works by Hottovy and co-authors, amongst others; see e.g. [BHVW17, HHV16, BW18, LWL19] for quantitative convergence results. See also [Sol95, Kol00, AHK12, Gli11] for other works on Langevin dynamics in a Riemannian manifold.

This kind of homogenization result certainly echoes Bismut's program about his hypoelliptic Laplacian [Bis05, Bis15], whose probabilistic starting point is a similar interpolation result for Langevin process in $\mathbb{R}^d$ and its Cartan development on a Riemannian manifold. The dynamics is lifted to a dynamics on the space of differential forms to take advantage of the correspondence between the cohomology of differential forms and homology of $M$, via index-type theorems. See [Bis11, Bis15, Bis16, She16] for a sample of the deep results obtained by Bismut and co-authors on the hypoelliptic Laplacian.

Note also that kinetic Brownian motion is the natural Riemannian analogue of its Lorentzian counterpart, introduced first by Dudley in [Dud66] in Minkowski spacetime in the 60's. See the far reaching related works [FLJ07, Bai10, FLJ11, BF12], on relativistic diffusions in a general Lorentzian setting. No homogenization result is expected for these purely geometric diffusion processes, unless one has an additional non-geometric ingredient, e.g. in the form of a relativistic fluid flow, like in [AF07].

The object of the present work is to define and study kinetic Brownian motion in the diffeomorphism group $\mathscr{M}$, or volume preserving diffeomorphism group $\mathscr{M}_0$, of a closed Riemannian manifold $M$. As in the finite dimensional setting, we prove in our main Theorems 4.3 and 4.4 below that it provides an interpolation between the geodesic flow and an explicit Brownian flow, as the noise intensity parameter $\sigma$ ranges from 0 to $\infty$.

For $\sigma = 0$, the motion in each diffeomorphism group is purely geodesic, and it corresponds to the flow of the solutions of Euler's equation in the case of $\mathscr{M}_0$, after the seminal works of Arnold [Arn66] and Ebin & Marsden [EM69]. When considered in the setting of volume preserving diffeomorphisms, the Eulerian picture of kinetic Brownian motion thus provides a natural family of random perturbations of Euler's equations for the hydrodynamics of an incompressible fluid. There has been much work recently on random perturbations of Euler's equations, following Holm's seminal article [Hol15]. See [GBH17, CHR18, CFH18, DH18, BdLHLT19] for a sample. The structure of the noise in these works is intrinsically linked to the group structure of the diffeomorphism group, and it amounts to perturb Euler's equation for the velocity field by an additive Brownian term, with values in a space of vector fields on the fluid domain $M$. Our point of view is purely Riemannian, and does not appeal to the group structure of the diffeomorphism group of the fluid domain $M$. As in the above finite dimensional setting, we define kinetic Brownian on the diffeomorphism group as the Cartan development of its 'flat' counterpart. Unlike the group-oriented point of view, where the running time diffeomorphism is sufficient to describe its infinitesimal increment from the noise, we need here a notion of frame of the tangent space of the running diffeomorphism to build its increment from the noise. We provide an explicit description of the invariant measure of the energy of the Eulerian velocity field.

On the technical side, we use rough paths theory to transport a weak convergence result for the flat kinetic Brownian motion taking values in the tangent space to the configuration space $\mathscr{M}$, or $\mathscr{M}_0$, to a weak convergence result for the solution of a differential equation controlled by that flat kinetic Brownian motion. We use for that purpose the continuity of the Itô-Lyons solution map to a controlled ordinary differential equation, in the present infinite dimensional setting. This allows to bypass a number of difficulties that would appear otherwise if using the classical martingale problem approach, as in [Li12, Li16a]. All the material needed about rough paths theory is recalled in Section 2.4.

From a geometric point of view, the tangent space to the configuration space can naturally be seen as an infinite dimensional Hilbert space. For this reason, we define and study in the next Section 2 kinetic Brownian motion on a generic infinite dimensional Hilbert space $H$. More precisely, we provide an explicit description of the invariant measure of the velocity process in Section 2.1, and we establish exponential decorrelation identities for the latter in Section 2.2. The invariance principle for the position process associated to the time-rescaled $H$-valued kinetic Brownian motion is then established in Section 2.3. With the rough paths tools introduced in Section 2.4, Section 2.5 is then devoted to the proof of the fact that the canonical rough path above the time-rescaled position process converges weakly as a rough path to the Stratonovich Brownian rough path of a Brownian motion with an explicit covariance. Elements of the geometry of the configuration spaces $\mathscr{M}$ and $\mathscr{M}_0$ are recalled in Section 3. We develop in particular in Section 3.3 and Section 3.4 the material needed to talk about Cartan development operation as solving an ordinary differential equation driven by smooth vector fields. The final homogenisation result, proving the interpolation between geodesic and Brownian flows on the configuration spaces, is proved in Section 4 using the robust tools of rough paths theory. Appendix 5 contains the proof of a technical result about Cartan development in $\mathscr{M}_0$.

**Notations.** We gather here a number of notations that are used throughout the article.

- The letter $\gamma$ stands for a Gaussian measure on a Hilbert space $H$, with covariance $C_\gamma : H^* \times H^* \to \mathbb{R}$, and associated operator $\overline{C}_\gamma : H \to H$. The scalar product and norm on $H$ are denoted by $(\cdot,\cdot)$ and $\|\cdot\|$, respectively.

- We denote by $\mathcal{H}$ the Cameron-Martin space of the measure $\gamma$.

- We endow the algebraic tensor space $H \otimes_a H$ with its natural Hilbert norm. This amounts to identify $H \otimes H$ with the space of Hilbert-Schmidt operators on $H$.

- We use the notation $A \lesssim_p B$ for an inequality of the form $A \leq cB$, with a constant $c$ depending only on $p$.

# 2 Kinetic Brownian motion in a Hilbert space

## 2.1 Brownian motion on a Hilbert sphere

We first recall basic results on Brownian motion in $H$, and refer the reader to the nice lecture notes [Hai09, Str93] for short and detailed accounts.

Recall that a **Gaussian probability measure on** $H$ is a Borel measure $\gamma$ such that $\ell^*\gamma$ is a real Gaussian probability on $\mathbb{R}$, for every continuous linear functional $\ell : H \to \mathbb{R}$. Fernique's theorem [Fer70] ensures that

$$\int_H \exp\left(a\|x\|^2\right)\gamma(dx) < \infty,$$

for a small enough positive constant $a$. It follows that the covariance

$$C_\gamma(\ell,\ell') := \int \ell(x)\ell'(x)\,\gamma(dx), \quad \ell,\ell' \in H^*$$

is a well-defined continuous bilinear operator on $H^* \times H^*$. One can then define a continuous symmetric operator $\overline{C}_\gamma : H \to H$, by the identity

$$\left(\overline{C}_\gamma(h),k\right) = C(h,k),$$

for all $h,k \in H$. It has finite trace equal to

$$\mathrm{tr}(\overline{C}_\gamma) = \int \|x\|^2\,\gamma(dx).$$

Conversely, one can associate to any trace-class symmetric operator $\overline{C} : H \to H$, a Gaussian measure $\gamma$ on $H$ whose covariance $\overline{C}_\gamma(\ell) = \overline{C}(\ell)$, for all $\ell \in H$. Since $\overline{C}_\gamma$ is compact, there exists an orthonormal basis $(e_n)$ of $H$, such that

$$\overline{C}_\gamma(e_n) = \alpha_n^2 e_n,$$

for non-negative and non-increasing eigenvalues $\alpha_n$ with $\sum \alpha_n^2 < \infty$. We define a Hilbert space $\mathcal{H}$ by choosing $\big(\alpha_n e_n\big)$ as an orthonormal basis for it. The space $\mathcal{H}$ is continuously embedded inside $H$. Let $(X^n)$ stand for a sequence of independent, identically distributed, real-valued Gaussian random variables with zero mean and unit variance, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then the series

$$\sum_n X^n \alpha_n e_n$$

converges in $L^2(\Omega, H)$, and has distribution $\gamma$.

Fix a positive time horizon $T \in (0, \infty]$. An **$\mathcal{H}$-Brownian motion in** $H$, on the time interval $[0, T)$ is a random $H$-valued continuous path $W$ on $[0, T)$, with stationary, independent increments such that the distribution of $W_1$ is a Gaussian probability measure $\gamma$ on $H$. A simple construction is provided by taking a sequence $(W_t^n)$ of independent, identically distributed, real-valued Brownian motions, and setting

$$W_t := \sum_n W_t^n \alpha_n e_n.$$

Denote by $S$ the unit sphere of $H$, and let $P_a : H \to H$ stand for the orthogonal projection on $\langle a\rangle^\perp$, for $a \neq 0$. The **$\mathcal{H}$-spherical Brownian motion** $v_t^\sigma$ on $S$ is defined as the solution to the Stratonovich stochastic differential equation

$$dv_t^\sigma = \sigma\, P_{v_t^\sigma}(\circ dW_t) \tag{2.1}$$

associated to a given initial condition $v_0^\sigma \in S$; it is defined for all times. The speed parameter $\sigma$ is a non-negative real number. Write $Z$ for $\int_H \frac{1}{\|u\|}\gamma(du)$.

**Theorem 2.1.** *The image under the projection $u \mapsto u/\|u\|$ of the measure $\frac{1}{Z}\frac{1}{\|u\|}\gamma(du)$ in the ambiant space $H$ is a probability measure $\mu$ on $S$ that is invariant for the dynamics of $v_t^\sigma$, for any positive speed parameter $\sigma$.*

This statement generalizes Proposition 1.1 of [Per18] to the present infinite dimensional setting. The above description of the invariant measure $\mu$ as an image measure under the projection map actually coincides with the finite dimensional description given in the latter reference.

*Proof.* When written in Itô form, the stochastic differential equation (2.1) defining the process $(v_t^\sigma)_{t\geq 0}$ reads

$$dv_t^\sigma = -\frac{\sigma^2}{2}\Big(\operatorname{tr}(\overline{C}_\gamma)v_t^\sigma + \overline{C}_\gamma(v_t^\sigma) - 2C_\gamma(v_t, v_t)v_t^\sigma\Big)dt + \sigma\, P_{v_t^\sigma}(dW_t), \tag{2.2}$$

and setting $v_t^{\sigma,i} := (v_t^\sigma, e_i)$, for any integer $i$, we have

$$\begin{aligned} dv_t^{\sigma,i} = &-\frac{\sigma^2}{2}\left[\sum_n \alpha_n^2 + \alpha_i^2 - 2\sum_n \alpha_n^2 |v_t^{\sigma,n}|^2\right] v_t^{\sigma,i}\, dt \\ &+ \sigma\left[\alpha_i d\,W_t^i - v_t^{\sigma,i}\sum_n \alpha_n v_t^{\sigma,n} dW_t^n\right]. \end{aligned}$$

As in the finite dimensional anisotropic case treated in [Per18], it is actually easier to work with an $H$-valued lift of this $S$-valued process. We introduce for that purpose the process $(u_t^\sigma)_{t\geq 0}$ solution of the Stratonovich stochastic differential equation

$$du_t^\sigma = -\frac{\sigma^2}{2}\|u_t^\sigma\|^2 u_t^\sigma dt + \sigma\|u_t^\sigma\|\circ dW_t;$$

equivalently, in Itô form and coordinate-wise, setting $u_t^{\sigma,i} := (u_t^\sigma, e_i)$ as above, we have

$$du_t^{\sigma,i} = \frac{\sigma^2}{2}\left(-\|u_t^\sigma\|^2 + \alpha_i^2\right) u_t^{\sigma,i} dt + \sigma\|u_t^\sigma\|\alpha_i dW_t^i.$$

A direct application of Itô's formula then shows that $u_t^{\sigma,i}/\|u_t^\sigma\|$ satisfies the same stochastic differential equation as $v_t^{\sigma,i}$, for all $i$, so the two $S$-valued processes $(v_t^\sigma)_{t\geq 0}$ and $(u_t^\sigma/\|u_t^\sigma\|)_{t\geq 0}$ have the same distributions. As in the finite dimensional case, one can then check by a direct computation that the measure $\|u\|^{-1}\gamma(du)$ on $H$ is invariant for the processes $(u_t^\sigma)$; this implies the statement of Theorem 2.1.

Alternatively, one can bypass computations and argue using Malliavin calculus as follows. Denote by $L$ the infinitesimal generator of the process $(u_t^\sigma)$. Set $V(u) := u/\|u\|^2$ for $u \neq 0$, and let $\Delta_\gamma$ denote the Laplace operator associated with the covariance $C_\gamma$ with weights $(\alpha_n^2)$. We then have for any test function $f$ and any $u \in H$

$$Lf(u) = \frac{\sigma^2}{2}\|u\|^2(L_0 f)(u),$$

with $(L_0 f)(u) := \Delta_\gamma f(u) - u\nabla f(u) + C_\gamma\big(V(u), \nabla f(u)\big)$. One then has for any test function $f$, with usual notations $D$ for the gradient and $\delta$ for the divergence in the Malliavin sense,

$$\begin{aligned}
\int_H Lf(u)\|u\|^{-1}\gamma(du) &= \frac{\sigma^2}{2}\int_H L_0 f(u)\|u\|\gamma(du) \\
&= \sigma^2\,\mathbb{E}\Big[\big(-\delta Df + \langle V, Df\rangle_{C_\gamma}\big)\,\|W\|\Big] \\
&= \mathbb{E}\Big[\big(-\delta\,\underbrace{D\|W\|}_{=\frac{W}{\|W\|}} + \delta\,\underbrace{V\|W\|}_{=\frac{W}{\|W\|}}\big) f\Big] = 0.
\end{aligned}$$

□

We prove in the next Section 2.2 that the velocity process $(v_t^\sigma)$ converges exponentially fast in Wasserstein distance to the invariant probability measure $\mu$ of Theorem 2.1, for any initial velocity $v_0$. An invariance principle for the time-rescaled position process $(x_{\sigma^2 t}^\sigma)$ is then obtained as a consequence in Section 2.3. We recall in Section 2.4 all the material we need from rough paths theory in this work, and prove in Section 2.5 that the canonical rough path associated to the time-rescaled process $(x_{\sigma^2 t}^\sigma)$ also converges weakly as a rough path to an explicit Stratonovich Brownian rough path.

## 2.2 Exponential mixing of the velocity process

We consider in this section the mixing properties of the spherical process $(v_t^\sigma)_{t\geq 0}$ with unit speed parameter $\sigma = 1$. To simplify the expressions, we drop momentarily the exponents $\sigma$ from all our notations. Our objective is to show that the spherical process $(v_t)_{t\geq 0} = (v_t^1)_{t\geq 0}$ is exponentially mixing. Recall that the 1 and 2-Wasserstein distances are defined for any probability measures $\mu, \nu$ on $S$ by the identities

$$\mathcal{W}_2(\lambda, \nu) = \inf \Big\{ \mathbb{E}\big[\|X - Y\|^2\big]; X \sim \lambda, Y \sim \nu \Big\},$$
$$\mathcal{W}_1(\lambda, \nu) = \inf \Big\{ \mathbb{E}\big[\|X - Y\|\big]; X \sim \lambda, Y \sim \nu \Big\} = \sup \left\{ \int f\, d(\lambda - \nu); |f|_{\mathrm{Lip}} \le 1 \right\},$$

where the infimum is taken over all couplings $\mathbb{P}$ of $X \sim \lambda$ and $Y \sim \nu$, and the last supremum over all 1-Lipscthiz functions $f : S \to \mathbb{R}$. The first two equalities are definitions, the last one is the Kantorovich-Rubinstein duality principle. Note that $\mathcal{W}_1 \le \mathcal{W}_2$.

**Proposition 2.2.** *Assume that*

$$3\alpha_0^2 < \mathrm{tr}(\overline{C}_\gamma). \tag{2.3}$$

*There exists a positive time $\tau$ such that for any probability measures $\lambda$ and $\nu$ on the unit sphere $S$ of $H$, we have*

$$\mathcal{W}_2(P_t^* \lambda, P_t^* \nu) \le e^{-t/\tau} \mathcal{W}_2(\lambda, \nu),$$

*for all $t \ge 0$. In particular, the invariant measure $\mu$ is unique, and for any probability measure $\lambda$ on the sphere $S$, and $t \ge 0$, we have*

$$\mathcal{W}_2(P_t^* \lambda, \mu) \le 2e^{-t/\tau}. \tag{2.4}$$

The role of the trace condition (2.3) will be clear from the proof. If we have the freedom to choose the covariance $C_\gamma$ of the Brownian noise, this is not a constraint. Note that the rougher the noise, that is the more slowly the sequence of the eigenvalues $\alpha_n$ converges to 0, the easier it is to satisfy condition (2.3). We shall see in Section 4 that it holds automatically in a number of relevant examples of random dynamics in the configuration space of a fluid flow.

*Proof.* Denote by $\mathbb{P}$ the law of the Brownian motion $(B_t)$ with covariance $C_\gamma$, and by $\mathbb{P}_v$ the law of the solution of Equation (2.1) with $\sigma = 1$, starting from $v \in S$. Denote by $\mathbb{E}$ and $\mathbb{E}_v$ the associated expectations operators. Recall that the notation $(a, b)$ stands for the scalar product of $a$ and $b$ in $H$. Fix $v_0, w_0 \in S$, and consider the two diffusion processes $(v_t)$ and $(w_t)$, started from $v_0$ and $w_0$, respectively, and solutions of the Itô stochastic differential equations

$$dv_t = -\frac{1}{2}\Big( \mathrm{tr}(\overline{C}_\gamma) v_t + \overline{C}_\gamma(v_t) - 2C_\gamma(v_t, v_t) v_t \Big) dt + P_{v_t}(dW_t),$$
$$dw_t = -\frac{1}{2}\Big( \mathrm{tr}(\overline{C}_\gamma) w_t + \overline{C}_\gamma(w_t) - 2C_\gamma(w_t, w_t) w_t \Big) dt + P_{w_t}(dW_t).$$

Comparing with Equation (2.2), it is clear that $(v_t)$ has law $\mathbb{P}_{v_0}$ and $(w_t)$ has law $\mathbb{P}_{w_0}$. Moreover, Itô's formula yields

$$d(v_t, w_t) = \Big( \mathrm{tr}(\overline{C}_\gamma) - C_\gamma(v_t, v_t) - C_\gamma(w_t, w_t) - C_\gamma(v_t, w_t) \Big) \big(1 - (v_t, w_t)\big) dt$$
$$+ \big(1 - (v_t, w_t)\big) \Big( (v_t, dW_t) + (w_t, dW_t) \Big),$$

or equivalently, setting

$$N_t := \frac{1}{2}\|w_t - v_t\|^2 = 1 - (v_t, w_t),$$

we get

$$\begin{aligned} dN_t = &-\Big( \mathrm{tr}(\overline{C}_\gamma) - C_\gamma(v_t, v_t) - C_\gamma(w_t, w_t) - C_\gamma(v_t, w_t) \Big) N_t dt \\ &- N_t\big( (v_t, dW_t) + (w_t, dW_t) \big). \end{aligned} \tag{2.5}$$

Now remark that since the sequence $(\alpha_n)$ is non-increasing, we have

$$C_\gamma(v,v) = \sum_{n\geq 0} \alpha_n^2 |v_n|^2 \leq \alpha_0^2,$$

for any $v \in S$. Taking the expectation under $\mathbb{P}$ in equation (2.5), we have from Grönwall inequality

$$\mathbb{E}[N_t] \leq e^{-t(\mathrm{tr}(\overline{C}_\gamma)-3\alpha_0^2)} \mathbb{E}[N_0],$$

that is

$$\mathbb{E}[\|v_t - w_t\|^2] \leq e^{-t(\mathrm{tr}(\overline{C}_\gamma)-3\alpha_0^2)} \|x-y\|^2.$$

The conclusion of the statement follows. □

Remark that $\mathbb{E}_\mu[v_t] = 0$, as a consequence of the symmetry properties of the invariant measure $\mu$.

**Corollary 2.3.** *For any $v_0 \in S$, we have*

$$\big\|\mathbb{E}_{v_0}[v_t]\big\| \leq 2e^{-t/\tau}.$$

The process $(v_t)$ is stationary if $v_0$ has distribution $\mu$; it can then be extended into a two sided process defined for all real times. Denote by $(\mathcal{F}_t)_{t\in\mathbb{R}}$ the complete filtration generated by $(v_t)$ on the probability space where it is defined. Set $\mathcal{F}_{\leq 0} := \sigma\big(\mathcal{F}_t\,;\, t\leq 0\big)$ and $\mathcal{F}_{\geq s} := \sigma\big(\mathcal{F}_t\,;\, t\geq s\big)$, for any real time $s$. Recall that the mixing coefficient $\alpha(s)$ of the velocity process $v$ is defined, for $s > 0$, by the formula

$$\alpha(s) := \sup_{A\in\mathcal{F}_{\leq 0}, B\in\mathcal{F}_{\geq s}} \big|\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)\big|.$$

The following fact will be useful to get for free the independence of the increments of the limit processes obtained after proper rescalings of functionals of $(v_t)$.

**Corollary 2.4.** *The mixing coefficient $\alpha(s)$ tends to $0$ as $s$ increases to $\infty$.*

*Proof.* As a preliminary remark, recall the definition of the lift $(u_t^\sigma)$ to $H$ of $(v_t^\sigma)$, introduced in the proof of Theorem 2.1. This process is strong Feller, as it can be seen to satisfy a Bismut-Li integration by parts formula. See e.g. Peszat and Zabczyk' seminal paper [PZ95], and Wang and Zhang's extension [WZ10] to unbounded drift and diffusivity. The velocity process $(v_t^\sigma)$ is thus itself a strong Feller diffusion, and if one denotes by $(P_t)$ its transition semigroup, the functions $P_1 g$, for $g$ measurable, bounded by 1, are all Lipschitz continuous, with a finite common upper bound $L$ for their Lipschitz constants.

Now, it follows from the Markovian character of the dynamics of $(v_t)$, and the Feller property of its semigroup, that it suffices to see that

$$\mathbb{E}\big[f(v_0)g(v_s)\big] \tag{2.6}$$

tends to 0 as $s$ goes to $\infty$, for any real-valued continuous functions $f, g$ on the unit sphere $S$, with null mean with respect to the invariant measure $\mu$, uniformly with respect to $f$ and $g$ with $L^\infty$-norm 1. Writing further

$$\mathbb{E}\Big[f(v_0)\mathbb{E}\big[g(v_s)|v_{s_1}\big]\Big] = \mathbb{E}\Big[f(v_0)\,(P_1 g)(v_{s-1})\Big],$$

for $s > 1$, and using the strong Feller property of the semigroup of the diffusion process $(v_t)$, we can further assume that the function $g$ in (2.6) is $L\|g\|_\infty$-Lipschitz continuous. Let $w_g$ stand for its uniform modulus of continuity. For each $s$, denote by $(v_s, \overline{v}_s)$ a $\mathcal{W}_1$-optimal coupling of the measures $P_s^*\delta_{v_0}$ and $\mu$, for a deterministic $v_0$, so we have

$$\mathbb{E}\big[|v_s - \overline{v}_s|\big] = \mathcal{W}_1\big(P_s^*\delta_{v_0}, \mu\big).$$

Using the fact that $\int g d\mu = 0$, one then has

$$\begin{aligned} \big|\mathbb{E}\big[f(v_0)g(v_s)\big]\big| &= \Big|\mathbb{E}\Big[f(v_0)\,\mathbb{E}\big[g(v_s)|v_0\big]\Big]\Big| \\ &\leq \|f\|_\infty\,\mathbb{E}\Big[w_g\big(|v_s - \overline{v}_s|\big)\Big] \\ &\leq L\|f\|_\infty\|g\|_\infty\,\mathbb{E}\big[|v_s - \overline{v}_s|\big], \end{aligned}$$

so the statement follows from Proposition 2.2. □

## 2.3 Invariance principle for the position process

*We assume in all of this section that the initial condition $v_0$ of the velocity process of kinetic Brownian motion is distribued according to its invariant probability measure $\mu$, from Theorem 2.1.*

Pick $1/3 < \alpha < 1/2$. The goal of this section is to establish the next Proposition 2.5 stated just below, which asserts that, as $\sigma$ goes to infinity, the distribution in $C^\alpha([0,1], H)$ of the time-rescaled position process $(x^\sigma_{\sigma^2 t})$ with values in $H$ converges to the distribution of a Brownian motion in $H$ with an explicit covariance. The usual invariance principles in Hilbert spaces often only consider weak convergence in $C^0([0,1], H)$, so we need an extra tightness estimate provided in Section 2.3 to complete the program. To make the most out of the convergence results from Section 2.2, set

$$X^\sigma_t := x^\sigma_{\sigma^2 t};$$

we have

$$X^\sigma_t - X^\sigma_s = \int_{\sigma^2 s}^{\sigma^2 t} v_{\sigma^2 u} du = \frac{1}{\sigma^2}\int_{\sigma^4 s}^{\sigma^4 t} v_u du,$$

with $(v_t) = (v^1_t)$, the spherical Brownian motion run at speed $\sigma^2 = 1$.

**Proposition 2.5.** *For every $0 < \alpha < 1/2$, the distribution in $\mathcal{C}^\alpha([0,1], H)$ of the process $(X^\sigma_t)$ converges as $\sigma$ goes to $\infty$ to the Brownian motion on $H$ with covariance operator*

$$C(\ell, \ell') := \int_0^\infty \mathbb{E}\Big[\ell(v_0)\ell'(v_t) + \ell'(v_0)\ell(v_t)\Big]\, dt, \tag{2.7}$$

*for $\ell, \ell' \in H^*$.*

### Tightness in Hölder spaces

We dedicate this section to proving the following uniform estimate.

**Proposition 2.6.** *For any $p \geq 2$, we have*

$$\sup_{\sigma>0} \mathbb{E}\big[\|X^\sigma_t - X^\sigma_s\|^p\big] \lesssim_p |t-s|^{p/2}.$$

It follows from Kolmogorov-Lamperti tightness criterion that the laws of $X^\sigma$ form a tight family in $\mathcal{C}^\alpha([0,1],H)$, for any $0<\alpha<1/2$. Note that for $T=\sigma^4(t-s)>0$, we have

$$\|X_t^\sigma - X_s^\sigma\| \overset{\mathcal{L}}{=} \frac{1}{\sigma^2}\left\|\int_0^{\sigma^4(t-s)} v_u\,du\right\| = |t-s|\cdot\frac{1}{\sqrt{T}}\left\|\int_0^T v_u\,du\right\|,$$

so Proposition 2.6 is a consequence of the estimate

$$\mathbb{E}\left[\left|\int_0^T v_t\,dt\right|^p\right] \lesssim_p T^{p/2}.$$

We translate our problem in discrete time, writing

$$\int_0^T = \sum_{k<T}\int_k^{k+1}$$

to work with the correlations between different integral slices, and compare this sequence to martingale differences. There is an abundant literature on the subject; we follow here the approach of C. Cuny [Cun17].

Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space with a filtration $(\mathcal{F}_n)_{n\geq n_0}$, where $-\infty\leq n_0\leq 0$, and let $(X_n)_{n\geq n_0}$ be $H$-valued random variables such that each $X_n$ is measurable with respect to $\mathcal{F}_n$. Recall that $(X_n)_{n\geq 0}$ is said to be a **martingale difference** with respect to $(\mathcal{F}_n)$ if each $X_n$ is integrable and $\mathbb{E}\big[X_{n+1}|\mathcal{F}_n\big]=0$, for all $n\geq n_0$. The following result is an elementary consequence of the Burkholder-Davis-Gundy and Jensen inequalities.

**Lemma 2.7.** *Let $X$ be an $H$-valued martingale difference with moments of order $p\geq 2$. Then*

$$\mathbb{E}\big[|X_0+\cdots+X_{n-1}|^p\big]^{\frac{1}{p}} \lesssim_p \sqrt{n}\left(\frac{1}{n}\Big(\mathbb{E}\big[|X_0|^p\big]+\cdots+\mathbb{E}\big[|X_{n-1}|^p\big]\Big)\right)^{\frac{1}{p}}.$$

*In particular, if $X$ is stationary, then*

$$\mathbb{E}\Big[|X_0+\cdots+X_{n-1}|^p\Big]^{\frac{1}{p}} \lesssim_p \sqrt{n}\,\|X_0\|_{L^p}.$$

Assume from now on that we are given a sequence $(X_n)_{\geq n_0}$ of integrable $H$-valued random variables on $(\Omega,\mathcal{F},\mathbb{P})$. For $j\in\mathbb{Z}$, and $k\geq 0$, define the $\sigma$-algebra

$$\mathcal{F}_j^{(k)} := \mathcal{F}_{j2^k},$$

and set

$$Y_j^{(k)} := \mathbb{E}\Big[X_{j2^k}+\cdots+X_{j2^k+(2^k-1)}\Big|\mathcal{F}_{j-1}^{(k)}\Big].$$

(It may not make sense for all $j,k$, depending on how far in the past the $\sigma$-algebras $(\mathcal{F}_n)$ are defined.) Note that

$$Y_j^{(\ell+1)} = \mathbb{E}\Big[Y_{2j}^{(\ell)}+Y_{2j+1}^{(\ell)}\Big|\mathcal{F}_{j-1}^{(\ell+1)}\Big],$$

so

$$M_j^{(\ell)} := Y_{2j}^{(\ell)}+Y_{2j+1}^{(\ell)}-Y_j^{(\ell+1)}$$

is a stationary martingale difference with respect to the filtration $\big(\mathcal{F}_j^{(\ell+1)}\big)_{j\geq 0}$. We use the classical martingale/co-boundary decomposition to prove the next result.

**Lemma 2.8.** *Fix $p \geq 2$, and assume that $\mathcal{F}_n$ is defined for $n \geq -2^{k+1}$, then*

$$\mathbb{E}\Big[\big|Y_0^{(0)} + \cdots + Y_{2^k-1}^{(0)}\big|^p\Big]^{\frac{1}{p}} \lesssim_p \sum_{0\leq j\leq k} 2^{\frac{k-j}{2}} \left(\frac{1}{2^{k-j}}\Big(\mathbb{E}\big[\big|Y_0^{(j)}\big|^p\big] + \cdots + \mathbb{E}\big[\big|Y_{2^{k-j}-1}^{(j)}\big|^p\big]\Big)\right)^{\frac{1}{p}}.$$

*In particular, if the sequence $(X_n)$ is stationary, then*

$$\mathbb{E}\big[\big|Y_0^{(0)} + \cdots + Y_{2^k-1}^{(0)}\big|^p\big]^{\frac{1}{p}} \lesssim_p 2^{k/2}\Big(\mathbb{E}\big[\big|Y_0^{(0)}\big|^p\big]^{\frac{1}{p}} + \cdots + 2^{-k/2}\mathbb{E}\big[\big|Y_0^{(k)}\big|^p\big]^{\frac{1}{p}}\Big).$$

*Proof.* For any $0 \leq j \leq k$, set $n_j := 2^{k-j}$; note that $n_k = 1$. We have for $j < k$ the identity

$$\begin{aligned} Y_0^{(j)} + \cdots + Y_{n_j-1}^{(j)} &= \big(Y_0^{(j)} + Y_1^{(j)}\big) + \cdots + \big(Y_{2n_{j+1}-2}^{(j)} + Y_{2n_{j+1}-1}^{(0)}\big) \\ &= M_0^{(j)} + \cdots + M_{n_{j+1}-1}^{(j)} + Y_0^{(j+1)} + \cdots + Y_{n_{j+1}-1}^{(j+1)}. \end{aligned}$$

By induction we get

$$Y_0^{(0)} + \cdots + Y_{n-1}^{(0)} = \big(M_0^{(0)} + \cdots + M_{n_1-1}^{(0)}\big) + \cdots + \big(M_0^{(k-1)}\big) + Y_0^{(k)}.$$

Because $M^{(j)}$ is a martingale difference, we know from Lemma 2.7 that

$$\begin{aligned} &\mathbb{E}\big[\big|M_0^{(j)} + \cdots + M_{n_{j+1}-1}^{(j)}\big|^p\big]^{\frac{1}{p}} \\ &\qquad \lesssim_p \sqrt{n_{j+1}} \cdot \left(\frac{1}{n_{j+1}}\Big(\mathbb{E}\big[\big|M_0^{(j)}\big|^p\big] + \cdots + \mathbb{E}\big[\big|M_{n_{j+1}}^{(j)}\big|^p\big]\Big)\right)^{\frac{1}{p}}. \end{aligned}$$

We also know that

$$M_{2^k}^{(j)} = Y_{2^{k+1}}^{(j)} + Y_{2^{k+1}+1}^{(j)} - \mathbb{E}\big[Y_{2^{k+1}}^{(j)} + Y_{2^{k+1}+1}^{(j)}\big|\mathcal{F}_{2^k-1}^{(j+1)}\big],$$

so we have

$$\begin{aligned} \mathbb{E}\big[\big|M_{2^k}^{(j)}\big|^p\big]^{\frac{1}{p}} &\leq \mathbb{E}\big[\big|Y_{2^{k+1}}^{(j)}\big|^p\big]^{\frac{1}{p}} + \mathbb{E}\big[\big|Y_{2^{k+1}+1}^{(j)}\big|^p\big]^{\frac{1}{p}} + \mathbb{E}\Big[\mathbb{E}\big[|Y_{2^{k+1}}^{(j)}|^p\big|\mathcal{F}_{-1}^{(j+1)}\big]\Big]^{\frac{1}{p}} \\ &\quad + \mathbb{E}\Big[\mathbb{E}\big[|Y_{2^{k+1}+1}^{(j)}|^p\big|\mathcal{F}_{-1}^{(j+1)}\big]\Big]^{\frac{1}{p}} \\ &\leq 2\mathbb{E}\big[\big|Y_{2^{k+1}}^{(j)}\big|^p\big]^{\frac{1}{p}} + 2\mathbb{E}\big[\big|Y_{2^{k+1}+1}^{(j)}\big|^p\big]^{\frac{1}{p}}. \end{aligned}$$

Putting it all together, we obtain

$$\begin{aligned} &\mathbb{E}\big[\big|Y_0^{(0)} + \cdots + Y_{2^k-1}^{(0)}\big|^p\big]^{\frac{1}{p}} \\ &\quad \lesssim_p \sum_{0\leq j<k} \sqrt{n_{j+1}} \cdot \left(\frac{1}{2n_{j+1}}\Big(\mathbb{E}\big[\big|Y_0^{(j)}\big|^p\big] + \cdots + \mathbb{E}\big[\big|Y_{2n_{j+1}}^{(j)}\big|^p\big]\Big)\right)^{\frac{1}{p}} \\ &\qquad + \mathbb{E}\big[\big|Y_0^{(k)}\big|^p\big]^{\frac{1}{p}} \\ &\quad \lesssim_p \sum_{0\leq j\leq k} 2^{(k-j)/2} \cdot \left(\frac{1}{2^{k-j}}\Big(\mathbb{E}\big[\big|Y_0^{(j)}\big|^p\big] + \cdots + \mathbb{E}\big[\big|Y_{2^{k-j}-1}^{(j)}\big|^p\big]\Big)\right)^{\frac{1}{p}}. \end{aligned}$$

□

*Proof of Proposition 2.6.* It is enough to prove that we have for any $T \geq 1$ and $p \geq 2$, the estimate

$$\mathbb{E}\left[\left|\int_0^T v_t\,dt\right|^p\right] \lesssim_p T^{p/2}.$$

Fix the integer $k$ such that $T/2 \leq 2^k < T$, and define

$$X_j := \int_{jT2^{-k}}^{(j+1)T2^{-k}} v_t\,dt, \qquad \mathcal{F}_j = \sigma\Big(v_s, s \leq (j+1)T2^{-k}\Big).$$

Since we assume that $v_0$ is distributed according to an invariant probability measure, we can actually have our process started for a time arbitrarily far in the past, so we can assume that $\mathcal{F}_j$ is well-defined for any $j \geq -2^{k+1}$. We can then write

$$\begin{aligned}\int_0^T v_t\,dt = \big(X_0 - \mathbb{E}[X_0|\mathcal{F}_{-1}]\big) + \cdots + \big(X_{2^k-1} - \mathbb{E}[X_{2^k-1}|\mathcal{F}_{2^k-2}]\big)\\ + \mathbb{E}[X_0|\mathcal{F}_{-1}] + \cdots + \mathbb{E}[X_{2^k-1}|\mathcal{F}_{2^k-2}].\end{aligned}$$

The first sum is a stationary martingale difference with respect to the $\sigma$-algebra $(\mathcal{F}_j)_{j\geq 0}$; the second is the subject of the previous lemma. One then has the estimate

$$\begin{aligned}\mathbb{E}\left[\left|\int_0^T v_t\,dt\right|^p\right]^{\frac{1}{p}} &\lesssim_p 2^{k/2}\,\mathbb{E}\Big[\big|X_0 - \mathbb{E}[X_0|\mathcal{F}_{-1}]\big|^p\Big]^{\frac{1}{p}}\\ &+ 2^{k/2}\left(\mathbb{E}\Big[\big|Y_0^{(0)}\big|^p\Big]^{\frac{1}{p}} + \cdots + 2^{-k/2}\,\mathbb{E}\Big[\big|Y_0^{(k)}\big|^p\Big]^{\frac{1}{p}}\right)\\ &\lesssim_p \sqrt{T}\left(\mathbb{E}\Big[\big|X_0\big|^p\Big]^{\frac{1}{p}} + \mathbb{E}\Big[\big|Y_0^{(0)}\big|^p\Big]^{\frac{1}{p}} + \cdots + 2^{-k/2}\,\mathbb{E}\Big[\big|Y_0^{(k)}\big|^p\Big]^{\frac{1}{p}}\right)\end{aligned}$$

with the notations of Lemma 2.8. In our setting,

$$\|X_0\|_{L^p} = \mathbb{E}\left[\left|\int_0^{T2^{-k}} v_t\,dt\right|^p\right]^{\frac{1}{p}} \leq \big(T2^{-k}\big)^{\frac{p}{p}} \leq 2$$

and

$$Y_0^{(j)} = \mathbb{E}\left[\int_{2^jT2^{-k}}^{2^{j+1}T2^{-k}} v_t\,dt\,\middle|\mathcal{F}_{-1}\right] = \mathbb{E}_{v_0}\left[\int_{2^jT2^{-k}}^{2^{j+1}T2^{-k}} v_t\,dt\right].$$

Note that we have from Corollary 2.3

$$\begin{aligned}\left|\mathbb{E}_{v_0}\left[\int_{2^jT2^{-k}}^{2^{j+1}T2^{-k}} v_t\,dt\right]\right| \leq \int_{2^jT2^{-k}}^{2^{j+1}T2^{-k}} \big|\mathbb{E}_{v_0}[v_t]\big|\,dt \lesssim \int_{2^jT2^{-k}}^{\infty} e^{-t/\tau}\,dt\\ \lesssim e^{-2^{j-1}/\tau}.\end{aligned}$$

We can insert this in the upper bound for the integral to obtain

$$\left\|\int_0^T v_t\,dt\right\|_{L^p} \leq \left(1 + \sum_{j\geq 0} 2^{-j/2} e^{-2^{j-1}/\tau}\right)\sqrt{T}. \tag{2.8}$$

□

**Convergence in Hölder spaces**

We are now ready to prove Proposition 2.5 on the weak convergence of $X^\sigma$ in any Hölder space $C^\alpha([0,1],H)$ to the Brownian motion in $H$ with covariance given by formula (2.7).

*Proof of Proposition 2.5.* From the tightness result in $C^\alpha([0,1],H)$ stated in Proposition 2.6, it is sufficient to show that $X^\sigma$ converges weakly in $C^0([0,1],H)$ to the above mentionned Brownian motion. Since we start the velocity process from its invariant measure, the position process $X^\sigma$ has stationary increments, and any weak limit will have the same property. From Corollary 2.4, the increments of a weak limit are independent on disjoint intervals; continuity of a limit process gives independence of the increments on adjacent intervals. Any weak limit of the $X^\sigma$ is thus a Brownian motion, and uniqueness will follow from identifying uniquely its covariance. The latter is identified writing

$$\begin{aligned}\mathbb{E}\big[\ell(X_T^\sigma)^2\big] &= \frac{2}{\sigma^4}\int_0^{\sigma^4 T}\int_s^{\sigma^4 T}\mathbb{E}\big[\ell(v_s)\ell(v_t)\big]\,dt\,ds\\ &= 2\int_0^\infty\left(\frac{1}{\sigma^4}\int_0^\infty \mathbf{1}_{s+u\le\sigma^4 T}ds\right)\mathbb{E}\big[\ell(v_0)\ell(v_u)\big]\,du.\end{aligned}$$

We see on this last expression that it has limit

$$2T\int_0^\infty \mathbb{E}\big[\ell(v_0)\ell(v_u)\big]\,du,$$

using the decorrelation estimate from Proposition 2.2 to justify dominated convergence. $\square$

We aim now at improving the weak invariance principle of Proposition 2.5 into a weak invariance principle for the canonical rough path associated with $X^\sigma$. This will be crucial in Section 4 when defining kinetic Brownian motion in a diffeomorphism space as the solution of a differential equation driven by $X^\sigma$, and proving the interpolation results of Theorem 4.3 and Theorem 4.4 by a continuity argument. For the sake of self containedness, we recall in the next section all we need to know from rough paths theory.

## 2.4 The flavour of rough paths theory

It is not our purpose here to give a detailled account of rough paths theory. We refer the reader to the lecture notes [LCL07, FH14, Bau14, Bai15b], for introductions to the subject from different point of views. The following will be sufficient for our needs here. Rough paths theory is a theory of ordinary differential equations

$$dz_t = \sum_{i=1}^{\ell} V_i(z_t)\,dh_t^i, \tag{2.9}$$

controlled by non-smooth signals $h\in C^\alpha([0,1],\mathbb{R}^\ell)$. The point $z_t$ moves here in $\mathbb{R}^d$, where we are given sufficiently regular vector fields $V_i$. Young integration theory [You36, Lyo94] allows to make sense of the integral $\int_0^\cdot V(y_s)dh_s$, for paths $y,h$ that are $\alpha$-Hölder, for $\alpha>\frac12$, as an $\mathbb{R}^d$-valued $\alpha$-Hölder path depending in locally Lipscthiz way on $y$ and $h$. This allows to formulate the differential equation (2.9) as a fixed point problem for a contracting map from $C^\alpha([0,1],\mathbb{R}^d)$ into itself, and to obtain as a consequence the continuous dependence of the solution path on the driving control $h$. Lyons-Young theory cannot be used for $\alpha$-Hölder controls with $\alpha<\frac12$, as even in $\mathbb{R}$, with one dimensional controls, there exists no *continuous* bilinear form on $C^\alpha([0,1],\mathbb{R})\times C^\alpha([0,1],\mathbb{R})$ extending the Riemann integral $\int_0^1 y_t dh_t$, of smooth paths $y,h$; see

Propositon 1.29 of [LCL07]. (This can be understood from a Fourier analysis point of view as a consequence of the fact that the resonant operator from Littlewood-Paley theory is unbounded on $C^\alpha([0,1],\mathbb{R}) \times C^{\alpha-1}([0,1],\mathbb{R})$, when $2\alpha - 1 < 0$; see [BCD11].) Lyons' deep insight was to realize that what really fixes the dynamics of a solution path to the controlled differential equation (2.9) is not only the increments $dh_t$, or $h_t - h_s$, of the control, but rather the increments of $h$ together with the increments of a number of its iterated integrals. This can be understood from the fact that for a smooth control, one has the Taylor-type expansion

$$\begin{aligned} f(z_t) = f(z_s) + \left(\int_s^t dh_u^i\right)(V_i f)(z_s) + \left(\int_{s \le u_2 \le u_1 \le t} dh_{u_2}^j dh_{u_1}^k\right)(V_j V_k f)(z_s) \\ + \int_{s \le u_3 \le u_2 \le u_1 \le t} (V_n V_j V_k f)(z_{u_3})\, dh_{u_3}^n dh_{u_2}^j dh_{u_1}^k, \end{aligned}$$

for any real-valued smooth function $f$ on $\mathbb{R}^d$. (We use Einstein's summation convention, with integer indices in $[1,\ell]$.) We consider here the vector fields $V_i$ as first order differential operators, so we have for instance

$$V_j V_k f = (D^2 f)(V_j, V_k) + (Df)\big((DV_k)(V_j)\big).$$

The usual first order Euler scheme

$$z_t \simeq z_s + (h_t^i - h_s^i) V_i(z_s),$$

is refined by the above second order Milstein scheme

$$z_t \simeq z_s + (h_t^i - h_s^i) V_i(z_s) + \left(\int_{s \le u_2 \le u_1 \le t} dh_{u_2}^j dh_{u_1}^k\right)(V_j V_k)(z_s),$$

whose one step error is given explicitly by the above triple integral, of order $|t-s|^3$, for a $C^1$ control $h$. The iterated integrals

$$\int_{s \le u_2 \le u_1 \le t} dh_{u_2}^j dh_{u_1}^k = \int_{s \le u_1 \le t} \left(h_{u_1}^j - h_s^j\right) dh_{u_1}^k,$$

are however meaningless for a control $h \in C^\alpha([0,1],\mathbb{R}^\ell)$, when $\alpha \le 1/2$. A $p$-rough path $\mathbf{X}$ above $h$, with $2 \le p < 3$, is exactly the datum of $h$ together with a quantity, indexed by $(s \le t)$, that plays the role of these iterated integrals. Set $[0,1]_\le := \{(s,t) \in [0,1]^2\,;\, s \le t\}$, and recall that $\left(\mathbb{R}^\ell\right)^{\otimes 2}$ stands for the set of $\ell \times \ell$ matrices.

**Definition 2.9.** Fix $2 \le p < 3$. A ***p*-rough path $\mathbf{X}$** over $\mathbb{R}^\ell$, is a map

$$\begin{aligned} [0,1]_\le &\to \mathbb{R}^\ell \times \left(\mathbb{R}^\ell\right)^{\otimes 2} \\ (s,t) &\mapsto \big(X_{ts}, \mathbb{X}_{ts}\big), \end{aligned}$$

such that

$$X_{ts} = h_t - h_s,$$

for a $C^\alpha([0,1],\mathbb{R}^\ell)$ path $h$, and $\mathbb{X}$ satisfies Chen's relations

$$\mathbb{X}_{ts} = \mathbb{X}_{tu} + X_{us} \otimes X_{tu} + \mathbb{X}_{us},$$

for all $0 \le s \le u \le t \le 1$. The $1/p$-Hölder norm on $X$, and the $2/p$-Hölder norm on $\mathbb{X}$, define jointly a complete metric on the nonlinear space $\mathsf{RP}(p)$ of $p$-rough paths. $\triangle$

Chen's relation accouts for the fact that for a $C^1$ path $h$, one has indeed

$$\int_{s\leq u_1\leq t}\big(h^j_{u_1}-h^j_s\big)\,dh^k_{u_1}=\int_{u\leq u_1\leq t}\big(h^j_{u_1}-h^j_u\big)\,dh^k_{u_1}+\big(h^j_u-h^j_s\big)\big(h^k_t-h^k_u\big)$$
$$+\int_{s\leq u_1\leq u}\big(h^j_{u_1}-h^j_s\big)\,dh^k_{u_1}$$

for any $0\leq s\leq u\leq t\leq 1$, and any indices $1\leq j,k\leq \ell$. One has also in that case, by integration by parts, the identiy

$$\int_{s\leq u_1\leq t}\big(h^j_{u_1}-h^j_s\big)\,dh^k_{u_1}+\int_{s\leq u_1\leq t}\big(h^k_{u_1}-h^k_s\big)\,dh^j_{u_1}$$
$$=\frac{1}{2}\big(h^j_t-h^j_s\big)\big(h^k_t-h^k_s\big).$$

A $p$-rough path $\mathbf{X}$ such that the symmetric part of $\mathbb{X}_{ts}$ is equal to $\frac{1}{2}X_{ts}\otimes X_{ts}$, for all times $0\leq s\leq t\leq 1$, is called **weakly geometric**. The set of weakly geometric $p$-rough paths is closed in $\mathsf{RP}(p)$. For a $C^1$ path $h$ defined on the time interval $[0,1]$, setting $X_{ts}:=h_t-h_s$ and

$$\mathbb{X}_{ts}:=\int_s^t X_{us}\otimes dX_u,$$

for all $0\leq s\leq t\leq 1$, defines a weak geometric $p$-rough path, for any $2\leq p<3$, called the **canonical rough path associated with** $h$**.** Let $B$ stand for an $\ell$-dimensional Brownian motion. The Stratonovich Brownian rough path $\mathbf{B}=(B,\mathbb{B})$ is defined by

$$\mathbb{B}_{ts}:=\int_{s\leq u\leq t}(B_u-B_s)\otimes\circ dB_u.$$

It is almost surely a weak geometric $p$-rough path, for any $2<p<3$.

**Definition 2.10.** Let $C^3_b$ vector fields $(V_i)_{1\leq i\leq \ell}$ on $\mathbb{R}^d$ be given, together with a weak geometric $p$-rough path $\mathbf{X}$ over $\mathbb{R}^\ell$. A path $(z_t)_{0\leq t\leq 1}$ is said to be a solution to the rough differential equation

$$dz_t=V(z_t)\,d\mathbf{X}_t \tag{2.10}$$

if there is an exponent $a>1$, such that one has

$$f(z_t)=f(z_s)+X^i_{ts}(V_if)(z_s)+\mathbb{X}^{jk}_{ts}\big(V_jV_kf\big)(z_s)+O\big(|t-s|^a\big), \tag{2.11}$$

for any smooth real-valued function $f$ on $\mathbb{R}^d$, and any times $0\leq s\leq t\leq 1$. $\triangle$

The above $O(\cdot)$ term is allowed to depend on $f$. Importantly, the solution of a rough differential equation driven by the Stratonovich Brownian rough path coincides almost surely with the solution of the corresponding Stratonovich differential equation; see e.g. the lecture notes [FH14, Bai].

**Theorem 2.11** (Lyons' universal limit theorem)**.** *The rough differential equation* (2.10) *has a unique solution. It is an element of* $C^{1/p}([0,1],\mathbb{R}^d)$ *that depends continuously on* $\mathbf{X}$.

The map that associates to the driving rough path the solution to a given rough differential equation, seen as an element of $C^{1/p}([0,1],\mathbb{R}^d)$, is called the **Itô-Lyons solution map**. If $(\mathbf{X}^n)$ is a sequence of random geometric $p$-rough path in $\mathbb{R}^\ell$, converging weakly to a limit random geometric $p$-rough path $\mathbf{X}$, the continuity of the Itô-Lyons solution map gives for free the weak convergence in $C^{1/p}([0,1],\mathbb{R}^d)$ of the laws of the solutions to Equation (2.10) driven by the $\mathbf{X}^n$, to the law of the solution of that equation driven by $\mathbf{X}$.

The theory works perfectly well for dynamics with values in Banach spaces or Banach manifolds, and driving rough paths $\mathbf{X} = (X, \mathbb{X})$, with $X$ taking values in a Banach space $E$. One needs to take care in that setting to the tensor norm used to define the completion of the algebraic tensor space $E \otimes_a E$, as this may produce non-equivalent norms, and that norm is used to define the norm of a rough path. Note that families of vector fields $(V_1, \dots, V_\ell)$ are then replaced in that setting by one forms on $E$ with values in the space of vector fields on the space where the dynamics takes place. See e.g. Lyons' original work [Lyo98] or Cass and Weidner's work [CW16] for the details. See e.g. [Bai15a] for a simple proof of Lyons' universal limit theorem in that general setting.

The vector fields in Definition 2.10 and Theorem 2.11 are required to be $C_b^3$. This is used to get solution of equation (2.10) that are defined on the whole time interval $[0,1]$. Only local in time existence results can be obtained when working with unbounded vector fields, or on a manifold. The Taylor-like expansion property (2.11) defining a solution path is then only required to hold for each time $s$, for $t$ sufficiently close to $s$. One still has continuity of the solution path with respect to the driving rough path, in an adapted sense. See e.g. Section 2.4.2 of [ABT15]. This continuity property is sufficient to obtain the local weak convergence of the laws of the solution path to the corresponding limit path, for random driving weak geometric $p$-rough paths converging weakly to a limit random weak geometric $p$-rough path. See Definition 4.2 for the definition of local weak convergence.

So far, we have defined kinetic Brownian motion $(x_t^\sigma, v_t^\sigma)$ in $H$ from its unit velocity process $v^\sigma$. We have seen in Proposition 2.5 that its time rescaled position process $(X_t^\sigma) := (x_{\sigma^2 t}^\sigma)$ is converging weakly in $C^\alpha\big([0,1], H\big)$ to a Brownian motion with explicit covariance (2.7), for any $\alpha < 1/2$. We prove in the next section that the canonical rough path $\mathbf{X}^\sigma$ associated with $X^\sigma$ converges weakly as a weak geometric $p$-rough path to the Stratonovich Brownian rough path associated with the Brownian motion with covariance (2.7), for any $2 < p < 3$. This convergence result will be instrumental in Section 4 to prove that the Cartan development in diffeomorphism spaces of the time rescaled kinetic Brownian motion in Hilbert spaces of vector fields converge to some limit dynamics as $\sigma$ increases to $\infty$. This will come as a direct consequence of the continuity of the Itô-Lyons solution map.

**Remark 2.12.** The idea of using rough paths theory for proving elementary homogenization results was first tested in the work [FGL15] of Friz, Gassiat and Lyons, in their study of the so-called *physical Brownian motion in a magnetic field.* That random process is described as a $C^1$ path $(x_t)_{0 \le t \le 1}$ in $\mathbb{R}^d$ modelling the motion of an object of mass $m$, with momentum $p = m\dot{x}$, subject to a damping force and a magnetic field. Its momentum satisfies a stochastic differential equation of Ornstein-Uhlenbeck form

$$dp_t = -\frac{1}{m} M p_t dt + dB_t,$$

for some matrix $M$, whose eigenvalues all have positive real parts, and $B$ is a $d$-dimensional Brownian motion. While the process $(Mx_t)_{0 \le t \le 1}$ is easily seen to converge to a Brownian motion $W$, its rough path lift is shown to converge in a rough paths sense in $L^q$, for any $q \ge 2$, to a random rough path *different from* the Stratonovich Brownian rough path associated to $W$.

A number of works have followed this approach to homogenization problems for fast-slow systems; see [ABT15, KM16, KM17, BC17, CFK+19] for a sample. $\triangle$

## 2.5 Rough paths invariance principle for the canonical lift

As in Section 2.3, *we assume in all of this section that the initial condition $v_0$ of the velocity process of kinetic Brownian motion is distribued according to its invariant probability measure $\mu$, from Theorem 2.1.*

Let $\mathbf{X}^\sigma = (X^\sigma, \mathbb{X}^\sigma)$ stand for the canonical rough path associated to the random $\mathcal{C}^1$ path $X^\sigma$, where we recall that

$$\mathbb{X}^\sigma_{ts} = \int_s^t (X^\sigma_u - X^\sigma_s) \otimes dX^\sigma_u = \frac{1}{\sigma^4} \int_{\sigma^4 s}^{\sigma^4 t} \int_{\sigma^4 s}^{u} v_r \otimes v_u \, drdu.$$

Recall that the tensor space $H \otimes H$ is equipped with its natural complete Hilbert–Schmidt norm.

### Tightness in rough paths space

We first establish the following tighness estimate.

**Proposition 2.13.** *For any $p \geq 2$, we have*

$$\sup_{\sigma>0} \mathbb{E}\big[|\mathbb{X}^\sigma_{t,s}|^p\big] \lesssim |t-s|^p.$$

It follows in particular from Proposition 2.6, the above Proposition 2.13 and the known Kolmogorov–Lamperti criterion for rough paths that the family of laws $(\mathcal{L}(\mathbf{X}^\sigma))_{\sigma>0}$ is tight in $\mathrm{RP}(\alpha^{-1})$, for any $1/3 < \alpha < 1/2$.

*Proof.* The statement of the Proposition is a consequence of the estimate

$$\mathbb{E}\left[\left| \int_0^T \int_0^t v_s \otimes v_t \, dsdt \right|^p\right] \lesssim_p T^p,$$

for $T \geq 1$; we prove the latter. We use for that purpose the same kind of multiscale martingale/coboundary decomposition as in the proof of Lemma 2.8. Let $k$ the unique integer such that

$$1 \leq \delta := T2^{-k} < 2.$$

Define

$$A_j := \int_{j\delta}^{(j+1)\delta} \int_0^t v_s \otimes v_t \, dsdt,$$

and

$$\widehat{\mathcal{F}}_j := \mathcal{F}_{(j+1)\delta} = \sigma\Big(v_s, s \leq (j+1)\delta\Big).$$

As above, we can assume without loss of generality that $\widehat{\mathcal{F}}_j$ is defined for all $j \geq -2^{k+1}$, as $v_0$ is assumed to be distributed according to the invariant probability measure of the velocity process. Then the integral rewrites as

$$\begin{aligned}\int_0^T \int_0^t v_s \otimes v_t \, dsdt \\ &= \Big(A_0 - \mathbb{E}\big[A_0|\widehat{\mathcal{F}}_{-1}\big]\Big) + \cdots + \Big(A_{2^k-1} - \mathbb{E}\big[A_{2^k-1}|\widehat{\mathcal{F}}_{2^k-2}\big]\Big) \\ &\quad + \mathbb{E}\big[A_0|\widehat{\mathcal{F}}_{-1}\big] + \cdots + \mathbb{E}\big[A_{2^k-1}|\widehat{\mathcal{F}}_{2^k-2}\big].\end{aligned} \tag{2.12}$$

The first sum is a martingale difference with respect to $(\widehat{\mathcal{F}}_n)_{n\geq 0}$, albeit not stationary,

$$\begin{aligned}\mathbb{E}\Big[\Big|\sum_{0\leq j<2^k}\Big(A_j-\mathbb{E}\big[A_j\big|\widehat{\mathcal{F}}_{j-1}\big]\Big)\Big|^p\Big]^{\frac{1}{p}}&\\ &\lesssim_p 2^{k/2}\bigg(2^{-k}\sum_{0\leq j<2^k}\mathbb{E}\Big[\big|A_j-\mathbb{E}\big[A_j\big|\widehat{\mathcal{F}}_{j-1}\big]\big|^p\Big]\bigg)^{\frac{1}{p}}\\ &\lesssim_p 2^{k/2}\bigg(2^{-k}\sum_{0\leq j<2^k}\mathbb{E}\big[\big|A_j\big|^p\big]\bigg)^{\frac{1}{p}}.\end{aligned}$$

Each term is controlled using Lemma 2.6, and the fact that $|v_t|=1$,

$$\mathbb{E}\big[\big|A_j\big|^p\big]\leq\delta^{p-1}\int_{j\delta}^{(j+1)\delta}\mathbb{E}\Big[\Big|\int_0^t v_s\,ds\Big|^p\Big]\,dt\lesssim_p\int_{j\delta}^{(j+1)\delta}t^{p/2}\,dt\lesssim(2^k)^{p/2},$$

so the $L^p$ norm of the first sum in (2.12) is bounded above by $2^k$, up to a constant depending only on $p$.

The second sum in 2.12 is treated as in the proof of Lemma 2.8. Set here

$$Z_j^{(n)}:=\mathbb{E}\Big[A_{j2^n}+\cdots+A_{j2^n+(2^n-1)}\Big|\widehat{\mathcal{F}}_{j-1}^{(n)}\Big],$$

with

$$\widehat{\mathcal{F}}_j^{(n)}:=\widehat{\mathcal{F}}_{(j-1)2^n}.$$

One has

$$\begin{aligned}\mathbb{E}\bigg[\bigg|\sum_{0\leq j<2^k}\mathbb{E}\big[A_j\big|\widehat{\mathcal{F}}_{j-1}\big]\bigg|^p\bigg]^{\frac{1}{p}}&\\ &\lesssim_p\sum_{0\leq n\leq k}2^{(k-n)/2}\bigg(\frac{1}{2^{k-n}}\Big(\mathbb{E}\big[\big|Z_0^{(n)}\big|^p\big]+\cdots+\mathbb{E}\big[\big|Z_{2^{k-n}-1}^{(n)}\big|^p\big]\Big)\bigg)^{\frac{1}{p}},\end{aligned}$$

and we are left with the study of the moments of the $Z_j^{(n)}$. These random variables are the conditional expectation of a double integral, which can be decomposed at time $(j-1)2^n\delta+\delta$ as follows.

$$\begin{aligned}Z_j^{(n)}&=\mathbb{E}\left[\int_{j2^n\delta}^{(j+1)2^n\delta}\int_0^t v_s\otimes v_t\,dsdt\,\Big|\widehat{\mathcal{F}}_{(j-1)2^n}\right]\\ &=\int_{j2^\ell\delta}^{(j+1)2^n\delta}\int_0^{(j-1)2^n\delta+\delta\vee 0}v_s\otimes\mathbb{E}\Big[v_t\,\big|\,\widehat{\mathcal{F}}_{(j-1)2^n}\Big]\,dsdt\\ &\quad+\int_{j2^n\delta}^{(j+1)2^n\delta}\mathbb{E}\left[\int_{(j-1)2^n\delta+\delta\vee 0}^t v_s\otimes\mathbb{E}\big[v_t|\mathcal{F}_s\big]\,ds\,\Big|\,\widehat{\mathcal{F}}_{(j-1)2^n}\right]\,dt\\ &=:R_j^{(n)}+S_j^{(n)}.\end{aligned}$$

Because the conditioning is from a distant past, the first term is controlled using the exponential mixing and the estimate of the proof of Proposition 2.6.

$$\begin{aligned}\mathbb{E}\Big[\big|R_j^{(n)}\big|^p\Big] &= \mathbb{E}\left[\left|\int_0^{(j-1)2^n\delta+\delta\vee 0} v_s\,ds\right|^p \left|\int_{j2^n\delta}^{(j+1)2^n\delta} \mathbb{E}\big[v_t \,\big|\, \widehat{\mathcal{F}}_{(j-1)2^n}\big]\,dt\right|^p\right] \\ &\lesssim \mathbb{E}\left[\left|\int_0^{(j-1)2^n\delta+\delta\vee 0} v_s\,ds\right|^p\right] \left(\int_{2^n\delta}^{2^{n+1}\delta} e^{-(t-\delta)/\tau}\,dt\right)^p \\ &\lesssim_p (2^{k-n})^{\frac{p}{2}}(2^n)^{\frac{p}{2}} e^{-p2^n/\tau}.\end{aligned}$$

When dealing with the second term, we use the stationarity of $v$ to write

$$\begin{aligned}\big|S_j^{(n)}\big| &\le \int_{j2^n\delta}^{(j+1)2^n\delta} \mathbb{E}\left[\int_{(j-1)2^n\delta+\delta}^t \Big|v_s\otimes \mathbb{E}\big[v_t|\mathcal{F}_s\big]\Big|\,ds \,\Big|\, \widehat{\mathcal{F}}_{(j-1)2^n}\right]dt \\ &\overset{\mathcal{L}}{=} \int_{2^n\delta}^{2^{n+1}\delta} \mathbb{E}\left[\int_\delta^t \Big|v_s\otimes \mathbb{E}\big[v_t|\mathcal{F}_s\big]\Big|\,ds \,\Big|\, \widehat{\mathcal{F}}_0\right]dt \\ &\lesssim \int_{2^n\delta}^{2^{n+1}\delta} \mathbb{E}\Big[\int_\delta^t e^{-(t-s)/\tau}\,ds \,\Big|\, \widehat{\mathcal{F}}_0\Big]\,dt \\ &\lesssim 2^n.\end{aligned}$$

Now we have, for each $0\le n\le k$ and $0\le j<2^{k-n}$,

$$\mathbb{E}\big[\big|Z_j^{(\ell)}\big|^p\big] \lesssim_p (2^{k-\ell})^{\frac{p}{2}}(2^\ell)^{\frac{p}{2}}\cdot e^{-p2^\ell/\tau} + 2^{p\ell}$$

so we eventually have

$$\begin{aligned}\mathbb{E}\left[\left|\sum_{0\le j<n}\mathbb{E}\big[A_j|\widehat{\mathcal{F}}_{j-1}\big]\right|^p\right]^{\frac{1}{p}} &\lesssim_p \sum_{0\le \ell\le k}\big(2^{k-\ell}2^{\ell/2}e^{-2^\ell/\tau} + 2^{(k-\ell)/2}2^\ell\big) \\ &= 2^k\sum_{0\le\ell\le k}\big(2^{-\ell/2}\,e^{-2^\ell/\tau} + 2^{-(k-\ell)/2}\big) \\ &= 2^k\sum_{0\le\ell\le k}2^{-\ell/2}\big(1+e^{-2^\ell/\tau}\big).\end{aligned}$$

This last sum is convergent, so the $L^p$ norm of the second term in (2.12) is no greater than a constant multiple of $2^k$. □

**Convergence in rough path space**

We are now ready to state and prove the main result of this section.

**Theorem 2.14.** *Pick $1/3<\alpha<1/2$. The processes $\mathbf{X}^\sigma$ converge in law in $\mathsf{RP}(\alpha^{-1})$, as $\sigma$ goes to $\infty$, to the Stratonovich Brownian rough path with covariance*

$$C(\ell,\ell') = \int_0^\infty \mathbb{E}\Big[\ell(v_0)\ell'(v_t)+\ell'(v_0)\ell(v_t)\Big]\,dt.$$

Let $\mathbf{X}$ be a random weak geometric $\alpha^{-1}$-rough path with distribution an arbitrary limit point of the family of laws of the $\mathbf{X}^\sigma$. Write $\mathbf{X} = (B, \mathbb{X})$, with $B$ a Brownian motion with the above covariance. Denote by $\underline{\mathbf{X}}$ the projection of $\mathbf{X}$ on the finite dimensional space generated by the first $d$ vectors of the basis $(e_i)$ from Section 2.1 – we use below the associated coordinate system. Using a monotone class argument and the tightness result stated in Proposition 2.13, the statement of Theorem 2.14 is a consequence of the following result, given that $d \geq 1$ is arbitrary.

**Lemma 2.15.** *The $d$-dimensional random rough path $\underline{\mathbf{X}}$ is a Stratonovich Brownian rough path with associated covariance matrix* $\mathrm{diag}(\gamma_1, \cdots, \gamma_d)$, *with*

$$\gamma_i := 2 \int_0^\infty \mathbb{E}\big[v_0^i v_t^i\big]\, dt.$$

*Proof.* Let $G_d^2$ stand for the step-2 nilpotent Lie group over $\mathbb{R}^d$. We prove that the process $(\underline{\mathbf{X}}_{t0})_{0\leq t\leq 1}$ is a $G_d^2$-valued Brownian motion by showing that it has stationary, independent, increments. The stationarity is inherited from the stationarity of the $\mathbf{X}^\sigma$. The independence of the increments of $\underline{\mathbf{X}}$ on disjoint closed intervals is a consequence of Corollary 2.4 on the convergence to 0 of the mixing coefficient of $(v_t)$. Continuity of $\underline{\mathbf{X}}$ allows to extend the result to adjacent time intervals.

We identify the generator of the $G_d^2$-valued Brownian motion $(\underline{\mathbf{X}}_t)$ as the generator of the $d$-dimensional Stratonovich Brownian rough path following the method of [Per18]. We recall the details for the reader's convenience. Note that we only need to consider the joint dynamics of $\underline{B}_t$ and the antisymmetric part $(\underline{\mathbb{A}}_t)$ of $(\underline{\mathbb{X}}_t)$; the former takes values in the Lie algebra $\mathfrak{g}_d^2$ of $G_d^2$ – a linear space. Denote by $\underline{\mathbb{A}}^B$ the antisymmetric part of Stratonovich Brownian rough path associated with $\underline{B}$. We then have, for any smooth real-valued function $f$ on $\mathbb{R}^d \times \mathfrak{g}_d^2$ with compact support, the identity

$$\begin{aligned}
&\Big(f\big(\underline{B}_t, \underline{\mathbb{A}}_t\big) - f(0)\Big) - \Big(f\big(\underline{B}_t, \underline{\mathbb{A}}_t^B\big) - f(0)\Big) \\
&= (\partial_2 f)(\underline{B}_t, 0)\big(\underline{\mathbb{A}}_t - \underline{\mathbb{A}}_t^B\big) + O\Big(\big|\underline{\mathbb{A}}_t - \underline{\mathbb{A}}_t^B\big|^2\Big) \\
&= \Big((\partial_2 f)(\underline{B}_t, 0) - (\partial_2 f)(0,0)\Big)\big(\underline{\mathbb{A}}_t - \underline{\mathbb{A}}_t^B\big) + (\partial_2 f)(0,0)\big(\underline{\mathbb{A}}_t - \underline{\mathbb{A}}_t^B\big) + O\Big(\big|\underline{\mathbb{A}}_t - \underline{\mathbb{A}}_t^B\big|^2\Big).
\end{aligned}$$

The conclusion follows by multiplying by $t^{-1}$ and taking expectation, sending $t$ to 0, after recalling that $\underline{\mathbb{A}}_t$ and $\underline{\mathbb{A}}_t^B$ are centered, and recalling the uniform estimates from Proposition 2.13 under the form

$$\big\|\underline{\mathbb{A}}_t\big\|_{L^2} \vee \big\|\underline{\mathbb{A}}_t^B\big\|_{L^2} \lesssim t.$$

□

# 3 Geometry of the configuration space

We now recall some elements of geometry of the configuration space associated with a compact Riemannian manifold. In particular, in Sections 3.3 and 3.4 below, we make explicit the notions of parallel transport and Cartan development in that context, as solving ordinary differential equations driven by smooth vector fields.

## 3.1 Configuration space

Let $(M, g)$ be a $d$-dimensional connected and oriented Riemannian manifold, and $\pi : F \to M$ a finite dimensional fiber bundle over $M$, with vertical bundle $VF \to M$. Think of the trivial bundles $M \times M \to M$, or $M \times TM \to M$, as typical examples. We collect from Palais' seminal work [Pal68] elementary results on the Hilbert manifold $H^s(F)$ of sections of $\pi$ with Sobolev regularity exponent $s > \frac{d}{2}$.

1. **Sobolev embedings** hold true, with in particular $H^s(F) \subset C^k(M, F)$, as soon as $s > k + \frac{d}{2}$ and $k \geq 0$.

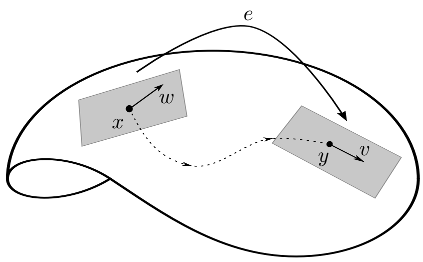
2. **Variations of $H^s$-sections of $F$.** The spaces $TH^s(F)$ and $H^s(VF)$ are isomorphic as Hilbert manifolds. This isomorphism accounts for the fact that an infinitesimal perturbation $(\delta f)$ of a section $f$ of $F$, reads as a collection of vertical tangent vectors $(\delta f)(x) \in V_{f(x)}F$, indexed by $x \in M$. As a particular example, for any finite dimensional manifold $N$, the spaces $TH^s(M, N)$ and $H^s(M, TN)$ are isomorphic.

3. For any two finite dimensional fiber bundles $F, G$ above $M$, the map
$$(f, g) \mapsto \big(x \in M \mapsto (f(x), g(x))\big)$$
is an isomorphism between $H^s(F) \times H^s(G)$ and $H^s(F \times_M G)$.

4. **Omega lemma.** Given a smooth fiber bundle morphism $\Phi : F \to G$, above $M$, set
$$\omega_\Phi(f) := \Phi \circ f,$$
for any section $f$ of $F$. Then $\omega_\Phi$ sends $H^s(F)$ in $H^s(G)$, and $d\omega_\Phi : TH^s(F) \to TH^s(G)$ is isomorphic to $\omega_{d\Phi} : H^s(VF) \to H^s(VG)$, via the isomorphisms $TH^s(F) \simeq H^s(VF)$ and $TH^s(F') \simeq H^s(VG)$.

For $s > \frac{d}{2}$, set
$$\mathscr{M} := H^s(M, M);$$
this will be the **configuration space** of our dynamics. Choosing $s > \frac{d}{2}$, ensures that $\mathscr{M} \subset C^0(M, M)$, by Sobolev embedings. The tangent space to this Hilbert manifold is given by
$$T\mathscr{M} \simeq H^s(M, TM),$$
from item (2) above. If $s > \frac{d}{2} + 1$, elements of $\mathscr{M}$ are $C^1$ maps from $M$ into itself. Recall in that case from Section 4 of [EM69] that the subset $\mathscr{M}_0$ of $\mathscr{M}$ of $H^s$ maps from $M$ into itself that preserve the volume form by pull-back is then a closed submanifold of $\mathscr{M}$, and that elements of $\mathscr{M}_0$ are diffeomorphisms. So $\mathscr{M}_0$ is a group. We shall always assume implicitly these constraints on the regularity exponent $s$, when talking about $\mathscr{M}$ or $\mathscr{M}_0$. We recall other elementary facts on $H^s(TM)$ at the end of this section.

To implement a version of Cartan's development machinery in the weak Riemannian setting of the next section, we introduce the following finite dimensional fiber bundles above $M$, seen below as the first component. Given $x, y \in M$, denote by $\mathsf{O}(T_xM, T_yM)$ the set of isometries from $T_xM$ to $T_yM$. Set
$$\begin{aligned}
F^{(e)} &:= \Big\{(x, y; e)\,;\, (x, y) \in M \times M, e \in \mathsf{O}(T_xM, T_yM)\Big\},\\
F^{(w)} &:= \Big\{(x, y; w)\,;\, (x, y) \in M \times M, w \in T_xM\Big\},\\
F^{(v)} &:= \Big\{(x, y; v)\,;\, (x, y) \in M \times M, v \in T_yM\Big\},\\
F^{(e,v)} &:= \Big\{\big(x, y; e, v\big)\,;\, (x, y) \in M \times M, e \in \mathsf{O}(T_xM, T_yM), v \in T_yM\Big\}.
\end{aligned}$$




*An infinitesimal rigid object $x$ is moving along a path. It has position $y$ and velocity $v$ at some time. Its orientation at that time is given by an isometry $e : T_xM \to T_yM$, and its velocity $v$ is given in its initial reference frame by $w$.*

FIGURE 3.3 – Description of the fibre bundles involved.

We understand $H^s(F^{(v)})$ as the set of $H^s$ maps from $M$ into $TM$, so $T\mathscr{M} \simeq H^s\big(F^{(v)}\big)$. We denote by $\big(\varphi(\cdot), v(\cdot)\big)$ a generic element of $H^s(F^{(v)})$. We have similar interpretations of the other $H^s$ spaces over the corresponding bundles, with similar notations. Since the map

$$\begin{aligned} F^{(e)} \times_{M\times M} F^{(w)} &\to F^{(v)} \\ \big((x,y;e),(x,y;w)\big) &\mapsto \big(x,y;e(w)\big), \end{aligned}$$

is a smooth bundle morphism, it follows from items 3 and 4 above, that it induces a *smooth map* from $H^s(F^{(w,e)})$ into $H^s(F^{(v)})$. Similarly, the smooth map

$$\begin{aligned} F^{(e,v)} &\to F^{(w)} \\ \big(x,y;e,v\big) &\mapsto \big(x,y;e^{-1}(v)\big), \end{aligned}$$

induces a *smooth map* from $H^s(F^{(e,v)})$ into $H^s(F^{(w)})$.

We refer the reader to the classic textbook [Ros97] for the following elementary facts from functional analysis about the Laplace operator $\Delta$ on vector fields on $M$. We take the convention that $-\Delta$ is a non-positive symmetric operator on $L^2(TM)$. This operator has compact resolvant, so one has an eigenspaces decomposition

$$L^2(TM) = \bigoplus_{n\geq 0} E_{\lambda_n}, \tag{3.13}$$

with finite dimensional eigenspaces $E_{\lambda_n}$, with corresponding non-positive eigenvalues $\lambda_n \downarrow -\infty$. Eigenvectors of $-\Delta$ are smooth, from elliptic regularity results. We recover the space $H^s(TM)$ described above setting

$$H^s(TM) = \left\{ f = \sum_{n\geq 0} f_n \in L^2(TM)\,;\, \sum_{n\geq 0} \lambda_n^s \|f_n\|_{L^2}^2 < \infty \right\}.$$

The 0-eigenspace is finite dimensional. Any choice of Euclidean norm $\|\cdot\|$ on it defines the topology of $H^s(TM)$, associated with the norm

$$\|f\|_s := \|f_0\| + \left( \sum_{n\geq 0} \lambda_n^s \|f_n\|_{L^2}^2 \right)^{1/2}.$$

## 3.2 Weak Riemannian structure on the configuration space

Denote by $\mathrm{vol}_g$ the Riemannian volume measure on $(M,g)$, and by $\exp : TM \to M$, its exponential map. The configuration space $\mathscr{M}$ is endowed with a smooth weak Riemannian structure, setting for any $\varphi \in \mathscr{M}$ and $X(\varphi), Y(\varphi) \in T_\varphi\mathscr{M}$,

$$\big(X(\varphi), Y(\varphi)\big)_\varphi := \int_M g_{\varphi(m)}\big(X(\varphi)(m), Y(\varphi)(m)\big)\,\mathrm{vol}_g(dm). \tag{3.14}$$

This formula defines by restriction a weak Riemannian metric on the space $\mathscr{M}_0$ of $H^s$ maps from $M$ into itself preserving the volume form. In that setting, notice that if $X(\varphi) = \mathbf{X}\circ\varphi$ and $Y(\varphi) = \mathbf{Y}\circ\varphi$, for some vector fields $\mathbf{X}, \mathbf{Y}$ on $M$, then the change of variable formula gives

$$\big(X(\varphi), Y(\varphi)\big)_\varphi = \int_M g_m\big(\mathbf{X}(m), \mathbf{Y}(m)\big)\,\mathrm{vol}_g(dm),$$

so the scalar product is in that case the $L^2$ scalar product of the vector fields $\mathbf{X}$ and $\mathbf{Y}$. The fact that the topology on $\mathscr{M}$ induced by the scalar product is weaker than the $H^s$-topology makes non-obvious the existence of a smooth Levi-Civita connection. Ebin and Marsden have proved that

- the $L^2$ metric (3.14) is a smooth function on $\mathscr{M}$,
- it has a smooth Levi-Civita connection $\overline{\nabla}$, with associated exponential map Exp well-defined and smooth in a neighbourhood of the zero section; it is explicitly given by

$$\mathrm{Exp}_\varphi(X)(m) = \exp_{\varphi(m)}\big(X(m)\big).$$

The geodesics of $(\mathscr{M}, \overline{\nabla})$ are defined for all times. Denote by $\nabla$ the Levi-Civita connection of $(M,g)$. For smooth right invariant vector fields $X, Y$ on $\mathscr{M}$, with $X(\varphi) = \mathbf{X}\circ\varphi$ and $Y(\varphi) = \mathbf{Y}\circ\varphi$, one has

$$(\overline{\nabla}_X Y)(\varphi) = (\nabla_{\mathbf{X}}\mathbf{Y})\circ\varphi.$$

The $L^2$-scalar product is right invariant on the group $\mathscr{M}_0$, from the change of variable formula. The Levi-Civita connection of the $L^2$ metric on the volume preserving configuration space $\mathscr{M}_0$ is explicitly given in terms of the Hodge projection operator $P$ on divergence-free vector fields on $M$. Denote by $R_\varphi$ the right composition by $\varphi$. For any $\varphi \in \mathscr{M}_0$, the map

$$P_\varphi := dR_\varphi \circ P \circ dR_\varphi^{-1}, \tag{3.15}$$

is indeed the orthogonal projection map from $T_\varphi\mathscr{M}$ into $T_\varphi\mathscr{M}_0$, and its depends smoothly on $\varphi \in \mathscr{M}_0$. So the Levi-Civita connection $\overline{\nabla}^0$ on $\mathscr{M}_0$ is given by

$$\overline{\nabla}^0 = P \circ \overline{\nabla};$$

it is a smooth map. Its associated exponential map is no longer given by the exponential map on $TM$, due to the non-local volume preserving constraint. Geodesics are not defined for all times anymore. Denote by Id the identity map on $M$. For smooth right invariant vector fields $X, Y$ on $\mathscr{M}$, with $X(\varphi) = \mathbf{X}\circ\varphi$ and $Y(\varphi) = \mathbf{Y}\circ\varphi$, for vector fields $\mathbf{X}, \mathbf{Y}$ on $M$, one has

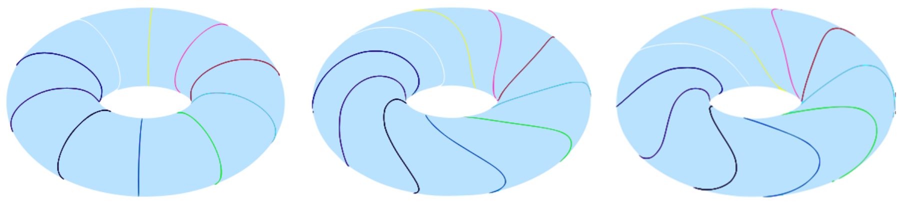
$$\big(\overline{\nabla}^0_X Y\big)(\mathrm{id}) = P\big(\nabla_{\mathbf{X}}\mathbf{Y}\big).$$

V.I. Arnol'd showed formally in his seminal work [Arn66] that the velocity field $u : [0,T] \to H^s(TM)$ of a geodesic $\varphi_t$ in $\mathscr{M}_0$, with $u_t := \dot\varphi_t \circ \varphi_t^{-1}$, is a solution to Euler's equation for the hydrodynamics of an incompressible fluid. Ebin and Marsden gave an analytical proof of that fact in their seminal work [EM69]. (Besides that classical reference, we refere the reader to Arnold and Khesin's book [AK98], or Smolentsev's thourough review [Smo07] for reference works on the weak Riemannian geometry of the configuration space.)

The flat two-dimensional torus $\mathbb{T}^2$ offers an interesting concrete example. Its symplectic structure allows to identify a Hilbert basis $(A_k, B_k)_{k\in\mathbb{Z}^2\setminus 0}$ of $T_{\mathrm{id}}\mathscr{M}_0$ from an eigenbasis for the Laplace operator on real-valued functions on $\mathbf{T}^2$; see e.g. Arnold and Khesin's book [AK98], Section 7 of Chap. 1. Denote by $\partial_1, \partial_2$ the constant vector fields in the coordinate directions, and $k = (k_1, k_2) \in \mathbb{Z}^2$. One has

$$A_k = |k|^{-1}\Big(k_2 \cos(k\cdot\theta)\partial_1 - k_1\cos(k\cdot\theta)\partial_2\Big),$$
$$B_k = |k|^{-1}\Big(k_2 \sin(k\cdot\theta)\partial_1 - k_1\sin(k\cdot\theta)\partial_2\Big).$$

One can see in the following simulations the image of axis circles by the time 1 map of the associated flow in $\mathbb{T}^2$, corresponding to different inital conditions for $u_0$, with $\varphi_0 = \mathrm{id}$. The simulations were done using an elementary finite dimensional approximation for the dynamics, using the explicit expressions for the Christoffel symbols first given by Arnold in [Arn66].



*Time 1 snapshots in the case $M = \mathbb{T}^2$ for different initial momenta.*

FIGURE 3.4 – The geodesic flow in $\mathscr{M}_0$ acting on great circles.

We will come back to this point in Section 3.4.

## 3.3 Parallel transport

We recast in this section the parallel transport operations in $\mathscr{M}$ and $\mathscr{M}_0$, using the bundles $F$ from Section 3.1. This allows to set the notations for the next section on Cartan development operation in $\mathscr{M}$ and $\mathscr{M}_0$. Recall $H^s(F)$ stands for $H^s$ sections from $M$ into the corresponding bundle $F$. We denote by $VF$ the vertical space in $TF$, for the canonical projection map $F \to M$. Recall also that $T_{\mathrm{id}}\mathscr{M}$ is simply the set of $H^s$ vector fields on $M$.

Denote by $K : TTM \to TM$, the connector associated with the Levi-Civita connection $\nabla$ on $M$. So, for a path $\gamma_t = (m_t, v_t)$ in $TM$, one has

$$\nabla_{\dot m_t} v_t = K(\dot\gamma_t),$$

and

$$\nabla_{\mathbf{X}}\mathbf{Y} = K\big((d\mathbf{Y})(\mathbf{X})\big),$$

for any smooth vector fields $\mathbf{X}, \mathbf{Y}$ on $M$. The second order tangent bundle $TT\mathscr{M}$ of $\mathscr{M}$ identifies with $H^s(M, TTM)$. The connector $\overline{K}$ associated with the $L^2$-Levi-Civita connection $\overline{\nabla}$ is given, for a section $Y$ of $TTM$ over an element of $\mathscr{M}$, by

$$\overline{K}(Y) := K \circ Y \in T\mathscr{M}.$$

Set

$$F^{(v,\dot{y})} := F^{(v)} \times_{M\times M} TM.$$

One defines a smooth one form on $F^{(v,\dot{y})}$, with values in $TF^{(v)}$, by requiring that $\nabla_{\dot{y}_t} v_t = 0$ iff

$$\frac{d}{dt}\,(y_t, v_t) = \mathfrak{H}^{(v)}(y_t, v_t; \dot{y}_t).$$

We choose the letter $\mathfrak{H}$, for this horizontal lift of the connection. In simple terms, for any fixed $(y,v) \in TM$, the linear map $\mathfrak{H}^{(v)}(y,v;\cdot)$ identifies the space $T_yM$ to the horizontal subspace of $T_{(y,v)}TM$, via the usual horizontal lift. Note that the definition of $\mathfrak{H}^{(v)}(y,v;\dot{y})$ does not depend on the base point $x \in M$, for a generic element $(x,y;v) \in F^{(v)}$ and $\dot{y} \in T_yM$.

Denote also by $\mathfrak{H}^{(e)}$ the smooth one form on $V_2F^{(v)}$ with values in the space of vector field on $F^{(e)}$, such that for any path $(x, y_t; e_t)$ in $F^{(e)}$, and any vector $w \in T_xM$, the vector $e_t(w) \in T_{y_t}M$ is transported parallely along the $M$-valued path $(y_t)$ iff

$$\frac{d}{dt}\,(y_t, e_t) = \mathfrak{H}^{(e)}(y_t, e_t; \dot{y}_t).$$

Here again, the base point $x \in M$ is not involved in the definition of the tangent vector $\mathfrak{H}^{(e)}(y,e;\dot{y})$, for a generic element $(x,y;e) \in F^{(e)}$ and $\dot{y} \in T_yM$. Pick

$$(x_0, y_0; e_0) \in F^{(e)},$$

and note that for any vertical vector

$$(\dot{y}, \dot{e}) \in V_{(x_0,y_0;e_0)}F^{(e)},$$

and $v_0 \in T_{y_0}M$, one has

$$(\dot{y}, \dot{e}) = \mathfrak{H}^{(e)}\big(y_0, e_0; v_0\big)$$

iff

$$\big(\dot{y}, \dot{e}(w)\big) = \mathfrak{H}^{(v)}\big(y_0, e_0(w); v_0\big) \in V_{(x_0,y_0;e_0(w))}F^{(v)},$$

for any $w \in T_yM$, with $\dot{e}(w)$ defined naturally. It follows from the Omega Lemma that one defines a *smooth* operator from $H^s\big(F^{(v,\dot{y})}\big)$ to $TH^s(F^{(v)})$, setting

$$\overline{\mathfrak{H}}^{(v)}\big(\varphi(\cdot), v(\cdot); \dot{\varphi}(\cdot)\big) := \mathfrak{H}^{(v)} \circ \big(\varphi(\cdot), v(\cdot); \dot{\varphi}(\cdot)\big).$$

Similarly, we define a *smooth one-form* on $T_{\mathrm{id}}\mathscr{M}$ with values in vector fields on $H^s(F^{(e)})$, setting

$$\overline{\mathfrak{H}}^{e}\big(\varphi(\cdot), e(\cdot); \mathbf{X}\big) := \mathfrak{H}^{(e)} \circ \big(\varphi(\cdot), e(\cdot); e(\mathbf{X})\big), \qquad \mathbf{X} \in T_{\mathrm{id}}\mathscr{M}.$$

**Proposition 3.1.** *Given a path $\big(\varphi_t(\cdot); e_t(\cdot), v_t(\cdot)\big)_{0\leq t\leq 1}$ in $H^s(F^{(e,v)})$, one has pointwise*

$$\frac{d}{dt}\,\big(\varphi_t(x), e_t(x)\big) = \mathfrak{H}^{(e)}\big(\varphi_t(x), e_t(x); v_t(x)\big),$$

*for all $x \in M$, iff*

$$\frac{d}{dt}\,\big(\varphi_t, e_t(\mathbf{X})\big) = \overline{\mathfrak{H}}^{(v)}\big(\varphi_t, e_t(\mathbf{X}); v_t\big),$$

*for every $\mathbf{X} \in T_{\mathrm{id}}\mathscr{M}$.*

The next two propositions give a description of parallel transport in $\mathscr{M}$ and $\mathscr{M}_0$, respectively, in terms of the vector field $\overline{\mathfrak{H}}^{(v)}$ on $H^s(F^{(v)})$.

**Proposition 3.2.** *Let $\big(\varphi_t(\cdot), v_t(\cdot)\big)_{0\le t\le 1}$ be a $T\mathscr{M}$-valued path. Then*

$$\overline{\nabla}_{\dot\varphi_t} v_t = 0,$$

*iff*

$$\frac{d}{dt}\big(\varphi_t, v_t\big) = \overline{\mathfrak{H}}^{(v)}\big(\varphi_t, v_t; \dot\varphi_t\big).$$

*Proof.* Given $(y, v) \in TM$, the following map identifies $T_yM$ with the vertical subspace of $T_{(y,v)}TM$

$$\mathfrak{V}^{(v)}(y, v; \cdot) : w \in T_yM \mapsto \frac{d}{dt}_{\big|t=0}(v + tw) \in T_{(y,v)}TM.$$

For any $(x, y; v) \in F^{(v)}$ and $u \in T_{(y,v)}\big(F_x^{(v)}\big)$, one then has

$$u = \mathfrak{H}^{(v)}(y, v; a) + \mathfrak{V}^{(v)}(y, v; b) \quad \text{ iff } \quad a = dp_2(u), \text{ and } b = K(u).$$

For an $H^s(F_v)$-valued path $\big(\varphi_t(\cdot), v_t(\cdot)\big)$, one then has the splitting

$$\begin{aligned} \frac{d}{dt}\,(\varphi_t, v_t) &= \mathfrak{V}^{(v)} \circ \big(\varphi_t, v_t; K \circ \dot v_t)\big) + \mathfrak{H}^{(v)} \circ \big(\varphi_t, v_t; \dot\varphi_t\big) \\ &= \mathfrak{V}^{(v)} \circ \big(\varphi_t, v_t; \overline{\nabla}_{\dot\varphi_t} v_t)\big) + \overline{\mathfrak{H}}^{(v)}\big(\varphi_t, v_t; \dot\varphi_t\big). \end{aligned} \tag{3.16}$$

The result follows because composition by $\mathfrak{V}_v(y, v; \cdot)$ is one-to-one. □

Recall that $P$ stands for Hodge projector on divergence-free vector fields.

**Proposition 3.3.** *Let $\big(\varphi_t(\cdot), v_t(\cdot)\big)_{0\le t\le 1}$ be a $T\mathscr{M}_0$-valued path. Then*

$$\overline{\nabla}^0_{\dot\varphi_t} v_t = 0,$$

*iff*

$$\frac{d}{dt}\big(\varphi_t, v_t\big) = (dP)\Big(\overline{\mathfrak{H}}^{(v)}\big(\varphi_t, v_t, ; \dot\varphi_t\big)\Big).$$

*Proof.* Write $T_{\mathscr{M}_0}\mathscr{M}$ for the section of $T\mathscr{M}$ above $\mathscr{M}_0$, and write $Q := \mathrm{id} - P : T_{\mathscr{M}_0}\mathscr{M} \to T_{\mathscr{M}_0}\mathscr{M}$, for the projection on the orthogonal in $T\mathscr{M}$ of $T\mathscr{M}_0$. Note that the differential $dP$ of $P$ identifies to $P$ in the fibers, since it is linear. The identification is up to an isomorphism which is exactly the composition by $\mathfrak{V}_v$, in the sense that

$$dP\big(\mathfrak{V}^{(v)}(\varphi, v; v')\big) = \mathfrak{V}^{(v)} \circ \big(\varphi, v; P(v')\big)$$

for any $v, v' \in T_\varphi\mathscr{M}$. As we work with a $T\mathscr{M}_0$-valued path $(\varphi_t, v_t)$, one has $Q(v_t) = 0$, at all times, so differentiating this identity with respect to $t$ gives

$$dQ(\dot v_t) = 0.$$

Since $P + Q = \mathrm{id}$, we can conclude with the decomposition (3.16), by rewriting the expression for the time derivative under the form

$$\begin{aligned} \frac{dv_t}{dt} &= dP(\dot v_t) + dQ(\dot v_t) \\ &= \mathfrak{V}^{(v)} \circ \Big(\varphi_t, v_t; P\big(K(\dot v_t)\big)\Big) + dP\big(\mathfrak{H}^{(v)} \circ \big(\varphi_t, v_t; \dot\varphi_t\big)\big) \\ &= \mathfrak{V}^{(v)} \circ \Big(\varphi_t, v_t; \overline{\nabla}^0_{\dot\varphi_t} v_t\Big) + dP\Big(\overline{\mathfrak{H}}^{(v)}\big(\varphi_t, v_t; \dot\varphi_t\big)\Big). \end{aligned}$$

□

## 3.4 Cartan and Lie developments

In a finite dimensional context, Cartan's moving frame method [Car01] provides a mechanics for constructing $C^1$ paths on $M$ from $C^1$ path on $\mathbb{R}^d$, giving something of a chart on pathspace in $M$. Its description requires the introduction of the orthonormal frame bundle $OM$ over $M$. It is made up of pairs $z = (m, e)$, with $m \in M$ and $e$ an isometry from $\mathbb{R}^d$ to $T_mM$. It has a natural finite dimensional manifold structure, and the Riemannian connection on $TM$ induces vector fields $H_1, \dots, H_d$ on $OM$ by parallel transport of a frame in the direction of its $i^{\text{th}}$ direction along the corresponding path in $M$.

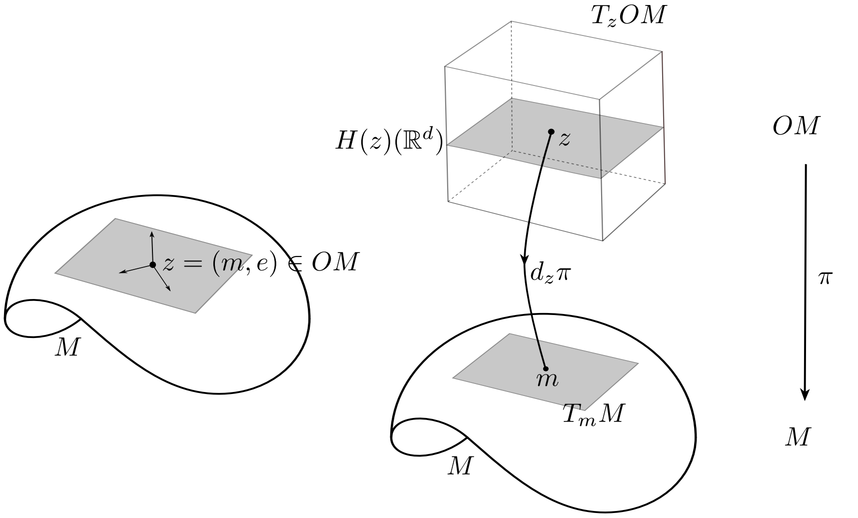



FIGURE 3.5 – For $z \in OM$ and $a = (a_1, \dots, a_d) \in \mathbb{R}^d$, we have $H(z)(a) := \sum_{i=1}^d a_i H_i(z) \in T_zOM$.

The development in $M$ of a path $(x_t)_{0\leq t\leq 1}$ in $\mathbb{R}^d$ is the natural projection $(m_t)$ in $M$ of the $OM$-valued path $(z_t)$ solution to the equation

$$\dot{z}_t = H(z_t)(\dot{x}_t).$$

Explosion may happen before time 1. This path in $M$ depends not only on $m_0$ but also on $e_0$. Conversely, given any $C^1$ path $(m_t)_{0\leq t\leq 1}$ in $M$ and $z_0 = (m_0, e_0) \in OM$ above $m_0$, parallel transport of $e_0$ along the path $(m_t)_{0\leq t\leq 1}$ defines a path $(z_t)_{0\leq t\leq 1}$ in $OM$, and setting $x_t := \int_0^t e_s^{-1}(\dot{m}_s)\, ds$, defines a path in $\mathbb{R}^d$ whose Cartan development is $(m_t)_{0\leq t\leq 1}$. Geodesics are Cartan's development of straight lines in $\mathbb{R}^d$.

We recast the definition of Cartan development given above in a finite dimensional setting in the following form well suited for the present infinite dimensional setting.

**Definition 3.4.** Let a $C^1$ path $(\mathbf{X}_t)$ in $T_{\mathrm{id}}\mathscr{M}$ be given. An $\mathscr{M}$-valued path $(\varphi_t)$ is the **Cartan development of** $(\mathbf{X}_t)$ if there exists a family

$$e_t : T_{\mathrm{id}}\mathscr{M} \to T_{\varphi_t}\mathscr{M},$$

of bounded linear maps, with $e_0 = \mathrm{id}$, such that

$$\begin{aligned} &\dot{\varphi}_t = e_t(\dot{\mathbf{X}}_t), \\ &\overline{\nabla}_{\dot{\varphi}_t} e_t(\mathbf{Y}) = 0, \qquad \text{for all } \mathbf{Y} \in T_{\mathrm{id}}\mathscr{M}, \end{aligned} \tag{3.17}$$

at all times where $\varphi_t$ is well-defined. $\triangle$

This definition conveys the same picture as above. The map $e_t$, named 'frame', is transported parallely along the path $(\varphi_t)$, while $\dot{\varphi}_t$ is given by the image by $e_t$ of $\dot{\mathbf{X}}_t$. The existence of a unique Cartan development for a path $(\mathbf{X}_t)$ in $T_{\mathrm{id}}\mathscr{M}$ is elementary in that case. It follows from Proposition 3.2 that equation (3.17) is equivalent to requiring that the $H^s\big(F^{(e)}\big)$-valued path $(\varphi_t, e_t)$ satisfies the equation

$$\frac{d}{dt}(\varphi_t, e_t) = \overline{\mathfrak{H}}^e\big(\varphi_t, e_t; \dot{\mathbf{X}}_t\big). \tag{3.18}$$

Since the one-form $\overline{\mathfrak{H}}^e$ is smooth, this equation has a unique solution until its possibly finite explosion time.

Here is now the form of Cartan development dynamics in $\mathscr{M}_0$. Recall $T_{\mathrm{id}}\mathscr{M}_0$ is the set of $H^s$ divergence-free vector fields on $M$.

**Definition 3.5.** Let a $C^1$ path $(\mathbf{X}_t)$ in $T_{\mathrm{id}}\mathscr{M}_0$ be given. An $\mathscr{M}_0$-valued path $(\varphi_t)$ is the **Cartan development of** $(\mathbf{X}_t)$ if there exists a family

$$e_t : T_{\mathrm{id}}\mathscr{M}_0 \to T_{\varphi_t}\mathscr{M}_0,$$

of bounded linear maps, with $e_0 = \mathrm{id}$, such that

$$\begin{aligned} &\dot{\varphi}_t = e_t(\dot{\mathbf{X}}_t), \\ &\overline{\nabla}^0_{\dot{\varphi}_t} e_t(\mathbf{Y}) = 0, \qquad \text{for all } \mathbf{Y} \in T_{\mathrm{id}}\mathscr{M}_0, \end{aligned} \tag{3.19}$$

at all times where $\varphi_t$ is well-defined. $\triangle$

The proof of existence of a unique solution to Cartan's development system (3.19) in $\mathscr{M}_0$ is not fundamentally different from the case of $\mathscr{M}$, and uses Proposition 3.3 instead of Proposition 3.2. It is however more technical, and full details are given in Appendix 5. The system is recast as a controlled ordinary differential equation in the state space

$$\mathscr{Z} := H^s(F^{(e)}) \times \mathsf{L}\big(H^s(TM)\big),$$

with generic element $\big((\varphi, e), f\big)$, and dynamics of the form

$$\frac{d}{dt}\,(\varphi_t, e_t) = \overline{\mathfrak{H}}^e\Big(\varphi_t, e_t; f_t(\dot{\mathbf{X}}_t)\Big), \quad \frac{d}{dt}\,f_t = \overline{\mathfrak{H}}^f\left(\frac{d}{dt}\,(\varphi_t, e_t), f_t\right),$$

driven by a *smooth* vector field-valued one form on $T_{\mathrm{id}}\mathscr{M}_0$. We use Cartan's development map in the configuration manifolds $\mathscr{M}$ and $\mathscr{M}_0$ in the next section. We conclude this section by a brief comparison between Cartan development and the Lie group notion of development, commonly used to define the stochastic Euler equation.

Let $G$ stand for a finite dimensional Lie group with Lie algebra $\textsc{Lie}(G)$. Lie's development operation provides another way of constructing paths $(g_t)_{0\leq t\leq 1}$ with values in $G$ from paths $(x_t)_{0\leq t\leq 1}$ in $\mathbb{R}^d$, by identifying $T_{g_0}G$ and $\mathbb{R}^d$ via a linear map $\iota_0$, and solving the ordinary differential equation $\dot{g}_t = \iota_0(\dot{x}_t)\, g_t$. In such a group setting, Malliavin and Airault [AM06] gave a correspondance between the Cartan and Lie notions of development, although this was certainly known to practitioners before; see also [CFM07]. Choose an orthonormal basis of the Lie algebra of $G$, and denote by $c^n_{k,\ell}$ the structure constants, so the Christoffel symbols are given by $\Gamma^n_{k,\ell} = \frac{1}{2}\big(c^n_{k,\ell} - c^k_{\ell,n} + c^\ell_{n,k}\big)$. Write $\Gamma_k$ for the antisymmetric endomorphism with matrix $\Gamma^{\cdot}_{k,\cdot}$ in the chosen basis, for $1 \leq k \leq d$, and consider $\Gamma$ as a linear map from $\mathbb{R}^d$ into the set of antisymmetric endomorphism of the Lie algebra. Denote by $O\textsc{Lie}(G)$ the orthonormal group of $\textsc{Lie}(G)$.

**Proposition 3.6.** *Let $(w_t)_{0\leq t\leq 1}$ be a $C^1$ path in the Lie algebra of $G$. The path $(g_t)_{0\leq t\leq 1}$ solution to the $\big(O\textsc{Lie}(G)\times G\big)$-valued equation*

$$\begin{aligned} dO_t &:= O_t\,\Gamma(\dot w_t)\,dt, \quad O_0 = \mathrm{id}, \\ dg_t &= O_t(\dot w_t)g_t, \end{aligned} \tag{3.20}$$

*is the Cartan development of the path $(w_t)$.*

(The system (3.20) is reminiscent of the equation in

$$H^s(F^{(e)})\times \mathsf{L}\big(H^s(TM)\big)$$

from Appendix 5, recasting Cartan's development dynamics in $\mathscr{M}_0$.) The geodesic started from the identity of $G$, with direction $\omega\in \textsc{Lie}(G)$, is in particular given in the Lie picture as the solution $(g_t)_{0\leq t\leq 1}$ to the equation

$$\dot g_t = \exp\big(t\Gamma(\omega)\big)(\omega)\,g_t.$$

Note that $\exp\big(t\Gamma(\omega)\big)(\omega)\in \textsc{Lie}(G)$. Note also that it is the fact that the Christoffel symbols are constants that allows to reduce the second order differential equation for the geodesics on a generic Riemannian manifold into a first order differential equation, in a Riemannian Lie group setting.

Following Euler's picture, it is this group-oriented point of view that has been considered so far in the geometric viewpoint on fluid hydrodynamics, deterministic or stochastic. The naive implementation of Cartan's machinery in terms of Lie development runs into trouble in the infinite dimensional setting of $\mathscr{M}$ or $\mathscr{M}_0$. This can be seen on the example of the two dimensional torus and the volume preserving diffeomorphism group as a consequence of the fact that Christoffel symbols define antisymmetric unbounded operators that have no good exponential in the orthonormal group of $T_{\mathrm{id}}\mathscr{M}_0$. The problem comes from the fact that $\mathscr{M}$ of $\mathscr{M}_0$ have a *fixed* regularity. See Malliavin's works [Mal99, CFM07] for a quantification of the loss of regularity of Brownian motion in the set of homeomorphisms of the circle, as time increases. The Lie development picture of Cartan's development map can however be used for numerical purposes for simulating kinetic Brownian motion in $\mathscr{M}_0$. In the context of Equation (3.20), it corresponds to having $\dot w_t$ a Brownian motion on the unit sphere of the $H^s$ space of divergence-free vector fields on $M$; see Section 4.2 below for various examples of such simulations.

# 4 Kinetic Brownian motion on the diffeomorphism group

Pick $s>\frac{d}{2}$, or $s>\frac{d}{2}+1$, depending on whether we work on $\mathscr{M}$ or $\mathscr{M}_0$.

## 4.1 Kinetic Brownian motion in $\mathscr{M}$

Set $H := H^s(TM)$. Pick another exponent $a>\frac{1}{2}$, and let $\mathcal{H}$ stand for the $L^2$-orthogonal of $\ker(\Delta)$ in $H^{s+a}(TM)$, with norm

$$\|f\|^2_{s+a} = \sum_{n\geq 1} |\lambda_n|^{s+a}\|f_n\|^2_{L^2},$$

inherited from the eigenspace decomposition (3.13) of $L^2(TM)$. Let $\iota$ stand for the continuous inclusion of $\mathcal{H}$ into $H$, and use freely the canonical identification of $H$ and $H^*$. The continuous symmetric operator $\iota\iota^* : H \to H$, is trace-class, as a consequence of Weyl's law on a closed manifold, so it is the covariance of an $H$-valued Brownian motion $W$. Note the correspondance $\overline{C} = \iota\iota^*$, and

$$\alpha_n^2 = |\lambda_n|^{-a},$$

with the notations of Section 2.1. *We assume that the trace condition* (2.3) *of section 2.2*

$$3\alpha_1^2 < \mathrm{tr}(\overline{C}), \tag{4.21}$$

*holds true*. Note that the faster $\lambda_i$ goes to $\infty$, the lesser there is noise in $W$. The extreme case corresponds to only finitely many non-null $\alpha_i$. On the other extreme, the bigger the multiplicity of $\alpha_1^2$ is, the more noise there is in $W$. The trace condition (2.3) holds automatically as soon as $\alpha_1^2$ has multiplicity three.

The Brownian motion $v_t^\sigma$ on the sphere $S$ of $H$, associated with the injection $\mathcal{H} \hookrightarrow H$, is defined as the solution to the stochastic differential equation

$$dv_t^\sigma = \sigma\, P_{v_t^\sigma}(\circ dW_t),$$

where $P_a : H \to H$, is the orthogonal projection on $\langle a\rangle^\perp$, for any $a \neq 0$, and the position process $x_t^\sigma$ of kinetic Brownian motion $\big(x_t^\sigma, v_t^\sigma\big)$ in $H$, given as its integral

$$x_t^\sigma = x_0 + \int_0^t v_s^\sigma\, ds.$$

Kinetic Brownian motion on $\mathscr{M}$ is then defined as Cartan development in $\mathscr{M}$ of the time rescaled kinetic Brownian motion $\big(x_{\sigma^2 t}^\sigma\big)$ in $H$.

**Definition 4.1.** Kinetic Brownian motion on $\mathscr{M}$ is the projection $\varphi_t^\sigma$ on the configuration space $\mathscr{M}$ of the solution $\big(\varphi_t^\sigma, e_t^\sigma\big)$ to the equation in $H^s(F^{(e)})$

$$\frac{d}{dt}\,(\varphi_t^\sigma, e_t^\sigma) = \overline{\mathfrak{H}}^e\Big(\varphi_t^\sigma, e_t^\sigma; \sigma^2 v_{\sigma^2 t}^\sigma\Big), \tag{4.22}$$

with initial condition $\varphi_0 = \mathrm{id}$ and $e_0 = \mathrm{id} \in \mathsf{L}\big(H^s(TM)\big)$. $\triangle$

This equation is only locally well-posed. We introduce the following definition to deal with weak convergence questions for possibly exploding solutions of random or stochastic differential equations. Add a cemetary point $\dagger$ to $H^s(F^{(e)})$, and endow the disjoint union $H^s(F^{(e)}) \sqcup \{\dagger\}$ with its natural topology. Denote by $\Omega_0$ the set of continuous paths $z : [0,1] \to H^s(F^{(e)}) \sqcup \{\dagger\}$, that start from a reference point $z_0 := (\mathrm{id}, e_0)$ above the identity map on $M$, and that stay at the cemetery point $\dagger$, if it leaves $H^s(F^{(e)})$. Let $\mathcal{F} := \bigvee_{t\in[0,1]} \mathcal{F}_t$ where $(\mathcal{F}_t)_{0\le t\le 1}$ stands for the filtration generated by the canonical coordinate process on pathspace. Let $B_R$ stand for the $H^s$ balls with center $z_0$ and radius $R$, for any $R > 0$. The first exit time from $B_R$ is denoted by $\tau_R$, and used to define a measurable map

$$T_R : \Omega_0 \to C\big([0,1], \overline{B}_R\big),$$

which associates to any path $(z_t)_{0\le t\le 1} \in \Omega_0$ the path which coincides with $z$ on the time interval $\big[0, \tau_R\big]$, and which is constant, equal to $z_{\tau_R}$, on the time interval $\big[\tau_R, 1\big]$. The following definition then provides a convenient setting for dealing with sequences of random process whose limit may explode.

**Definition 4.2.** A sequence $(\mathbb{Q}_n)_{n\geq 0}$ of probability measures on $\big(\Omega_0, \mathcal{F}\big)$ is said to **converge locally weakly** to some limit probability $\mathbb{Q}$ if the sequence $\mathbb{Q}_n \circ T_R^{-1}$ of probability measures on $C([0,1], \overline{B}_R)$ converges weakly to $\mathbb{Q} \circ T_R^{-1}$, for every $R > 0$. $\triangle$

We proved in Theorem 2.14 that the canonical rough path lift $\mathbf{X}^\sigma$ of $\big(x^\sigma_{\sigma^2 t}\big)_{0\leq t\leq 1}$, converges weakly in the space of weak geometric $p$-rough paths in $H$, to the Stratonovich Brownian rough path $\mathbf{B} = (B, \mathbb{B})$, with covariance operator

$$C_B(\ell, \ell') = \int_0^\infty \mathbb{E}\Big[\ell(v_0)\ell'(v_t) + \ell'(v_0)\ell(v_t)\Big]\,dt, \quad \ell, \ell' \in H^*.$$

Since one can rewrite Equation (4.22) as a rough differential equation driven by the rough path $\mathbf{X}^\sigma$

$$\frac{d}{dt}\,(\varphi^\sigma_t, e^\sigma_t) = \overline{\mathfrak{H}}^e\Big(\varphi^\sigma_t, e^\sigma_t; d\mathbf{X}^\sigma_t\Big),$$

the continuity of the Itô-Lyons solution map gives the following theorem. Recall that the solution of a rough differential equation driven by the Stratonovich Brownian rough path coincides almost surely with the solution of the corresponding Stratonovich differential equation.

**Theorem 4.3.** *The $\mathscr{M}$-valued part $(\varphi^\sigma_t)$ of kinetic Brownian motion is converging locally weakly to the projection on $\mathscr{M}$ of the $H^s(F^{(e)})$-valued Brownian motion $(\varphi_t, e_t)$ solution to the stochastic differential equation*

$$\frac{d}{dt}\,(\varphi_t, e_t) = \overline{\mathfrak{H}}^e\Big((\varphi_t, e_t); \circ dB_t\Big).$$

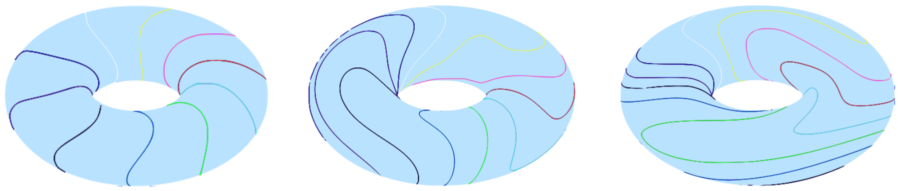
The motion of $\varphi_t$ itself is not given as the solution of a stochastic differential equation. This happens already in finite dimension, when defining anisotropic Brownian motion on a $d$-dimensional Riemannian manifold $M$ as Cartan development of an anisotropic Brownian motion in $\mathbb{R}^d$. One needs the moving orthonormal frame attached to the running point on $M$, to define the position increment in $M$ from the increment of the driving anisotropic Brownian motion in $\mathbb{R}^d$. The motion in $M$ is in particular non-Markovian, while the motion in $OM$ is Markovian. The same phenomenon happens in the present infinite dimensional setting, and we do not get here classical semimartingale flows in $H^s(M,M)$ [Kun90], or Brownian flows in critical spaces, such as in Malliavin's work on the canonical Brownian motion on the diffeomorphism group of the circle [Mal99, Fan02, AR02].

We remark here that the stochastic homogenization methods that X.-M. Li used in [Li16a] to prove the homogenization result for kinetic Brownian motion in a finite dimensional, complete, Riemannian manifold, require a positive injectivity radius and a uniform control on the gradient of the distance function over the whole manifold. It is unclear that anything like that is available in the present infinite dimensional setting, or in the setting of volume-preserving diffeomorphisms investigated in the next section, especially given the fact that $\mathscr{M}$ or $\mathscr{M}_0$ have infinite negative curvature in some directions. The robust pathwise approach of rough paths allows to circumvent these potential issues.

## 4.2 Kinetic Brownian motion in $\mathscr{M}_0$

Let $H_0$ stand for the closed subspace of $H$ of divergence-free vector fields on the fluid domain $M$. It is the tangent space at the identity map of the closed submanifold $\mathscr{M}_0$ of $\mathscr{M}$ of diffeomorphisms that leave invariant the Riemannian volume form of $M$. The intersection $\mathcal{H}^{s+a}_0$ of $\mathcal{H}^{s+a}$ with $H_0$, is continuously embedded into $H_0$. If $\iota_0$ stands for this injection, the continuous symmetric operator $\iota_0\iota_0^* : H_0 \to H_0$, is trace-class, so it is the covariance of an $H_0$-valued Brownian motion $W$. The spectrum of $\overline{C}_0 := \iota_0\iota_0^*$ is explicit in the example of the 2-dimensional torus, with maximal eigenvalue 1, with multiplicity 4. The trace condition (2.3) thus holds true for any $a > \frac{1}{2}$, in that case. Similarly, the spectrum of the Laplacian operator on vector fields on the 2-dimensional sphere is obtained from the spectrum of the Laplacian operator on real-valued functions on the 2-sphere, as a consequence of its canonical symplectic structure [AS89, Yos97]. Eigenvectors are constant multiples of the complex spherical harmonics, so eigenvalues have multiplicity at least two. Here as well, symmetry properties of the 2-dimensional sphere imply that they have actually multiplicity four, so the trace condition (2.3) holds for free. More generally, divergence-free vector fields on a simply connected $d$-dimensional manifold $M$ are gradients of functions, so one gets the spectrum of the covariance operator $C$ from the spectrum of the Laplacian operator on real-valued functions on $M$. One needs to assume the trace condition (2.3) in this generality.

Kinetic Brownian motion $(x_t^\sigma, v_t^\sigma)$ in $H_0$ is naturally defined as above from the associated Brownian motion $(v_t^\sigma)$ on the sphere $S_0$ of $H_0$, and its integral. With the help of the Cartan development machinery, one can then mimic Definition 4.1 above to extend the construction to the whole space $\mathscr{M}_0$ of volume preserving diffeomorphisms. The next figure, to be compared with its purely geodesic analogue Figure 3.4, illustrates for example the time evolution of great circles on the two dimensional flat torus $\mathbb{T}^2$ under the associated volume preserving kinetic Brownian flow with $\sigma = 1$.



*The flow acts on a family of great circles in the case $M = \mathbb{T}^2$.*
*From left to right, snapshots at time $t = 1$, $t = 2$ and $t = 5$, with $\sigma = 1$.*

FIGURE 4.6 – Time evolution of the volume preserving kinetic Brownian flow on $\mathscr{M}_0$.

To illustrate the fact that elements of area are indeed preserved under this random flow, the next figure represents the time evolution of such an element of area at different times. As indicated at the end of Section 3.4, the different simulations of this section were obtained using a finite dimensional approximation of the dynamics and a discretized version of Equation (3.20) describing the latter from the Lie algebra point of view.

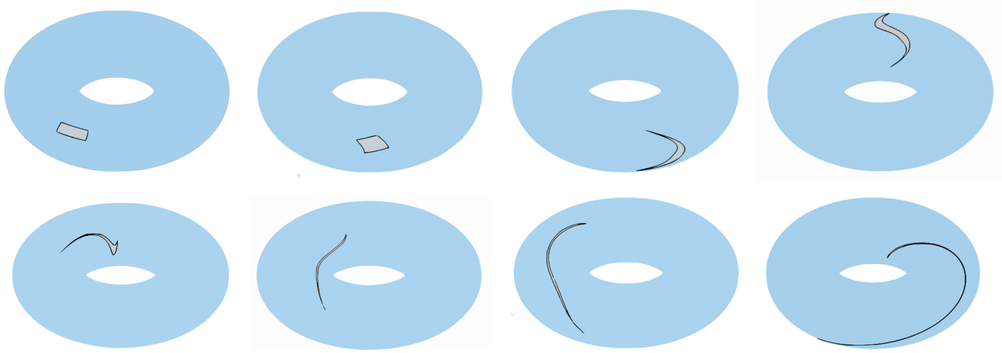
We prove in Theorem 5.3 of Appendix 5 that the Cartan development $\varphi_t^\sigma$ in $\mathscr{M}_0$, of the time rescaled kinetic Brownian motion in $H_0$ is the $\mathscr{M}_0$-part of the solution $(\varphi_t^\sigma, e_t^\sigma, f_t^\sigma)$, to a controlled ordinary differential equation on

$$\mathscr{X} = H^s(F^{(e)}) \times \mathsf{L}\big(H^s(TM)\big)$$

driven by a *smooth* vector field

$$\begin{aligned} \frac{d}{dt}\,(\varphi_t^\sigma, e_t^\sigma) &= \overline{\mathfrak{H}}^e\Big(\varphi_t^\sigma, e_t^\sigma; f_t^\sigma\,(\sigma^2 v_t^\sigma)\Big), \\ \frac{d}{dt}\,f_t^\sigma &= \overline{\mathfrak{H}}^f\left(\frac{d}{dt}\,(\varphi_t^\sigma, e_t^\sigma), f_t^\sigma\right). \end{aligned}$$

Here again, one can rewrite that equation as a rough differential equation driven by the canonical rough path $\mathbf{X}^\sigma$ above the time rescalled position process of kinetic Brownian motion

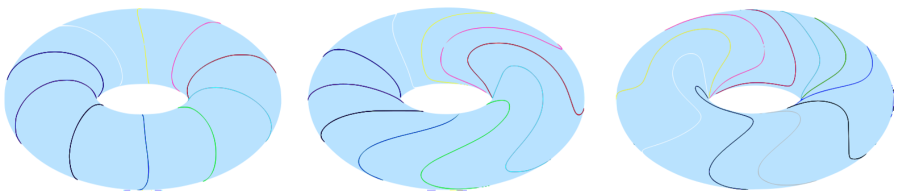


*Snapshots of the action of the flow on an area element, initially a small square. Here $M = \mathbb{T}^2$ and $\sigma = 1$.*

FIGURE 4.7 – Area element under the volume preserving kinetic Brownian flow on $\mathscr{M}_0$.

in $H_0$. Associated with the rough invariance principle established in Theorem 2.14 above, the continuity of the Itô-Lyons solution map then gives the following theorem.

**Theorem 4.4.** *The $\mathscr{M}_0$-valued part $(\varphi_t^\sigma)$ of kinetic Brownian motion in $\mathscr{Z}$ is converging locally weakly to the projection $(\varphi_t)$ on $\mathscr{M}_0$ of a $\mathscr{Z}$-valued Brownian motion.*



*The flow acts on a family of great circles in the case $M = \mathbb{T}^2$. From left to right, snapshots for $\sigma = 0.1$, $\sigma = 1$ and $\sigma = 4$, at time $t = 1$.*

FIGURE 4.8 – Influence of the noise parameter $\sigma$ on the volume preserving kinetic Brownian flow on $\mathscr{M}_0$.

Here again, the dynamics of $\varphi_t^\sigma$ is non-Markovian. Note that since kinetic Brownian motion on $\mathscr{M}_0$ is defined by Cartan development, using the $L^2$ metric (3.14), the $L^2$-size of $\dot\varphi_t^\sigma$ is equal to the $L^2$-norm of $v_t^\sigma$. The metric being right invariant on the group $\mathscr{M}_0$, the Eulerian velocity

$$u_t^\sigma := \dot\varphi_t^\sigma \circ (\varphi_t^\sigma)^{-1},$$

also has the same $L^2$-norm as $v_t^\sigma$. The latter is not preserved a priori; neither is the $H^s$-norm of $u_t^\sigma$, as mentioned above after Proposition 3.6.

Denote by $Q^0$ the quadratic form on $H^s(TM)$, with matrix

$$\operatorname{diag}\left(|\lambda_n|^{-s}\right)_{n\geq 0},$$

in the orthonormal basis of $H^s(TM)$ associated with the eigenvector decomposition (3.13) for $-\Delta$ on $L^2(TM)$. For each $v$ in the unit sphere $S$ of $H^s(TM)$, one has $Q^0(v) = \|v\|_{L^2}^2$, and

$$\|v\|_{L^2}^2 \leq \lambda_0^{-s}\|v\|_{H^s}.$$

Since the $S$-valued diffusion $(v_t^\sigma)$ is ergodic, each component $(v_t^\sigma)_n$ of $v_t^\sigma$, in the decomposition (3.13), is an ergodic process in the interval $\big(-\lambda_n^{-s/2}, \lambda_n^{-s/2}\big)$. The squared $L^2$-norm of $v_t^\sigma$ is also an ergodic process in the interval $(0, \lambda_0^{-s})$. It has invariant measure the image of a constant multiple of the measure with density $1/\|u\|$ with respect to the Gaussian measure in $H$ with covariance $\iota_0\iota_0^*$, by the map

$$u \in H \mapsto Q^0\big(u/\|u\|\big),$$

from Proposition 2.1. This is the invariant measure of the squared $L^2$-norm of the Eulerian velocity process $u_t^\sigma$. We emphasize that this invariant measure is independent of the interpolation parameter $\sigma \in (0,\infty)$. We record part of these facts in the following statement.

**Corollary 4.5.** *Fix $\sigma \in (0,\infty)$. The $L^2$-norm of the velocity field $u^\sigma$ of kinetic Brownian motion is an ergodic process taking values in the interval $(0, \lambda_0^{-s})$, with invariant probability measure the image of a constant multiple of the measure with density $1/\|u\|$ with respect to the Gaussian measure in $H$ with covariance $\iota_0\iota_0^*$, by the map*

$$u \in H \mapsto Q^0\big(u/\|u\|\big).$$

It is desirable to study the homogenization problem for other intrinsically randomly perturbated partial differential equations of geometric nature, such as the KdV, (modified) Camassa-Holm equations, or equations with non-local inertia operator, such as the modified Constantin-Lax-Majda equation [Kol17]. The core technical problem, from the geometric/analytic point of view, is the definition of Cartan development map as the solution map of an ordinary differential equation driven by sufficiently regular vector fields on the configuration space. We took advantage, in the present $L^2$ setting, of the 'pointwise' character of the associated geometric objects to recast things in terms of the $F$-bundles of Section 3.1. One may have to proceed differently for other weak metrics. We expect the homogenization results proved in Theorem 4.3 and Theorem 4.4 to have analogues in the setting of the strong, complete, Riemannian metrics of [BV20]. Global in time existence results for kinetic Brownian motion and its limit Brownian motion are expected. We leave these questions for a forthcoming work.

We worked here in the Sobolev setting to make things easier and concentrate on the probabilistic problems, and the implementation of the rough path approach in this infinite dimensional setting. It is a natural question to ask whether one can run the analysis in the Fréchet setting of smooth diffeomorphisms of $M$, asking for preservation of the regularity of the initial condition and velocity, as in Ebin-Marsden seminal work – Section 12 in [EM69], under proper assumptions on the noise.

# 5 Cartan development in $\mathscr{M}_0$

We prove in this last part that Cartan's development system (3.19) on $\mathscr{M}_0$ can be recast as an ordinary differential equation in $H^s\big(F^{(e)}\big) \times \mathsf{L}\big(H^s(TM)\big)$, driven by a smooth vector field. It has, as a consequence, a unique solution, up to a possibly finite explosion time.

Let $\overline{P} : T\mathscr{M} \to T\mathscr{M}$, stand for a smooth vector bundle morphism that coincides with the Hodge projector $P$ from (3.15) on $T\mathscr{M}_0$. The existence of such a map follows from the following elementary partition of unity result.

**Proposition 5.1.** *Let $(\mathcal{O}_i)_{i\in I}$ be an open cover of $\mathscr{M}$. Then there exists a smooth partition of unity subordinated to $(\mathcal{O}_i)_{i\in I}$.*

Set

$$\overline{\mathfrak{H}}^f : T\mathrm{H}^s\big(F^{(e)}\big) \times \mathsf{L}\big(H^s(TM)\big) \to T\mathsf{L}\big(H^s(TM)\big)$$
$$\left(\frac{d}{dt}\Big|_{t=0}\big(\varphi_t(\cdot), e_t(\cdot)\big),\, f\right) \mapsto \frac{d}{dt}\Big|_{t=0}\Big(\mathbf{X} \mapsto e_t^{-1}\Big(\overline{P}\big(e_t(f(\mathbf{X}))\big)\Big)\Big).$$

The letter $\mathbf{X}$ stands for a generic element of $H^s(TM)$, and

$$T\mathsf{L}\big(H^s(TM)\big) = \mathsf{L}\big(H^s(TM)\big).$$

We give the details of the following elementary result.

**Lemma 5.2.** *The map $\overline{\mathfrak{H}}^f$ is well-defined and smooth.*

*Proof.* It is enough to prove that the map

$$H^s(F^{(e)}) \times \mathsf{L}\big(H^s(TM)\big) \to \mathsf{L}\big(H^s(TM)\big)$$
$$\Big(\big(\varphi(\cdot), e(\cdot)\big), f\Big) \mapsto \Big(\mathbf{X} \mapsto e^{-1}\Big(\overline{P}\big(e(f(\mathbf{X}))\big)\Big)\Big)$$

is smooth. Since the map

$$H^s(F^{(e)}) \times \mathsf{L}\big(H^s(TM)\big) \times H^s(TM) \to H^s(TM)$$
$$\Big(\big(\varphi(\cdot), e(\cdot)\big), f, \mathbf{X}\Big) \mapsto e^{-1}\Big(\overline{P}\big(e(f(\mathbf{X}))\big)\Big)$$

is smooth, the problem reduces to the following question. Let a Banach manifold $A$ and a Hilbert space $H$, be given together with a smooth map $F : A \times H \to H$, that is linear with respect to its second argument. Denote by $a$ and $b$ generic elements of $A$. Prove that the curryfication $\mathrm{Cur}\, F : a \in A \mapsto F(a, \cdot) \in \mathsf{L}(H)$ is well-defined and smooth.

Write $d$ for the differential operator. We show that $d(\mathrm{Cur}\, F) = \mathrm{Cur}\,(\partial_a F)$. This will be enough, since we can then bootstrap the construction to show that $d^n(\mathrm{Cur}\, F) = \mathrm{Cur}\,(\partial_a^n F)$, is differentiable for any $n$. Because the result is local, we can assume without loss of generality that $A$ an open set of a Banach space. Fix $a \in M$, and let $\mathcal{U} \times B(0, \varepsilon)$ be a convex neighbourhood of $(a, 0)$ in $A \times H$, such that $\|\partial_a^2 F\|_\infty < 1 + \|\partial_a^2 F(a, 0)\|$. Then for all $b \in \mathcal{U}$ and $|w| < 1$, one has

$$\Big|F(b, w) - F(a, w) - \partial_a F(a, w)(b - a)\Big| \le \frac{|b - a|^2}{2}\|\partial_a^2 F\|_\infty\, |w|/\epsilon.$$

The conclusion follows from the fact that we have in particular the estimate

$$\Big\|\mathrm{Cur} F(b) - \mathrm{Cur} F(a) - \mathrm{Cur}(\partial_a F)(a; b - a)\Big\| \le c\, |b - a|^2,$$

for a positive constant $c$ independent of $b$. □

Choose now a $\mathcal{C}^1$ path $(\mathbf{X}_t)$ with values in $T_{\mathrm{id}}\mathscr{M}_0$, and zero initial condition. Let $\big((\varphi_t, e_t), f_t\big)$ be the solution in $H^s(F^{(e)}) \times \mathsf{L}\big(H^s(TM)\big)$ of the equation

$$\begin{aligned} \frac{d}{dt}\,(\varphi_t, e_t) &= \overline{\mathfrak{H}}^e\Big(\varphi_t, e_t; e_t\big(f_t(\dot{\mathbf{X}}_t)\big)\Big), \\ \frac{d}{dt}\, f_t &= \overline{\mathfrak{H}}^f\left(\frac{d}{dt}\,(\varphi_t, e_t), f_t\right), \end{aligned} \tag{5.23}$$

with initial condition $e_0 = \mathrm{id}_{TM}$, and $f_0 = \mathrm{id}_{H^s(TM)}$. Since the vector field $(\overline{\mathfrak{H}}^e, \overline{\mathfrak{H}}^f)$ is smooth, equation (5.23) is locally well-posed, possibly up to a finite explosion time $\zeta$.

**Theorem 5.3.** *The path* $(\varphi_t)$ *takes values in* $\mathscr{M}_0$*, and coincides with the Cartan development of* $(\mathbf{X}_t)$*. We further have* $\dot{\varphi}_t = e_t\big(f_t(\dot{\mathbf{X}}_t)\big)$*, so the dynamics* (5.23) *does not depend on the extension* $\overline{P}$ *of the Hodge projector* $P$ *used in the definition of* $\overline{\mathfrak{H}}^f$*.*

*Proof.* Let $\mathbf{Y} \in T_{\mathrm{id}}\mathscr{M}_0$, be a fixed divergence-free vector field on $M$. We need to show that

$$\overline{\nabla}^0_{\dot{\varphi}_t} e_t(\mathbf{Y}) = 0,$$

on the whole time interval $[0,\zeta)$. From Proposition 3.3, this is equivalent to showing that we have

$$\frac{d}{dt}\Big(\varphi_t, e_t\big(f_t(\mathbf{Y})\big)\Big) = dP\Big(\overline{\mathfrak{H}}^{(v)}\Big(\varphi_t, e_t\big(f_t(\mathbf{Y})\big); \dot{\varphi}_t\Big)\Big).$$

Look at the function $(\varphi, e, \mathbf{Z}) \mapsto \big(\varphi, e(\mathbf{Z})\big)$ from $H^s(F^{(e)}) \times T_{\mathrm{id}}\mathscr{M}$ to $H^s(F^{(v)})$, and set

$$\mathfrak{F} := \partial_{(\varphi,e)}\Big\{(\varphi, e, \mathbf{Z}) \mapsto \big(\varphi, e(\mathbf{Z})\big)\Big\}.$$

We have

$$\begin{aligned}
&\frac{d}{dt}\Big(\varphi_t, e_t\big(f_t(\mathbf{Y})\big)\Big)\\
&\quad= \mathfrak{F}\left(\frac{d}{dt}(\varphi_t, e_t), f_t(\mathbf{Y})\right) - \mathfrak{F}\left(\frac{d}{dt}(\varphi_t, e_t), e_t^{-1}\Big(\overline{P}\big(e_t(f_t(\mathbf{Y}))\big)\Big)\right)\\
&\qquad+ d\overline{P}\left(\mathfrak{F}\Big(\frac{d}{dt}(\varphi_t, e_t), f_t(\mathbf{Y})\Big)\right).
\end{aligned}$$

We prove that $e_t(\mathbf{Y})$ is divergence-free. Define for that purpose the subset $I \subset [0,\zeta)$ of times $t$ such that $e_t(\mathbf{Z})$ is divergence-free for all $\mathbf{Z} \in T_{\mathrm{id}}\mathscr{M}_0$, and $\varphi_t$ preserves the volume form. It is a non-empty closed subset of $[0,\zeta)$. Fix $t_0 \in I$. It suffices to prove that $t_0$ is in the interior of $I$ for a well-chosen extension $\widehat{P}$ of $P$, possibly different from $\overline{P}$. We choose for $\widehat{P}$ any smooth extension of $P$ defined on a neighbourhood of $\varphi_{t_0}$, such that $\widehat{P} \circ \widehat{P} = \widehat{P}$. Set $\widehat{Q} := \mathrm{id} - \widehat{P} : T\mathscr{M} \to T\mathscr{M}$, so for a fixed $\mathbf{Z} \in T_{\mathrm{id}}\mathscr{M}_0$, the quantity

$$Z_t := \widehat{Q}\big(e_t(f_t(\mathbf{Z}))\big)$$

satisfies the equation

$$\begin{aligned}
\frac{d}{dt} Z_t &= d\widehat{Q}\left(\mathfrak{F}\left(\frac{d}{dt}(\varphi_t, e_t), e_t^{-1}\Big(\widehat{Q}(e_t[f_t(\mathbf{Z})])\Big)\right)\right)\\
&= d\widehat{Q}\left(\mathfrak{F}\Big(\frac{d}{dt}(\varphi_t, e_t), e_t^{-1}(Z_t)\Big)\right).
\end{aligned}$$

This differential equation satisfies the classical Picard-Lindelöf assumptions, so it has a unique solution with given initial condition. Since $Z_0 = 0$ and the constant zero vector field is a solution to the equation, $Z_t$ is identically zero, and $e_t(\mathbf{Z})$ is divergence-free.

This holds true for any $\mathbf{Z}$, in a time interval independent of $\mathbf{Z}$. It follows in particular that $\dot{\varphi}_t = e_t\big(f_t(\dot{\mathbf{X}}_t)\big)$ is locally divergence-free, and $\varphi_t$ preserves the volume form, in a neighbourhood of the time $t_0$. The interval $I$ is thus both closed and open, so $I = [0,\zeta)$. The statement of

Theorem 5.3 follows, since $P\big(e_t(f_t(\mathbf{Y}))\big) = e_t\big(f_t(\mathbf{Y})\big)$, so we get

$$\begin{aligned}\frac{d}{dt}\Big(\varphi_t, e_t\big(f_t(\mathbf{Y})\big)\Big) &= d\overline{P}\left(\mathfrak{F}\Big(\frac{d}{dt}(\varphi_t, e_t), f_t(\mathbf{Y})\Big)\right)\\ &= d\overline{P}\left(\frac{d}{ds}\Big|_{s=t}\Big(\varphi_s, e_s\big(f_t(\mathbf{Y})\big)\Big)\right)\\ &= dP\Big(\overline{\mathfrak{H}}^{(v)}\Big(\varphi_t, e_t\big(f_t(\mathbf{Y})\big); \dot{\varphi}_t\Big)\Big),\end{aligned}$$

using Proposition 3.1 in the last equality. □

# Chapter IV

# Heat kernel estimates

In this chapter, we study the small time asymptotic of the density of the kinetic Brownian motion on the euclidean plane $\mathbb{R}^2$. In probabilistic terms, this diffusion is the solution in $\mathbb{R}^3$ of the stochastic differential equation

$$\begin{cases} \mathrm{d}x_t & = & \cos(\phi_t)\mathrm{d}t \\ \mathrm{d}y_t & = & \sin(\phi_t)\mathrm{d}t \\ \mathrm{d}\phi_t & = & \mathrm{d}W_t \end{cases}$$

for $W$ a standard Brownian motion in $\mathbb{R}$ and initial condition $X_0 = (x_0, y_0, \phi_0)$. It can be thought in $\mathbb{C} \times \mathbb{R}$ by setting $z_t = x_t + \mathrm{i}y_t$, and considering the equation

$$\begin{cases} \mathrm{d}z_t & = & \mathrm{e}^{\mathrm{i}\phi_t}\mathrm{d}t \\ \mathrm{d}\phi_t & = & \mathrm{d}W_t. \end{cases}$$

It is maybe clearer in this form that the isometries of $\mathbb{C}$ carry a solution with initial condition $(z_0, \phi_0)$ to the solution started at $(0, 0)$, so the initial condition $(x_0, y_0, \phi_0)$ is not overly relevant. From the analytic point of view, the density $u$ of the kinetic Brownian motion satisfies the PDE

$$\begin{cases} \partial_t u & = -\cos(\phi)\partial_x u - \sin(\phi)\partial_y u + \frac{1}{2}\partial_\phi^2 u \\ & =: Lu \\ u_0 & = \delta_{X_0}. \end{cases} \tag{0.1}$$

From February to June 2018, I visited Vassili Kolokoltsov at the university of Warwick to work on this problem. During this time, V. Kolokoltsov provided very valuable insight for all the results described below, and I learned a lot from the experience. I would like to thank him and the university warmly.

The French-speaking reader may find in section I.5 a discussion of the challenges of such a task, and a quick overview of the standard techniques we will use in the following. Namely, two powerful methods for similar problems are the parametrix approach and the WKB approximation. Although there are problems where they are seen to work hand in hand, it does not seem to be the case here ; the first part 1 in what follows is devoted to the former, and the second part 2 to the latter. Unfortunately, I was only able to give conclusive results using the first approach ; however, I believe that the second provides a window to the underlying geometry of the problem, and that the variational results, proved or to come, are interesting in their own right.

In section 1, we fix an approximation $\widetilde{u}$ of $u$ of order 2. Using the error $E := (-\partial_t + L)\widetilde{u}$ and a good notion of convolution, we show that

$$u = \widetilde{u} + E * \widetilde{u} + E * E * \widetilde{u} + \cdots$$

in a certain sense. This is done by showing that $E^{*n} * \widetilde{u} \in \Psi^{n+1}$, for some function spaces $\Psi^a$ adapted to the geometry of the problem, and $a$ is a regularity parameter. For more details on this approach, see the introduction 1.1.

In section 2, we study the following minimisation problem. For fixed $X_0, X \in \mathbb{R}^3$ and $t > 0$, find the curve $h : [0;t] \to \mathbb{R}^3$ such that the solution $X^h$ of

$$\mathrm{d}x_s^h = \cos(\phi_s^h)\mathrm{d}s, \qquad \mathrm{d}y_s^h = \sin(\phi_s^h)\mathrm{d}s, \qquad \mathrm{d}\phi_s^h = \mathrm{d}h_s$$

with initial condition $X_0^h = X_0$ has endpoint $X_t^h = X$ and the energy

$$\frac{1}{2}\int_0^t |\dot{h}_s|\mathrm{d}s$$

is minimal. The link between this problem and the study of $u$ is discussed in 2.1. This question can be restated as a Hamiltonian problem with Hamiltonian $H$, and it is singular in some sense. By studying the associated Hamilton equations, we prove that the minimisers, if regular enough, are in fact characteristics for $H$ ; this is our main Theorem 2.6. Some notations and preliminary results are needed to give an overview of the strategy; the reader will have to wait until the remark after Proposition 2.2.

**Notations.** We denote by $X$ a point $(x, y, \phi) \in \mathbb{R}^3$, and similarly for its variants : $\widehat{X}$, $X_0$, etc. Its dual variable in the phase space, for instance when we consider the Fourier transform, is $P = (p, q, \psi) \in \mathbb{R}^3$.

# 1 Quadratic approximation

## 1.1 General strategy

**Parametrix.** This section describes an approach based on the parametrix method. Formally, the method is very simple. We try to approximate the kernel $u$, seen as a function of time $t$, initial condition $X_0$, and evaluation point $X_1$, by some approximation $\widetilde{u}$, that we call a parametrix.[1] If the parametrix $\widetilde{u}(t, X_0, \cdot)$ tends to $\delta_{X_0}$ in some appropriate sense, the Duhamel principle asserts that

$$\widetilde{u}_t(X_0, X_1) = u_t(X_0, X_1) + \int_0^t \int_{\mathbb{R}^3} (\partial_t - L)\widetilde{u}_s(X_0, \Xi) \cdot u_{t-s}(\Xi, X_1) \mathrm{d}s\, \mathrm{d}\Xi.$$

Setting $E = (-\partial_t + L)\widetilde{u}$, and $\mathcal{E}$ the convolution operator characterised by

$$\mathcal{E}v : t, X_0, X_1 \mapsto \int_0^t \int_{\mathbb{R}^3} E(s, X_0, \Xi) \cdot v(t-s, \Xi, X_1) \mathrm{d}s\, \mathrm{d}\Xi$$

whenever the integral is well-defined, the Duhamel principle rewrites as

$$\widetilde{u} = (\mathrm{id} - \mathcal{E})u.$$

This relation is easily inverted, provided we stay strictly at a formal level. For now, we imposed very few constraints on $\widetilde{u}$. Such function will be a good parametrix provided we have

$$u = \sum_{k \geq 0} \mathcal{E}^k \widetilde{u}$$

in some sense, the stronger the better. See [BKH15] for a probabilistic take on this approach.

These are in essence the fundamentals of the parametrix method. It has become a standard tool for constructing the heat kernel on Riemannian manifolds, in particular with views on the small time study of the eigenvalues of the Laplacian; see for instance [Mel93, Ros97]. To the best of our knowledge, our work is unusual in the sense that the parametrix is not modelled on a Gaussian distribution. Although it means that the usual estimates we know for Gaussian functions are not available, and that the description of the parametrix is less explicit, we argue that any approach to this problem should face the same difficulties, since this hypoelliptic diffusion has a very different behaviour than, say, processes associated to regular Hamiltonians of [Kol00].

**Main result.** In the construction of the heat kernels on manifolds, the chosen parametrix $\widetilde{u}_t(X_0, \cdot)$ is the solution of the equation with coefficients frozen at $X_0$ (up to a suitable cutoff). In our case, this would lead to a strongly degenerate diffusion. In fact, even linearisation at $X_0$ would kill any transport in the direction $\mathsf{e}^{\mathsf{i}\phi_0}$ in $\mathbb{C}$, so we must at least consider the second order. Setting $\mathfrak{exp} : x \mapsto 1 + x + x^2/2$, we choose as a parametrix the solution $\widetilde{u}$ of the partial differential equation

$$\partial_t \widetilde{u} = -\Re\big(\mathsf{e}^{\mathsf{i}\phi_0} \mathfrak{exp}(\mathsf{i}\phi - \mathsf{i}\phi_0)\big)\partial_x \widetilde{u} - \Im\big(\mathsf{e}^{\mathsf{i}\phi_0} \mathfrak{exp}(\mathsf{i}\phi - \mathsf{i}\phi_0)\big)\partial_y \widetilde{u} + \frac{1}{2}\partial_\phi^2 \widetilde{u}$$

with initial condition $\widetilde{u}_0(X_0, \cdot) = \delta_{X_0}$. As mentioned above, this $\widetilde{u}$ is not modelled on a Gaussian, and in fact its study is not easy either. See for instance [Pao17] for an treatment of the study on the diagonal of nilpotent diffusions, closely related to this one. The objective of this work is the following result.

[1] Note that we use the term *parametrix* in a very broad sense. Some authors give a precise definition for these functions, see e.g. [Hör90, Definition 7.1.21.].

**Theorem 1.1.** *The series*

$$\sum_{k\geq 0} \mathcal{E}^k \widetilde{u}$$

*converges to $u$ over $[t_-;t_+]\times\mathbb{R}^3\times\mathbb{R}^3$ in the $\mathcal{C}^\ell$ norm, for any $\ell\geq 0$ and $0<t_-<t_+$.*

Combined with estimates on $\widetilde{u}$, and using the methods described below, we can extract interesting information about the kernel.

**Corollary 1.2.** *Let $D$ be a domain where $\widetilde{u}_1(0,\cdot+(1,0,0))$ is bounded away from zero. Then*

$$u_t(0,X+(t,0,0)) = \widetilde{u}_t(0,X+(t,0,0))(1+O(t))$$

*for all $(x/t^2, y/t\sqrt{t}, \phi/\sqrt{t})\in D$.*

The proof of the theorem goes in three steps. First, we introduce some convolution calculus, meant to control the iterated convolution arising from the terms $\mathcal{E}^k\widetilde{u}$. It is heavily inspired by the notes [Gri04] of D. Grieser, which restricts to the simpler case of second order elliptic equations. The backbone of this calculus consists in Theorem 1.6, and we discuss some more properties in Subsection 1.3. Then, we use it to show that the series converges in Theorem 1.11. At this point, it is sufficient to identify the limit to conclude: this is done in Theorem 1.14 and Proposition 1.15 by checking that it satisfies the PDE (0.1) and has the expected initial condition, respectively. Because the study of $\widetilde{u}$ uses entirely different techniques, we postpone it in our last Section 1.5; thanks to our framework, it is very easy to state the property of $\widetilde{u}$ that we need; we do so in Proposition 1.12.

**Caveat.** This work is still in progress, and in particular it is conditional on the positivity of two functions. This is discussed in the proof of Proposition 1.16, along with numerical evidence.

## 1.2 Convolution calculus

As discussed above, define $\widetilde{u}$ as the solution of

$$\begin{aligned}\partial_t\widetilde{u} &= -\Re\big(\mathsf{e}^{\mathsf{i}\phi_0}\mathfrak{exp}(\mathsf{i}\phi-\mathsf{i}\phi_0)\big)\partial_x\widetilde{u} - \Im\big(\mathsf{e}^{\mathsf{i}\phi_0}\mathfrak{exp}(\mathsf{i}\phi-\mathsf{i}\phi_0)\big)\partial_y\widetilde{u} + \frac{1}{2}\partial_\phi^2\widetilde{u} \\ &=: L_{\phi_0}\widetilde{u},\end{aligned} \tag{1.2}$$

with initial condition $\widetilde{u}_0(X_0,\cdot)=\delta_{X_0}$. It is sometimes easier to write it under the form

$$\partial_t\widetilde{u} = -\Big(1-\frac{(\phi-\phi_0)^2}{2}\Big)\big(\cos(\phi_0)\partial_x+\sin(\phi_0)\partial_y\big)\widetilde{u} - (\phi-\phi_0)\big(-\sin(\phi_0)\partial_x+\cos(\phi_0)\partial_y\big)\widetilde{u} + \frac{1}{2}\partial_\phi^2\widetilde{u}.$$

The uniqueness of the solution is not discussed here in the PDE setting. It is well-known that $\widetilde{u}$ is the density of $X_t$, where under some probability $\widetilde{\mathbb{P}}_{X_0}$, $X=(z,\phi)$ is the (unique strong) solution of the stochastic differential equation

$$\begin{cases} \mathrm{d}z_t &= \mathsf{e}^{\mathsf{i}\phi_0}\mathfrak{exp}(\mathsf{i}\phi_t-\mathsf{i}\phi_0)\mathrm{d}t \\ \mathrm{d}\phi_t &= \mathrm{d}W_t,\end{cases}$$

for some standard Brownian motion $W$.

Before defining the functions spaces $\Psi^a$ we will need, we give some more intuition about the structure of the functions $\mathcal{E}^k\widetilde{u}$, which will belong to $\Psi^{k+1}$. As is easily seen, for fixed $X_0$, the map

$$v: t,X\mapsto \widetilde{u}_t\big(X_0, X_0+(\mathsf{e}^{\mathsf{i}\phi_0}(z+t),\phi)\big)$$

is solution of the equation

$$\partial_t v = \frac{\phi^2}{2}\partial_x v - \phi\partial_y v + \frac{1}{2}\partial_\phi^2 v.$$

As with $u$ and $\widetilde{u}$, it is easy to express $v_t$ as the density at time $t$ of some stochastic process $X$ under some probability measure $\mathbb{Q}$. Indeed, working in the same coordinates, the solution of

$$\begin{cases} \mathrm{d}x_t &= -\frac{\phi_t^2}{2}\mathrm{d}t \\ \mathrm{d}y_t &= \phi_t \mathrm{d}t \\ \mathrm{d}\phi_t &= \mathrm{d}W_t \end{cases}$$

with initial condition $X = 0$ is such a process. Because of the scaling invariance $W_{at} \overset{\mathcal{L}}{=} \sqrt{a}W_t$, it is easily seen that for a fixed $a > 0$, the following equality holds in distribution, as processes indexed by $t$.

$$(x_{at}, y_{at}, \phi_{at}) \overset{\mathcal{L}}{=} (a^2 x_t, a\sqrt{a}y_t, \sqrt{a}\phi_t)$$

In terms of its density $v$, it means that $v_t$ is entirely described by $v_1$ through the relation

$$v_t\big(t^2 x, t\sqrt{t}y, \sqrt{t}\phi\big) = \frac{1}{t^4}v_1(x, y, \phi').$$

In some sense, $v$ can be seen as the product of a normalised function $\check{v} = v_1$ applied to normalised variables $\check{x} = x'/t^2$, $\check{y} = y'/t\sqrt{t}$ and $\check{\phi} = \phi'/\sqrt{t}$, with a time singularity $t^{-4}$. This can be compared to the Gaussian case, where $1/\sqrt{2\pi t}^d$ is the time singularity, whereas $\exp(-|x|^2/2t)$ is smooth with respect to $x/\sqrt{t}$. The point here is that the singularity is very simple, a power of $t$, and the remaining function very well-behaved; in the Gaussian case and that of $v$, it depends only on $\check{X}$ and is smooth.

These are roughly the conditions describing the spaces $\Psi^a$ defined in Definition 1.4, which are the basis of our convolution calculus. They consist of smooth functions $A$ that can be expressed as the product of a smooth rapidly decaying function $\check{A}$ applied to normalised variables $\check{X}$ with a time singularity $t^{-5+a}$. It can be shown that both $\widetilde{u}$ and $E$ are in $\Psi^1$. The hope is of course that $\mathcal{E}^k\widetilde{u}$ has increasing regularity as $k$ grows larger, so that its influence is asymptotically negligible for small $t > 0$. We will see this as a consequence of the central Theorem 1.6, which states that $A * B \in \Psi^{a+b}$ for all $A \in \Psi^a$, $B \in \Psi^b$ with $a, b > 0$; indeed, we have directly $\mathcal{E}^k\widetilde{u} \in \Psi^{k+1}$. This means that such functions vanish at high orders as $t$ tends to zero, and will allow us to state strong convergence results.

**Function spaces.** We begin this section by introducing the spaces $\Psi^a$ described above. They are inspired by the notes [Gri04], and similar considerations can be found in the earlier work [Mel93]. For any time $t \geq 0$ and initial position $X_0 \in \mathbb{R}^3 \simeq \mathbb{C}\times\mathbb{R}$, define the dilation operator $T_{t,X_0}$ as

$$T_{t,X_0} : (\check{z}, \check{\phi}) \mapsto (z_0 + \mathrm{e}^{\mathrm{i}\phi_0}(t + t^2\check{x}) + \mathrm{i}\mathrm{e}^{\mathrm{i}\phi_0}t\sqrt{t}\check{y}, \phi_0 + \sqrt{t}\check{\phi}) \in \mathbb{C}\times\mathbb{R},$$

and its endpoint version $S_{t,X_1}$ as

$$S_{t,X_1} : (\check{z}, \check{\phi}) \mapsto (z_1 - \mathrm{e}^{\mathrm{i}(\phi_1 - \sqrt{t}\check{\phi})}(t + t^2\check{x}) - \mathrm{i}\mathrm{e}^{\mathrm{i}(\phi_1 - \sqrt{t}\check{\phi})}t\sqrt{t}\check{y}, \phi_1 - \sqrt{t}\check{\phi}) \in \mathbb{C}\times\mathbb{R}.$$

They are the right normalising dilations for our problem; indeed, we will show in section 1.5 that there exists a smooth function $\check{u}$ such that

$$\widetilde{u}_t(X_0; T_{t,X_0}(\check{X})) = \frac{1}{t^4}\check{u}(\check{X}), \qquad \widetilde{u}_t(S_{t,X_1}(\check{X}), X_1) = \frac{1}{t^4}\check{u}(\check{X}).$$

In fact, the proof consists in writing down exactly what are the manipulations described at the end of section 1.1.

For any multi-index $\alpha = (a, b, c) \in \mathbb{N}^3$ and smooth function $f : \mathbb{R}^3 \to \mathbb{R}$ (or $\mathbb{R}^3 \times E \to F$), set $D_X^\alpha f = \partial_x^a \partial_y^b \partial_\phi^c f$.

**Definition 1.3.** We denote by $\check{\Psi}$ the set of smooth functions $\check{A} : [0; \infty) \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ satisfying the following boundedness conditions. Denoting by $\tau$, $X_0$ and $\check{X}$ the variables of $\check{A}$,

$$\sup_{0 \le \tau \le T} \sup_{X_0, \check{X}} |\check{X}|^k \cdot \left| \partial_\tau^n D_{X_0}^\alpha D_{\check{X}}^\beta \check{A}(\tau, X_0, \check{X}) \right| < \infty \tag{1.3}$$

for any time horizon $T > 0$, any multi-indices $\alpha, \beta \in \mathbb{N}^3$ and any indices $n, k \in \mathbb{N}$. △

Note that they are smooth up to the boundary, so we can see them as restrictions of smooth functions defined on $\mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^3$.

The transformations $T_{t,X_0}$ and $S_{t,X_1}$ are well-defined and smooth as functions of $\sqrt{t}$ up to $t \ge 0$. If $t > 0$, they are in fact invertible, with

$$\begin{aligned} T_{t,X_0}^{-1}(z_1, \phi_1) &= S_{t,X_1}^{-1}(z_0, \phi_0) \\ &= \left( \frac{1}{t^2} \Re\big((z_1 - z_0)\mathrm{e}^{-\mathrm{i}\phi_0} - t\big) + \frac{\mathrm{i}}{t\sqrt{t}} \Im\big((z_1 - z_0)\mathrm{e}^{-\mathrm{i}\phi_0}\big), \frac{1}{\sqrt{t}}(\phi_1 - \phi_0) \right). \end{aligned}$$

Both these functions, with respectively $(t, X_0)$ and $(t, X_1)$ fixed, have Jacobian $t^{-4}$.

**Definition 1.4.** For a fixed parameter $a \in \mathbb{R}$, a function $A : (0; \infty) \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ can be written as

$$A(t, X_0, X_1) = t^{-5+a} \check{A}^0\big(\sqrt{t}, X_0, T_{t,X_0}^{-1}(X_1)\big), \qquad \check{A}^0 \in \check{\Psi} \tag{1.4 a}$$

if and only if it can be written as

$$A(t, X_0, X_1) = t^{-5+a} \check{A}^1\big(\sqrt{t}, X_1, S_{t,X_1}^{-1}(X_0)\big), \qquad \check{A}^1 \in \check{\Psi}. \tag{1.4 b}$$

We write $\Psi^a$ for the set of such functions, and $\Psi^{>a}$ for the union of the spaces $\Psi^b$ for $b > a$. △

In other words, $A$ is in $\Psi^a$ if

$$(\tau, X_0, \check{X}) \mapsto (\tau^2)^{5-a} A(\tau^2, X_0, T_{\tau^2, X_0}(\check{X}))$$

extends to a smooth function on $\mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^3$ rapidly decaying with respect to $\check{X}$ — alternatively, if

$$(\tau, X_1, \check{X}) \mapsto (\tau^2)^{5-a} A(\tau^2, S_{\tau^2, X_1}(\check{X}), X_1)$$

extends to such a function.

*Proof.* Write $A$ as

$$A(t, X_0, X_1) = t^{-5+a} \check{A}^0\big(\sqrt{t}, X_0, T_{t,X_0}^{-1}(X_1)\big) = t^{-5+a} \check{A}^1\big(\sqrt{t}, X_1, S_{t,X_1}^{-1}(X_0)\big).$$

We must show that $\check{A}^0$ extends to a function in $\check{\Psi}$ if and only if $\check{A}^1$ does. But

$$\check{A}^0(\tau, X_0, \check{X}) = \check{A}^1(\tau, T_{\tau^2, X_0}(\check{X}), \check{X}), \qquad \check{A}^1(\tau, X_1, \check{X}) = \check{A}^0(\tau, S_{\tau^2, X_1}(\check{X}), \check{X}).$$

Since $T_{\tau^2, X_0}(\check{X})$ and $S_{\tau^2, X_0}(\check{X})$ are smooth up to the boundary, and all their derivatives with respect to any of their variables are bounded above by polynomials in $\tau$ and $\check{X}$, the result follows easily. □

It is elementary from the definition that functions in $\Psi^a$ are smooth on $(0;\infty)\times\mathbb{R}^3\times\mathbb{R}^3$. In fact, one can easily bound the derivatives of $T^{-1}_{t,X_0}(X_1)$ to get the following closedness result by induction.

**Proposition 1.5.** *Let $A$ be in $\Psi^a$. For any $\alpha_0,\alpha_1\in\mathbb{N}^3$, $n\in\mathbb{N}$,*

$$\partial_t^n D^{\alpha_0}_{X_0} D^{\alpha_1}_{X_1} A\in\Psi^{a-K},$$

*for $\alpha_i=(a_i,b_i,c_i)$, $K=n+2(a_0+b_0+c_0)+2(a_1+b_1+c_1)$.*

**Convolutions.** These spaces are introduced because they facilitate a simple control of convolutions. The following result is central in our study.

**Theorem 1.6.** *For any $A\in\Psi^a$ and $B\in\Psi^b$, if $a>0$ and $b>0$ the convolution*

$$A*B:(t,X_0,X_1)\mapsto\int_0^t\int_{\mathbb{R}^3}A(s,X_0,\Xi)B(t-s,\Xi,X_1)\mathrm{d}s\,\mathrm{d}\Xi$$

*is well defined and belongs to $\Psi^{a+b}$. In symbols, $\Psi^a*\Psi^b\subset\Psi^{a+b}$.*

*Proof.* We begin by performing the change of variables $s=t\sigma$, and make functions of $\check\Phi$ appear. Note that the integrals are not defined yet, so a formal proof would go from the bottom up.

There exists $\check A,\check B\in\check\Psi$ such that $A*B(t,X_0,X_1)$ equals

$$t^{-10+a+b}\int_0^1\int_{\mathbb{R}^3}\sigma^{-5+a}(1-\sigma)^{-5+b}\check A\big(\sqrt{\sigma t},X_0,T^{-1}_{\sigma t,X_0}(\Xi)\big)\check B\big(\sqrt{(1-\sigma)t},X_1,S^{-1}_{(1-\sigma)t,X_1}(\Xi)\big)t\mathrm{d}\sigma\,\mathrm{d}\Xi.$$

Because there are two singularities in the above integral, for $\sigma=0$ and $\sigma=1$ respectively, we separate the integral at $\sigma=1/2$, and get

$$t^{5-a-b}(A*B)(t,X_0,X_1)=\frac{1}{t^4}\big(I^0(t,X_0,X_1)+I^1(t,X_0,X_1)\big),$$

where $I^0$ (resp. $I^1$) is the integral of $\sigma^{-5+a}(1-\sigma)^{-5+b}F_{\sigma,\Xi}(t,X_0,X_1)$ over $(0;1/2]\times\mathbb{R}^3$ (resp. $[1/2;1)\times\mathbb{R}^3$), with

$$F_{\sigma,\Xi}(t,X_0,X_1):=\check A\big(\sqrt{\sigma t},X_0,T^{-1}_{\sigma t,X_0}(\Xi)\big)\check B\big(\sqrt{(1-\sigma)t},X_1,S^{-1}_{(1-\sigma)t,X_1}(\Xi)\big).$$

We are to prove that $t^{-4}I^0(t,X_0,T_{t,X_0}(\check X))$ and $t^{-4}I^1(t,S_{t,X_1}(\check X),X_1)$ are well-defined for $t>0$ and extend to smooth functions of $(\sqrt t,X_0,\check X)$ and $(\sqrt t,X_1,\check X)$ up to the boundary. Let us consider $I^0$ first. The change of variables $\Xi=T_{\sigma t,X_0}(\check\Xi)$ yields

$$\begin{aligned}I^0\big(t,X_0,T_{t,X_0}(\check X)\big)&=t^4\int_0^{\frac12}\int_{\mathbb{R}^3}\sigma^{-1+a}(1-\sigma)^{-5+b}F_{\sigma,T_{\sigma t,X_0}(\check\Xi)}\big(t,X_0,T_{t,X_0}(\check X)\big)\mathrm{d}\sigma\,\mathrm{d}\check\Xi\\&=:t^4\int_0^{\frac12}\int_{\mathbb{R}^3}\sigma^{-1+a}(1-\sigma)^{-5+b}\check F^0_{\sigma,\check\Xi}(\sqrt t,X_0,\check X)\mathrm{d}\sigma\,\mathrm{d}\check\Xi.\end{aligned}$$

Similarly, setting $\Xi=S_{(1-\sigma)t,X_1}(\check X)$,

$$\begin{aligned}I^1\big(t,S_{t,X_1}(\check X),X_1\big)&=t^4\int_{\frac12}^1\int_{\mathbb{R}^3}\sigma^{-5+a}(1-\sigma)^{-1+b}F_{\sigma,S_{(1-\sigma)t,X_1}(\check\Xi)}\big(t,S_{t,X_1}(\check X),X_1\big)\mathrm{d}\sigma\,\mathrm{d}\check\Xi\\&=:t^4\int_{\frac12}^1\int_{\mathbb{R}^3}\sigma^{-5+a}(1-\sigma)^{-1+b}\check F^1_{\sigma,\check\Xi}(\sqrt t,X_1,\check X)\mathrm{d}\sigma\,\mathrm{d}\check\Xi.\end{aligned}$$

We see that these integrals will behave well if their integrands do. More precisely, the singularities in $\sigma$ and $(1-\sigma)$ are tame, since $\sigma^{-1+a}$ is integrable in a neighbourhood of zero, and similarly for $(1-\sigma)$. It would then be sufficient to see that the derivatives of $F^i$ are well-defined, smooth up to the boundary, and bounded by integrable functions of $\sigma$ and $\check{\Xi}$, locally uniformly in $\sqrt{t}$, $X_0$, $X_1$ and $\check{X}$. The fact that they do is at the heart of our approach, and will be used throughout this work.

Setting

$$h^0_{\sigma,\check{\Xi}}(\sqrt{t}, X_0, \check{X}) = S^{-1}_{(1-\sigma)t, T_{t,X_0}(\check{X})} \circ T_{\sigma t, X_0}(\check{\Xi}),$$

the function $\check{F}^0$ rewrites as

$$\check{F}^0_{\sigma,\check{\Xi}}(\sqrt{t}, X_0, \check{X}) = \check{A}(\sqrt{\sigma t}, X_0, \check{\Xi})\check{B}\big(\sqrt{(1-\sigma)t}, T_{t,X_0}(\check{X}), h^0_{\sigma,\check{\Xi}}(\sqrt{t}, X_0, \check{X})\big)$$

Similarly, for

$$h^1_{\sigma,\check{\Xi}}(\sqrt{t}, X_1, \check{X}) = T^{-1}_{\sigma t, S_{t,X_1}(\check{X})} \circ S_{(1-\sigma)t, X_1}(\check{\Xi}),$$

we get

$$\check{F}^1_{\sigma,\check{\Xi}}(\sqrt{t}, X_1, \check{X}) = \check{A}\big(\sqrt{\sigma t}, S_{t,X_1}(\check{X}), h^1_{\sigma,\check{\Xi}}(\sqrt{t}, X_1, \check{X})\big)\check{B}(\sqrt{(1-\sigma)t}, X_1, \check{\Xi}).$$

We will prove the following bounds: for any $\varepsilon > 0$, $k, \ell \in \mathbb{N}$, and $\alpha, \beta \in \mathbb{N}^3$, we have

$$\sup_{0<\sigma<1-\varepsilon} \sup_{\tau^2 \le T} \sup_{X_0, \check{X}, \check{\Xi}} |\check{X}|^k |\check{\Xi}|^\ell \cdot \left|\partial_\tau^n D^\alpha_{X_0} D^\beta_{\check{X}} \check{F}^0_{\sigma,\check{\Xi}}(\tau, X_0, \check{X})\right| < \infty, \tag{1.5 a}$$

$$\sup_{\varepsilon<\sigma<1} \sup_{\tau^2 \le T} \sup_{X_1, \check{X}, \check{\Xi}} |\check{X}|^k |\check{\Xi}|^\ell \cdot \left|\partial_\tau^n D^\alpha_{X_1} D^\beta_{\check{X}} \check{F}^1_{\sigma,\check{\Xi}}(\tau, X_1, \check{X})\right| < \infty. \tag{1.5 b}$$

Using these estimates, the reader will have no difficulty in proving that $I^0$ is in $\check{\Psi}$, when seen as a function of $\sqrt{t}$, $X_0$ and $\check{X}$. Since $I^1$ is in $\check{\Psi}$ using similar arguments, $A * B$ is indeed an element of $\Psi^{a+b}$.

It is elementary to reduce the proof of said bounds to the two facts below.

1. There exists a constant $C > 0$ large enough so that

$$|\check{X}| \le C\Big(1 + \sqrt{T} + \check{\Xi} + h^0_{\sigma,\check{\Xi}}(\tau, X_0, \check{X})\Big)^3,$$

   uniformly for $X_0, \check{X}, \check{\Xi} \in \mathbb{R}^3$, $0 < \tau \le T$, and $0 \le \sigma \le 1-\varepsilon$. The analogous result holds for $h^1$, uniformly for $\varepsilon \le \sigma \le 1$.

2. For fixed $0 \le \sigma \le 1-\varepsilon$ and $\check{\Xi} \in \mathbb{R}^3$, $h^0_{\sigma,\check{\Xi}}$ is smooth up to the boundary. Moreover, all its derivatives with respect to $(\tau, X_0, \check{X})$ are bounded by polynomials in $\tau$, $\check{X}$ and $\check{\Xi}$, uniformly for $X_0, \check{X}, \check{\Xi} \in \mathbb{R}^3$, $\tau \ge 0$, and $0 \le \sigma \le 1-\varepsilon$. The analogous result holds for $h^1$, uniformly for $\varepsilon \le \sigma \le 1$.

Let us first find explicit expressions for $h^0$ and $h^1$. We write sinc : $x \mapsto \sin(x)/x$ for the cardinal sine function, and cosc : $x \mapsto (\cos(x)-1)/x^2$. The function sinc is bounded by $C/x$ at infinity, as well as all of its derivatives, for different constants $C > 0$. The same holds for cosc, with a bound of the form $C/x^2$. Then, if $\check{\Xi} = (\check{\alpha}, \check{\beta}, \check{\psi})$, direct computations lead to

$$h^0_{\sigma,\check{\Xi}}(\sqrt{t}, X_0, \check{X}) = \left(\frac{(a1)}{(1-\sigma)^2}, \frac{(a2)}{(1-\sigma)^{\frac{3}{2}}}, \frac{(a3)}{(1-\sigma)^{\frac{1}{2}}}\right),$$

$$\begin{aligned}
(a1) &= (\check{x}-\sigma^2\check{\alpha})\cos(-\sqrt{t}\phi^0) + (1-\sigma)|\phi^0|^2\operatorname{cosc}(-\sqrt{t}\phi^0) + (\check{y}-\sigma\sqrt{\sigma}\check{\beta})\phi^0\operatorname{sinc}(-\sqrt{t}\phi^0)\\
(a2) &= \sqrt{t}(\check{x}-\sigma^2\check{\alpha})\sin(-\sqrt{t}\phi^0) - (1-\sigma)\phi^0\operatorname{sinc}(-\sqrt{t}\phi^0) + (\check{y}-\sigma\sqrt{\sigma}\check{\beta})\cos(-\sqrt{t}\phi^0)\\
(a3) &= \check{\phi}-\phi^0 \qquad \text{for} \qquad \phi^0 = \sqrt{\sigma}\check{\psi},
\end{aligned}$$

$$h^1_{\sigma,\check{\Xi}}(\sqrt{t}, X_1, \check{X}) = \left(\frac{(b1)}{\sigma^2}, \frac{(b2)}{\sigma^{\frac{3}{2}}}, \frac{(b3)}{\sigma^{\frac{1}{2}}}\right),$$

$$\begin{aligned}
(b1) &= \check{x} - (1-\sigma)^2\check{\alpha}\cos(\sqrt{t}\phi^1) - (1-\sigma)|\phi^1|^2\operatorname{cosc}(\sqrt{t}\phi^1) + (1-\sigma)^{\frac{3}{2}}\phi^1\check{\beta}\operatorname{sinc}(\sqrt{t}\phi^1)\\
(b2) &= \check{y} - (1-\sigma)^2\check{\alpha}\sqrt{t}\sin(\sqrt{t}\phi^1) - (1-\sigma)\phi^1\operatorname{sinc}(\sqrt{t}\phi^1) - (1-\sigma)^{\frac{3}{2}}\check{\beta}\cos(\sqrt{t}\phi^1)\\
(b3) &= \phi^1 \qquad \text{for} \qquad \phi^1 = \check{\phi} - \sqrt{1-\sigma}\check{\psi}.
\end{aligned}$$

It is noteworthy that they do not depend on $X_0$ or $X_1$; invariance arguments could have shown that property also. In this form, it is clear that both functions are smooth up to the boundary, and that the derivatives satisfy the bounds described in the second fact above.

Concerning the first point, we consider first the easiest case of $i=1$. Denote by $(v,w,p)$ the components of $h^1_{\sigma,\check{\Xi}}$. Then

$$|\check{\phi}| \le |\phi^1| + |\check{\phi}-\phi^1| = \sqrt{\sigma}|p| + \sqrt{1-\sigma}|\check{\psi}|,$$

which is enough to control $|\check{\phi}|$. Using the same line of reasoning,

$$|\check{x}| \le \sigma^2|v| + \big(|\check{\alpha}| + |\phi^1|^2 + |\phi^1|\cdot|\check{\beta}|\big) \le |v| + |\check{\alpha}| + |p|^2 + |p|\cdot|\check{\beta}|,$$

$$|\check{y}| \le \sigma^{\frac{3}{2}}|w| + \big(|\check{\alpha}|\sqrt{T} + |\phi^1| + |\check{\beta}|\big) \le |w| + |\check{\alpha}|\sqrt{T} + |p| + |\check{\beta}|.$$

For the case $i=0$, setting again $(v,w,p)$ the components of $h^0_{\sigma,\check{\Xi}}$, we find an upper bound on $\check{\phi}$ in a similar way:

$$|\check{\phi}| \le |\check{\phi} - \sqrt{\sigma}\check{\psi}| + \sqrt{\sigma}|\check{\psi}| = \sqrt{1-\sigma}|p| + \sqrt{\sigma}|\check{\psi}|.$$

Regarding $\check{x}$ and $\check{y}$, note that

$$\begin{aligned}
&\begin{pmatrix} \cos(-\sqrt{t}\phi^0) & \phi^0\operatorname{sinc}(-\sqrt{t}\phi^0)\\ \sqrt{t}\sin(-\sqrt{t}\phi^0) & \cos(-\sqrt{t}\phi^0)\end{pmatrix}\begin{pmatrix}\check{x}-\sigma^2\check{\alpha}\\ \check{y}-\sigma\sqrt{\sigma}\check{\beta}\end{pmatrix}\\
&\qquad\qquad = \begin{pmatrix}(1-\sigma)^2 v\\ (1-\sigma)^{\frac{3}{2}}w\end{pmatrix} + (1-\sigma)\phi^0\begin{pmatrix}\phi^0\operatorname{cosc}(-\sqrt{t}\phi^0)\\ -\operatorname{sinc}(-\sqrt{t}\phi^0)\end{pmatrix}.
\end{aligned}$$

The norm of the right hand side is clearly bounded by a constant multiple of

$$|v| + |w| + |\phi^0|(1+|\phi^0|).$$

Since $|\phi^0|$ is controlled by $|p| + |\check{\phi}|$, to conclude, it would be sufficient to control the size of the inverse the $2\times 2$-matrix on the left hand side.

Call $M$ this matrix. Then, expressing $M^{-1}$ in terms of the adjugate of $M$ and denoting $\|\cdot\|$ some norm on the space of 2 by 2 matrices, we have

$$\big\|M^{-1}\big\| \le \frac{C\|M\|}{|\det M|} \le \frac{C'\big(|\phi^0| + \sqrt{T}\big)}{\cos(-\sqrt{t}\phi^0)^2 + \sin(-\sqrt{t}\phi^0)^2} = C'\big(|\phi^0| + \sqrt{T}\big)$$

for some $C, C' > 0$ large enough. Using again the bound on $|\phi^0|$, we are done with the proof of the first item, hence that of Theorem 1.6. □

The techniques introduced in the proof of the theorem are of interest in other parts of the work. We isolate the following result, which should be seen as a corollary of the above proof, and will be used in various arguments.

**Corollary 1.7.** *Let $A \in \Psi^a$, $B \in \Psi^b$, for some $a, b \in \mathbb{R}$. Fix some $0 < \varepsilon < 1/2$, and a compact subset $K$ of $(0;\infty) \times \mathbb{R}^3 \times \mathbb{R}^3$. Then over $\varepsilon < \sigma < 1-\varepsilon$, $\Xi \in \mathbb{R}$, the quantity*

$$|A(\sigma t, X_0, \Xi) B((1-\sigma)t, \Xi, X_1)|$$

*is bounded by an integrable function $M(\sigma, \Xi)$ over $\varepsilon < \sigma < 1-\varepsilon$, $\Xi \in \mathbb{R}^3$, for all $(t, X_0, X_1) \in K$.*

*If moreover $a > 0$, resp. $b > 0$, the bound actually holds for sigma in $(0; 1-\varepsilon)$, resp. $(\varepsilon; 1)$.*

Of course, if both $a$ and $b$ are positive, the bound is valid over $0 < \sigma < 1$.

*Proof.* The first inequality is an easy consequence of the definition of $\check{\Psi}$. Since $\sigma t$ is bounded away from zero, there exists large constants $M$ and $M'$ depending on $K$, $\varepsilon$ and $A$ such that

$$|A(\sigma t, X_0, \Xi)| \leq \frac{M(\sigma t)^{-5+a}}{1 + |T^{-1}_{\sigma t, X_0}(\Xi)|^2} \leq \frac{M'}{1 + |\Xi|^2}.$$

The same type of estimate holds for $B$,

$$|B((1-\sigma)t, \Xi, X_1)| \leq \frac{M'}{1 + |\Xi|^2},$$

and the first statement is proved.

We consider then the case $a > 0$; the case $b > 0$ is similar and left to the reader. The quantity of interest will satisfy the conclusion of the corollary if and only if for any compact $\check{K}$ of $(0;\infty) \times \mathbb{R}^3 \times \mathbb{R}^3$, the function

$$(\sigma t)^4 \big| A\big(\sigma t, X_0, T_{\sigma t, X_0}(\check{\Xi})\big) B\big((1-\sigma)t, T_{\sigma t, X_0}(\check{\Xi}), T_{t, X_0}(\check{X})\big) \big|$$

is bounded by an integrable function of $\sigma \in (0; 1-\varepsilon)$ and $\Xi \in \mathbb{R}^3$, uniformly in $(\sqrt{t}, X_0, \check{X}) \in \check{K}$. Indeed, the function

$$(\tau, X_0, \check{X}) \mapsto \big(\tau^2, X_0, T_{\tau^2, X_0}(\check{X})\big)$$

is a diffeomorphism from $(0;\infty) \times \mathbb{R}^3 \times \mathbb{R}^3$ to itself, and the Jacobian of $T_{\sigma t, X_0}$ is $(\sigma t)^4$. Defining $\check{F}^0$ as in the proof of Theorem 1.6, this last quantity is

$$t^{-6+a+b} \sigma^{-1+a} (1-\sigma)^{-5+b} \check{F}^0_{\sigma, \check{\Xi}}(\sqrt{t}, X_0, \check{X}).$$

Since $t$ is bounded away from zero, $a$ positive, and $\sigma$ away from one, it would suffice to have, for instance,

$$\big| \check{F}^0_{\sigma, \check{\Xi}}(\sqrt{t}, X_0, \check{X}) \big| \leq \frac{M}{1 + |\check{\Xi}|^4}$$

for some large constant $M$. But this is a lot weaker than the bound (1.5 a) that we established in the proof of Theorem 1.6, and so we are done with the second item of Corollary 1.7. $\square$

## 1.3 Iterated convolutions and differentiation rules

**Iterated convolutions.** Let $A_1, A_2, A_3$ be functions of $\Psi^a$ for some fixed $a > 0$. Then the convolution Theorem 1.6 asserts that $A_1 * (A_2 * A_3)$ and $(A_1 * A_2) * A_3$ are well-defined and belong to $\Psi^{3a}$. In the following, we would like to reduce the study of iterated convolutions on the left (the functions $\mathcal{E}^k\widetilde{u}$) to that of convolution of smooth functions arising as instances of $E^{*5}$; of course, we would like the convolution operation to be associative whenever well-defined. The method of proof of Theorem 1.6 is in fact able to prove the following result.

**Proposition 1.8.** *Let $A_0, \cdots, A_n$ be elements of $\Psi^{a_0}, \cdots, \Psi^{a_n}$ respectively, with each $a_i$ positive. Then, for any choice of parentheses for the expression $A_0 * \cdots * A_n$, the convolution is well-defined, belongs to $\Psi^{a_0+\cdots+a_n}$, and equals*

$$\iint_{[0,t]^n \times (\mathbb{R}^3)^n} \mathbf{1}_{t_1 \le \cdots \le t_n} A_0(t_1, X_0, \Xi_1) \cdots A_i(t_{i+1} - t_i, \Xi_i, \Xi_{i+1}) \cdots A_n(t - t_n, \Xi_n, X_1) \mathrm{d}t\, \mathrm{d}\Xi. \quad (1.6)$$

*In particular, we can write $A_0 * \cdots * A_n$ unambiguously.*

*Informally, $*$ is associative on $\Psi^{>0}$.*

*Proof.* It is clear that if the integrand is in $\mathrm{L}^1$, then we can swap integrals around and show that it equals an expression with parentheses of our liking. On the other hand, such an $\mathrm{L}^1$ estimate would follow if we were to show that some choice of parentheses makes $|A_0| * \cdots * |A_n|$ well-defined.

But this is clear from the proof of Theorem 1.6; let us review it quickly. Set $\check{\Psi}_0$ the space of functions satisfying (1.3) for $\alpha = \beta = 0$, $n = 0$, and $\Psi_0^a$ that of functions $A$ that admits a representation as in (1.4 a) or (1.4 b) with $\check{A}^i \in \check{\Psi}_0$. Then it is clear that $|A|$ is in $\Psi_0^a$ whenever $A$ is in $\Psi^a$. In the theorem, the convergence of the integral defining $A * B$ follows from (1.5 a) and (1.5 b), with $\alpha = \beta = 0$, $n = 0$. For this to hold, it suffices that $A$ be in $\Psi_0^a$, and $B$ be in $\Psi_0^a$. In other words, the proof of Theorem 1.6 gives $\Psi_0^a * \Psi_0^b \subset \Psi_0^{a+b}$. Since $|A_i|$ belongs to $\Psi_0^{a_i}$ for all $i$, the convolution $|A_0| * \cdots * |A_n|$ is well defined in $\Psi_0^{a_0+\cdots+a_n}$ for each choice of parentheses, and the integral in the Proposition is convergent, which concludes as explained. □

As described above, this representation is particularly useful when we consider the $\ell$-fold convolution of functions smooth up to the boundary, as we illustrate below. For technical reasons, we introduce the operator $\mathcal{T} : \Psi^a \to \Psi^{a+1}$ characterised by

$$\mathcal{T}v : (t, X_0, X_1) \mapsto tv(t, X_0, X_1). \quad (1.7)$$

Of course, there is an inverse $\mathcal{T}^{-1} : \Psi^a \to \Psi^{a-1}$ with obvious definition. Note also that $\mathcal{T}(A*B) = (\mathcal{T}A) * B + A * (\mathcal{T}B)$ whenever each term is well-defined. We have seen in Proposition 1.5 that $\partial_t : \Psi^{a+1} \to \Psi^a$ is well-defined, so that both $\mathcal{T}\partial_t$ and $\partial_t\mathcal{T}$ are well-defined as operators on $\Psi^a$.

**Proposition 1.9.** *Let $A_0, \cdots, A_\ell \in \Psi^a$ for some $a > 0$ and $\ell \ge 1$, and fix $T > 0$. Suppose that there exists a constant $M > 0$ such that*

$$\left|A_i\big(t, X_0, T_{t,X_0}(\check{X})\big)\right| \le \frac{M}{(1 + |\check{X}|)^4}$$

*for all $i$, uniformly in $t \le T$, $X_0, \check{X} \in \mathbb{R}^3$. Then for any $p \le 5\ell$,*

$$|\mathcal{T}^{-p}(A_0 * \cdots * A_\ell)|_\infty \le \frac{C^\ell}{\ell!} M^{\ell+1} T^{5\ell - p} \quad (1.8)$$

*for a universal constant $C > 0$ large enough, and $|A|_\infty$ the supremum of $A$ over $(0;T] \times \mathbb{R}^3 \times \mathbb{R}^3$.*

Remark that for $a \geq 5$, the existence of such a constant $M$ follows from the definition of $\Psi^a$.

*Proof.* We use representation (1.6) to get

$$\begin{aligned}
&|(A_0 * \cdots * A_\ell)(t, X_0, X_1)| \\
&\quad \leq \iint_{[0,t]^\ell \times (\mathbb{R}^3)^\ell} \mathbf{1}_{t_1 \leq \cdots \leq t_\ell} M^{\ell+1} \big(1 + |T^{-1}_{t_1, X_0}(\Xi_1)|\big)^{-4} \cdots \big(1 + |T^{-1}_{t-t_\ell, \Xi_\ell}(X_1)|\big)^{-4} \mathrm{d}t \,\mathrm{d}\Xi \\
&\quad \leq \iint_{[0,t]^\ell \times (\mathbb{R}^3)^\ell} \mathbf{1}_{t_1 \leq \cdots \leq t_\ell} t_1^4 \cdots (t_\ell - t_{\ell-1})^4 M^{\ell+1} \big(1 + |\check{\Xi}_1|\big)^{-4} \cdots \big(1 + |\check{\Xi}_\ell|\big)^{-4} \cdot 1^{-4} \mathrm{d}t \,\mathrm{d}\check{\Xi},
\end{aligned}$$

where we used the change of variables $\check{\Xi}_i = T_{t_i - t_{i-1}, \Xi_{i-1}}(\Xi_i)$, with the convention $\Xi_0 = X_0$, $\Xi_{\ell+1} = X_1$, $t_0 = 0$, $t_{\ell+1} = t$. This gives

$$\begin{aligned}
|(A_0 * \cdots * A_\ell)(t, X_0, X_1)| &\leq M^{\ell+1} \Big( \int_{[0,t]^\ell} \mathbf{1}_{t_1 \leq \cdots \leq t_\ell} t_1^4 \cdots (t_\ell - t_{\ell-1})^4 \mathrm{d}t \Big) \cdot \Big( \int_{\mathbb{R}^3} \frac{\mathrm{d}\check{\Xi}}{(1 + |\check{\Xi}|)^4} \Big)^\ell \\
&\leq M^{\ell+1} t^{4\ell} \frac{t^\ell}{\ell!} C^\ell
\end{aligned}$$

for some universal constant $C > 0$, from which the proposition follows. □

**Derivatives.** Recall that $\mathcal{T}$ is defined by (1.7). The estimates of Corollary 1.7 give easy descriptions of the derivatives of the convolutions $A * B$.

**Proposition 1.10.** *Let $A \in \Psi^a$ and $B \in \Psi^b$ be some functions of regularity $a, b > 0$, and $\alpha = (\alpha_x, \alpha_y, \alpha_\phi) \in \mathbb{N}^3$. We set $K = 2(\alpha_x + \alpha_y + \alpha_\phi)$. Then, over $(0; \infty) \times \mathbb{R}^3 \times \mathbb{R}^3$, the derivatives of $A * B$ are given by*

$$\begin{aligned}
D^\alpha_{X_0}(A * B) &= (D^\alpha_{X_0} A) * B \text{ whenever } a > K, \\
D^{\alpha_1}_{X_1}(A * B) &= A * (D^{\alpha_1}_{X_1} B) \text{ whenever } b > K, \\
\partial_t(A * B) &= \mathcal{T}^{-1}\big(A * B + (\mathcal{T}\partial_t A) * B + A * (\mathcal{T}\partial_t B)\big).
\end{aligned}$$

For the sake of conciseness, we will use that last formula in the form

$$(\partial_t \mathcal{T})(A * B) = \big((\partial_t \mathcal{T})A\big) * B + A * \big((\partial_t \mathcal{T})B\big).$$

*Proof.* Recall that the convolution $(A * B)(t, X_0, X_1)$ rewrites as

$$t \int_0^1 \int_{\mathbb{R}^3} A(\sigma t, X_0, \Xi) B((1-\sigma)t, \Xi, X_1) \mathrm{d}\sigma \,\mathrm{d}\Xi.$$

We consider the first case, and fix $t > 0$, $X_1 \in \mathbb{R}^3$. If we prove that, for all $\beta \leq \alpha$ coordinate-wise and all $R > 0$, the quantity

$$|(D^\beta_{X_0} A)(\sigma t, X_0, \Xi) B((1-\sigma)t, \Xi, X_1)| \tag{1.9}$$

is bounded by an integrable function $f(\sigma, \Xi)$, uniformly in $|X_0| \leq R$, then we can differentiate inside the integral, and get

$$\begin{aligned}
(A * B)(t, X_0, X_1) &= \frac{\partial^\alpha}{\partial X_0^\alpha} \Big( t \int_0^1 \int_{\mathbb{R}^3} A(\sigma t, X_0, \Xi) B((1-\sigma)t, \Xi, X_1) \mathrm{d}\sigma \,\mathrm{d}\Xi \Big) \\
&= t \int_0^1 \int_{\mathbb{R}^3} (D^\alpha_{X_0} A)(\sigma t, X_0, \Xi) B((1-\sigma)t, \Xi, X_1) \mathrm{d}\sigma \,\mathrm{d}\Xi \\
&= \big((D^\alpha_{X_0} A) * B\big)(t, X_0, X_1).
\end{aligned}$$

But according to Proposition 1.5, $D^{\beta}_{X_0}A$ is in $\Psi^{>0}$ when $a > K$. Then Corollary 1.7 gives the expected estimate.

The other two relations are treated in the same way, once one notices that

$$\frac{\partial}{\partial t}\Big(A(\sigma t, X_0, \Xi)B((1-\sigma)t, \Xi, X_1)\Big)$$
$$= \frac{1}{t}(\mathcal{T}\partial_t A)(\sigma t, X_0, \Xi)B((1-\sigma)t, \Xi, X_1) + \frac{1}{t}A(\sigma t, X_0, \Xi)(\mathcal{T}\partial_t B)((1-\sigma)t, \Xi, X_1). \quad \square$$

## 1.4 Convergence of the series

Recall that $\widetilde{u}$ is the approximate kernel defined by equation (1.2), and $\mathcal{E}$ is the convolution operator $\mathcal{E} : v \mapsto ((-\partial_t + L)\widetilde{u}) * v =: E * \widetilde{u}$. We are ready for the proof of the main Theorem 1.1: the series

$$\sum_{k\geq 0} \mathcal{E}^k \widetilde{u}$$

converges to $u$ in the smooth topology over the sets $[t_-; t_+] \times \mathbb{R}^3 \times \mathbb{R}^3$.

**Convergence.** The first part of this result is the convergence of the series.

**Theorem 1.11.** *For all $0 < t_- < t_+$ and $k \geq 0$, the formal sum*

$$\widetilde{u} + \mathcal{E}\widetilde{u} + \mathcal{E}^2\widetilde{u} + \cdots$$

*is in fact convergent in the $\mathcal{C}^k$ topology, uniformly over the sets $[t_-; t_+] \times \mathbb{R}^3 \times \mathbb{R}^3$.*

Of course, we will need to study $\widetilde{u}$ to be able to prove such a result. The following property is all we need to make our approach work, and is proved in Section 1.5.

**Proposition 1.12.** *There exists $(\tau, X_0, \check{X}) \mapsto \check{u}(\check{X}) \in \check{\Psi}$ depending only on $\check{X}$ such that for any $t > 0$, $X_0, \check{X} \in \Psi^1$,*

$$\widetilde{u}_t(X_0; T_{t,X_0}(\check{X})) = \frac{1}{t^4}\check{u}(\check{X}) \qquad \textit{and} \qquad \widetilde{u}_t(S_{t,X_1}(\check{X}), X_1) = \frac{1}{t^4}\check{u}(\check{X}).$$

Of course, this implies that $\widetilde{u}$ belongs to $\Psi^1$. According to Proposition 1.5, a very rough argument gives $E \in \Psi^{-3}$. In fact, one can state a much stronger result.

**Proposition 1.13.** *The error term $E$ belongs to $\Psi^1$.*

*Proof.* Recall that $L$ and $L_\phi$ are the operators defining $u$ and $\widetilde{u}$ — see (0.1) and (1.2). Note that by definition of $E$ and $\widetilde{u}$,

$$E_t(X_0, X_1) = (L - L_{\phi_0})\widetilde{u}_t(X_0, X_1),$$

where $L$ and $L_{\phi_0}$ act on the second variable. For some fixed angle $\psi \in \mathbb{R}$, set $e_\psi$ the first order operator $\cos(\psi)\partial_x + \sin(\psi_0)\partial_y$ on $\mathbb{R}^3$. Then it is straightforward that

$$L = -\cos(\phi - \phi_0)e_{\phi_0} - \sin(\phi - \phi_0)e_{\phi_0 + \frac{\pi}{2}} + \frac{1}{2}\partial^2_\phi$$
$$L_{\phi_0} = -\big(1 - (\phi - \phi_0)^2/2\big)e_{\phi_0} - (\phi - \phi_0)e_{\phi_0 + \frac{\pi}{2}} + \frac{1}{2}\partial^2_\phi,$$

and that

$$e_{\phi_0}\widetilde{u}_t(X_0, X) = \frac{1}{t^4} \cdot \frac{1}{t^2}(\partial_{\check{x}}\check{u})\big(T^{-1}_{t,X_0}(X)\big), \qquad e_{\phi_0 + \frac{\pi}{2}}\widetilde{u}_t(X_0, X) = \frac{1}{t^4} \cdot \frac{1}{t\sqrt{t}}(\partial_{\check{y}}\check{u})\big(T^{-1}_{t,X_0}(X)\big).$$

All in all, setting $\mathcal{C}$ (resp. $\mathcal{S}$) the analytic continuations of $(\cos(x)-1+x^2/2)/x^4$ (resp. $(\sin(x)-x)/x^3$), we get

$$t^4 E_t\big(X_0, T_{t,X_0}(X)\big) = -\mathcal{C}(\sqrt{t}\check{\phi})\check{\phi}^4(\partial_{\check{x}}\check{u})(\check{X}) - \mathcal{S}(\sqrt{t}\check{\phi})\check{\phi}^3(\partial_{\check{y}}\check{u})(\check{X}).$$

Because $\check{u}$ is in $\check{\Psi}$ and the derivatives of $\mathcal{C}$ and $\mathcal{S}$ are bounded by polynomials, $E$ does indeed belong to $\Psi^1$. $\square$

We can now turn to the proof of Theorem 1.11.

*Proof of Theorem 1.11.* We show that the following stronger result holds. For any $U \in \Psi^a$, $a > 0$, and any $\alpha_0, \alpha_1 \in \mathbb{N}^3$, $n, p \in \mathbb{N}$, there exists $K \geq 0$ such that each term of the series

$$\sum_{k \geq K} \mathcal{T}^{-p}(\partial_t \mathcal{T})^n D^{\alpha_0}_{X_0} D^{\alpha_1}_{X_1}(\mathcal{E}^k U) \tag{1.10}$$

extends to a function continuous up to the boundary, and the series converges uniformly over the sets $[0;T] \times \mathbb{R}^3 \times \mathbb{R}^3$. It is indeed stronger, since the operator $(\partial_t \mathcal{T})^n$ decomposes as a linear combination of $\mathcal{T}^m \partial_t^m$ for $m \leq n$, with leading coefficient $\mathcal{T}^n \partial_t^n$. Hence, $\partial_t^n$ is a linear combination of terms of the form $\mathcal{T}^{-q}(\partial_t \mathcal{T})^m$, for $q, m \leq n$.

Fix exponents $\alpha_i = (a_i, b_i, c_i) \in \mathbb{N}^3$ for $i = 0, 1$ and $n \in \mathbb{N}$, as well as a time horizon $T > 0$ and a function $U \in \Psi^a$, $a > 0$. Set $K = 2\max(a_0 + b_0 + c_0, a_1 + b_1 + c_1)$. Then there exists a finite set of functions $\mathcal{A} \subset \Psi^{5+K}$ and a constant $c > 0$ such that for all $k$ large enough, one can write

$$\mathcal{E}^k U = A_0 * \cdots * A_\ell$$

with $A_i \in \mathcal{A}$ and $\ell \geq \max(1, k/c - 1)$. For instance, one can take

$$\mathcal{A} = \big\{\mathcal{E}^k U, 5 + K \leq k \leq 9 + 2K\big\} \cup \big\{E^{*(5+K)}\big\}$$

for $k \geq 10 + 2K$.

Fix some $M > 0$ such that

$$|(\partial_t \mathcal{T})^m D^{\beta_0}_{X_0} A(t, X_0, T_{t,X_0}(\check{X}))| \leq \frac{M}{(1 + |\check{X}|)^4}, \quad |(\partial_t \mathcal{T})^m D^{\beta_1}_{X_1} A(t, X_0, T_{t,X_0}(\check{X}))| \leq \frac{M}{(1 + |\check{X}|)^4},$$

for all $\beta_i \leq \alpha_i$ component-wise, $m \leq n$, and $A \in \mathcal{A}$. The existence of such $M$ follows from the fact that these derivatives are finitely many and belong to $\Psi^5$, by Proposition 1.5. Proposition 1.10 and its following remark, by induction, yields

$$\begin{aligned}
&\mathcal{T}^{-p}(\partial_t \mathcal{T})^n D^{\alpha_0}_{X_0} D^{\alpha_1}_{X_1}(A_0 * \cdots * A_\ell) \\
&= \sum_{n_0 + \cdots + n_\ell = n} \binom{n}{n_0, \cdots, n_\ell} \mathcal{T}^{-p}\big[(\partial_t \mathcal{T})^{n_0}(D^{\alpha_0}_{X_0} A_0) * \cdots * (\partial_t \mathcal{T})^{n_i} A_i * \cdots * (\partial_t \mathcal{T})^{n_\ell}(D^{\alpha_1}_{X_1} A_\ell)\big].
\end{aligned}$$

For $k$ sufficiently large, we can decompose $\mathcal{E}^k U$ into convolution product of $\ell + 1$ functions in $\mathcal{A}$, where $\ell$ is large as well, and the terms in the above sum are controlled by Proposition 1.9. Specifically, for all $\ell \geq p/5$, they extend to smooth functions up to the boundary, hence $\mathcal{T}^{-p}(\partial_t \mathcal{T})^n D^{\alpha_0}_{X_0} D^{\alpha_1}_{X_1}(\mathcal{E}^k U)$ does as well, and we have

$$\big|\mathcal{T}^{-p}(\partial_t \mathcal{T})^n D^{\alpha_0}_{X_0} D^{\alpha_1}_{X_1}(\mathcal{E}^k U)\big|_\infty \leq \sum_{n_0 + \cdots + n_\ell = n} \binom{n}{n_0, \cdots, n_\ell} \frac{C^{\ell+1}}{\ell!} = \frac{(\ell+1)^n C^{\ell+1}}{\ell!},$$

for some constant $C > 0$ large enough, independently of $k$ and $\ell$ (the constant $C$ depends on $M$ and $T$). In particular, since $\ell \geq k/c - 1$, there exists some $K_0 \geq 0$ such that the series

$$\sum_{k \geq K_0} \mathcal{T}^{-p}(\partial_t \mathcal{T})^n D^{\alpha_0}_{X_0} D^{\alpha_1}_{X_1}(\mathcal{E}^k U)$$

extends to a series of continuous functions over $[0;T] \times \mathbb{R}^3 \times \mathbb{R}^3$ that converges uniformly (at rate at least $\exp(-\varepsilon k \ln k)$ for some $\varepsilon > 0$). $\square$

**Limit.** Denote by $U$ the limit of the sum $\sum \mathcal{E}^k \widetilde{u}$. We want to show that $U = u$; we expect that it suffices to show that $U$ satisfies the defining equation (0.1) for $u$, as well as the initial condition $U_t(X_0, \cdot) \to \delta_{X_0}$. For this reasoning to hold, we need some uniqueness argument. We only give an outline of how such an argument may go; it is borrowed from [Mel93, page 271]. Assuming Theorem 1.14 and Proposition 1.15, which correspond to the two criteria aforementioned, and defining for fixed $X_0 \in \mathbb{R}^3$ the distribution

$$w : f \mapsto \int_0^t \big(u_t(X_0, X) - U_t(X_0, X)\big) f(t, X) \mathrm{d}X,$$

we see that $(\partial_t - L)w = 0$ as a distribution over $\mathbb{R} \times \mathbb{R}^3$. Then, for any smooth compactly supported $f : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}$ that vanishes, say, for $t \geq T$, there exists a smooth function $g$ such that $(-\partial_t - L^*)g = f$ and $g$ also vanishes for $t \geq T$. Indeed, the operator $L$ is the generator associated to a hypoelliptic diffusion $X^L$ (a noteworthy point is that it has no term of order zero),[2] and as such, stochastic processes provide tools to construct such a $g$:

$$g : t, x \mapsto \int_0^\infty \int_{\mathbb{R}^3} f(t - s, X_0) u^L_s(X_0, x) \mathrm{d}X_0 \mathrm{d}s,$$

where $u^L_t(X_0, \cdot)$ is the (smooth) density of $X^L_t$ with initial condition $X^L_0 = X_0$. Then, the following reasoning can be made rigorous:

$$\int_{\mathbb{R}\times\mathbb{R}^3} wf = \int_{\mathbb{R}\times\mathbb{R}^3} w(-\partial_t - L^*)g = \int_{\mathbb{R}\times\mathbb{R}^3} (\partial_t - L)wg = 0$$

for all $f \in \mathcal{D}$, hence $w = 0$ and $U = u$.

**Theorem 1.14.** *Over $\mathbb{R}^*_+ \times \mathbb{R}^3 \times \mathbb{R}^3$, $U$ satisfies*

$$\partial_t U = LU.$$

*Proof.* We show that $(\partial_t - L)(\mathcal{E}^{k+1}\widetilde{u}) = \mathcal{E}^k E - \mathcal{E}^{k+1} E$. The sum then telescopes to give, say in the pointwise sense,

$$(\partial_t - L)U = -E + \sum_{k \geq 0} \big(\mathcal{E}^k E - \mathcal{E}^{k+1} E\big) = \lim_{k\to\infty} \mathcal{E}^k E.$$

Note that we used the fact that the series defining $U$ converges in the $\mathcal{C}^2$ topology. Since, as stated at the beginning of the proof of Theorem 1.11, the series $\sum_{k\geq 0} \mathcal{E}^k E$ is convergent in $\mathcal{C}^0$, the limit is in fact zero and $U$ satisfies the partial differential equation.

[2] In fact, $X^L$ is in essence a kinetic Brownian motion.

It remains to show the property stated above. We will prove a slightly more general result: for any $A \in \Psi^a$, $a > 0$,

$$\partial_t(A * \widetilde{u}) = A + L(A * \widetilde{u}) - A * E. \tag{1.11}$$

For $A = \mathcal{E}^k E$, this gives as expected

$$\partial_t(\mathcal{E}^{k+1}\widetilde{u}) = \mathcal{E}^k E + L(\mathcal{E}^{k+1}\widetilde{u}) - \mathcal{E}^{k+1}E.$$

For a few paragraphs, we write integrals and limits without checking if they are defined or if the exchanges are licit. The reader can rest assured that they are, as we will prove after these formal manipulations. Let $A$ be a function in $\Psi^a$ for some $a > 0$. Fix some $X_0 \in \mathbb{R}^3$ and $t > 0$: we see maps like $A$ as functions of their third variable $X_1$. Then for any $\varepsilon > 0$,

$$\begin{aligned}\frac{\partial}{\partial t}\Big(\int_0^{(1-\varepsilon)t}\int_{\mathbb{R}^3} A(s, X_0, \Xi)\widetilde{u}(t-s, \Xi, X_1)\mathrm{d}s\,\mathrm{d}\Xi\Big) \qquad\qquad \\ = (1-\varepsilon)\int_{\mathbb{R}^3} A((1-\varepsilon)t, X_0, \Xi)\widetilde{u}(\varepsilon t, \Xi, X_1)\mathrm{d}\Xi \\ + \int_0^{(1-\varepsilon)t}\int_{\mathbb{R}^3} A(s, X_0, \Xi)(\partial_t\widetilde{u})(t-s, \Xi, X_1)\mathrm{d}s\,\mathrm{d}\Xi.\end{aligned} \tag{1.12}$$

Performing the change of variables $s = \sigma t$, the left hand term is equal to

$$\begin{aligned}\int_0^{1-\varepsilon}\int_{\mathbb{R}^3} A(\sigma t, X_0, \Xi)\widetilde{u}((1-\sigma)t, \Xi, X_1)\mathrm{d}\sigma\,\mathrm{d}\Xi \\ + t\int_0^{1-\varepsilon}\int_{\mathbb{R}^3}\frac{\partial}{\partial t}\Big(A(\sigma t, X_0, \Xi)\widetilde{u}((1-\sigma)t, \Xi, X_1)\Big)\mathrm{d}\sigma\,\mathrm{d}\Xi,\end{aligned} \tag{1.13}$$

which converges uniformly for bounded $X_1$ to the same expression with $\varepsilon = 0$, which is but $\partial_t(A * \widetilde{u})$. Similarly, since $E = (-\partial_t + L)\widetilde{u}$, the last term is equal to

$$\begin{aligned}Z_t^{(\varepsilon)}(X_0, X_1) := L\Big(\int_0^{(1-\varepsilon)t}\int_{\mathbb{R}^3} A(s, X_0, \Xi)\widetilde{u}(t-s, \Xi, X_1)\mathrm{d}s\,\mathrm{d}\Xi\Big) \\ - \int_0^{(1-\varepsilon)t}\int_{\mathbb{R}^3} A(s, X_0, \Xi)E(t-s, \Xi, X_1)\mathrm{d}s\,\mathrm{d}\Xi,\end{aligned}$$

which converges to the same expression with $\varepsilon = 0$, in a weak sense that we define later. We recognise the limit as $L(A * \widetilde{u}) - A * E$.

To conclude the proof of (1.11), it remains to show that the first term in the right hand side of equation (1.12) converges in some sense to $A$. Let $f : \mathbb{R}^3 \to \mathbb{R}$ be continuous compactly supported. Then

$$\begin{aligned}\int_{\mathbb{R}^3}\Big(\int_{\mathbb{R}^3} A((1-\varepsilon)t, X_0, \Xi)\widetilde{u}(\varepsilon t, \Xi, X_1)\mathrm{d}\Xi\Big)f(X_1)\mathrm{d}X_1 \qquad\qquad \\ = \int_{\mathbb{R}^3}\Big(\int_{\mathbb{R}^3}\widetilde{u}(\varepsilon t, \Xi, X_1)f(X_1)\mathrm{d}X_1\Big)A((1-\varepsilon)t, X_0, \Xi)\mathrm{d}\Xi \\ \to \int_{\mathbb{R}^3} A(t, X_0, \Xi)f(\Xi)\mathrm{d}\Xi.\end{aligned}$$

This holds for any such $f$; we say that the first term of the right hand side of (1.12) converges to $A$ in $(\mathcal{C}_c^0)^*$. In the same mindset, the convergence of $Z^{(\varepsilon)}$ to $L(A * \widetilde{u}) - A * E$ holds in $(\mathcal{C}_c^2)^*$:

for any $f : \mathbb{R}^3 \to \mathbb{R}$ twice continuously differentiable with compact support, integration by parts gives

$$\int_{\mathbb{R}^3} Z_t^{(\varepsilon)}(X_0, X_1) f(X_1) \mathrm{d}X_1 \to \int_{\mathbb{R}^3} \big(L(A * \widetilde{u}) - A * E\big)_t (X_0, X_1) f(X_1) \mathrm{d}X_1.$$

Hence, both sides of (1.12) converge to both sides of (1.11) in the $(\mathcal{C}_c^2)^*$ sense. Since the terms appearing in the equation (1.11) are actually continuous in $X_1$, this proves that said equality holds pointwise in $(0;\infty) \times \mathbb{R}^3 \times \mathbb{R}^3$.

There is a lot left to justify here; we will reduce it all to the following facts. Fix $0 < t_- < t_+$, and a large radius $R > 0$. For brevity, we write $A(s)$ for $A(s, X_0, \Xi)$, and $\widetilde{u}(s)$ for $\widetilde{u}(s, \Xi, X_1)$. When there is no confusion hazard, we also write $(\partial_t \widetilde{u})(s)$ for $(\partial_t \widetilde{u})(s, \Xi, X_1)$, etc.

1. For any $A \in \Psi^{>0}$, $B \in \Psi^b$, $b \in \mathbb{R}$ and $0 < \delta \leq 1$, there exists a constant $M$ such that
$$\int_0^{(1-\delta)t} \int_{\mathbb{R}^3} |A(s, X_0, \Xi) B(t - s, \Xi, X_1)| \mathrm{d}s\, \mathrm{d}\Xi \leq M$$
for all $t_- < t < t_+$, $|X_0| + |X_1| \leq R$.
2. For any $A \in \Psi^a$ and $B \in \Psi^b$, $a, b \in \mathbb{R}$, and any $0 < \delta < 1/2$, there exists a constant $M$ such that
$$|A(s, X_0, \Xi) B(t - s, \Xi, X_1)| \leq \frac{M}{1 + |\Xi|^4}$$
over $\delta t < s < (1 - \delta)t$, for all $t_- < t < t_+$, $|X_0| + |X_1| \leq R$.

The first fact follows directly from the change of variables $s = \sigma t$ and Corollary 1.7. The second is a consequence of the fact that when $t$ is bounded away from zero,

$$|A(t, X_0, \Xi)| \leq \frac{M t^{-5+a}}{1 + |T_{t,X_0}^{-1}(\Xi)|^2} \leq \frac{M' t^{-5+a}}{1 + |\Xi|^2}$$

for constants $M$ and $M'$ large enough, and similarly

$$|B(t, \Xi, X_1)| \leq \frac{M t^{-5+b}}{1 + |S_{t,X_0}^{-1}(\Xi)|^2} \leq \frac{M' t^{-5+b}}{1 + |\Xi|^2}.$$

*Time derivative.* First, we need to show that the representation (1.12) is valid. Fix $X_0, X_1 \in \mathbb{R}^3$ such that $|X_0| + |X_1| \leq R$. The integrals appearing in the right hand side are well-defined according to the above two facts. To prove that the integral in the left hand side is differentiable, we must control, say for any $t$ and $\tau$ such that $t_- < t < t_+$ and $|\tau| < \varepsilon t_-/2$,

$$\Bigg| \int_0^{(1-\varepsilon)(t+\tau)} \int_{\mathbb{R}^3} A(s) \widetilde{u}(t + \tau - s) \mathrm{d}s\, \mathrm{d}\Xi - \int_0^{(1-\varepsilon)t} \int_{\mathbb{R}^3} A(s) \widetilde{u}(t - s) \mathrm{d}s\, \mathrm{d}\Xi$$
$$- (1 - \varepsilon)\tau \int_{\mathbb{R}^3} A((1 - \varepsilon)t) \widetilde{u}(\varepsilon t) \mathrm{d}s\, \mathrm{d}\Xi - \tau \int_0^{(1-\varepsilon)t} \int_{\mathbb{R}^3} A(s) (\partial_t \widetilde{u})(t - s) \mathrm{d}s\, \mathrm{d}\Xi \Bigg|.$$

Precisely, we want to show that it is negligible when compared to $\tau$. A bound on this quantity is given by

$$\int_0^{(1-\varepsilon)t} \int_{\mathbb{R}^3} |A(s)| \cdot \Big| \widetilde{u}(t + \tau - s) - \widetilde{u}(t - s) - \tau (\partial_t \widetilde{u})(t - s) \Big| \mathrm{d}s\, \mathrm{d}\Xi$$
$$+ \int_{(1-\varepsilon)t}^{(1-\varepsilon)(t+\tau)} \int_{\mathbb{R}^3} \Big| A(s) \widetilde{u}(t + \tau - s) - A((1 - \varepsilon)t) \widetilde{u}(\varepsilon t) \Big| \mathrm{d}s\, \mathrm{d}\Xi.$$

Since

$$\widetilde{u}(t+\tau-s)-\widetilde{u}(t-s)-\tau(\partial_t\widetilde{u})(t-s)=\int_0^\tau(\tau-\theta)(\partial_t^2\widetilde{u})(t+\theta-s)\mathrm{d}\theta,$$

the integrand in the first term is less than

$$\int_0^\tau(\tau-\theta)\int_0^{(1-\varepsilon)t}\int_{\mathbb{R}^3}|A(s)|\cdot\Big|(\partial_t^2\widetilde{u})(t+\theta-s)\Big|\mathrm{d}s\,\mathrm{d}\Xi\,\mathrm{d}\theta.$$

By the hypotheses on $t$ and $\tau$, we have

$$(1-\varepsilon)t/(t+\theta)<\frac{1-\varepsilon}{1-\varepsilon/2}<1,$$

so that the first fact above yields

$$\int_0^{(1-\varepsilon)t}\int_{\mathbb{R}^3}|A(s)|\cdot\Big|\widetilde{u}(t+\tau-s)-\widetilde{u}(t-s)-\tau(\partial_t\widetilde{u})(t-s)\Big|\mathrm{d}s\,\mathrm{d}\Xi\le M\int_0^\tau(\tau-\theta)\mathrm{d}\theta\le Mt_+\tau^2.$$

Hence, the first term is negligible compared to $\tau$, as needed.

The second term in the computation of the time derivative is bounded by the integral over $[(1-\varepsilon)t;(1-\varepsilon)(t+\tau)]\times\mathbb{R}^3$ of

$$|A(s,X_0,\Xi)|\cdot\Big|\widetilde{u}(t+\tau-s,\Xi,X_1)-\widetilde{u}(\varepsilon t,\Xi,X_1)\Big|+\Big|A(s,X_0,\Xi)-A((1-\varepsilon)t,X_0,\Xi)\Big|\cdot|\widetilde{u}(\varepsilon t,\Xi,X_1)|.$$

For any $s\in[(1-\varepsilon)t;(1-\varepsilon)(t+\tau)]$ and $\Xi\in\mathbb{R}^3$, there exists some $\theta$ and $\nu$ in $[0;\tau]$ (or $[\tau;0]$ is $\tau$ is negative) such that this last expression is less than

$$|A(s,X_0,\Xi)\cdot\tau(\partial_t\widetilde{u})(\varepsilon t+\theta,\Xi,X_1)|+|(1-\varepsilon)\tau(\partial_tA)((1-\varepsilon)(t+\nu),X_0,\Xi)\cdot\widetilde{u}(\varepsilon t,\Xi,X_1)|.$$

We can apply the second fact to show that

$$\int_{(1-\varepsilon)t}^{(1-\varepsilon)(t+\tau)}\int_{\mathbb{R}^3}\Big|A(s)\widetilde{u}(t+\tau-s)-A((1-\varepsilon)t)\widetilde{u}(\varepsilon t)\Big|\mathrm{d}s\,\mathrm{d}\Xi\le(1-\varepsilon)\tau\int_{\mathbb{R}^3}\frac{M\tau}{1+|\Xi|^4}\mathrm{d}\Xi,$$

and the second term is also negligible with respect to $\tau$.

*Left hand side.* The derivative in the left hand side of (1.12) can be expressed as (1.13) if the integrands in this last expression are uniformly bounded for $t_-<t<t_+$ by integrable functions of $0<\sigma<1-\varepsilon$ and $\Xi\in\mathbb{R}^3$. Moreover, they will converge to the same expression with $\varepsilon=0$ uniformly for $|X_1|\le R$ if we can in fact found such a bound over $0<\sigma<1$. But this is the case because of Theorem 1.6: the integrands are the same that arise in the definition of the convolution products $A*\widetilde{u}$ and

$$(\mathcal{T}\partial_tA)*\widetilde{u}+A*(\mathcal{T}\partial_t\widetilde{u}),$$

where each of the functions involved are in $\Psi^{>0}$. The limit as $\varepsilon$ tends to zero is then

$$\mathcal{T}^{-1}\big(A*\widetilde{u}+(\mathcal{T}\partial_tA)*\widetilde{u}+A*(\mathcal{T}\partial_t\widetilde{u})\big),$$

which equals $\partial_t(A*\widetilde{u})$ according to Proposition 1.10.

*Space derivatives.* In the same fashion as above, the second term in the right hand side of (1.12) can be expressed as $Z^{(\varepsilon)}$ provided the integrands $A\times(D^\alpha_{X_1}\widetilde{u})$ and $A\times E$, for $\alpha$ in a finite set (say $\alpha\in\{0;1;2\}^3$) are bounded by integrable functions of $s\in(0;(1-\varepsilon)t)$ and $\Xi\in\mathbb{R}^3$, uniformly in $|X_0| + |X_1| \leq R$ and $t_- < t < t_+$. This is again a consequence of the first fact. Moreover, since $\widetilde{u}$ and $E$ are in $\Psi^{>0}$, the same fact also shows that both integrals involved in the definition of $Z^{(\varepsilon)}$ converge uniformly in $t_- < t < t_+$ and $|X_0| + |X_1| \leq R$ to the same integrals with $\varepsilon = 0$. This justifies the informal reasoning used to show convergence in the $(\mathcal{C}_c^2)^*$ topology described above.

*Approximate unity.* There remains to back up the claim that the integral exchange and convergence inside the integral holds, in the treatment of the first term of the right hand side of (1.12). Because $\widetilde{u}_s(X_0, \cdot)$ is the density of a probability law for any $s$ and $X_0$, for any $f : \mathbb{R}^3 \to \mathbb{R}$ continuous with compact support we have

$$\int_{\mathbb{R}^3} \Big( \int_{\mathbb{R}^3} \big|\widetilde{u}(\varepsilon t, \Xi, X_1) f(X_1)\big| \mathrm{d}X_1 \Big) \cdot |A((1-\varepsilon)t, X_0, \Xi)| \mathrm{d}\Xi \leq |f|_\infty \int_{\mathbb{R}^3} |A((1-\varepsilon)t, X_0, \Xi)| \mathrm{d}\Xi$$

with $|f|_\infty$ the supremum of $|f|$ over $\mathbb{R}^3$. Since this last integral is uniformly bounded according to the second fact, and because of the point-wise convergence, the convergence in $(\mathcal{C}_c^0)^*$ described above does indeed hold. $\square$

**Proposition 1.15.** *For any $f : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}$ continuous with compact support,*

$$\int_{\mathbb{R}^3} U(t, X_0, X_1) f(t, X_1) \mathrm{d}X_1 \xrightarrow[t \to 0]{} f(X_0).$$

*Proof.* Since $\widetilde{u}(t, X_0, \cdot)$ is the density of a continuous process starting at $X_0$ and $\sum_{k \geq K} \mathcal{E}^k \widetilde{u}$ converges uniformly on $[0;T] \times \mathbb{R}^3 \times \mathbb{R}^3$ for $K$ large enough (see the fact stated at the beginning of the proof of Theorem 1.11), it suffices to show that

$$\int_{\mathbb{R}^3} \mathcal{E}^k \widetilde{u}(t, X_0, X_1) f(t, X_1) \mathrm{d}X_1 \to 0$$

for any $k \geq 1$ and any such $f$, as $t$ tends to zero. In fact, it holds for any $A \in \Psi^{1+a}$, $a > 0$ in lieu of $\mathcal{E}^k \widetilde{u}$. Indeed, let $A$ be such a function; there exists $\check{A} \in \check{\Psi}$ such that

$$A(t, X_0, T_{t,X_0}(\check{X})) = t^{-4+a} \check{A}(\sqrt{t}, X_0, \check{X}).$$

Set $M > 0$ be a constant large enough so that

$$|\check{A}(\sqrt{t}, X_0, \check{X})| \leq \frac{M}{1 + |\check{X}|^4}, \qquad |f(t, X_1)| \leq M$$

uniformly in $t \leq 1$ and $X_0, \check{X}, X_1 \in \mathbb{R}^3$. Then

$$\begin{aligned} \Big| \int_{\mathbb{R}^3} A(t, X_0, X_1) f(t, X_1) \mathrm{d}X_1 \Big| &= \Big| \int_{\mathbb{R}^3} t^{-4+a} \check{A}(\sqrt{t}, X_0, \check{X}) f(t, T_{t,X_0}(\check{X}))\, t^4 \mathrm{d}\check{X} \Big| \\ &\leq t^a \int_{\mathbb{R}^3} \frac{M^2}{1 + |\check{X}|^4} \mathrm{d}\check{X}, \end{aligned}$$

which clearly vanishes as $t$ tends to zero. $\square$

This, according to the discussion before Theorem 1.14, concludes the proof of the main Theorem 1.1. Its following Corollary 1.2 is an easy consequence of its proof. Indeed, as announced in the first lines of the proof of Theorem 1.11, we know that the series

$$\sum_{k \geq 1} \mathcal{T}^3(\mathcal{E}^k \tilde{u})$$

converges uniformly over the sets $[0;T]\times\mathbb{R}^3\times\mathbb{R}^3$ to an error function $f$; we can take $K=1$ in equation (1.10) since each term of the sum is in $\Psi^{(k+1)+3}$, and functions of $\Psi^5$ are continuous up to the boundary. Each term $\mathcal{T}^3(\mathcal{E}^k\tilde{u})$ is moreover uniformly bounded on the same sets, so the same holds for $f$. In other words, there exists a continuous function $f:\mathbb{R}_+\times\mathbb{R}^3\times\mathbb{R}^3\to\mathbb{R}$ bounded on the sets $[0;T]\times\mathbb{R}^3\times\mathbb{R}^3$ such that

$$u_t(X_0,X_1)=\widetilde{u}_t(X_0,X_1)+\frac{f_t(X_0,X_1)}{t^3}=\frac{1}{t^4}\big(\check{u}(T^{-1}_{t,X_0}(X_1))+tf_t(X_0,X_1)\big).$$

If we set $M>0$ a bound on $f$ over $[0;T]\times\mathbb{R}^3\times\mathbb{R}^3$, and $D$ a domain of $\mathbb{R}^3$ over which $\check{u}$ is larger than $c>0$, then for $t\in[0;T]$ and $\check{X}\in D$, $|f|\le M/c\cdot\check{u}(\check{X})$ and

$$u_t\big(X_0,T_{t,X_0}(\check{X})\big)=\frac{1}{t^4}\check{u}(\check{X})(1+O(t))=\widetilde{u}_t\big(X_0,T_{t,X_0}(\check{X})\big)(1+O(t)).$$

Corollary 1.2 follows from the special case $X_0=0$. Similar estimates can be extracted for the derivatives of $u$.

## 1.5 Study of the approximation

In this section, we turn to the study of the approximation $\widetilde{u}$, solution to the quadratic approximation, and in particular prove Proposition 1.12: there exists $\check{u}\in\check{\Psi}$ depending only on $\check{X}$ such that for any $t>0$, $X_0,\check{X}\in\Psi^1$,

$$\widetilde{u}_t(X_0;T_{t,X_0}(\check{X}))=\frac{1}{t^4}\check{u}(\check{X})\qquad\text{and}\qquad\widetilde{u}_t(S_{t,X_1}(\check{X}),X_1)=\frac{1}{t^4}\check{u}(\check{X}).$$

**Normalisation.** We make precise the arguments at the end of section 1.2, to express $\widetilde{u}$ in terms of a simpler function $v$. Recall that $\widetilde{u}$ is the solution of (1.2). Using for instance the reformulation below it, we see that the function

$$v^{X_0}:(t,X)\mapsto\widetilde{u}_t\big(X_0,X_0+(\mathsf{e}^{\mathsf{i}\phi_0}(z+t),\phi)\big)$$

is the solution of

$$\partial_t v^{X_0}=\frac{\phi^2}{2}\partial_x v^{X_0}-\phi\,\partial_y v^{X_0}+\frac{1}{2}\partial_\phi^2 v^{X_0}.$$

Moreover, because $X\mapsto X_0+(\mathsf{e}^{\mathsf{i}\phi_0}(z-t),\phi)$ is an isometry for fixed $t$, and tends to $X\mapsto X_0+(\mathsf{e}^{\mathsf{i}\phi_0}z,\phi)$ as $t$ tends to zero, we see that $v^{X_0}$ has initial condition a Dirac distribution at the point whose image by the latter application is $X_0$, which is zero. Hence it does not depend on $X_0$, and we can write $v:=v^{X^0}$ for any $X_0$.

A probabilistic way to see to this result is that $\widetilde{u}_t(X_0,\cdot)$ is the density of the solution of

$$\begin{cases}\mathrm{d}z_t=\mathsf{e}^{\mathsf{i}\phi_0}\mathfrak{exp}(\mathsf{i}\phi_t-\mathsf{i}\phi_0)\mathrm{d}t\\ \mathrm{d}\phi_t=\mathrm{d}W_t\end{cases}\qquad\text{with fixed deterministic } z_0,\ \phi_0,$$

for $\mathfrak{exp}:x\mapsto 1+x+x^2/2$ the quadratic approximation of the exponential map, whereas $v$ is the density of

$$\begin{cases}\mathrm{d}z_t=\mathfrak{exp}(\mathsf{i}\phi_t)\mathrm{d}t\\ \mathrm{d}\phi_t=\mathrm{d}W_t\end{cases}\qquad\text{with}\qquad z_0=0,\ \phi_0=0.\tag{1.14}$$

Hence, for a given solution $(z,\phi)$ of the second equation, the stochastic process

$$t\mapsto\big(z_0+\mathsf{e}^{\mathsf{i}\phi_0}(z_t-t),\ \phi_0+\phi_t\big)$$

is solution to the first, which yields the same change of variables formula.

Note that since $\widetilde{u}(\cdot, X_0, \cdot)$ is a smooth function on $\mathbb{R}_+^* \times \mathbb{R}^3$, $v$ is smooth as well. Alternatively, it can be seen from the fact that the equation it satisfies is also hypoelliptic.

**Support.** The support of $v_t$ is $\{y^2 \leq 2tx\}$. Indeed, if $(x_t, y_t, \phi_t)$ is the solution of the SDE (1.14) with density $v$, we have

$$y_t^2 = \Big(\int_0^t \phi_s \mathrm{d}s\Big)^2 \leq t \int_0^t \phi_s^2 \mathrm{d}s = 2tx_t$$

almost surely. In the opposite direction, let $(x, y, \phi)$ be a point such that $y^2 \leq 2tx$. Let $w_n$ be a sequence of smooth functions $[0;t] \to \mathbb{R}$ such that $w_n(0) = 0$, $w_n(t) = \phi$, and $w_n \to w$ in the $L^1$ and $L^2$ topologies, where

$$w(s) := -\frac{y}{\alpha}\mathbf{1}_{s<\alpha}, \qquad \alpha := \frac{y^2}{2x}.$$

Then, according to the Stroock-Varadhan theorem,

$$\Big(\int_0^t \frac{w_n(s)^2}{2}\mathrm{d}s, -\int_0^t w_n(s)\mathrm{d}s, w_n(t)\Big) \in \mathrm{supp}(v_t),$$

hence $(x, y, \phi)$ is in the support as well.

**Fourier transform.** Since, for $t$ fixed, $v_t$ is a probability measure, we may consider its Fourier transform $\hat{v}_t$ with respect to space:

$$\hat{v}_t : P = (p, q, \psi) \mapsto \int_{\mathbb{R}^3} v_t(sX)\, \mathrm{e}^{-\mathrm{i}P\cdot X}\mathrm{d}X.$$

According to the semigroup approach of stochastic processes, we know that $\frac{1}{s}(v_{t+s} - v_s)$ converges to $\partial_t v$ in a mild sense, for instance as continuous linear functionals on the space of twice continuously derivable functions with all derivatives rapidly decreasing. In particular, the derivative of $t \mapsto v_t(X)$ coincides pointwise with the derivative of $t \mapsto v$ as a curve with values in the topological space of tempered distributions. Then $\hat{v}_t$ is the solution, as a tempered distribution-valued curve, of

$$\partial_t \hat{v}_t = -\frac{1}{2}\mathrm{i}p\partial_\psi^2 \hat{v}_t + q\partial_\psi \hat{v}_t - \frac{1}{2}\psi^2 \hat{v}_t.$$

We look for a solution of the form $\exp(\frac{1}{2}a\psi^2 + b\psi + c)$, for (genuine) functions $a$, $b$ and $c$ of $(t, p, q)$. Plugging it in the above equation, and arranging it according to the powers of $\psi$, we get

$$\frac{1}{2}\partial_t a\psi^2 + \partial_t b\psi + \partial_t c = -\frac{1}{2}(\mathrm{i}pa^2 + 1)\psi^2 + (-\mathrm{i}pab + qa)\psi - \frac{1}{2}\mathrm{i}p(a + b^2) + qb,$$

so we are looking for $a$, $b$ and $c$ satisfying

$$\partial_t a = -\mathrm{i}pa^2 - 1, \tag{1.15}$$

$$\partial_t b = -\mathrm{i}pab + qa, \tag{1.16}$$

$$\partial_t c = -\frac{1}{2}\mathrm{i}p(a + b^2) + qb, \tag{1.17}$$

with vanishing initial conditions: $a_0 = b_0 = c_0 \equiv 0$ as functions of $P$.

In order to get convenient notations to express the solutions of said equations, define the following *ad hoc* trigonometric quantities, both of which being obviously holomorphic over $\mathbb{C}$.

$$\mathrm{cah} : \zeta \mapsto \sum_{k\geq 0} \frac{\zeta^k}{(2k)!} \qquad \mathrm{sah} : \zeta \mapsto \sum_{k\geq 0} \frac{\zeta^k}{(2k+1)!}$$

They are the analytic continuations of, respectively, $\cosh(\sqrt{\zeta})$ and $\sinh(\sqrt{\zeta})/\sqrt{\zeta}$, say on $\mathbb{R}_+^*$. It is clear that

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} \mathrm{cah}(\omega\alpha^2) = \omega\alpha\,\mathrm{sah}(\omega\alpha^2), \qquad \frac{\mathrm{d}}{\mathrm{d}\alpha} \alpha\,\mathrm{sah}(\omega\alpha^2) = \mathrm{cah}(\omega\alpha^2).$$

An important property is that both functions have no zero on $\{\Re\zeta > -\pi^2/4\}$, so $\zeta \mapsto \ln(\mathrm{cah}(\zeta))$ has a unique continuation on that open set, and similarly for, e.g., $\zeta \mapsto \sqrt{\mathrm{sah}(\zeta)}$. Indeed,

$$\mathrm{cah}(\alpha^2) = 0 \Leftrightarrow \cosh(\alpha) = 0 \Leftrightarrow \alpha \in \mathrm{i}\pi\Big(\mathbb{Z} + \frac{1}{2}\Big),$$

$$\sin(\alpha^2) = 0 \Leftrightarrow \sinh(\alpha) = 0 \text{ and } \alpha \neq 0 \Leftrightarrow \alpha \in \mathrm{i}\pi\mathbb{Z}^*.$$

Hence, $\zeta$ is a zero of either cah or sah if and only if it can be written as $-k^2\pi^2/4$, with $k \in \mathbb{Z}^*$.

One can check that

$$a_t := -\frac{t\,\mathrm{sah}(-\mathrm{i}pt^2)}{\mathrm{cah}(-\mathrm{i}pt^2)}$$

is the solution of equation (1.15). Then $b_t$ must be

$$\begin{aligned} b_t &= \int_0^t q a_s \exp\Big(-\int_s^t \mathrm{i}pa_u \mathrm{d}u\Big)\mathrm{d}s \\ &= q\int_0^t a_s \exp\Big(-[\ln \mathrm{cah}(-\mathrm{i}pu^2)]_s^t\Big)\mathrm{d}s \\ &= -q\int_0^t \frac{s\,\mathrm{sah}(-\mathrm{i}ps^2)}{\mathrm{cah}(-\mathrm{i}ps^2)} \cdot \frac{\mathrm{cah}(-\mathrm{i}ps^2)}{\mathrm{cah}(-\mathrm{i}pt^2)}\mathrm{d}s \\ &= \frac{q}{\mathrm{cah}(-\mathrm{i}pt^2)}\Big[\frac{1}{\mathrm{i}p}\mathrm{cah}(-\mathrm{i}ps^2)\Big]_0^t \\ &= \frac{q}{\mathrm{i}p}\Big(1 - \frac{1}{\mathrm{cah}(-\mathrm{i}pt^2)}\Big). \end{aligned}$$

The last coefficient $c_t$ is but the integral

$$\begin{aligned} c_t &= -\frac{1}{2}\mathrm{i}p\int_0^t (a_s + b_s^2)\mathrm{d}s + q\int_0^t b_s \mathrm{d}s \\ &= -\frac{1}{2}[\ln \mathrm{cah}(-\mathrm{i}ps^2)]_0^t + \frac{q^2}{2\mathrm{i}p}\int_0^t \Big(1 - \frac{1}{\mathrm{cah}(-\mathrm{i}ps^2)^2}\Big)\mathrm{d}s \\ &= -\frac{1}{2}\ln \mathrm{cah}(-\mathrm{i}pt^2) + \frac{q^2}{2\mathrm{i}p}\Big[s - \frac{s\,\mathrm{sah}(-\mathrm{i}ps^2)}{\mathrm{cah}(-\mathrm{i}ps^2)}\Big]_0^t \\ &= -\frac{1}{2}\ln \mathrm{cah}(-\mathrm{i}pt^2) + \frac{q^2 t}{2\mathrm{i}p}\Big(1 - \frac{\mathrm{sah}(-\mathrm{i}pt^2)}{\mathrm{cah}(-\mathrm{i}pt^2)}\Big). \end{aligned}$$

We have used in the last equality the identity

$$\mathrm{cah}(\alpha^2)^2 - \alpha^2\,\mathrm{sah}(\alpha^2)^2 = 1. \tag{1.18}$$

In conclusion, we found

$$\hat{v}_t = \frac{1}{\sqrt{\mathrm{cah}}} \exp\Big(\frac{t}{2\mathrm{i}p}\Big(1-\frac{\mathrm{sah}}{\mathrm{cah}}\Big)q^2 + \frac{1}{\mathrm{i}p}\Big(1-\frac{1}{\mathrm{cah}}\Big)q\psi + \frac{-t\,\mathrm{sah}}{2\,\mathrm{cah}}\psi^2\Big),$$

where the argument of cah and sah is $-\mathrm{i}pt^2$.

**Contour integral.** It is clear that $\hat{v}_t$ depends on $(q,\psi)$ in a Gaussian manner. Setting the temporary variables

$$\lambda_0 := \frac{t\,\mathrm{sah}(-\mathrm{i}pt^2)}{\mathrm{cah}(-\mathrm{i}pt^2)}, \qquad A_0 := \frac{1}{\mathrm{i}p}\Big(\frac{1}{\mathrm{cah}(-\mathrm{i}pt^2)}-1\Big), \qquad \mu_0 := \frac{t}{\mathrm{i}p}\Big(\frac{\mathrm{sah}(-\mathrm{i}pt^2)}{\mathrm{cah}(-\mathrm{i}pt^2)}-1\Big),$$

the Fourier transform $\hat{v}_t$ rewrites as

$$\hat{v}_t(P) = \frac{1}{\sqrt{\mathrm{cah}(-\mathrm{i}pt^2)}}\exp\Big(-\frac{1}{2}(\lambda_0\psi^2 + 2A_0 q\psi + \mu_0 q^2)\Big).$$

Whenever the determinant

$$D_0 := \det\begin{pmatrix}\lambda_0 & A_0\\ A_0 & \mu_0\end{pmatrix} = \frac{1}{(\mathrm{i}p)^2\,\mathrm{cah}(-\mathrm{i}pt^2)}\big(-\mathrm{i}pt^2\,\mathrm{sah}(-\mathrm{i}pt^2) + 2 - 2\,\mathrm{cah}(-\mathrm{i}pt^2)\big)$$

does not vanish, we can take the inverse of the Fourier transform on $(q,\psi)$ to get

$$\begin{aligned} v_t(X) &= \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3}\hat{v}_t(P)\mathrm{e}^{\mathrm{i}X\cdot P}\mathrm{d}X\\ &= \frac{1}{4\pi^2}\int_{\mathbb{R}}\frac{1}{\sqrt{\mathrm{cah}(-\mathrm{i}pt^2)D_0}}\exp\Big(-\frac{1}{2}\big(\frac{\mu_0}{D_0}\phi^2 - 2\frac{A_0}{D_0}y\phi + \frac{\lambda_0}{D_0}y^2\big)\Big)\mathrm{e}^{\mathrm{i}px}\mathrm{d}p, \end{aligned}$$

using identity (1.18) again.

In fact, $D_0$ does not vanish on the values we are interested in. To see this, we need to state a 'doubling of the angle' formula; namely, we can check using the expression in terms of cosh and sinh and the unique analytic continuation that

$$\mathrm{cah}(\zeta) = 2\,\mathrm{cah}(\zeta/4)^2 - 1 = \frac{\zeta}{2}\,\mathrm{sah}(\zeta/4)^2 + 1 \qquad\text{and}\qquad \mathrm{sah}(\zeta) = \mathrm{cah}(\zeta/4)\,\mathrm{sah}(\zeta/4).$$

Then the numerator of $D_0$ rewrites as

$$-\mathrm{i}pt^2\,\mathrm{sah}(-\mathrm{i}pt^2/4)\big(\mathrm{cah}(-\mathrm{i}pt^2/4) - \mathrm{sah}(-\mathrm{i}pt^2/4)\big),$$

and in fact

$$D_0 = \frac{\mathrm{sah}(\zeta/4)}{\mathrm{cah}(\zeta)}\,\frac{\mathrm{cah}(\zeta/4)-\mathrm{sah}(\zeta/4)}{\zeta}t^4$$

with $\zeta = -\mathrm{i}pt^2$. We know that sah does not vanish on $\{\Re\zeta > -4\pi^2\}$. On the other hand, $\mathrm{cah}-\mathrm{sah}$ has a zero of order one at 0, but this singularity is compensated by the denominator. It is possible to get an expression for the other zeroes of $\mathrm{cah}-\mathrm{sah}$:

$$\mathrm{cah}(\alpha^2) - \mathrm{sah}(\alpha^2) = 0 \Leftrightarrow \alpha\cosh(\alpha) - \sinh(\alpha) = 0 \Leftrightarrow \alpha - \tanh(\alpha) = 0 \Leftrightarrow \alpha \in \{\mathrm{i}\theta, \tan\theta = \theta\}.$$

In other words, the zeroes of $D_0$ are the $-4\theta^2$, for $\theta$ the non-zero roots of $\tan\theta - \theta$. These roots are real, symmetrical with respect to zero, and the smallest positive one is close to 4.49, so $D_0$ does not vanish on $\{\Re\zeta > -4\pi^2\}$.

The change of variable $\zeta = -\mathrm{i}pt^2$ seems appropriate as this point. We get an expression of the integrand in terms of an amplitude $I_\zeta$ and a quadratic form $Q_\zeta$ independent of $t$ and $P$. Indeed,

$$v_t(X) = \frac{1}{4\mathrm{i}\pi^2 t^4}\int_{\mathrm{i}\mathbb{R}} I_\zeta \exp\Big(-\frac{1}{2}Q_\zeta\Big(\frac{y}{t\sqrt{t}}, \frac{\phi}{\sqrt{t}}\Big)\Big)\mathsf{e}^{-\zeta x/t^2}\mathrm{d}\zeta,$$

for

$$I_\zeta := \frac{1}{\sqrt{\operatorname{sah}(\zeta/4)}}\cdot\sqrt{\frac{\zeta}{\operatorname{cah}(\zeta/4)-\operatorname{sah}(\zeta/4)}}\,, \qquad Q_\zeta : (y,\phi) \mapsto \mu_\zeta\phi^2 + 2A_\zeta y\phi + \lambda_\zeta y^2,$$

$$\begin{aligned}
\mu_\zeta &:= \frac{1}{\operatorname{sah}(\zeta/4)}\frac{\operatorname{cah}(\zeta)-\operatorname{sah}(\zeta)}{\operatorname{cah}(\zeta/4)-\operatorname{sah}(\zeta/4)},\\
A_\zeta &:= -\frac{1}{2}\frac{\zeta\operatorname{sah}(\zeta/4)}{\operatorname{cah}(\zeta/4)-\operatorname{sah}(\zeta/4)},\\
\lambda_\zeta &:= \frac{\zeta\operatorname{cah}(\zeta/4)}{\operatorname{cah}(\zeta/4)-\operatorname{sah}(\zeta/4)}.
\end{aligned}$$

Set $\check{u} := v_1$. It is clear from the above expression that

$$v_t(X) = \check{u}\Big(\frac{x}{t^2}, \frac{t}{t\sqrt{t}}, \frac{\phi}{\sqrt{t}}\Big)$$

Regarding our initial $\widetilde{u}$, direct computations show that indeed,

$$\widetilde{u}_t\big(X_0; T_{t,X_0}(\check{X})\big) = \frac{1}{t^4}\check{u}(\check{X}), \qquad\qquad \widetilde{u}_t\big(S_{t,X_1}(\check{X}), X_1\big) = \frac{1}{t^4}\check{u}(\check{X}).$$

**Estimates.** It remains to show that $(\tau, X_0, \check{X}) \mapsto \check{u}(\check{X})$ is in $\check{\Psi}$, i.e. that it is smooth, and rapidly decaying in the $\check{X}$ variable. Since $\check{u}$ depends only on $\check{X}$, we only need to prove

$$\sup_{\check{X}} |\check{X}|^k \big|D^\alpha_{\check{X}}\check{u}(\check{X})\big| < \infty \tag{1.19}$$

for all $k \in \mathbb{N}$, $\alpha \in \mathbb{N}^3$. Based on the definition of $\check{u}$, we can hope to get the derivatives inside the integral, and to write

$$D^\alpha_{\check{X}}\check{u}(\check{X}) = \frac{1}{4\mathrm{i}\pi^2}\int_{\mathrm{i}\mathbb{R}} P^\alpha_\zeta(\check{y},\check{\phi}) I_\zeta \exp\Big(-\frac{1}{2}Q_\zeta(\check{y},\check{\phi})\Big)\mathsf{e}^{-\zeta\check{x}}\mathrm{d}\zeta,$$

where $P^\alpha_\zeta(\check{y},\check{\phi})$ is a polynomial in $\check{y}$, $\check{\phi}$ and the coefficients of $Q_\zeta$. In fact, it will be the case if we show that the above integrand is bounded by an integrable function, uniformly over the balls $|\check{X}| \le R$. The following relations introduce the bounds we will need to show these estimates.

**Proposition 1.16.** *As $|\zeta|$ tends to infinity under the constraint $\Re\zeta > -4\pi^2$, we have*

$$\mu_\zeta^2 \sim \zeta, \qquad A_\zeta^2 \sim \zeta, \qquad \lambda_\zeta \sim \zeta.$$

*Moreover, on a neighbourhood of the imaginary axis $[-\varepsilon,\varepsilon] + \mathrm{i}\mathbb{R}$, with $\varepsilon > 0$ small enough, we have*

$$|I_\zeta| \le C\exp(-\sqrt{|\zeta|}/C), \qquad \Re Q_\zeta(y,\phi) \ge \frac{y^2+\phi^2}{C}$$

*for any $(y,\phi) \in \mathbb{R}^2$ and some constant $C > 0$ large enough.*

*Proof.* There exists a constant $R > 0$ such that, under the constraint $\Re\zeta > -4\pi^2$ and $|\zeta| \geq R$, there is a unique $\alpha$ of positive real part such that $\alpha^2 = \zeta$. Moreover, such $\alpha$ have arguments in $(-\pi/4 - \varepsilon; \pi/4 + \varepsilon)$, where $\varepsilon > 0$ may be chosen arbitrarily small, provided we enlarge $R$. With this in mind, we easily prove that

$$|\operatorname{cah}(\zeta)|, |\operatorname{sah}(\zeta)| \geq \exp\big((\sqrt{2} - \varepsilon)|\zeta|\big)/C$$

for all $\varepsilon > 0$ and $\Re\zeta > -4\pi^2$, $|\zeta| \geq R$, provided $R$ and $C$ are chosen according to $\varepsilon$.

Using the 'sum of squares' formula

$$\operatorname{cah}(\zeta)^2 - \zeta \operatorname{sah}(\zeta)^2 = 1,$$

we see that $\operatorname{sah}(\zeta/4) = o\big(\operatorname{cah}(\zeta/4)\big)$, from whence the equivalents follow, using both the above identity and the 'doubling of the angle' relation

$$\operatorname{sah}(\zeta) = \operatorname{sah}(\zeta/4)\operatorname{cah}(\zeta/4).$$

The first inequality is a consequence of our lower bound for sah, and the fact that it does not vanish on $\Re\zeta > -4\pi^2$. Regarding the second, direct computations show that

$$Q_0 = \begin{pmatrix} 12 & -6 \\ -6 & 4 \end{pmatrix},$$

which is definite positive, hence $Q_\zeta$ is (real) definite positive for $\zeta \in [-\varepsilon; \varepsilon]$, provided $\varepsilon > 0$ is small enough. To conclude, it would suffice to show that $\Re Q_\zeta(y, \phi) \geq Q_{\Re\zeta}(y, \phi)$ for all $\Re\zeta \in [-\varepsilon; \varepsilon]$. Because $Q$ is holomorphic with respect to $\zeta$, we would want to show that $\Re(\mathrm{i}\partial_\zeta Q_\zeta)$ is definite positive for all $\Im\zeta > 0$, $\Re\zeta \in [-\varepsilon; \varepsilon]$. The fact that its opposite is definite positive for $\Im\zeta < 0$ would follow from the fact that $Q_\zeta$ is real on the real axis.

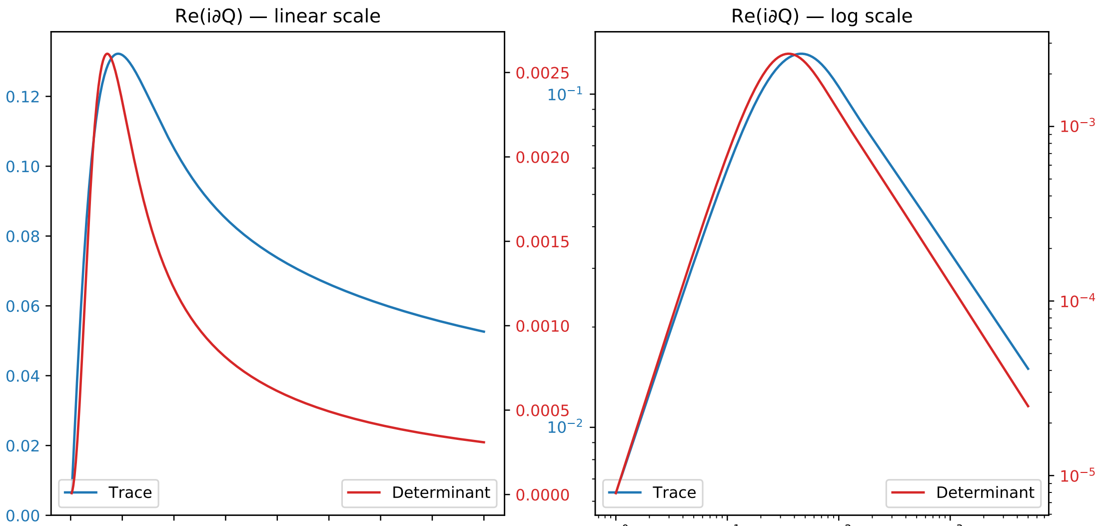



FIGURE 1.1 – Plot of $\Re(\mathrm{i}\partial_\zeta Q_\zeta)$, for $\zeta \in \mathrm{i}\mathbb{R}_+$.

In theory, it reduces to prove that the trace and determinant of $\Re(\mathrm{i}\partial_\zeta Q_\zeta)$ are positive, say for all $\Re\zeta > -4\pi^2$:

$$\Re(\mathrm{i}\partial_\zeta\lambda + \mathrm{i}\partial_\zeta\mu) > 0, \qquad \Re(\mathrm{i}\partial_\zeta\lambda)\Re(\mathrm{i}\partial_\zeta\mu) - \Re(\mathrm{i}\partial_\zeta A)^2 > 0.$$

In practice, these expressions are involved, and I have not been able to give a full proof yet. I do believe however that the numerical evidence is in favour of the result. The expressions for the coefficients of $Q$ are explicit, and there is no difficulty for a computer to evaluate them. The reader will find in Figure 1.1 the curves of the trace and determinant mentioned above. $\square$

Let $N$ be the total degree of $P^\alpha$. According to the proposition, we have

$$\begin{aligned}\Big|P^\alpha_\zeta(\check{y},\check{\phi})I_\zeta \exp\Big(-\frac{1}{2}Q_\zeta(\check{y},\check{\phi})\Big)\mathsf{e}^{-\zeta\check{x}}\Big| \\ \leq C\big(1+|\check{y}|+|\check{\psi}|+|\zeta|\big)^N \exp\big(-\sqrt{|\zeta|}/C\big)\exp\Big(-\frac{\check{y}^2+\check{\phi}^2}{C}-\Re\zeta\check{x}\Big)\end{aligned}$$

for a constant $C>0$ large enough, on a neighbourhood $[-\varepsilon,\varepsilon]+\mathrm{i}\mathbb{R}$ of the imaginary axis. This upper bound is integrable uniformly in $\check{X}$ along $\mathrm{i}\mathbb{R}$, so the formula for the derivative is indeed valid.

If $\check{x}\geq 0$, then $\check{u}$ is identically zero on open sets arbitrarily close to $\check{X}$, so the supremum in (1.19) can be considered over the set of $\check{X}$ such that $\check{x}<0$. In this case, the upper bound for the integrand shows that the integration contour can be moved to $-\varepsilon+\mathrm{i}\mathbb{R}$, which yields

$$|\check{X}|^k\big|D^\alpha_{\check{X}}\check{u}(\check{X})\big| \leq C(1+|\check{X}|)^{N+k}\exp\Big(-\frac{\check{y}^2+\check{\phi}^2}{C}-\varepsilon|\check{x}|\Big)\int_{\mathbb{R}}(1+|\zeta+\mathrm{i}\varepsilon|)^N\exp\big(-\sqrt{|\zeta+\mathrm{i}\varepsilon|}/C\big)\mathrm{d}\zeta$$

for some constant $C>0$ large enough. This last quantity is uniformly bounded in $\check{X}$, which completes the proof of $\check{u}\in\check{\Psi}$.

# 2 Variational approach

## 2.1 Introduction

Recall that we are interested in the equation

$$\begin{cases} \partial_t u &= -\cos(\phi)\partial_x u - \sin(\phi)\partial_y u, + \frac{1}{2}\partial_\phi^2 u \\ u_0 &= \delta_{X_0}. \end{cases} \tag{2.20}$$

In this section, we consider another approach, related to semiclassical analysis. An overview of these techniques is given in the book [Kol00] of V. Kolokoltsov. More precisely, we were inspired by the WKB method.

**The WKB method.** The first step towards these techniques is to rewrite our partial differential equation as a Schrödinger equation with imaginary phase, one might say. The rule of thumb for semiclassical analysis is that the solutions of the Schrödinger equation

$$\mathrm{i}h\partial_t u = H\big(X, -\mathrm{i}h\partial_X\big)u$$

should behave, for small $h$, like the solutions of the Hamilton equations

$$\dot{X} = \frac{\partial H}{\partial P}, \; \dot{P} = -\frac{\partial H}{\partial X}.$$

Here, $H : (\mathbb{R}^d)^2 \to \mathbb{R}$ is a function called the Hamiltonian, and the definition of $H(X, \mathrm{i}h\partial_X)$, called the quantization of $H$, is part of the problem. A standard example is given by

$$H(X,P) = P^2 + V(X), \qquad\qquad H(X, \mathrm{i}h\partial_X) = -h^2\Delta_X + V(X)$$

for some potential $V : \mathbb{R}^d \to \mathbb{R}$. To give precise meaning to this limit behaviour is difficult, because the correspondence between the finite dimensional classical space and the infinite dimensional quantum space is not clear. Heuristically, points $(X_0, P_0)$ in classical space correspond to a wave packet $u$ concentrated in an infinitesimal neighbourhood of $X_0$, with wave vector $P_0$; think $X \mapsto \exp(-(|X - X_0|^2/\varepsilon + P_0 \cdot (X - X_0))$. It would take us too far to try and draw this link, but the interested reader may find an introduction to semiclassical analysis in the classical book [Zwo12] of M. Zworski, as well as an instance of such a treatment of $h \to 0$ in the form of the Egorov Theorem, Theorem 11.1.

A fundamental tool in the study of these equations is the action. Associated to a suitable Hamiltonian $H$ is a Lagrangian $L : (\mathbb{R}^d)^2 \to \mathbb{R}$, defined by the Legendre transform:

$$L(X,V) := \sup_P \big(P \cdot V - H(X,P)\big).$$

A regular curve $\gamma$ with values in $\mathbb{R}^d$ can be assigned an action $I[\gamma]$, the time integral of $L(\gamma, \dot{\gamma})$. On the classical side, it turns out that the minimisers of this action, for fixed endpoints and duration, are likely (depending on the Hamiltonian) to be the space part of a solution $(\gamma, P^\gamma)$ of Hamilton equations. Such curves are called characteristics, and appear naturally in various situations in semiclassical and microlocal analysis as well. On the quantum side, if we define the so-called two-point function $S_t(X_0, X)$ as the infimum of the action $I[\gamma]$ over the curves $\gamma$ joining $X_0$ to $X$ in time $t$, the solutions $u$ of the Schödinger equation are likely to be well-approximated by the ansatz

$$X \mapsto A(X_0, X) \exp\Big(\frac{\mathrm{i}}{h} S_t(X_0, X)\Big),$$

with $A$ slowly varying in $X$. This is in essence the first observation of the WKB method.

Let us see what we can try in our case. Defining the Hamiltonian

$$H : (X; P) = (x, y, \phi; p, q, \psi) \mapsto \cos(\phi)p + \sin(\phi)q + \frac{\sigma^2}{2}\psi^2,$$

some natural quantization for $H$ gives the PDE (2.20):

$$\partial_t u = H\Big(X, -\frac{\partial}{\partial X}\Big)u_t.$$

In similar problems, the Hamiltonian can be seen to encode the heat kernel asymptotics in the following sense. There exists functions $A_0, A_1, \cdots$ such that

$$u(X_0, X) \simeq (A_0(X_0, X) + \cdots + A_n(X_0, X))\mathrm{e}^{-S_t(X_0,X)}.$$

Moreover, natural candidates for $S$ and $A_0$ are given in terms of the Hamiltonian flow. For $P_0$ small enough, the infimum $S_t(X_0, X_t(P_0))$, is attained by the action $I[\gamma]$ of the curve $\gamma : t \mapsto X_t(X_0, P_0)$, whereas $A_0$ is the Jacobian of $X_t(X_0, \cdot)^{-1}$ at $X$.

This holds for instance in the case of elliptic diffusions. However, in our case, we will show that $X_t(X_0, \cdot)$ is not a diffeomorphism over any neighbourhood of 0, so we need to find other methods to tackle this problem.

**Large deviations.** Another apparently unrelated method for the treatment of the densities of diffusions is the theory of large deviations. In our case, the study of the density is closely related to the control problem

$$\begin{aligned} \mathrm{d}z_t^h &= \mathrm{e}^{\mathrm{i}\phi}\mathrm{d}t \\ \mathrm{d}\phi_t^h &= \sigma^2\mathrm{d}h_t \end{aligned} \qquad \text{with} \qquad X_0^h = X_0,$$

where $h$ is a control with enough regularity. Remarkably, it turns out that a $\mathcal{C}^1$ curve $\gamma$ with values in $\mathbb{R}^3$ can be written as $X^h$ for some control $h$ if and only if the action $I[\gamma]$ defined above using variational techniques is finite. In this case, this action is actually equal to the energy

$$\frac{1}{2}\int_0^t |\dot{h}_s|^2\mathrm{d}s$$

of $h$. In other words, finding the curves with least action $\gamma$ from $X_0$ to $X$ in the variational problem is the same as finding the minimising controls $h$ for the energy such that $X^h$ joins $X_0$ to $X$.

An example of the usefulness of this control approach is the Strook-Varadhan support theorem. It states that all points $X_t^h$ that can be reach by such controls $h$ must lie in the support of $u_t(X_0, \cdot)$ — see the introduction for a proof that its support is $\{|z| \leq t\}$ using these methods. Other results are based on the identification of the minimal energy required to get from $X_0$ to $X$ by curves solving the control problem above. For instance, here is a toy result making use of minimisers.

Let $H$ be a normed vector space endowed with a family of probability measures $\mathbb{P}_t$, whose canonical random variable we call $h$. Think of $H$ as the space $\mathrm{H}^1$ with the Wiener probability law — of course this is not rigorous because the Brownian motion is not in that space almost surely, but one can work with for instance a finite dimensional subspace of piecewise linear functions, so that $h$ is an approximation of $s \mapsto W_{st}$. We assume that $\mathbb{P}_t$ satisfies a large deviation principle with rate $\frac{1}{2}|k|^2$ in the sense

$$- \inf_{k \in \mathring{A}} \frac{1}{2}|k|^2 \leq \liminf_{t\to 0} t \ln \mathbb{P}_t\big(h \in A\big) \leq \limsup_{t\to 0} t \ln \mathbb{P}_t\big(h \in A\big) \leq - \inf_{k \in \overline{A}} \frac{1}{2}|k|^2$$

for any Borel set $A \subset H$.

Let $\Phi : H \to \mathbb{R}^d$ be a continuous function. For instance, $\Phi(h)$ is the value at time 1 of a differential equation controlled by $h$; again, this example is unrealistic, since such functions are not continuous in most cases. We want to get a large deviation result on $\Phi(h)$. Suppose that for all $x \in \mathbb{R}^d$, we have found minimisers $h^x$ of the norm $|h|$ amongst elements such that $\Phi_t(h) = x$, unique in a sense that we define later. We assume that $h^x$ does not depend on $t$, which is unattainable in most cases. Then $\Phi(h)$ satisfies the large deviation principle

$$- \inf_{y \in \mathring{A}} \frac{1}{2}|h^y|^2 \leq \liminf_{t\to 0} t \ln \mathbb{P}_t\big(\Phi(h) \in A\big) \leq \limsup_{t\to 0} t \ln \mathbb{P}_t\big(\Phi(h) \in A\big) \leq - \inf_{k \in \overline{A}} \frac{1}{2}|h^y|^2.$$

The first inequality is fairly simple. Given some point $x \in \mathring{A}$ such that $|h^x| < \inf_{y\in\mathring{A}} |h^x| + \delta$, we find a small ball $B(h^x, \varepsilon)$ whose elements $k$ satisfy $\Phi(k) \in A$, and

$$\mathbb{P}_t(\Phi(h) \in A) \geq \mathbb{P}_t(h \in B(h^x, \varepsilon)) \geq \exp\Big( - \Big( \inf_{k \in B(h^x,\varepsilon)} \frac{1}{2}|k|^2 + \eta\Big)/t\Big)$$

for any $\eta$ provided $t$ is small enough, up to maybe choosing some smaller $\delta$ and $\varepsilon$.

The second relies on strong assumptions on the minimisers. We assume that for all $\delta > 0$ fixed, we may find some $\varepsilon > 0$ such that any curve $\varepsilon$-close to be a minimiser for the problem $\Phi(h) = x$ is in fact $\delta$-close to $h^x$. In mathematical terms, if $|h| < |h^x| + \delta$ with $\Phi(h) = x$, then $|h - h^x| < \varepsilon$. If this holds, we split the probability in two parts to get

$$\mathbb{P}_t(\Phi(h) \in A) \leq \mathbb{P}_t\big(|h| < \inf_{k\in\overline{A}} |k| + \varepsilon \text{ and } \Phi(h) \in A\big) + \mathbb{P}_t\big(|h| \geq \inf_{k\in\overline{A}} |k|\big).$$

The second term is controlled using the large deviation principle on $h$. Regarding the first, we note that the event is included in $\{|h - h^{\Phi(h)}| < \delta\} \subset \{|h| > \inf_{k\in\overline{A}} |k| - \delta\}$. Using again the large deviation principle on $h$, we conclude.

There are many point in that naive approach that leave no hope for kinetic Brownian motion. However, it is not impossible that more sophisticated tools may bring interesting results, provided we have a sufficient understanding of the behaviour of the minimisers. The type of results I would hope for is that

$$\ln \mathbb{P}(X_t \in A) \underset{t\to 0}{\sim} \inf_{X\in A} d_t(X_0, X)^2,$$

for some notion of distance $d_t$ depending on time. Such a notion of distance may appear more or less explicitly in the study of the minimisers, or be linked to the normalisations used in the last section 1.

**Main result.** Unfortunately, the results presented here do not begin to touch the probabilistic questions described above. Nevertheless, given the achievements of the study of the characteristics in similar problems, I believe that the results of this work are an interesting step towards a better understanding of these equations. Namely, this work proves the following.

**Theorem 2.1.** *Let $\gamma : [0;t] \to \mathbb{R}^3$ be a global minimum for the action. If $\gamma$ is $\mathcal{C}^2$, then $\gamma$ is a characteristic for $H$.*

## 2.2 Calculus of variations

A quick review of the calculus of variations can be found in section 2 of [Kol00]. Here, we only need elementary results and the section stays self-contained.

**Lagrangian.** Recall that we consider the Hamiltonian

$$H : (X; P) = (x, y, \phi; p, q, \psi) \mapsto \cos(\phi)p + \sin(\phi)q + \frac{\sigma^2}{2}\psi^2.$$

The associated Lagrangian $L(X, V)$ is defined by the Legendre transform:

$$L(x, y, \phi; v, w, \nu) := \sup_P \big(P \cdot V - H(X, P)\big) = \sup_P \Big((v - \cos\phi)p + (w - \sin\phi)q + \nu\psi - \frac{\sigma^2}{2}\psi^2\Big).$$

If $\cos\phi \neq v$ or $\sin\phi \neq w$, the Lagrangian is infinite. Otherwise, we maximise the quadratic form; all in all,

$$L(X, V) = \begin{cases} \frac{\nu^2}{2\sigma^2} & \text{if } v = \cos(\phi) \text{ and } w = \sin(\phi), \\ \infty & \text{else.} \end{cases}$$

Of course, we have for any $X, V, P \in \mathbb{R}^3$

$$L(X, V) \geq PV - H(X, P).$$

The equality holds if and only if $\cos\phi = v$, $\sin\phi = w$ and $\nu = \psi\sigma^2$. The Hamiltonian is also the Legendre transform of the Lagrangian:

$$H(X, P) = \sup_V \big(P \cdot V - L(X, V)\big) = \sup_\nu \Big(\cos(\phi)p + \sin(\phi)q + \nu\psi - \frac{\nu^2}{2\sigma^2}\Big),$$

and the equality in

$$H(X, P) \geq P \cdot V - L(X, V)$$

holds of course under the same conditions.

**Characteristics.** The Hamilton equations associated to some Hamiltonian $H$ are

$$\dot{X} = \frac{\partial H}{\partial P}(X, P) \qquad\qquad \dot{P} = -\frac{\partial H}{\partial X}(X, P).$$

A curve $\gamma : [0; t] \to \mathbb{R}^3$ is called a characteristic if there exists some $P^\gamma : [0; t] \to \mathbb{R}^3$ such that $(\gamma, P^\gamma)$ is a solution of the above equation. Denoting $t \mapsto (X_t(X_0, P_0), P_t(X_0, P_0))$ the solution with initial condition $(X_0, P_0)$, it is equivalent to say that $\gamma$ is a characteristic if there exists $P_0$ such $\gamma_t = X_t(P_0)$. In the case of the kinetic Brownian motion, our expression for $H$ unveils

$$\begin{aligned} \dot{x}_t &= \cos(\phi_t) & \dot{p}_t &= 0 \\ \dot{y}_t &= \sin(\phi_t) & \dot{q}_t &= 0 \\ \dot{\phi}_t &= \sigma^2\psi_t & \dot{\psi}_t &= \sin(\phi_t)p_t - \cos(\phi_t)q_t. \end{aligned}$$

The initial conditions can always be reduced to $(0, 0, 0, p_0', q_0', \psi_0')$ using isometries of $\mathbb{R}^2$. Indeed, for any point $X_0$, the transformation

$$T : (x + \mathrm{i}y, \phi; p + \mathrm{i}q, \psi) \mapsto \big((x_0 + \mathrm{i}y_0) + \mathrm{e}^{\mathrm{i}\phi_0}(x + \mathrm{i}y); \mathrm{e}^{\mathrm{i}\phi_0}(p + \mathrm{i}q), \psi\big)$$

sends solutions from $(0, P_0)$ to solutions from $(X_0, P_0')$:

$$T\big((X_t, P_t)(0, P_0)\big) = (X_t, P_t)\big(T(0, P_0)\big).$$

When convenient, in particular in sections 2.3 and 2.5, we stop writing the initial point, so that $S(X)$ denotes $S(0, X)$, $x_t(P_0)$ is $x_t(0, P_0)$, and so on. In this situation, we can reduce the above system to the following one.

$$\begin{aligned}
x_t &= \int_0^t \cos(\phi_\tau)\mathrm{d}\tau & p_t &= p_0 \\
y_t &= \int_0^t \sin(\phi_\tau)\mathrm{d}\tau & q_t &= q_0 \\
\ddot{\phi}_t &= -\sigma^2 \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} \cdot \begin{pmatrix} -\sin\phi_t \\ \cos\phi_t \end{pmatrix} & \psi_t &= \frac{\dot{\phi}_t}{\sigma^2}.
\end{aligned}$$

We see that understanding the behaviour of the system essentially reduces to the study of $\phi$. Note also that the flow does not explode for any $t > 0$, since it is well-defined whenever $\phi$ is, and $\ddot{\phi}$ is bounded by a constant for bounded $\sigma$, $p_0$ and $q_0$.

**Proposition 2.2.** *Let $c : [0; t] \to \mathbb{R}^3$ be a path with values in the image of $X$, in the sense that there exists a path $\zeta : [0; t] \to \mathbb{R}^3$ of initial momenta such that $c_\tau = X_\tau(c_0, \zeta_\tau)$. Then the characteristic $\gamma : \tau \mapsto X_\tau(c_0, \zeta_t)$ has the same endpoints, and smaller action:*

$$I[\gamma] \leq I[c].$$

*The equality holds if and only if $DX_\tau(\zeta_\tau)\dot{\zeta}_\tau = 0$ for all $\tau \leq t$ — in particular, if $DX$ is invertible along $\zeta$, then the equality holds if and only $\gamma$ is a characteristic.*

In view of the above result, the way to Theorem 2.1 should be to prove that local minimisers have values in the image of the flow $X$, preferably where $DX$ is not singular. In the following section, we describe domains of $\mathbb{R}^3$ attained by $X$, where $DX$ is invertible, and in Lemma 2.7, we show that most $\mathcal{C}^2$ curve have values in such domains. We then conclude with Theorem 2.6.

*Proof.* It is clear that $\gamma$ has the same endpoints. We suppose $I[c] < \infty$; otherwise, there is nothing to prove. From the definition of $L$, it is clear that

$$\begin{aligned}
I[c] &= \int_0^t L(c_\tau, \dot{c}_\tau)\mathrm{d}\tau \\
&\geq \int_0^t \Big(P_\tau^*(\zeta_\tau)\dot{c}_\tau - H\big(c_\tau, P_\tau(\zeta_\tau)\big)\Big)\mathrm{d}\tau \\
&= \int_0^t \Big(P_\tau^*(\zeta_\tau)DX_\tau(\zeta_\tau)\mathrm{d}\zeta_\tau + \big(P_\tau^*(\zeta_\tau)(\partial_\tau X)_\tau(\zeta_\tau) - H(X_\tau(\zeta_\tau), P_\tau(\zeta_\tau))\big)\mathrm{d}\tau\Big) \\
&=: \mathcal{I},
\end{aligned}$$

with equality if and only if for all $\tau \leq t$,

$$\dot{c}_\tau = \frac{\partial H}{\partial P}\big(c_\tau, P_\tau(\zeta_\tau)\big).$$

We will show that the integral $\mathcal{I}$ is in fact $I[\gamma]$. Assuming this fact, and since $c = X \circ \zeta$, the equality indeed criterion rewrites as

$$DX_\tau(\zeta_\tau)\dot{\zeta}_\tau + (\partial_\tau X)_\tau(\zeta_\tau) = \frac{\partial H}{\partial P}\big(X_\tau(\zeta_\tau), P_\tau(\zeta_\tau)\big) = (\partial_\tau X)_\tau(\zeta_\tau).$$

We prove the fact stated above. Define on an open set containing $\mathbb{R}_+ \times \mathbb{R}^3$ the differential 1-form
$$\lambda_{\tau,\zeta} = P_\tau^*(\zeta) DX_\tau(\zeta)\mathrm{d}\zeta + \big(P_\tau^*(\zeta)(\partial_\tau X)_\tau(\zeta) - H(X_\tau(\zeta), P_\tau(\zeta))\big)\mathrm{d}\tau.$$
Note that its restriction to the submanifold $\{\tau = 0\}$ is zero since $X_0(\zeta) = 0$. Then the $\mathcal{I}$ is but the integral of $\lambda$ over $\tau \mapsto (\tau, \zeta_\tau)$. Denoting by $(\mathcal{T}, \zeta)$ this curve, $\mathcal{I} = \int_{\mathcal{T},\zeta} \lambda$.

Suppose for a second that we have shown $\mathrm{d}\lambda = 0$. Then according to Stokes formula,
$$\int_{\mathcal{T},\zeta} \lambda = \int_{\mathcal{T},P_0} \lambda + \int_{0,\zeta^0} \lambda$$
for some curve $\zeta^0$ with values in $\mathbb{R}^3$. As discussed above, this last integral is zero, so we have
$$\begin{aligned}\mathcal{I} = \int_{\mathcal{T},P_0} \lambda &= \int_0^t \big(P_\tau^*(P_0)(\partial_\tau X)_\tau(P_0) - H(X_\tau(P_0), P_\tau(P_0))\big)\mathrm{d}\tau \\ &= \int_0^t \Big(\frac{\partial H}{\partial P}\big(X_\tau(P_0), P_\tau(P_0)\big)P_\tau(P_0) - H\big(X_\tau(P_0), P_\tau(P_0)\big)\Big)\mathrm{d}\tau = I[\gamma].\end{aligned}$$
This will conclude the proof, provided we show that $\lambda$ is indeed closed.

The differential of $\lambda$, using Einstein's summation conventions and omitting the $\zeta$ and $\tau$ arguments, is
$$\mathrm{d}\lambda = \partial_j(P^*\partial_i X)\mathrm{d}\zeta^j \wedge \mathrm{d}\zeta^i + \partial_\tau(P^*\partial_i X)\mathrm{d}\tau \wedge \mathrm{d}\zeta^i + \partial_i\big(P^*\partial_\tau X - H(X,P)\big)\mathrm{d}\zeta^i \wedge \mathrm{d}\tau.$$
It is clear that, since $X_0(\zeta) = 0$, $\partial_j(P^*\partial_i X) = 0$ at $\tau = 0$. We show that its time derivative is symmetric.
$$\partial_\tau(P^*\partial_j X) = -\frac{\partial H}{\partial X}(X,P)\partial_j X + \partial_j\Big(\frac{\partial H}{\partial P}(X,P)\Big)P,$$
hence, there exists a symmetric quantity $S_{ij} = S_{ji}$
$$\begin{aligned}\partial_\tau\partial_i(P^*\partial_j X) = &-(\partial_j X)^*\frac{\partial^2 H}{\partial X^2}(X,P)\partial_i X - (\partial_j P)^*\frac{\partial^2 H}{\partial P\partial X}(X,P)\partial_i X \\ &+ \partial_i\partial_j\Big(\frac{\partial H}{\partial P}(X,P)\Big)P + \partial_j\Big(\frac{\partial H}{\partial P}(X,P)\Big)\partial_i P \\ = &-(\partial_j P)^*\frac{\partial^2 H}{\partial P\partial X}(X,P)\partial_i X + (\partial_j X)^*\frac{\partial^2 H}{\partial X\partial P}(X,P)\partial_i P \\ &+ (\partial_j P)^*\frac{\partial^2 H}{\partial P^2}(X,P)\partial_i P + S_{ij}\end{aligned}$$
and $\partial_i(P^*\partial_j X)$ is symmetric.

It remains to show that
$$\partial_\tau(P^*\partial_i X) = \partial_i\big(P^*\partial_\tau X - H(X,P)\big)$$
for all $i$. The left hand side is
$$-\frac{\partial H}{\partial X}(X,P)\partial_i X + \partial_i\Big(\frac{\partial H}{\partial P}(X,P)\Big)P,$$
while the right hand one is
$$\frac{\partial H}{\partial P}(X,P)\partial_i P + \partial_i\Big(\frac{\partial H}{\partial P}(X,P)\Big)P - \partial_i\big(H(X,P)\big) = \partial_i\Big(\frac{\partial H}{\partial P}(X,P)\Big)P - \frac{\partial H}{\partial X}(X,P)\partial_i X,$$
which concludes. $\square$

## 2.3 Image of the flow

In this section, we give explicit neighbourhoods $U_t$ of 0 such that $X_t(X_0, \cdot)$ is a diffeomorphism on $U_t$, and describe its image.

One can see directly from the equations that $(t, 0, 0, p_0, 0, 0)$ is solution of the Hamiltonian system; in particular, $X_t$ will not be one to one until restricted to a domain containing at most one point of the form $(p_0, 0, 0)$. Then we set $0 < \varepsilon \leq 1$, and define

$$U_t = \left(-\frac{\varepsilon^2}{t^2}, \frac{\varepsilon^2}{t^2}\right) \times \left(-\frac{\varepsilon^2}{t^2}, \frac{\varepsilon^2}{t^2}\right) \times \left(-\frac{\varepsilon}{t}, \frac{\varepsilon}{t}\right) \cap \{(q_0, \psi_0) \neq (0,0)\}\,.$$

**Proposition 2.3.** *There exist $\varepsilon > 0$ sufficiently small such that $X_t : U_t \to \mathbb{R}^3$ is a diffeomorphism onto its image $V_t$ for all $0 < t \leq 1$, and*

$$\Psi_{t,p_0} : \left(-\frac{\varepsilon^2}{t^2}, \frac{\varepsilon^2}{t^2}\right) \times \left(-\frac{\varepsilon}{t}, \frac{\varepsilon}{t}\right) \to \mathbb{R}^3$$
$$q_0, \psi_0 \mapsto (y_t(P_0), \psi_y(P_0))$$

*is a diffeomorphism for all $|p_0| < \varepsilon^2/t^2$.*

The proof of this lemma will rely on the following technical result, which deals with the above considerations regarding the singularity along $\{(q_0, \psi_0) = 0\}$. We postpone its (tedious but straightforward) proof to part 2.5.

**Lemma 2.4.** *Suppose that $0 < t \leq 1$ and $0 < \varepsilon \leq 1$. Define $d = |(tq_0, \psi_0)|$.*

*There exists universal $2 \times 2$ matrices $Q$, $A$, $B$, $M$ and a constant $C > 0$ depending only on $\sigma$ such that for any $|p_0| \leq \varepsilon^2/t^2$, $|q_0| \leq \varepsilon^2/t^2$, $|\psi_0| \leq \varepsilon/t$, we have*

$$DX_t = \sigma^2 t \begin{pmatrix} t^2\sigma^2 & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 1 \end{pmatrix} (T_t + R_t) \begin{pmatrix} t^2\sigma^2 & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

*with*

$$T_t(p_0, q_0, \psi_0) = \begin{pmatrix} \begin{pmatrix} tq_0 & \psi_0 \end{pmatrix} Q \begin{pmatrix} tq_0 \\ \psi_0 \end{pmatrix} & \begin{pmatrix} tq_0 & \psi_0 \end{pmatrix} A \\ \\ B \begin{pmatrix} tq_0 \\ \psi_0 \end{pmatrix} & M \end{pmatrix}$$

*and*

$$|R_t(p_0, q_0, \psi_0)| \leq C\varepsilon^2 \begin{pmatrix} d^2 & d & d \\ d & 1 & 1 \\ d & 1 & 1 \end{pmatrix},$$

*where the bound in the last statement is thought coefficient-wise. Moreover, $M$ is invertible and the symmetric part of $Q - AM^{-1}B$ is definite positive.*

*Proof of Proposition 2.3.* For a fixed $P_0$, let us set $Q_0 = (tq_0, \psi_0)$ and $d = |Q_0|$, and decompose $T_t + R_t$ as

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

where $\alpha$ (resp. $\beta$, $\gamma$, $\delta$) is $1 \times 1$ (resp. $1 \times 2$, $2 \times 1$, $2 \times 2$) and are close to $Q_0^* Q Q_0$ (resp. $Q_0^* A$, $BQ_0$, $M$) when $\varepsilon$ is small.

To prove the proposition, let us show first that $\det DX_t$ does not vanish on $U_t$, so that $X_t$ is a local diffeomorphism, and then that it is one-to-one on the same region.

**Local diffeomorphism.** According to the technical Lemma 2.4, $DX_t$ is invertible if and only if $T_t + R_t$ is. Let us note that if $\delta$ is invertible, then

$$\begin{pmatrix} 1 & -\beta\delta^{-1} \\ 0 & I_2 \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \delta^{-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\gamma & I_2 \end{pmatrix} = \begin{pmatrix} \alpha - \beta\delta^{-1}\gamma & 0 \\ 0 & I_2 \end{pmatrix}.$$

Because $\delta = M + O(\varepsilon^2)$, there exists some constant $C > 0$ depending only on $\sigma$ such that $\|\delta M^{-1} - I\| \leq C\varepsilon^2$. Whenever $C\varepsilon^2 \leq 1/2$, we thus have

$$\|\delta^{-1} - M^{-1}\| \leq \|M^{-1}\| \cdot \sum_{k\geq 1} (C\varepsilon^2)^k \leq 2\|M^{-1}\|C\varepsilon^2$$

Using the above formula, and taking the determinant, we get

$$\det(T_t + R_t) = \det(M + O(\varepsilon^2)) \times \Big(Q_0^* \left(Q - AM^{-1}B + O(\varepsilon^2)\right) Q_0 + O(\varepsilon^2 d^2)\Big),$$

which is non zero provided $\varepsilon$ is small enough and $d$ is positive, since $M$ is invertible and $Q - AM^{-1}B$ has definite positive symmetric part.

Note that if $p_0$ is fixed, $D\Psi_{t,p_0}(q_0, \psi_0)$ is invertible if and only if $\delta$ is. In particular, the $\varepsilon$ we have chosen is small enough to ensure that $D\Psi_{t,p_0}$ is a local diffeomorphism.

**Global injectivity.** Suppose we have two sets $P_0$ and $P_0'$ of parameters in $U_t$ such that $X_t(P_0) = X_t(P_0')$, and define $d$ and $d'$ as in Lemma 2.4. If we wanted to prove that $X_t$ is one to one on $\mathbb{R}^3$, then we would have to prove that $P_0 = P_0'$. Here, because of the singularity, there is no hope of such a result; we will instead follow a two parts reasoning. First, $Q_0 = (tq_0, \psi_0)$ cannot be too different from $Q_0'$; then, if $Q_0$ and $Q_0'$ are close, they must be close to 0, so the parameters will eventually lie outside $U_t$.

To make this reasoning explicit, define for $P_0$ and $P_0'$ the quantities $D = \max(d, d')$ and $\delta = |Q_0' - Q_0|$. Suppose that $d, d' > 0$. We discuss two cases, according to the values of the ratio $0 < D/\delta \leq \infty$.

1. For any $r > 0$, there exists $\varepsilon_r > 0$ such that for any $0 < \varepsilon \leq \varepsilon_r$, $0 < t \leq 1$ and $P_0$, $P_0'$ in the corresponding $U_t$ satisfying $D/\delta \leq r$, we have $(y_t, \phi_t)(P_0) \neq (y_t, \phi_t)(P_0')$. This $\varepsilon_r$ depends only on $\sigma$ and $r$.

2. There exists $r_0 > 0$ and $\varepsilon_0 > 0$ such that for any $0 < \varepsilon \leq \varepsilon_0$, $0 < t \leq 1$ and $P_0 \neq P_0'$ in the corresponding $U_t$ satisfying $D/\delta \geq r_0$, we have $X_t(P_0) \neq X_t(P_0')$. The constants $r_0$ and $\varepsilon_0$ depend only on $\sigma$.

The flow $X_t$ is then one-to-one on $U_t$ whenever $\varepsilon \leq \varepsilon_0, \varepsilon_{2r_0}$.

Fix $r > 0$, and let us consider the first point. Choose some $P_0$ and $P_0'$ in $\mathbb{R} \times (\mathbb{R}^2 \setminus \{(0,0)\})$ and $0 < t \leq 1$, and suppose that $(y_t, \phi_t)(P_0) = (y_t, \phi_t)(P_0')$ and $D/\delta \leq r$; in particular, $\delta > 0$. According to the technical lemma, and using notations described below, we get

$$\begin{aligned} 0 = \frac{1}{\sigma^2 t^2} \begin{pmatrix} y_t(P_0') - y_t(P_0) \\ t\phi_t(P_0') - t\phi_t(P_0) \end{pmatrix} &= \int_0^1 \left(T_t^b + R_t^b\right)\left(P_0 + s(P_0' - P_0)\right) \begin{pmatrix} \sigma^2 t^2 p_0' - \sigma^2 t^2 p_0 \\ tq_0' - tq_0 \\ \psi_0' - \psi_0 \end{pmatrix} \mathrm{d}s \\ &= \sigma^2 \int_0^1 \left(B\mathcal{Q}_s + R_t^B(\mathcal{P}_s)\right) t^2 (p_0' - p_0) \mathrm{d}s \\ &\quad + \int_0^1 (M + R_t^M(\mathcal{P}_s))(Q_0' - Q_0) \mathrm{d}s \end{aligned}$$

where $T_t^b$ (resp. $\tilde{R}_t^b$) is the bottom $2{\times}3$ submatrix of $T_t$ (resp. $R_t$), $R_t^M$ (resp. $\tilde{R}_t^B$) is the submatrix of $R_t$ whose corresponding submatrix in $T_t$ is M (resp. $BQ_0$), and $\mathcal{P}_s = P_0 + s(P_0' - P_0)$ (resp. $\mathcal{Q}_s = Q_0 + s(Q_0' - Q_0)$). In particular, if $P_0, P_0' \in U_t$ for some $\varepsilon > 0$, then

$$\begin{aligned}
|M(Q_0' - Q_0)| &\leq \sigma^2 \left( \|B\|D + \sup_{0\leq s\leq 1} |R_t^B(\mathcal{P}_s)| \right) \cdot 2\varepsilon^2 + \sup_{0\leq s\leq 1} \|R_t^M(\mathcal{P}_s)\| \cdot \delta \\
&\leq 2\sigma^2\|B\|\varepsilon^2 D + 4\sigma^2 C\varepsilon^4 D + \sqrt{2}C\varepsilon^2\delta \\
&\leq (2\sigma^2\|B\|r + 4\sigma^2 C\varepsilon^2 r + \sqrt{2}C)\varepsilon^2\delta
\end{aligned}$$

where $C > 0$ is the constant given by the technical lemma. On the other hand,

$$|M(Q_0' - Q_0)| \geq \|M^{-1}\|^{-1} \cdot |Q_0' - Q_0| \geq \|M^{-1}\|^{-1}\delta.$$

This implies

$$\|M^{-1}\|(2\sigma^2\|B\|r + 4\sigma^2 C\varepsilon^2 r + \sqrt{2}C) \cdot \varepsilon^2 \geq 1,$$

and we have proven the first point, for any $\varepsilon_r > 0$ not satisfying the above condition.

Note also that in the case $p_0' = p_0$, the upper bound becomes $\sqrt{C}\varepsilon^2\delta$, and is established without the hypothesis on $D/\delta$. In particular, $X_t(p_0, \cdot)$ is one-to-one for $\varepsilon < \varepsilon_r$ for any $r$.

Let us introduce a few constructions before getting into the details of the second point. Choose some $P_0$ and $P_0'$ in $\mathbb{R} \times (\mathbb{R}^2 \setminus \{(0,0)\})$ and $0 < t \leq 1$. Define

$$\mathcal{T} = \int_0^s T_t(P_0 + s(P_0' - P_0))\mathrm{d}s,$$

and $\mathcal{R}$ the same way. Using the technical lemma, we get

$$\begin{aligned}
&\frac{1}{\sigma^2 t}\begin{pmatrix} 1/\sigma^2 t^2 & 0 & 0 \\ 0 & 1/t & 0 \\ 0 & 0 & 1 \end{pmatrix} (X_t(P_0') - X_t(P_0)) \\
&\qquad = \mathcal{T}\begin{pmatrix} \sigma^2 t^2 & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 1 \end{pmatrix} (P_0 - P_0') + \mathcal{R}\begin{pmatrix} \sigma^2 t^2 & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 1 \end{pmatrix} (P_0 - P_0').
\end{aligned}$$

Define the norm $|\cdot|_\lambda$ on $\mathbb{R}^3$ by $|(a,b,c)|_\lambda = \max(\lambda|a|, |b|, |c|)$, and denote by $\|\cdot\|_{\lambda\to\mu}$ the operator norm from $(\mathbb{R}^3, |\cdot|_\lambda)$ to $(\mathbb{R}^3, |\cdot|_\mu)$. The proof of the second point will follow from the statement below.

*There exists constants $C' > 0$ and $r_0 > 0$ depending only on $\sigma$ such that for any $0 < \varepsilon \leq 1$, $0 < t \leq 1$ and $P_0$, $P_0'$ belonging to the associated $U_t$ and satisfying $D/\delta \geq r_0$, we have $\mathcal{T}$ invertible and*

$$\begin{aligned}
\|\mathcal{R}\|_{D\to 1/D} &\leq C'\varepsilon^2, \\
\|\mathcal{T}^{-1}\|_{1/D\to D} &\leq C'.
\end{aligned}$$

Let us assume the result for now, and conclude the proof. Fix some $0 < \varepsilon_0 < 1/C'$. Choose $\varepsilon$, $t$, $P_0$ and $P_0'$ satisfying the hypotheses of the second point, and suppose by contradiction that $X_t(P_0) = X_t(P_0')$. Then we have $0 = \mathcal{T}u + \mathcal{R}u$ for

$$u = \begin{pmatrix} \sigma^2 t^2 & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 1 \end{pmatrix} (P_0 - P_0') \neq 0.$$

Also,

$$|\mathcal{R}u|_{1/D} \leq \|\mathcal{R}\|_{D\to 1/D} \cdot |u|_D \leq \|\mathcal{R}\|_{D\to 1/D} \cdot \|\mathcal{T}^{-1}\|_{1/D\to D} \cdot |\mathcal{T}u|_{1/D} \leq C'^2\varepsilon^2 |\mathcal{T}u|_{1/D}.$$

Because $C'\varepsilon < 1$ and $|\mathcal{R}u| = |\mathcal{T}u|$, we must have $\mathcal{R}u = \mathcal{T}u = 0$. But this is impossible, because $u \neq 0$ and $\mathcal{T}$ is invertible; $X_t(P_0)$ must be different from $X_t(P_0')$.

Let us turn to the proof of the statement. We can see directly from the technical lemma that the first inequality holds with $C' = 3C$. The difficulty lies in the second inequality.

We define $\overline{Q}_0 = (t\overline{q}_0, \overline{\psi}_0) = (Q_0 + Q_0')/2$ and $\mathcal{Q}_s = Q_0 + s(Q_0' - Q_0)$, so that

$$\begin{aligned}
\mathcal{T} &= \begin{pmatrix} \int_0^1 \mathcal{Q}_s^* Q \mathcal{Q}_s \mathrm{d}s & \overline{Q}_0^* A \\ B\overline{Q}_0 & M \end{pmatrix} \\
&= \begin{pmatrix} 1 & \overline{Q}_0^* A M^{-1} \\ 0 & I_2 \end{pmatrix} \begin{pmatrix} \int_0^1 \mathcal{Q}_s^* Q \mathcal{Q}_s \mathrm{d}s - \overline{Q}_0^* A M^{-1} B \overline{Q}_0 & 0 \\ 0 & I_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ B\overline{Q}_0 & I_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & M \end{pmatrix} \\
&=: \mathcal{T}_A \mathcal{T}_Q \mathcal{T}_B \mathcal{T}_M
\end{aligned}$$

It is clear that

$$\|\mathcal{T}_A^{-1}\|_{1/D\to 1/D} \leq C_A \qquad \|\mathcal{T}_B^{-1}\|_{D\to D} \leq C_B \qquad \|\mathcal{T}_M^{-1}\|_{D\to D} \leq C_M$$

with $C_A, C_B, C_M > 0$ depending only on the coefficients of $A$, $B$ and $M$. Because the latter do not depend on anything, so do the constants. It remains to show that the top left coefficient of $\mathcal{T}_Q$ is not too small provided the ratio $r := D/\delta$ is big enough; it will be sufficient to show that it is bounded below by $D^2$ up to a constant factor. But this coefficient, call it $\alpha$, satisfies

$$\begin{aligned}
\alpha &= \int_0^1 (\mathcal{Q}_s - \overline{Q}_0)^* Q (\mathcal{Q}_s - \overline{Q}_0) \mathrm{d}s + \overline{Q}_0^* (Q - AM^{-1}B) \overline{Q}_0 \\
&\geq - C_Q \cdot \int_0^1 |\mathcal{Q}_s - \overline{Q}_0|^2 \mathrm{d}s + 1/C_Q \cdot |\overline{Q}_0|^2 \\
&\geq - C_Q \cdot \int_0^1 \frac{\delta^2}{4}(1-2s)^2 \mathrm{d}s + 1/C_Q \cdot (D - \delta/2)^2 \\
&= \left( -\frac{C_Q}{12r^2} + \frac{1}{C_Q}\left(1 - \frac{1}{2r}\right)^2 \right) D^2
\end{aligned}$$

for some constant $C_Q > 0$ big enough, because the symmetric part of $Q - AM^{-1}B$ is definite positive. Moreover, $C_Q$ depends only on the coefficients of $Q$ $A$, $B$ and $M$, i.e. on nothing. In particular, for some universal $r_0 > 0$, we have $\alpha \geq D^2/2C_Q$, $\mathcal{T}_Q$ invertible and

$$\|\mathcal{T}_Q^{-1}\|_{1/D\to D} \leq 2C_Q$$

for all $r \geq r_0$.

Finally, composition gives $\|\mathcal{T}^{-1}\|_{1/D\to D} \leq 2C_A C_Q C_B C_M$ as expected. □

**Proposition 2.5.** *Fix some $\sigma > 0$. For any $\varepsilon > 0$ small enough, there exists $c > 0$ such that the following holds. For any $0 < t \leq 1$, the image of $\Psi_{t,p_0}$ contains all points $(y, \phi)$ satisfying*

$$|y| \leq ct, \qquad |\phi| \leq c.$$

*If we choose $x$ such that moreover*

$$\left| x - x_t\big(0, \Psi_{t,0}^{-1}(y,\phi)\big) \right| < \frac{c}{t}(y^2 + t^2\phi^2),$$

*then $(x, y, \phi)$ is in the image $V_t = X_t(U_t)$.*

*Proof.* We will show that the image of $\overline{U_t}$ by $X_t$ surrounds, in some sense, any point satisfying the above conditions, and similarly for $\Psi_{t,p_0}$. We claim that it suffices to show that this point is in the image of $U_t$ — in fact, we will reduce the problem until we can apply Lemma 2.10 and 2.11. In this proof, for some $\varepsilon > 0$ we write $\widetilde{U}_t = \widetilde{U}_t^\varepsilon$ for the set obtained from $U_t$ by adding the open interval we removed:

$$\widetilde{U}_t = \left(-\frac{\varepsilon^2}{t^2}; \frac{\varepsilon^2}{t^2}\right) \times \left(-\frac{\varepsilon^2}{t^2}; \frac{\varepsilon^2}{t^2}\right) \times \left(-\frac{\varepsilon}{t}; \frac{\varepsilon}{t}\right).$$

It is the interior of $\overline{U_t}$. We also set

$$\pi_1 : (p, q, \psi) \in \mathbb{R}^3 \mapsto p \in \mathbb{R}, \qquad \pi_{2,3} : (p, q, \psi) \in \mathbb{R}^3 \mapsto (q, \psi) \in \mathbb{R}^2.$$

The proof will involve a lot of different $\varepsilon > 0$. To keep track of them all, we will define a sequence

$$\varepsilon_0' > \varepsilon_0 > \varepsilon_F' > \varepsilon_F > \cdots > 0$$

such that, for instance, some function $F$ is well-defined for all $\varepsilon < 2\varepsilon_F'$,[3] and well-behaved for $\varepsilon < 2\varepsilon_F$.

Let $0 < \varepsilon_0' \leq 1$ be as small as one wishes, and in any case small enough so that Proposition 2.3 holds for all $\varepsilon < 2\varepsilon_0'$. Set $d = |(tq_0, \psi_0)|$ and

$$\begin{pmatrix} a_t \\ b_t \end{pmatrix} := \frac{1}{t} M^{-1} \begin{pmatrix} y_t \\ t\phi_t \end{pmatrix},$$

seen as functions of $t$ and $P_0 \in U_t$. We can express the latter as

$$\begin{aligned}
\begin{pmatrix} a_t \\ b_t \end{pmatrix} &= \int_0^t \frac{1}{t} M^{-1} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & t \end{pmatrix} DX_t(sP_0) P_0 \mathrm{d}s \\
&= \sigma^2 t \int_0^t M^{-1} \begin{pmatrix} \gamma_t(sP_0) & \delta_t(sP_0) \end{pmatrix} \begin{pmatrix} \sigma^2 t^2 & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 1 \end{pmatrix} P_0 \mathrm{d}s,
\end{aligned}$$

where we wrote again the decomposition of $DX_t$ given by the technical Lemma 2.4 as

$$T_t + R_t = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \tag{2.21}$$

with, for instance, $\|\delta - M\| \leq C\varepsilon^2$ for some constant $C > 0$. We get a bound

$$\begin{aligned}
\left| \begin{pmatrix} a_t \\ b_t \end{pmatrix} - \sigma^2 t \begin{pmatrix} tq_0 \\ \psi_0 \end{pmatrix} \right| &\leq \sigma^2 t \int_0^1 \|M^{-1}\delta_t(sP_0) - I_2\| \cdot |(tq_0, \psi_0)| \,\mathrm{d}s + \sigma^4 t^3 \int_0^1 |M^{-1}\gamma_t(sP_0)| \cdot |p_0| \mathrm{d}s \\
&\leq C\varepsilon^2 td + C\varepsilon^4 td \\
&\leq 2C\varepsilon^3 td
\end{aligned}$$

for some constant $C > 0$ large enough. In particular, for all $\varepsilon$ small enough and $P_0 \in \overline{U_t}$,

$$\pm a_t(p_0, \pm\varepsilon^2/t^2, \psi_0) \geq \frac{\sigma^2}{2}\varepsilon^2, \qquad \pm b_t(p_0, q_0, \pm\varepsilon/t) \geq \frac{\sigma^2}{2}\varepsilon^2. \tag{2.22}$$

[3]The doubling condition ensures that we need not overly worry about considering either $U_t$ or its closure $\overline{U_t}$.

In this case, according to the first (elementary) topological Lemma 2.10, for any $|p_0| \leq \varepsilon^2/t^2$, the image of $(q_0, \psi_0) \mapsto (a_t(P_0), b_t(P_0))$ contains all $(a, b)$ such that $|a|, |b| < \sigma^2\varepsilon^2/2$, when restricted to $\pi_{2,3}(\widetilde{U}_t)$. Since $M$ is constant, there is a small constant $c > 0$ such that $|y| \leq ct$ and $|\phi| \leq c$ collectively imply that $M^{-1}(y, t\phi)/t$ is less than $\sigma^2\varepsilon^2/2$, so in fact $(y, \phi)$ is in the image of $\Psi_{t,p_0}$. The first part of the proposition is set; however we do not fix such a $c$ just yet.

We will also need

$$\frac{\sigma^2}{2} td \leq |(a_t, b_t)| \leq 2\sigma^2 td \tag{2.23}$$

for $\varepsilon > 0$ small enough, a direct consequence of the above. We fix some $\varepsilon_0 < \varepsilon_0'$ such that (2.22) and (2.23) holds for any $\varepsilon < 2\varepsilon_0$. This $\varepsilon_0$ will correspond to the $\varepsilon$ chosen in the proposition.

We try to define

$$\begin{aligned} &F_t : \mathbb{R} \times \widetilde{U}_t \to \mathbb{R}^4 \\ &\qquad x, P_0 \mapsto \big(p_0, x_t(P_0) - x_t\big(0, \Psi_{t,0}^{-1}(y_t(P_0), \phi_t(P_0))\big) - x, a_t(P_0), b_t(P_0)\big). \end{aligned}$$

We need to choose $\varepsilon > 0$ small enough so that $y_t(P_0)$ and $\phi_t(P_0)$ can be inverted by $\Psi_{t,0}$, seen as a function $\pi_{2,3}(\widetilde{U}_t^{\varepsilon_0}) \to \mathbb{R}^3$. According to the remark following equation (2.22) and (2.23), it is enough to have $\varepsilon < \varepsilon_0$ such that $2\sigma^2 t \cdot \varepsilon/t < \sigma^2\varepsilon_0^2/2$. Fix some $\varepsilon_F < \varepsilon_0$ such that it holds for all $\varepsilon < 2\varepsilon_F$, which ensures that $F_t$ is well-defined for any such $\varepsilon$, when $\Psi_{t,p_0}$ has domain $\pi_{2,3}(\widetilde{U}_t^{\varepsilon_0})$; from now on, we in fact fix the domain of $F_t$ to be $\mathbb{R} \times \widetilde{U}_t^{\varepsilon_F}$. For brevity, in the following we write

$$P_0^{y,\phi} := (0, \Psi_{t,0}^{-1}(y, \phi))$$

whenever well-defined.

According to Proposition 2.3, $F_t$ is one-to-one, since $(q_0, \psi_0)$ is characterised by $(a_t, b_t)$ when $\varepsilon < 2\varepsilon_0'$, and after that we can deduce $x$. Moreover, and using the decomposition of Lemma 2.4 and the notations of (2.21), we can show that its differential in blocks is

$$DF_t = \begin{pmatrix} 0 & 1 & \begin{pmatrix} 0 & 0 \end{pmatrix} \\ -1 & \partial_{p_0}(F_t)_2 & \partial_{\binom{q_0}{\psi_0}}(F_t)_2 \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \sigma^4 t^3 M^{-1}\gamma_t(P_0) & \sigma^2 t M^{-1}\delta_t(P_0)\begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix} \end{pmatrix},$$

$$\partial_{p_0}(F_t)_2 = \sigma^6 t^5 \Big(\alpha_t(P_0) - \beta_t(P_0^{y,\phi})\delta_t^{-1}(P_0^{y,\phi})\gamma_t(P_0)\Big),$$

$$\partial_{\binom{q_0}{\psi_0}}(F_t)_2 = \sigma^4 t^3 \Big(\beta_t(P_0) - \beta_t(P_0^{y,\phi})\delta_t(P_0^{y,\phi})^{-1}\delta_t(P_0)\Big)\begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix}.$$

Indeed, the first and last columns are clear from the representation of $DX_t$ and the definition of $(a_t, b_t)$. It remains to compute the partial derivatives of the second component $(F_t)_2$ of $F_t$. Of course $\partial_x(F_t)_2 = -1$. Moreover,

$$\begin{aligned} \partial_{p_0}(F_t)_2(x, P_0) &= \sigma^2 t \cdot \sigma^2 t^2 \alpha_t(P_0)\sigma^2 t^2 \\ &\qquad - \sigma^2 t \cdot \sigma^2 t^2 \beta_t(P_0^{y,\phi})\begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix} \cdot D(\Psi_{t,0}^{-1}) \cdot \sigma^2 t \begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix}\gamma_t(P_0)\sigma^2 t^2 \\ &= \sigma^6 t^5 \alpha_t(P_0) - \sigma^8 t^6 \beta_t(P_0^{y,\phi})\begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix} D(\Psi_{t,0}^{-1})\begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix}\gamma_t(P_0), \end{aligned}$$

where $D(\Psi_{t,0}^{-1})$ and $P_0^{y,\phi}$ are evaluated at $(y_t(P_0),\phi_t(P_0))$. Since the differential of $\Psi_{t,0}$ is

$$D\Psi_{t,0}(q_0,\psi_0)=\sigma^2 t\begin{pmatrix}t&0\\0&1\end{pmatrix}\delta_t(0,q_0,\psi_0)\begin{pmatrix}t&0\\0&1\end{pmatrix},$$

we find

$$D(\Psi_{t,0}^{-1})(y,\phi)=\frac{1}{\sigma^2 t^3}\begin{pmatrix}1&0\\0&t\end{pmatrix}\delta_t(P_0^{y,\phi})^{-1}\begin{pmatrix}1&0\\0&t\end{pmatrix},$$

which yields the given expression for $\partial_{p_0}(F_t)_2$. Similarly,

$$\begin{aligned}\partial_{\left(\begin{smallmatrix}q_0\\ \psi_0\end{smallmatrix}\right)}(F_t)_2&=\sigma^2 t\cdot\sigma^2t^2\beta_t(P_0)\begin{pmatrix}t&0\\0&1\end{pmatrix}\\&\qquad-\sigma^2 t\cdot\sigma^2t^2\beta_t(P_0^{y,\phi})\begin{pmatrix}t&0\\0&1\end{pmatrix}\cdot D(\Psi_{t,0}^{-1})\cdot\sigma^2 t\begin{pmatrix}t&0\\0&1\end{pmatrix}\delta_t(P_0)\begin{pmatrix}t&0\\0&1\end{pmatrix}\\&=\sigma^4t^3\beta_t(P_0)\begin{pmatrix}t&0\\0&1\end{pmatrix}-\sigma^4t^3\beta(P_0^{y,\phi})\delta(P_0^{y,\phi})^{-1}\delta(P_0)\begin{pmatrix}t&0\\0&1\end{pmatrix}.\end{aligned}$$

As can be seen from the determinant, the differential is invertible whenever $\delta_t(P_0)$ is. Since $\delta_t(P_0)$ is but the differential $D\Psi_{t,p_0}(q_0,\psi_0)$, our hypotheses imply that it is invertible for all $\varepsilon<2\varepsilon_0$, and $DF_t$ is invertible since we assume $\varepsilon<2\varepsilon_F$. In particular, $F_t$ is a diffeomorphism onto its image. Because of the estimates (2.22) on $a_t$ and $b_t$ and the following remark, it contains

$$\left(-\frac{\varepsilon_F^2}{t^2};\frac{\varepsilon_F^2}{t^2}\right)\times\mathbb{R}\times\left(-\frac{\sigma^2}{2}\varepsilon_F^2;\frac{\sigma^2}{2}\varepsilon_F^2\right)\times\left(-\frac{\sigma^2}{2}\varepsilon_F^2;\frac{\sigma^2}{2}\varepsilon_F^2\right).$$

We finally define the map $G$ that will satisfy the hypotheses of Lemma 2.11. Its expression is given by

$$G:P_0\in\overline{U_t}\mapsto\Big((F_t^{-1})_1\big(p_0,0,a_t(P_0),b_t(P_0)\big),\ a_t(P_0),\ b_t(P_0)\Big)\in\mathbb{R}^3.$$

According to the form of the subset of the image given above, it will be well-defined for initial conditions $P_0$ small enough so that $|p_0|$ is less than $\varepsilon_2^2/t^2$, and $|a_t(P_0)|$ and $|b_t(P_0)|$ are less than $\sigma^2\varepsilon_2^2/2$. According to the bound on $(a_t,b_t)$ given by (2.23), there exists $\varepsilon_G'<\varepsilon_F$ such that $G$ is well-defined on the whole of $\overline{U_t^\varepsilon}$, for all $\varepsilon<2\varepsilon_G'$.

From the estimate (2.22) and Lemma 2.10, we already know that there exists $\eta>0$ such that the second and third components of $G$ satisfy

$$\pm G_2\Big(p_0,\pm\frac{(\varepsilon_G')^2}{t^2},\psi_0\Big)\geq\eta,\qquad\qquad \pm G_3\Big(p_0,q_0,\pm\frac{\varepsilon_G'}{t}\Big)\geq\eta.$$

Indeed, any $\eta<\sigma^2(\varepsilon_G')^2/2$ would do. However, we do not fix such an $\eta$ just yet, because the first condition of Lemma 2.11 will force us to choose some smaller $\varepsilon$. Assume that this first assumption holds, and precisely that we can prove the following. There exists $\eta_0',c_0'>0$ and $\varepsilon_\star<\varepsilon_G'$ sufficiently small such that for all $\varepsilon<2\varepsilon_\star$ and $|a|,|b|<c_0'$,

$$\pm(F_t^{-1})_1\left(\pm\frac{\varepsilon^2}{t^2},0,a,b\right)\geq\varepsilon^2\eta_0't(a^2+b^2).$$

Then, according to (2.23), we can find $\varepsilon_G<\varepsilon_\star$ such that $|a_t(P_0)|$ and $|b_t(P_0)|$ are less than $c_0'$ for all $\varepsilon<2\varepsilon_G$ and $P_0\in\overline{U_t}$. We can thus fix $\eta_0$ less than $\sigma^2\varepsilon_G^2/2$ and $\varepsilon_G^2\eta_0'$, which implies

$$\pm G_1\Big(\pm\frac{\varepsilon_G^2}{t^2},q_0,\psi_0\Big)\leq\eta_0t\Big(G_2\big(\pm\frac{\varepsilon_G^2}{t^2},q_0,\psi_0\big)^2+G_3\big(\pm\frac{\varepsilon_G^2}{t^2},q_0,\psi_0\big)^2\Big),\tag{2.24}$$

$$\pm G_2\Big(p_0, \pm\frac{\varepsilon_G^2}{t^2}, \psi_0\Big) \geq \eta_0, \qquad \pm G_3\Big(p_0, q_0, \pm\frac{\varepsilon_G}{t}\Big) \geq \eta_0 \tag{2.25}$$

for all $P_0 \in \overline{U_t^\varepsilon}$, $\varepsilon = \varepsilon_G$. Let us see how it is sufficient to prove the proposition for $\varepsilon = \varepsilon_0$.

Fix some $c_0 > 0$ such that for all and $(\delta x, y, \phi)$ satisfying

$$|y| \leq c_0 t, \qquad |\phi| \leq c_0, \qquad |\delta x| \leq \frac{c_0}{t}(y^2 + t^2\phi^2),$$

we have $|a|, |b| \leq \eta_0$ and $|\delta x| \leq \eta_0 t(a^2 + b^2)$, where $(a, b) = M^{-1}(y, t\phi)/t$. Since $\eta_0 < \sigma^2\varepsilon_0^2/2$, and according to the remark following equation (2.22), for any $|p_0| < \varepsilon_0^2/t^2$, there exists $(q_0, \psi_0)$ such that $P_0 \in \widetilde{U}_t^{\varepsilon_0}$ and $\Psi_{t,p_0}(q_0, \psi_0) = (y, \phi)$. There goes the first point of the proposition.

We now forget about this previous $P_0$. According to Lemma 2.11 and equations (2.24) and (2.25), there exists another $P_0 \in \overline{U_t^\varepsilon}$, $\varepsilon = \varepsilon_G$ such that $G(P_0) = (\delta x, a, b)$.

We have $(a_t(P_0), b_t(P_0)) = (a, b)$, so $(y_t(P_0), \phi_t(P_0)) = (y, \phi)$. Moreover, since

$$\delta x = (F_t)^{-1}(p_0, 0, a, b),$$

there exists some $P_0' \in \widetilde{U}_t^\varepsilon$, $\varepsilon = \varepsilon_F$ such that $F(\delta x, P_0') = (p_0, 0, a, b)$. Clearly, $p_0' = p_0$. But $(a_t(P_0), b_t(P_0)) = (a_t(P_0'), b_t(P_0'))$, and

$$(\widetilde{q}_0, \widetilde{\psi}_0) \mapsto \big(a_t(p_0, \widetilde{q}_0, \widetilde{\psi}_0), b_t(p_0, \widetilde{q}_0, \widetilde{\psi}_0)\big)$$

is one-to-one on $U_t^\varepsilon$ for all $\varepsilon < \varepsilon_0$ because $\Psi_{t,p_0}$ is. This implies

$$0 = F_2(\delta x, P_0) = x_t(P_0) - x_t\big(0, \Psi_{t,0}^{-1}(y, \phi)\big) - \delta x.$$

This will conclude the proof of the proposition, provided we show that the strict inequality $|\delta x| < \frac{c_0}{t}(y^2 + t^2\phi^2)$ implies we have $(q_0, \psi_0) \neq 0$. But if, on the contrary, $(q_0, \psi_0) = 0$, then $X_t(P_0) = (t, 0, 0)$ and $y = \phi = 0$, and in this case the strict inequality is impossible.

We are left with the proof of the assumed fact. Let us state it again. We want to find $\eta_0', c_0' > 0$ and $\varepsilon_\star < \varepsilon_G'$ such that for all $\varepsilon < 2\varepsilon_\star$ and $|a|, |b| < c_0'$,

$$\pm(F_t^{-1})_1\left(\pm\frac{\varepsilon^2}{t^2}, 0, a, b\right) \geq \varepsilon^2\eta_0' t(a^2 + b^2).$$

Since, for $P_0 \in \widetilde{U}_t^{\varepsilon_2}$, we have

$$F_t(x, 0, q_0, \psi_0) = (0, -x, a_t(0, q_0, \psi_0), b_t(0, q_0, \psi_0)),$$

and because $F_t$ is one-to-one, we can write, wherever it is well-defined,

$$(F_t^{-1})_1(0, 0, a, b) = 0.$$

For instance, it holds when $|a|, |b| < \sigma^2\varepsilon_2^2/2$. We prove that, for some $\eta_F > 0$ and $|a|, |b|, |p_0|$ sufficiently small, the derivative $\partial_x(F_t^{-1})_1$ satisfies a lower bound

$$\partial_x(F_t^{-1})_1(p_0, 0, a, b) \geq t^3\eta_F(a^2 + b^2),$$

which will be enough to conclude.

We can find an explicit expression for the inverse matrix $(DF_t(x,P_0))^{-1}$. Namely, it equals

$$\begin{pmatrix} \sigma^6 t^5\big(\alpha_t(P_0)-\beta_t(P_0)\delta_t^{-1}(P_0)\gamma_t(P_0)\big) & -1 & \sigma^2 t^2\big(\beta_t(P_0)\delta_t^{-1}(P_0)-\beta_t(P_0^{y,\phi})\delta_t^{-1}(P_0^{y,\phi})\big)M \\ 1 & 0 & \begin{pmatrix}0 & 0\end{pmatrix} \\ -\sigma^2 t\begin{pmatrix}1 & 0\\ 0 & t\end{pmatrix}\delta_t^{-1}(P_0)\gamma_t(P_0) & \begin{pmatrix}0\\0\end{pmatrix} & \frac{1}{\sigma^2 t^2}\begin{pmatrix}1 & 0\\ 0 & t\end{pmatrix}\delta_t^{-1}(P_0)M \end{pmatrix}.$$

Our goal is then to show that

$$\alpha_t(P_0)-\beta_t(P_0)\delta_t^{-1}(P_0)\gamma_t(P_0) \geq \frac{\eta}{t^2}\big(a_t(P_0)^2+b_t(P_0)^2\big)$$

for some $\eta>0$ small enough, and all $P_0$ arising as the inverse image by $F_t$ of $(p_0,0,a,b)$ with $|a|,|b|,|p_0|$ small to be made precise. In fact, using the comparison between $|(a_t(P_0),b_t(P_0))|$ and $|(tq_0,\psi_0)|$ given by equation (2.23), it is equivalent to show

$$\alpha_t(P_0)-\beta_t(P_0)\delta_t^{-1}(P_0)\gamma_t(P_0) \geq \eta(t^2q_0^2+\psi_0^2),$$

under the same constraints for $P_0$.

Using Lemma 2.4, we see directly that for any $P_0\in\mathbb{R}^3$,

$$\Big|\alpha_t(P_0)-\beta_t(P_0)\delta_t^{-1}(P_0)\gamma_t(P_0)-Q_0^*(Q+AM^{-1}B)Q_0\Big| \leq C\varepsilon^2 d^2$$

for some large constant $C>0$, where we wrote $Q_0$ for $(q_0,\psi_0)$ and $d$ for $|(q_0,t\psi_0)|$. Since the symmetric part of $Q-AM^{-1}B$ is definite positive, there exists $\varepsilon_Q>0$ and $\eta_Q>0$ small enough so that for any $\varepsilon<2\varepsilon_Q$ and $P_0\in\widetilde{U}_t^\varepsilon$,

$$\alpha_t(P_0)-\beta_t(P_0)\delta_t^{-1}(P_0)\gamma_t(P_0) \geq \eta_Q d^2.$$

Let us get the other way around, from this solid result to the fact to prove. Fix $\varepsilon_\star<\varepsilon_Q,\varepsilon'_G/2$. According to equation (2.23), there exists $\eta_a>0$ such that

$$\alpha_t(P_0)-\beta_t(P_0)\delta_t^{-1}(P_0)\gamma_t(P_0) \geq \frac{\eta_a}{t^2}\big(a_t(P_0)^2+b_t(P_0)^2\big)$$

for all $P_0\in\widetilde{U}_t^\varepsilon$, $\varepsilon<2\varepsilon_\star$. This gives an $\eta'_F>0$ such that

$$\partial_x(F_t^{-1})_1(p_0,0,a_t(P_0),b_t(P_0)) \geq t^3\eta_F(a_t(P_0)^2+b_t(P_0)^2),$$

for all $P_0$ in the same $\widetilde{U}_t^\varepsilon$. We know that the image of $F_t$, restricted to some $\mathbb{R}\times\widetilde{U}_t^\varepsilon$ for some $\varepsilon<\varepsilon_F$, contains all points $(p_0,0,a,b)$ with $|p_0|<\varepsilon^2/t^2$ and $|a|,|b|<\sigma^2\varepsilon^2/2$. Fix some positive $c'_0<\sigma^2\varepsilon_\star^2/2$. Let $\varepsilon<2\varepsilon_\star$, and $|p_0|<\varepsilon^2/t^2$. Then, since $2\varepsilon_\star<\varepsilon_F$, the tuple $(p_0,0,a,b)$ has an inverse for $F$ belonging to $\mathbb{R}\times\widetilde{U}_t^\varepsilon$ whenever $|a|,|b|<c'_0$. Moreover, for such data, $F_t$ satisfies

$$\partial_x(F_t^{-1})_1(p_0,0,a,b)\geq t^3\eta_F(a^2+b^2), \qquad (F_t^{-1})_1(0,0,a,b)=0.$$

Of course, this implies

$$\pm(F_t^{-1})_1(\pm|p_0|,0,a,b)\geq t^3|p_0|\eta_F(a^2+b^2)$$

for all such $p_0$ and $(a,b)$, which concludes the proof of the fact for $\eta'_0=\eta_F$, and with it that of the proposition. ∎

## 2.4 Minimising curves

We are now ready to prove the main theorem. We give here a (slightly stronger) precise version. For this section, $x_t$ will denote the first coordinate of an arbitrary curve $\gamma_t$ (and similarly with $y_t$ and $\phi_t$); we will use $x_t(P_0)$, and always write down that argument, to denote the solution to the Hamilton equations.

**Theorem 2.6.** *Let $\gamma : [0;t] \to \mathbb{R}^3$ be a curve. Suppose that $\gamma$ is $\mathcal{C}^2$ has finite action, and is locally strongly minimising in the following sense. For all $s \in (0;t)$, there is a $\delta > 0$ such that $\gamma_{|[s-\delta;s+\delta]}$ is minimising for the action, among all piecewise $\mathcal{C}^2$ curves.*
*Then there exists $P_0 \in \mathbb{R}^3$ such that $\gamma_\tau = X_\tau(\gamma_0, P_0)$ for all $\tau \in [0;t]$.*

The main tool in the proof of the theorem is the following.

**Lemma 2.7.** *Fix some $\varepsilon, \sigma > 0$. Let $\gamma : I \to \mathbb{R}$ be a $\mathcal{C}^2$ curve of finite action, and suppose that $(\dot\gamma_{t_0}, \ddot\gamma_{t_0}) \neq 0$ for some interior point $t_0$ of $I$. Then for all $t < t + \delta t$ close enough to $t_0$,*

$$\gamma_{t+\delta t} \in X_{\delta t}(\gamma_t, U^\varepsilon_{\delta t}).$$

We postpone its proof to the end of the section, and turn to that of the theorem.

*Proof.* Let $\gamma$ be such a curve. If $\dot\phi_\tau = 0$ over $[0;t]$, then

$$\gamma_\tau = (x_0 + \cos(\phi_0)\tau, y_0 + \sin(\phi_0)\tau, \phi_0) = X_\tau(\gamma_0, p_0(\cos(\phi_0), \sin(\phi_0), 0))$$

for any $p_0 \in \mathbb{R}$, and we are done.

On the contrary, there may exist a point $\tau_0 \in [0;t]$ such that $|(\dot\phi_{\tau_0}, \ddot\phi_{\tau_0})| > 0$. Since the condition is open, we may also assume that $\tau_0$ lies in $(0;t)$. According to Lemma 2.7, there is a neighbourhood $(\tau_-;\tau_+)$ of $\tau_0$ and a path $\zeta : [\tau_-;\tau_+] \to \mathbb{R}^3$ such that for all $\tau_- < \tau < \tau+$, $\gamma_\tau = X_{\tau-\tau_-}(\gamma_{\tau_-}, \zeta_\tau)$ and $(\dot\phi_\tau, \ddot\phi_\tau) \neq 0$. Moreover, we can apply the lemma with $\varepsilon > 0$ small enough for Proposition 2.3 to apply, and we find $\zeta_\tau$ in some $U^\varepsilon_{\tau-\tau_-}$ where $DX$ is invertible. Since $\gamma$ is locally strongly minimising, we can further reduce the interval $(\tau_-;\tau_+)$ so that it minimises the action between all times $\tau_- \leq a < b \leq \tau_+$, among $\mathcal{C}^2$ curves. According to Proposition 2.2, $\gamma$ must in fact be a characteristic over $[\tau_-;\tau_+]$, hence there exists $P^\gamma : [\tau_-;\tau_+] \to \mathbb{R}^3$ such that $\tau \mapsto (\gamma_\tau, P^\gamma_\tau)$ satisfies the Hamilton equations.

The notation $P^\gamma$ suggests that it does not depend on the construction, but only on the path. That it indeed the case, in particular $P^\gamma$ does not depend on $\tau_0$, provided the non-degeneracy condition $(\dot\phi_\tau, \ddot\phi_\tau) \neq 0$ holds. From the Hamilton equations we know that $p^\gamma$ and $q^\gamma$ are constant. Moreover,

$$\ddot\phi_\tau = -\sigma^2 \begin{pmatrix} p^\gamma_\tau \\ q^\gamma_\tau \end{pmatrix} \cdot \begin{pmatrix} -\sin\phi_\tau \\ \cos\phi_\tau \end{pmatrix}$$

over $(\tau_-;\tau_+)$, where $\phi$ does vary, and continuously so. This proves that $p^\gamma$ and $q^\gamma$ are characterised by $\gamma$. Moreover, $\psi^\gamma = \dot\phi/\sigma^2$, so $P^\gamma$ depends on $\gamma$ only.

Let $I$ be a maximal (open) interval of $(0;t)$ over which $|(\dot\phi_\tau, \ddot\phi_\tau)|$ does not vanish. According to the above, there exists $P^\gamma : I \to \mathbb{R}^3$ such that $(\gamma, P^\gamma)$ is a solution to the Hamilton equations. Because $p^\gamma$ is constant, we also write $p^\gamma$ for its value at any point, and do this for $q^\gamma$ as well. We show that $I = (0;t)$; continuity of $\gamma$ and well-posedness of the Hamilton flow then concludes. Suppose that $b := \sup I < t$; the case $\inf I > 0$ is similar and not treated. Then,

$$P^\gamma_\tau \xrightarrow[\tau\to b]{} \begin{pmatrix} p^\gamma \\ q^\gamma \\ 0 \end{pmatrix} \qquad \text{with} \qquad \begin{pmatrix} p^\gamma \\ q^\gamma \end{pmatrix} \in \mathbb{R} \begin{pmatrix} \cos\phi_b \\ \sin\phi_b \end{pmatrix}.$$

Indeed, $(\dot\phi_b, \ddot\phi_b) = 0$, so $\psi^\gamma_\tau = \dot\phi_\tau/\sigma^2$ must tend to zero, and

$$\begin{pmatrix} p^\gamma \\ q^\gamma \end{pmatrix} \cdot \begin{pmatrix} -\sin\phi_\tau \\ \cos\phi_\tau \end{pmatrix} = -\ddot\phi_\tau/\sigma^2 \xrightarrow[\tau\to b]{} 0.$$

Moreover, since $\dot\psi = \ddot\phi/\sigma^2$, $\dot P^\gamma_\tau$ tends to zero as $\tau$ tends to $b$. In particular, the curve

$$\tau \mapsto \begin{cases} (\gamma_\tau, P^\gamma_\tau) & \text{for } \tau < b \\ X_{\tau-b}\big(\gamma_b, (p^\gamma, q^\gamma, 0)\big) & \text{for } \tau \geq b \end{cases}$$

is of class $\mathcal{C}^1$ in a neighbourhood of $b$, and a solution to the Hamilton equation. But because of the collinearity of $(p^\gamma, q^\gamma)$ and $(\cos\phi_b, \sin\phi_b)$, a solution to this equation with the same value at $b$ is

$$\tau \mapsto \big(\gamma_b + (\tau - b)(\cos\phi_b, \sin\phi_b, 0), (p^\gamma, q^\gamma, 0)\big).$$

Hence $\phi_\tau = \phi_b$ for all $\tau < b$ close enough to $b$, and this is impossible because of the non-degeneracy condition $(\dot\phi_\tau, \ddot\phi_\tau) \neq 0$. □

*Proof of Lemma 2.7.* Up to considering a smaller $\varepsilon > 0$, we can assume that Propositions 2.3 and 2.5 hold, for some fixed $c > 0$. Let $\gamma : I \to \mathbb{R}$ be a curve with values in $\mathbb{R}^3$ with finite energy, and assume, up to reparametrisation, that 0 is in the interior of $\mathring{I}$, and satisfies the regularity condition $(\dot\gamma_0, \ddot\gamma_0) \neq 0$. As described at the beginning of section 2.2, we can assume, up to a Hamiltonian transformation, that $\gamma_t = 0$.

We will in fact prove a slightly stronger statement; namely, there exists $\eta > 0$ such that the following holds. If there exists and $0 < t \leq 1$ such that for all $0 < \tau \leq t$,

$$\left|\phi_\tau - a\tau - \frac{b}{2}\tau^2\right| \leq \Phi\tau^2 \tag{2.26}$$

with some constants $|a| \leq \eta/t$, $|b| \leq \eta/t^2$ and $\Phi \leq \eta^2 D/t$ where $D = |(a, tb)| > 0$, then $\gamma_\tau \in V_\tau$ for all $\tau \leq t$.

Let us see how it implies the lemma. If $\phi$ is $\mathcal{C}^2$, we find a bound

$$\left|\phi_\tau - \dot\phi_0\tau - \ddot\phi_0\frac{\tau^2}{2}\right| \leq \sup_{0\leq s\leq\tau} |\ddot\phi_s - \ddot\phi_0| \cdot \frac{\tau^2}{2}$$

for all $\tau > 0$ such that $\gamma_\tau$ is well-defined. Since $\eta > 0$ is fixed and $D/t = |(\dot\phi_0, t\ddot\phi_0)|/t$ is bounded below, there exists some $t > 0$ such that the above estimate holds. Moreover, since $\ddot\phi$ is uniformly continuous around 0, a sufficiently small $t$ will allow for an equivalent of this estimate to hold for the increments $\phi_{s+\tau} - \phi_s$ instead of $\phi_\tau - \phi_0$, for all times $s$ in a small neighbourhood of zero and $\tau \leq t$.

We turn to the proof that (2.26) implies $\gamma_\tau \in V_\tau$. We suppose that the estimate holds for some $0 < \eta \leq 1$, and make stronger and stronger assumptions on $\eta$, until $(x_\tau, y_\tau, \phi_\tau)$ is compelled to satisfy the hypotheses of Proposition 2.5. Set $p_0 = 0$, $q_0 = -b/\sigma^2$ and $\psi_0 = a/\sigma^2$, so that $\dot\phi_0(P_0) = a$ and $\ddot\phi_0(P_0) = b$; then, because the curve has finite action,

$$|\phi_\tau| \leq D\tau + \frac{D}{2}\tau + \eta^2 D\tau \leq 2D\tau, \qquad |y_\tau| \leq \int_0^\tau |\sin(\phi_s)| \mathrm{d}s \leq D\tau^2,$$

$$|x_\tau - x_\tau(P_0)| \leq \int_0^\tau |\cos(\phi_s) - \cos(\phi_s(P_0))| \mathrm{d}u \leq \frac{1}{2}\int_0^\tau |\phi_s^2 - \phi_s^2(P_0)| \mathrm{d}u.$$

Note that for $D_\tau := |(a, \tau b)|$, we have $D\tau \le D_\tau t$. Since, for any $s \le \tau$, we have

$$|\dot\phi_s(P_0)| \le |\dot\phi_0(P_0)| + \int_0^s \sigma^2 |(p_0, q_0)| \mathrm{d}u = \sigma^2(|\psi_0| + |(p_0, q_0)|s) \le |a| + |b|s \le 2D_\tau,$$

$$\phi_s^{(3)} = \sigma^2 \dot\phi_s \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} \cdot \begin{pmatrix} \cos\phi_s \\ \sin\phi_s \end{pmatrix},$$

we deduce that

$$\begin{aligned} |\phi_s - \phi_s(P_0)| &\le \left|\phi_s - \left(as + \frac{b}{2}s^2\right)\right| + \left|\left(as + \frac{b}{2}s^2\right) - \phi_s(P_0)\right| \\ &\le \Phi s^2 + \int_0^s \frac{(s-u)^2}{2} \left|\phi_u^{(3)}(P_0)\right| \mathrm{d}u \\ &\le \eta^2 D\tau^2/t + \sigma^2 2D_\tau |q_0 \sin(\phi_s)| \frac{s^3}{6} \\ &\le \frac{4}{3}\eta^2 D_\tau \tau \end{aligned}$$

and

$$\begin{aligned} |\phi_\tau + \phi_\tau(P_0)| &\le 2D_\tau\tau + \sigma^2(|\psi_0|\tau + |(p_0, q_0)|\tau^2/2) \\ &\le 2D_\tau\tau + (|a|\tau + |b|\tau^2/2) \\ &\le 2D_\tau\tau + \frac{3}{2}D_\tau\tau = \frac{7}{2}D_\tau\tau; \end{aligned}$$

all in all, we find a large constant $C > 0$ such that

$$|x_\tau - x_\tau(P_0)| \le \frac{1}{2}\int_0^\tau |\phi_s^2 + \phi_s^2(P_0)| \cdot |\phi_s - \phi_s(P_0)| \mathrm{d}s \le C\eta D_\tau^2 \tau^3.$$

Let us try to express $D_\tau = |(a, \tau b)|$ as a function of $y_\tau$ and $\phi_\tau$. With elementary computations like the above, we can show

$$\left|\begin{pmatrix} y_\tau \\ \tau\phi_\tau \end{pmatrix} - \begin{pmatrix} \frac{\tau^2}{2} & \frac{\tau^3}{6} \\ \tau^2 & \frac{\tau^3}{2} \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}\right| \le C D_\tau \eta \tau^2 \tag{2.27}$$

where $C > 0$ is universal. Indeed, our first hypothesis (2.26) gives

$$\left|\phi_\tau - a\tau^2 + \frac{b\tau^2}{2}\right| \le \Phi\tau^2 \le \eta^2 D_\tau \tau,$$

whereas

$$\begin{aligned} \left|y_\tau - \frac{a\tau^2}{2} + \frac{b\tau^3}{6}\right| &\le \int_0^\tau |\sin(\phi_s) - \phi_s| \mathrm{d}s + \int_0^\tau \left|\phi_s - as^2 + \frac{bs^2}{2}\right| \mathrm{d}s \\ &\le \int_0^\tau |\phi_s|^3/6 \, \mathrm{d}s + \eta^2 D_\tau \tau \\ &\le C\eta^2 D_\tau \tau^2. \end{aligned}$$

In particular, from (2.27) we deduce

$$\frac{1}{C'} D_\tau \tau^2 \le \left|\begin{pmatrix} y_\tau \\ \tau\phi_\tau \end{pmatrix}\right| \le C' D_\tau \tau^2 \tag{2.28}$$

where we supposed $\eta$ small enough, and $C' > 0$ is a large universal constant. In summary, we have for some other universal constant $C > 0$:

$$|\phi_\tau| \leq C\eta, \qquad |y_\tau| \leq C\eta\tau, \qquad |x_\tau - x_\tau(P_0)| \leq \frac{C\eta}{\tau}\left(y_\tau^2 + \tau^2\phi_\tau^2\right).$$

This is close to the condition described in Proposition 2.5, but not quite it. Provided $\eta > 0$ is small enough, it does however imply that $(y_\tau, \phi_\tau)$ is in the image of $\Psi_{\tau,0}$; we set $P_0^\tau := (0, \Psi_{\tau,0}^{-1}(y_\tau, \phi_\tau))$. To conclude, it would be enough to show

$$\left|x_\tau(P_0) - x_\tau(P_0^\tau)\right| \leq \frac{C\eta^2}{\tau}\left(y_\tau^2 + \tau^2\phi_\tau^2\right)$$

for some constant $C > 0$. In the case $|y_\tau(P_0)| \leq ct$, $|\phi_\tau(P_0)| \leq c$ and $P_0 \in U_t$,[4] we can write

$$\left|x_\tau(P_0) - x_\tau(P_0^\tau)\right| = \left|x_\tau\big(0, \Psi_{\tau,0}^{-1}(y_\tau(P_0), \phi_\tau(P_0))\big) - x_\tau\big(0, \Psi_{\tau,0}^{-1}(y_\tau, \phi_\tau)\big)\right|.$$

We will show the following facts.

1. There is some $\eta > 0$ such that $|y_\tau(P_0)| \leq ct$, $|\phi_\tau(P_0)| \leq c$ and $P_0 \in U_t$.
2. The trajectories $\gamma$ and $X(P_0)$ satisfy, for $C > 0$ large enough,
$$|y_\tau|, |y_\tau(P_0)| \leq CD_\tau\tau^2, \qquad |\phi_\tau|, |\phi_\tau(P_0)| \leq CD_\tau\tau.$$
3. These trajectories are close in the following sense. For some large $C > 0$,
$$|y_\tau - y_\tau(P_0)| \leq C\eta^2 D_\tau\tau^2, \qquad |\phi_\tau - \phi_\tau(P_0)| \leq C\eta^2 D_\tau\tau.$$

We conclude the proof assuming the three of them. Because of equation (2.28), the problem reduces to

$$\left|x_\tau(P_0) - x_\tau(P_0^\tau)\right| \leq C\eta^2 D_\tau^2\tau^3.$$

We can use the technical Lemma 2.4, together with the block decomposition

$$T_t + R_t = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

($\alpha$ is $1 \times 1$, $\delta$ is $2 \times 2$), to get

$$\partial_{\binom{y}{\phi}}\left(x_\tau(0, \Psi_{\tau,0}^{-1}(y,\phi))\right) = \sigma^2\tau\beta_\tau\delta_\tau^{-1}\begin{pmatrix} 1 & 0 \\ 0 & \tau \end{pmatrix},$$

where $\beta_\tau$ and $\delta_\tau$ are evaluated at $(0, \Psi_{\tau,0}^{-1}(y,\phi))$. Setting

$$\mathcal{P}_s := \left(0, \Psi_{\tau,0}^{-1}(\mathcal{Y}_s)\right) \qquad \text{with} \qquad \mathcal{Y}_s := \begin{pmatrix} y_\tau \\ \phi_\tau \end{pmatrix} + s\begin{pmatrix} y_\tau(P_0) - y_\tau \\ \phi_\tau(P_0) - \phi_\tau \end{pmatrix},$$

we derive

$$\left|x_\tau(P_0) - x_\tau(P_0^\tau)\right| \leq \int_0^1 \left|\sigma^2\tau\beta_\tau(\mathcal{P}_s)\delta_\tau(\mathcal{P}_s)^{-1}\begin{pmatrix} 1 & 0 \\ 0 & \tau \end{pmatrix}\begin{pmatrix} y_\tau(P_0) - y_\tau \\ \phi_\tau(P_0) - \phi_\tau \end{pmatrix}\right| \mathrm{d}s.$$

[4] This gives raise to a notion of uniqueness for $P_0$: it is the one arising as $\Psi^{-1}(y,\phi)$ in Proposition 2.5.

We recall from the proof of Proposition 2.5 that

$$(D\Psi_{\tau,0}(P_0'))^{-1} = \frac{1}{\sigma^2\tau^3}\begin{pmatrix}1 & 0\\ 0 & \tau\end{pmatrix}\delta_\tau(P_0')^{-1}\begin{pmatrix}1 & 0\\ 0 & \tau\end{pmatrix}.$$

Because of the bounds on $\gamma_\tau$ and $X_\tau(P_0)$, and those of Lemma 2.4, we have

$$\left|\begin{pmatrix}0 & \tau & 0\\ 0 & 0 & 1\end{pmatrix}\mathcal{P}_s\right| \leq \int_0^1 \frac{C}{\sigma^2\tau^2}\left\|\delta\big(0,\Psi_{\tau,0}^{-1}(u\mathcal{Y}_s)\big)^{-1}\right\| D_\tau\tau^2\mathrm{d}u \leq C' D_\tau.$$

This, together with the estimate for $\gamma_\tau - X_\tau(P_0)$, gives the bound

$$\left|x_\tau(P_0) - x_\tau(P_0^\tau)\right| \leq C\tau D_\tau\cdot C\eta^2 D_\tau\tau^2 \leq C'\eta^2 D_\tau^2\tau^3,$$

which is the announced estimate, and concludes the proof of the lemma.

Now to the proof of the three facts. From the proof of Proposition 2.5, more precisely equation (2.23), we recall

$$\left|\begin{pmatrix}y_\tau(P_0)\\ \tau\phi_\tau(P_0)\end{pmatrix}\right| \leq C\tau^2\left|\begin{pmatrix}\tau q_0\\ \psi_0\end{pmatrix}\right| = CD_\tau\tau^2 \leq 2C\eta\tau$$

for $\varepsilon$ small enough, and $C>0$ large. This shows the first point.

Regarding the second point, the estimates on $y_\tau$ and $\phi_\tau$, very similar to those we established at the start of the proof, are found in the exact same way. We then consider $y_\tau(P_0)$ only, the case of $\phi_\tau(P_0)$ can be treated in the same way, and is arguably easier. Using the Hamilton equation to expand $y_\tau(P_0)$ in Taylor series, we find

$$|y_\tau(P_0)| \leq 0 + 0\tau + a\frac{\tau^2}{2} + b\frac{\tau^3}{6} + \int_0^\tau \frac{(\tau-s)^3}{6}|y_s^{(3)}(P_0)|\mathrm{d}s.$$

Direct computations using again the Hamilton equations yield

$$\begin{aligned} y_\tau^{(3)} = & -\cos(\phi_\tau(P_0))\cdot\dot{\phi}_\tau(P_0)^3 \\ & + 3\sigma^2\sin(\phi_\tau(P_0))\begin{pmatrix}p_0\\ q_0\end{pmatrix}\cdot\begin{pmatrix}-\sin(\phi_\tau(P_0))\\ \cos(\phi_\tau(P_0))\end{pmatrix}\dot{\phi}_\tau(P_0) \\ & + \sigma^2\cos(\phi_\tau(P_0))\begin{pmatrix}p_0\\ q_0\end{pmatrix}\cdot\begin{pmatrix}\cos(\phi_\tau(P_0))\\ \sin(\phi_\tau(P_0))\end{pmatrix}\dot{\phi}_\tau(P_0). \end{aligned}$$

Recall that we found $|\dot{\phi}_s(P_0)| \leq 2D_\tau$ for all $s\leq\tau$. This gives $|\phi_s(P_0)| \leq 2D_\tau\tau$, then $|y_\tau^{(3)}| \leq 16D_\tau^3$. Plugging this in the Taylor expansion, we have indeed

$$|y_\tau(P_0)| \leq CD_\tau\tau^2 + CD_\tau^3\tau^4 \leq C(1+\eta^2)D_\tau\tau^2.$$

Half of the third point is treated in the same way:

$$\left|y_\tau(P_0) - \frac{a\tau^2}{2} + \frac{b\tau^3}{6}\right| \leq CD_\tau^3\tau^4 \leq C\eta^2 D_\tau\tau^2, \qquad \left|\phi_\tau(P_0) - a\tau^2 + \frac{b\tau^2}{2}\right| \leq CD_\tau^3\tau^3 \leq C\eta^2 D_\tau\tau.$$

The other half was already shown, in equation (2.27) □

## 2.5 Infinitesimal description of the flow

Fix $0 < \varepsilon < 1$ and $0 < T < 1$. Suppose $t \leq T$, $|p_0| \leq \frac{\varepsilon^2}{t^2}$, $|q_0| \leq \frac{\varepsilon^2}{t^2}$ and $|\psi_0| \leq \frac{\varepsilon}{t}$. Let $\tau$ denote some time ranging from 0 to $t$, and define $d = |(tq_0, \psi_0)|$.

In the following, any use of $C$ denotes a positive constant, depending only on $\sigma$, and allowed to change from one line to the other. For the remainder of the section, we use the $O$ notation in an unusual manner: we say that $f = O(g)$ for some functions $f$ and $g$ depending on the above variables ($t$, $\varepsilon$, $p_0$, etc.) if $|f| \leq C|g|$ holds for some constant $C > 0$ depending only on $\sigma$ *for all values of the parameters satisfying the above conditions.* For example, for $f = 1/(1-2t)$, we do *not* have $f = O(1)$ using our conventions. On the other hand, if $f = O(g)$ in our sense, then $f = O(g)$ in the usual sense.

Let us find an approximate expression for $DX_t$. As stated above, the singularity around the line $d = 0$ needs to be taken into consideration; we will develop the differential in terms of $(tq_0, \psi_0)$ in the following lemma.

**Lemma 2.8.** *The singularity of $DX_\tau$ on the line $q_0 = \psi_0 = 0$ exhibits the following behaviour.*

$$DX_\tau(p_0, 0, 0) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{pmatrix}$$

$$\partial_{q_0}\partial_{p_0} x_t(p_0, 0, 0) = \partial_{\psi_0}\partial_{p_0} x_t(p_0, 0, 0) = 0$$

*Moreover, we also get the following expansions*

$$DX_\tau(p_0, q_0, \psi_0) = \begin{pmatrix} * & * & * \\ * & -\frac{\sigma^2}{6}\tau^3 & \frac{\sigma^2}{2}\tau^2 \\ * & -\frac{\sigma^2}{2}\tau^2 & \sigma^2\tau \end{pmatrix} + O\begin{pmatrix} * & * & * \\ * & \frac{\varepsilon^2}{t^2}\tau^5 & \frac{\varepsilon^2}{t^2}\tau^4 \\ * & \frac{\varepsilon^2}{t^2}\tau^4 & \frac{\varepsilon^2}{t^2}\tau^3 \end{pmatrix}$$

$$\partial_{q_0} DX_\tau(p_0, q_0, \psi_0) = \begin{pmatrix} * & -\frac{\sigma^4}{20}\tau^5 & \frac{\sigma^4}{8}\tau^4 \\ -\frac{\sigma^4}{120}\tau^5 & * & * \\ -\frac{\sigma^4}{24}\tau^4 & * & * \end{pmatrix} + O\begin{pmatrix} * & \frac{\varepsilon^2}{t^2}\tau^7 & \frac{\varepsilon^2}{t^2}\tau^6 \\ \frac{\varepsilon^2}{t^2}\tau^7 & * & * \\ \frac{\varepsilon^2}{t^2}\tau^6 & * & * \end{pmatrix}$$

$$\partial_{\psi_0} DX_\tau(p_0, q_0, \psi_0) = \begin{pmatrix} * & \frac{\sigma^4}{8}\tau^4 & -\frac{\sigma^4}{3}\tau^3 \\ \frac{\sigma^4}{24}\tau^4 & * & * \\ \frac{\sigma^4}{6}\tau^3 & * & * \end{pmatrix} + O\begin{pmatrix} * & \frac{\varepsilon^2}{t^2}\tau^6 & \frac{\varepsilon^2}{t^2}\tau^5 \\ \frac{\varepsilon^2}{t^2}\tau^6 & * & * \\ \frac{\varepsilon^2}{t^2}\tau^5 & * & * \end{pmatrix}$$

$$\partial^2_{q_0}\partial_{p_0} x_\tau(p_0, q_0, \psi_0) = -\frac{\sigma^6}{168}\tau^7 + O\left(\frac{\varepsilon^2}{t^2}\tau^9\right)$$

$$\partial_{\psi_0}\partial_{q_0}\partial_{p_0} x_\tau(p_0, q_0, \psi_0) = \frac{\sigma^6}{48}\tau^6 + O\left(\frac{\varepsilon^2}{t^2}\tau^8\right)$$

$$\partial^2_{\psi_0}\partial_{p_0} x_\tau(p_0, q_0, \psi_0) = -\frac{\sigma^6}{15}\tau^5 + O\left(\frac{\varepsilon^2}{t^2}\tau^7\right)$$

*where the $O$ notation is used in the unusually strong sense described at the beginning of the section.*

*Proof.* We will treat each point through an example, and comment about the other cases along the way.

**Singularity** Let us consider $\tau \mapsto \partial_{p_0}\phi_\tau(p_0,0,0)$. It is the only function satisfying the differential equation

$$\ddot{u}_\tau = \sigma^2 \sin(\phi_\tau) + \sigma^2 \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} \begin{pmatrix} \cos\phi_\tau \\ \sin\phi_\tau \end{pmatrix} u_\tau$$

with initial conditions $u_0 = \dot{u}_0 = 0$. As discussed above, $\phi_\tau(p_0,0,0) \equiv 0$ because it satisfies the non linear Cauchy problem given by the Hamilton equations, so $\partial_{p_0}\phi_\tau(p_0,0,0)$ is in fact solution of

$$\ddot{u}_\tau = \sigma^2 p_0 u_\tau;$$

uniqueness of the solution implies $\partial_{p_0}\phi_\tau(p_0,0,0) \equiv 0$.

The other coefficients are treated the same way.

**Taylor expansions** Let us discuss the example of $\partial_{q_0}\phi_\tau$. We will use the Taylor expansion at time 0 up to the first non zero coefficient, and find bounds on the remainder.

Computing derivatives of $x$, $y$ and $\phi$ with respect to $\tau$ or the initial conditions is algorithmic: using commutativity of the derivatives and the differential relations, one reduces to expressions involving only spacial derivatives of $x$, $y$, $\phi$ and $\dot{\phi}$. We get

$$\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}\partial_{q_0}\phi_\tau = \partial_{q_0}\ddot{\phi}_\tau = -\sigma^2\cos(\phi_\tau) + \sigma^2 \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} \cdot \begin{pmatrix} \cos\phi_\tau \\ \sin\phi_\tau \end{pmatrix} \partial_{q_0}\phi_\tau,$$

$$\frac{\mathrm{d}^3}{\mathrm{d}\tau^3}\partial_{q_0}\phi_\tau = \partial_{q_0}\phi^{(3)}_\tau = \sigma^2\sin(\phi_\tau)\dot{\phi}_\tau + \sigma^2 \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} \cdot \begin{pmatrix} -\sin\phi_\tau \\ \cos\phi_\tau \end{pmatrix} \dot{\phi}_\tau\partial_{q_0}\phi_\tau + \sigma^2 \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} \cdot \begin{pmatrix} \cos\phi_\tau \\ \sin\phi_\tau \end{pmatrix} \partial_{q_0}\dot{\phi}_\tau.$$

We know the values of $x$, $y$, $\phi$ and $\dot{\phi}$ at time zero, namely 0 except for $\dot{\phi}_0 = \sigma^2\psi_0$, hence their spacial derivatives at time zero. Proceeding with our example, we get

$$\partial_{q_0}\phi_\tau = \partial_{q_0}\phi_0 + \partial_{q_0}\dot{\phi}_0\tau + \partial_{q_0}\ddot{\phi}_0\frac{\tau^2}{2} + \int_0^\tau \frac{(\tau-s)^2}{2}\partial_{q_0}\phi^{(3)}_s \mathrm{d}s = -\frac{\sigma^2}{2}\tau^2 + \int_0^\tau \frac{(\tau-s)^2}{2}\partial_{q_0}\phi^{(3)}_s \mathrm{d}s.$$

We estimate the remainder using the Gronwall lemma, in the following from: *if $f_\tau \le a_\tau + \int_0^\tau b_s f_s \mathrm{d}s$ for all $0 \le \tau \le t$ and $a$ is non decreasing, then $f_\tau \le a_\tau \exp\int_0^\tau b_s \mathrm{d}s$ for all $0 \le \tau \le t$.* This leads us to consider

$$|\partial_{q_0}\ddot{\phi}_\tau| \le \sigma^2 + \sigma^2\frac{\varepsilon^2}{t^2}\int_0^\tau (\tau-s)|\partial_{q_0}\ddot{\phi}_s|\mathrm{d}s,$$

thus $|\partial_{q_0}\ddot{\phi}_\tau| \le \sigma^2\exp(\sigma^2\varepsilon^2\tau^2/2t^2) \le C$. Integrating, we get $|\partial_{q_0}\dot{\phi}_\tau| \le C\tau$ and $|\partial_{q_0}\phi_\tau| \le C\tau^2$. More generally, the Gronwall lemma provides a bound on $\partial_\alpha\ddot{\phi}$ using bounds on $\partial_\beta\phi$ and $\partial_\beta\dot{\phi}$ for multindices $\beta < \alpha$. We can iterate this process to get estimates on any derivative of $\phi$.

Because $|\ddot{\phi}_\tau| \le \sigma^2\frac{\varepsilon^2}{t^2}$, we also have

$$|\dot{\phi}_\tau| \le |\sigma^2\psi_0| + \int_0^\tau \sigma^2\frac{\varepsilon^2}{t^2}\mathrm{d}\tau \le C\frac{\varepsilon}{t}$$

and $|\phi_\tau| \le C\frac{\varepsilon}{t}\tau$. We finally get

$$\partial_{q_0}\phi_\tau = -\frac{\sigma^2}{2}\tau^2 + \int_0^\tau \frac{(\tau-s)^2}{2}\left(O\left(\frac{\varepsilon^2}{t^2}\tau\right) + O\left(\frac{\varepsilon^3}{t^3}\tau^2\right) + O\left(\frac{\varepsilon^2}{t^2}\tau\right)\right)\mathrm{d}s = -\frac{\sigma^2}{2}\tau^2 + O\left(\frac{\varepsilon^2}{t^2}\tau^4\right)$$

as announced. □

**Corollary 2.9.** *We can express $DX_t$ as $U_t + V_t$, where*

$$U_t(p_0,q_0,\psi_0) = \begin{pmatrix} -\frac{\sigma^6}{336}t^7q_0^2 + \frac{\sigma^6}{48}t^6q_0\psi_0 - \frac{\sigma^6}{30}t^5\psi_0^2 & -\frac{\sigma^4}{20}t^5q_0 + \frac{\sigma^4}{8}t^4\psi_0 & \frac{\sigma^4}{8}t^4q_0 - \frac{\sigma^4}{3}t^3\psi_0 \\ -\frac{\sigma^4}{120}t^5q_0 + \frac{\sigma^4}{24}t^4\psi_0 & -\frac{\sigma^2}{6}t^3 & \frac{\sigma^2}{2}t^2 \\ -\frac{\sigma^4}{24}t^4q_0 + \frac{\sigma^4}{6}t^3\psi_0 & -\frac{\sigma^2}{2}t^2 & \sigma^2 t \end{pmatrix},$$

$$V_t(p_0,q_0,\psi_0) = O\begin{pmatrix} \varepsilon^2t^5d^2 & \varepsilon^2t^4d & \varepsilon^2t^3d \\ \varepsilon^2t^4d & \varepsilon^2t^3 & \varepsilon^2t^2 \\ \varepsilon^2t^3d & \varepsilon^2t^2 & \varepsilon^2t \end{pmatrix}$$

*where the $O$ notation is used in the unusually strong sense described at the beginning of the section.*

*Proof.* Let us discuss the most tedious example of $\partial_{p_0}x_t$. Define $Q_0 = (q_0,\psi_0)$ and $F : Q_0 = (q_0,\psi_0) \mapsto \partial_{p_0}x_t(p_0,q_0,\psi_0)$.

The Taylor expansion at $Q_0 = 0$ reads

$$F(Q_0) = F(0) + \mathrm{d}F_0(Q_0) + \int_0^1 (1-s)\mathrm{d}^2F_{sQ_0}(Q_0,Q_0)\mathrm{d}s.$$

According to the first part of the lemma, $F(0) = 0$ and $\mathrm{d}F_0 = 0$; using its second part, we get

$$\begin{aligned}
\partial_{p_0}x_t(P_0) &= \int_0^1 (1-s)\begin{pmatrix} q_0 & \psi_0\end{pmatrix}\left(\begin{pmatrix} -\frac{\sigma^6}{168}t^7 & \frac{\sigma^6}{48}t^6 \\ \frac{\sigma^6}{48}t^6 & -\frac{\sigma^6}{15}t^5\end{pmatrix} + O\begin{pmatrix}\varepsilon^2t^7 & \varepsilon^2t^6 \\ \varepsilon^2t^6 & \varepsilon^2t^5\end{pmatrix}\right)\begin{pmatrix} q_0 \\ \psi_0\end{pmatrix}\mathrm{d}s \\
&= t^5\int_0^1 (1-s)\begin{pmatrix} tq_0 & \psi_0\end{pmatrix}\left(\begin{pmatrix} -\frac{\sigma^6}{168} & \frac{\sigma^6}{48} \\ \frac{\sigma^6}{48} & -\frac{\sigma^6}{15}\end{pmatrix} + O(\varepsilon^2)\right)\begin{pmatrix} tq_0 \\ \psi_0\end{pmatrix}\mathrm{d}s \\
&= -\frac{\sigma^6}{336}t^7q_0^2 + \frac{\sigma^6}{48}t^6q_0\psi_0 - \frac{\sigma^6}{30}t^5\psi_0^2 + O(\varepsilon^2t^5d^2)
\end{aligned}$$

as expected. □

Let us rephrase the corollary to highlight the properties of $DX_t$. Define

$$T_t(p_0,q_0,\phi_0) = \frac{1}{\sigma^2 t}\begin{pmatrix} t^2\sigma^2 & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 1\end{pmatrix}^{-1} U_t(p_0,q_0,\psi_0)\begin{pmatrix} t^2\sigma^2 & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 1\end{pmatrix}^{-1}$$

and $R_t(p_0,q_0,\psi_0)$ in the same way, with $V$ instead of $U$. Then

$$T_t = \begin{pmatrix} \begin{pmatrix} tq_0 & \psi_0\end{pmatrix} Q\begin{pmatrix} tq_0 \\ \psi_0\end{pmatrix} & \begin{pmatrix} tq_0 & \psi_0\end{pmatrix} A \\ B\begin{pmatrix} tq_0 \\ \psi_0\end{pmatrix} & M\end{pmatrix}$$

and

$$R_t = O\begin{pmatrix} \varepsilon^2d^2 & \varepsilon^2d & \varepsilon^2d \\ \varepsilon^2d & \varepsilon^2 & \varepsilon^2 \\ \varepsilon^2d & \varepsilon^2 & \varepsilon^2\end{pmatrix}$$

where $Q$, $A$, $B$ and $M$ do not depend on anything. In fact,

$$Q = \begin{pmatrix} -\frac{1}{336} & \frac{1}{96} \\ \frac{1}{96} & -\frac{1}{30} \end{pmatrix}, \qquad A = \begin{pmatrix} -\frac{1}{20} & \frac{1}{8} \\ \frac{1}{8} & -\frac{1}{3} \end{pmatrix}, \qquad B = \begin{pmatrix} -\frac{1}{120} & \frac{1}{24} \\ -\frac{1}{24} & \frac{1}{6} \end{pmatrix}, \qquad M = \begin{pmatrix} -\frac{1}{6} & \frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}.$$

It is then straightforward to see that $M$ is invertible and $Q - AM^{-1}B$ has definite positive symmetric part, which was the content of Lemma 2.4.

## 2.6 An elementary topological appendix

**Lemma 2.10.** *Let $G : C := [-a_0, a_0] \times [-b_0, b_0] \to \mathbb{R}^2$ be a continuous function with $a_0, b_0 > 0$. Suppose that there exists constants $a_1, b_1 > 0$ such that*

$$\pm G_1(\pm a_0, y_0) \geq a_1, \qquad \pm G_2(x_0, \pm b_0) \geq b_1$$

*for all $(x_0, y_0) \in C$, where $\pm G(\pm u) \geq 1$ means $G(u) \geq 1$ and $-G(-u) \geq 1$.*

*Then every point $(x_1, y_1) \in \mathbb{R}^3$ satisfying*

$$|x_1| \leq a_1, \qquad |y_1| \leq b_1$$

*is in the image of $G$. Any point satisfying the corresponding strict inequalities is in the image of the interior of $C$.*

**Lemma 2.11.** *Let $G : C := [-a_0, a_0] \times [-b_0, b_0] \times [-c_0, c_0] \to \mathbb{R}^3$ be a continuous function with $a_0, b_0, c_0 > 0$. Suppose that there exists constants $a_1, b_1, c_1 > 0$ such that*

$$\pm G_1(\pm a_0, y, z_0) \geq a_1 \left(G_2(\pm a_0, y_0, z_0)^2 + G_3(\pm a_0, y_0, z_0)^2\right),$$

$$\pm G_2(x_0, \pm b_0, z_0) \geq b_1 \qquad \pm G_3(x_0, y_0, \pm c_0) \geq c_1$$

*for all $(x_0, y_0, z_0) \in C$, where $\pm G(\pm u) \geq 1$ means $G(u) \geq 1$ and $-G(-u) \geq 1$.*

*Then every point $(x_1, y_1, z_1) \in \mathbb{R}^3$ satisfying*

$$|x_1| \leq a_1(y_1^2 + z_1^2) \qquad |y_1| \leq b_1 \qquad |z_1| \leq c_1$$

*is in the image of $G$. Any point satisfying the corresponding strict inequalities is in the image of the interior of $C$.*

*Proof.* We prove only the second lemma, the first being in every way easier.

Define $U$ as the open set of points satisfying the hypotheses as the lemma, and such that the inequalities are in fact strict. Suppose by contradiction that there exists a point $q$ in $U$ that is not in the image of $G$; in other words, $G$ has values in $\mathbb{R}^3 \setminus \{q\}$. Then the application

$$\begin{aligned} \partial C \times [0,1] &\to \mathbb{R}^3 \setminus \{q\} \\ p, t &\mapsto G(tp) \end{aligned}$$

is a homotopy between $G_{|\partial C}$ and a constant function; in particular, the former is null-homotopic in $\mathbb{R}^3 \setminus \{q\}$.

Set $A := a_1(b_1^2 + c_1^2)$, and

$$\begin{aligned} M : \partial C &\to \mathbb{R}^3 \\ (x, y, z) &\mapsto \left(\frac{A}{a_0} x, \frac{b_1}{b_0} y, \frac{c_1}{c_0} z\right). \end{aligned}$$

Then $M$ has in fact values in $\mathbb{R}^3 \setminus \{q\}$, and is homotopic to $G_{|\partial C}$ through such maps, a possible homotopy being $p, t \mapsto tM(p) + (1-t)G(p)$. Indeed,

$$\pm M_1(\pm a_0, y, z) = A \geq a_1\big(M_2(\pm a_0, y, z)^2 + M_3(\pm a_0, y, z)^2\big)$$

for all $(x, y, z) \in C$, so neither $M(a_0, y, z)$ nor $M(-a_0, y, z)$ can belong to $U$. Other inequalities show that neither of the other faces are sent to points in $U$; namely, $M$ satisfies the hypotheses of the lemma. Those conditions are convex, so the said homotopy has indeed values in $\mathbb{R}^3 \setminus U$ as well.

This proves that $M : \partial C \to \mathbb{R} \setminus \{q\}$ is null-homotopic. But up a linear transformation $T$ at the target, $T \circ M$ is but the inclusion of $\partial C$ into $\mathbb{R}^3 \setminus \{T(q)\}$, with $T(q) \in \mathring{C}$. This inclusion is a homotopy equivalence; for the simplest case $T(q) = 0$ and $C = [-1; 1]^3$, $p \mapsto p/|p|_\infty$ is a homotopy inverse, and all other cases may reduce to this one. Since $M$ is both a homotopy equivalence and null-homotopic, $\mathbb{R}^3 \setminus \{q\}$ must be contractible, a well-known contradiction. The point $q$ was in fact in the image of $G$, and so was the open set $U$.

Because $C$ is compact, the image contains in fact the closure of $U$. Moreover, because of the hypotheses on $F$, a point in $U$ cannot be in the image of $\partial C$, so it must be the image of $\mathring{C}$. $\square$

# Chapter V

# Perspectives

I hope the reader has found, while reading this thesis or its table of contents, that kinetic Brownian motion has numerous fascinating aspects, and that many others are yet to be understood. I would argue that this particular diffusion, although it is not a fundamental object in stochastic analysis nor in stochastic differential geometry, is a great toy model for general homogenisation and hypoelliptic diffusions. Its behaviour has proven to be very rich, and is an invitation to ergodic theory, rough path theory, Riemannian geometry, infinite-dimensional geometry, martingale techniques, the parametrix method, calculus of variations, and so on. In this last part, I would like to give a few directions in which these objects can be explored, be it kinetic Brownian motion or similar objects.

Firstly, I want to isolate, among the restrictions that have been made during this work, two of them that feel particularly unnatural to me. The first one is the trace condition (III.2.3). Recall that it is a condition on the covariance of some underlying Brownian motion, ensuring the mixing property of the associated spherical Brownian motion, hence the convergence of kinetic Brownian motion. This condition does not feel particularly natural, and in fact the factor 3 can be weakened to a factor 2, using weak coupling methods such as the one described in [BKS18]. We do not know at the time of writing if the mixing properties hold under no non-trivial assumptions; however, based on the absence of any obvious geometric argument (or my being unable to find it), I feel like it should.

The second drawback is the conjecture made in Part IV.1. A strong form of it is concerned with the positivity of some function on $\mathbb{R}_+$, and ensures that the approximation $\tilde{u}$ of the kernel $u$ of the 2-dimensional kinetic Brownian motion is exponentially decreasing. This function is reasonably explicit, and an easy proof of the fact that it is positive at some point is to actually compute it. There is strong numerical evidence for it, as displayed for instance in figure IV.1.1. It is not clear that there is any difficulty beyond heavy computations; however, much to my frustration, I have not been able to tackle them.

These are the points that stand out to me within the thesis. However, there is much to be discussed outside the scope of this work, in a variety of directions.

**Homogenisation.** A by-product of Chapter III is the fact that any kinetic Brownian motion in a Hilbert space whose covariance satisfies condition (III.2.3) must converge, as the noise diverges and up to normalisation, to a Brownian motion. This leave open a whole world of curved spaces, out of which we have treated two general classes, namely when the (weak) Riemannian Hilbert manifold is either the space of diffeomorphisms or volume-preserving diffeomorphisms, endowed with the $L^2$ metric. As discussed, the difficulty is to construct a Cartan development in the form of a controlled differential equation. In our case, I was able to define it using the fact that we were dealing with manifolds of maps, and the differential properties of the so-called Leray operator, which projects vector fields onto their divergence-free component.

I would argue that a unified treatment of kinetic Brownian motion on Riemannian Hilbert manifolds is out of reach. The issue is that on a single Hilbert manifold, there are uncountably many weak Riemannian metrics which induce different topologies: for instance, on the space $H^s(M,N)$ of maps between two manifolds $M$ and $N$ with Sobolev regularity $H^s$, one finds metrics associated to the $H^\varsigma$ (or even $W^{p,k}$) topology for appropriate values of $\varsigma$, to spaces with weight, etc. The different behaviours of their topologies should lead to different constructions of the Cartan development.

However, some natural families are of particular interest. Many metrics and geodesic equations, like the ones we described, have physical significance or appear in some mathematical context: they include the (inviscid) Burgers and Camassa-Holm equation in fluid dynamics [HMR98], the Fisher-Rao metric in statistics [KLMP13], the EPDiff equation in image processing [Kol17]. Some of them have promising geometric properties, which make the possibility of constructing the Cartan development, hence the kinetic Brownian motion, believable. As described in Chapter III, this would directly show the convergence. One could also consider the case of strong Riemannian metrics; in other words, when the topology of the manifold coincides with that induced by the Riemannian metric. A typical example is the set $\mathrm{Diff}^s(M)$ of diffeomorphisms of a compact manifold $M$, of Sobolev regularity $s$. Although I do not currently have an idea of a systematic approach for strong Riemannian metrics, it seems more tractable than the general case. If such a method were to exist, a treatment of the Cartan development in the example of $\mathrm{Diff}^s(M)$ should describe its relation to the geometry of $M$.

**Simulations.** It is difficult to represent kinetic Brownian motion on the group of volume-preserving diffeomorphisms of a manifold. It is all the more unsatisfactory that it is a complex motion, hence difficult to convey to a large audience. The images found in this manuscript, due to J. Angst, take some time to produce, and are dependent of the fact that the simulation are done on the torus. Some other geometries, such as the sphere, can be considered with these methods, but it relies on the fact that we perfectly understand the geodesics on these spaces. Of course, kinetic Brownian motion, in the deterministic limit $\sigma = 0$, follows the Euler equations, and can be no easier to simulate than fluid dynamics.

In a sense, the Lagrangian point of view adopted in our study, centred around diffeomorphisms, is less suitable than the Eulerian one, centred around vector fields, when it comes to numerical analysis. A strong argument in this direction is that the latter is in essence linear. However, recent work has been successful in adapting the former approach to computer simulations, see for instance the work [PMT+11]. In this particular article, the authors successfully manage to describe the incompressible Euler equation on an arbitrary closed triangulated manifold. A few other people have developed these ideas, and it seems to me that an efficient treatment of kinetic Brownian motion can be reached through these methods.

**Heat kernels.** In part IV.1, I expressed the kernel of kinetic Brownian motion on $\mathbb{R}^2$ as a series involving the kernel of a simplified diffusion. I have no doubt that the methods described here generalise to the case of a closed surface. In fact, I am confident that they can treat the case of any Riemannian manifold, although there are some expected complications. Most notably, the velocity of the kinetic Brownian motion in dimension $d \geq 3$ has values in the sphere of dimension $d-1 \geq 2$, so we must trade the model of $\mathbb{S}^1 \simeq \mathbb{R}/\mathbb{Z}$, which enabled us to see the velocity as a standard Brownian motion on $\mathbb{R}$, for another approach. In the case of elliptic diffusions on compact manifolds, cutoff functions answer similar concerns, and I think they can play the same role in our case. The relativistic Brownian motion, defined as the equivalent of the kinetic Brownian motion on Lorentzian manifolds,[1] should exhibit a very similar structure, and, provided the approximation is well-behaved, I would expect its treatment to be no harder.

More generally, given a hypoelliptic diffusion, the work [Pao17] of E. Paoli shows how to construct a nilpotent approximation of it. Then it is natural to ask if the solution to this simplified equation is a good model for the kernel of the non-polynomial diffusion, in the sense that we can express the latter as a series involving convolutions of the former. Since only few universal methods are known to treat hypoelliptic problems, this would ironically be very interesting and quite useless, because in general we do not know how to tackle even the nilpotent case. In short, it would reduce an insolvable problem to another easier unsolvable problem; still, it would provide much insight into which structure is indeed relevant in the considered situation. Of course, the remedy to this contradiction is to come up with new results in the nilpotent case. This will not be achieved by the tools described here, and is a compelling challenge that has attracted much attention, see for instance [BP17, Hab18, Fra19].

In a different direction, heat kernel asymptotics are often a way to trace estimates. In the case of Brownian motion on a closed manifold, we can show the following; see for instance [Ros97]. Let $u$ be the kernel of the Brownian motion on a closed Riemannian manifold $M$, say of dimension $d$. With the convention that the Laplace operator $-\Delta$ is non-negative, we call $(\lambda_i)_{i\geq 0}$ its eigenvalues, counted with multiplicity. Then we have

$$\int_M u_t(x,x)\mathrm{d}x = \sum_i \mathsf{e}^{-t\lambda_i}.$$

We are interested in the behaviour of this equation for times $t>0$ very small. The left hand side can be approximated by the integral of a local quantity. Indeed, it is known that

$$u_t(x,x) = \frac{1}{\sqrt{4\pi t}^d}\Big(1+\frac{1}{6}s(x)t + O(t^2)\Big),$$

where $s(x)$ is the scalar curvature. It contains meaningful geometric and topological information: for instance, if $M$ is a surface, this yields

$$\sum_i \mathsf{e}^{-t\lambda_i} = \frac{1}{\sqrt{4\pi t}^d}\Big(\mathrm{vol}_g(M) + \frac{2\pi}{3}\chi(M) + O(t^2)\Big),$$

where $\mathrm{vol}_g(M)$ is the Riemannian volume of $M$, and $\chi(M)$ is its Euler characteristic, which can be used to describe the asymptotic behaviour of the eigenvalues.

Is there a kinetic equivalent of the above reasoning? In other words, does kinetic Brownian motion relate global quantities (such as the spectrum of the Laplacian) to integrals of local quantities (such as the curvature)? And if so, what geometrical meaning do these expressions convey? My guess would be that partial answers to these questions are within reach of the methods discussed in this work.

**Variational problem.** In part IV.2, we set to investigate the kernel of kinetic Brownian motion using variational methods. In a few words, we introduced a Lagrangian action $\gamma\mapsto I[\gamma]$ related to the geometry of the problem, and studied the curves $\gamma$ minimising its value for fixed endpoints. In some sense, the chosen Hamiltonian was a dequantisation of the hypoelliptic generator of

[1] Historically, kinetic Brownian motion was in fact defined as the equivalent of the relativistic Brownian motion on Riemannian manifolds.

kinetic Brownian motion. Theorem IV.2.6, the main result in this direction, stated that such minimisers must be solutions of the Hamiltonian equations associated to $I$ (in particular, they must be smooth).

In some sense, this has nothing to do with probability. Although we motivated this variational problem in the section's introduction by discussing how, in simpler cases, semiclassical analysis or large deviations techniques can be used to translate similar results to heat kernel estimates, the link is still conjectured in this particular instance. It has become clear to V. Kolokoltsov and myself that the application of the parametrix method to an ansatz given by semiclassical analogy would require a substantial new idea, to deal with difficulties that the elliptic case, as well as that of regular hypoelliptic equations described in [Kol00], do not seem to bear. Such an insight has eluded us, for what I believe are serious reasons. However, I hope that weaker results such as the rate of exponential decay are within reach of large deviation methods. Indeed, it seems to me that the geometrical difficulties on the way are close to the ones already treated in proving Theorem IV.2.6; time will tell in which particular way I am getting ahead of myself.

**Microlocal analysis.** In the introduction, we mentioned the fact that although the defining equation for kinetic Brownian motion has analytic coefficients, its kernel $u$ is not described by analytic functions: indeed, it is identically zero out of the cylinder $\{(x, \dot{x}), |x| \leq t\}$. This should be compared to Brownian motion, whose density is the Gaussian in Euclidean space, hence analytic in the whole space. However, the partial differential equation associated to kinetic Brownian motion is hypoelliptic, meaning that $u$ is smooth. I have hopes that the general theory of Hörmander, in this specific case, can describe how the analyticity breaks down. Is $u_t$ analytic everywhere outside $\{|x| = t\}$? is it analytic in no open set of $\{|x| < t\}$? is is analytic with respect to $\dot{x}$ but not with respect to $x$? are questions to which I believe answers can be found. In microlocal analysis, they can be rephrased in terms of the so-called analytic wave from set, which has been introduced around the seventies. Although many beautiful results are known, particularly the so-called propagation of singularities in the elliptic case, it is still not fully understood for hypoelliptic equations. As far as I am aware, our diffusion does not fit general models, and would be a nice addition to an interesting body of work.

# Bibliographie


[ABP19] J. Angst, I. Bailleul, and P. Perruchaud. Kinetic Brownian motion on the diffeomorphism group of a closed Riemannian manifold. *arXiv e-prints*, page arXiv :1905.04103, May 2019.

[ABT15] J. Angst, I. Bailleul, and C. Tardif. Kinetic Brownian motion on Riemannian manifolds. *Electron. J. Probab.*, 20 :no. 110, 40, 2015.

[AC12] M. Arnaudon and A.B. Cruzeiro. Lagrangian navier–stokes diffusions on manifolds : Variational principle and stability. *Bulletin des Sciences Mathématiques*, 136(8) :857 – 881, 2012.

[ACT08] M. Arnaudon, K.A. Coulibaly, and A. Thalmaier. Brownian motion with respect to a metric depending on time : definition, existence and applications to Ricci flow. *C. R. Math. Acad. Sci. Paris*, 346(13-14) :773–778, 2008.

[AF07] J. Angst and J. Franchi. A central limit theorem for a class of relativistic diffusions. *J. Math. Phys.*, 48(3) :083101, 2007.

[AHK12] S. Albeverio, A. Hilbert, and V.N. Kolokoltsov. Uniform asymptotic bounds for the heat kernel and the trace of a stochastic geodesic flow. *Stochastics*, 84(2-3) :315–333, 2012.

[AK98] V.I. Arnold and B.A. Khesin. *Topological methods in hydrodynamics*, volume 125 of *Applied Mathematical Sciences*. Springer, 1998.

[AM06] H. Airault and P. Malliavin. Quasi-invariance of Brownian measures on the group of circle homeomorphisms and infinite-dimensional Riemannian geometry. *J. Funct. Anal.*, 241(1) :99–142, 2006.

[App14] D. Applebaum. *Probability on compact Lie groups*, volume 70 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2014. With a foreword by Herbert Heyer.

[AR02] H. Airault and J. Ren. Modulus of continuity of the canonic Brownian on the group of diffeomorphisms of the circle. *J. Funct. Anal.*, 196 :395–446, 2002.

[Arn66] V.I. Arnold. Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier (Grenoble)*, 16(fasc. 1) :319–361, 1966.

[AS89] T.A. Arakelyan and G.K. Savvidy. Geometry of a group of area-preserving diffeomorphisms. *Physics Letters B*, 223(1), 1989.

[BA88] G. Ben Arous. Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus. *Annales scientifiques de l'École Normale Supérieure*, 4e série, 21(3) :307–331, 1988.

[BA89] G. Ben Arous. Développement asymptotique du noyau de la chaleur hypoelliptique sur la diagonale. *Annales de l'Institut Fourier*, 39(1) :73–99, 1989.

[Bai] I. Bailleul. A flow-based approach to rough differential equations. Lecture notes.

[Bai10] I. Bailleul. A stochastic approach to relativistic diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(3) :760–795, 2010.

[Bai15a] I. Bailleul. Flows driven by Banah space valued rough paths. *Séminaire Probab.*, XLVI :195–205, 2015.

[Bai15b] I. Bailleul. Flows driven by rough paths. *Rev. Mat. Iberoam.*, 31(3) :901–934, 2015.

[Bau14] F. Baudoin. *Diffusion processes and stochastic calculus.* EMS Textbooks in Mathematics. European Mathematical Society, 2014.

[BBCN16] D. Barilari, U. Boscain, G. Charlot, and R.W. Neel. On the Heat Diffusion for Generic Riemannian and Sub-Riemannian Structures. *International Mathematics Research Notices*, 2017(15) :4639–4672, 07 2016.

[BBN12] D. Barilari, U. Boscain, and R.W. Neel. Small-time heat kernel asymptotics at the sub-riemannian cut locus. *J. Differential Geom.*, 92(3) :373–416, 11 2012.

[BC17] I. Bailleul and R. Catellier. Rough flows and homogenization in stochastic turbulence. *J. Diff. Eq.*, 263(8) :4894–4928, 2017.

[BCD11] H. Bahouri, J.-Y. Chemin, and R. Danchin. *Fourier analysis and nonlinear partial differential euqations.*, volume 343 of *Grundlehren des mathemtischen Wissenschaften.* Springer, 2011.

[BdLHLT19] A. Bethencourt de Léon, D. Holm, E. Luesink, and S. Takao. Implications of Kunita-Itô-Wentwell formula for k-forms in stochastic fluid dynamics. *arXiv*, 1903.07201v1 :1–25, 2019.

[BF12] I. Bailleul and J. Franchi. Non-explosion criteria for relativistic diffusions. *Ann. Probab.*, 40(3) :2168–2196, 2012.

[BFH09] E. Breuillard, P. Friz, and M. Huesmann. From random walks to rough paths. *Proc. Amer. Math. Soc.*, 137(10) :3487–3496, 2009.

[BHVW17] J. Birrell, S. Hottovy, G. Volpe, and J. Wehr. Small mass limit of a Langevin equation on a manifold. *Ann. Henri Poincaré*, 18(2) :707–755, 2017.

[Bis05] J.-M. Bismut. The hypoelliptic Laplacian on the cotangent bundle. *J. Am. Math. Soc.*, 18(2) :379–476, 2005.

[Bis11] J.-M. Bismut. *Hypoelliptic Laplacian and Orbital Integrals.* Annals of Mathematical Studies. Princeton University Press, 2011.

[Bis15] J.-M. Bismut. Hypoelliptic Laplacian and probability. *J. Math. Soc. Japan*, 67(4) :1317–1357, 2015.

[Bis16] J.-M. Bismut. Eta invariants and the hypoelliptic Laplacian. *arXiv :1603.05103*, pages 1–155, 2016.

[BKH15] V. Bally and A. Kohatsu-Higa. A probabilistic interpretation of the parametrix method. *Ann. Appl. Probab.*, 25(6) :3095–3138, 2015.

[BKS18] O. Butkovsky, A. Kulik, and M. Scheutzow. Generalized couplings and ergodic rates for SPDEs and other Markov models. *ArXiv e-prints*, page arXiv :1806.00395, June 2018.

[BP17] D. Barilari and E. Paoli. Curvature terms in small time heat kernel expansion for a model class of hypoelliptic hörmander operators. *Nonlinear Analysis*, 164 :118 – 134, 2017.

[BV20] M. Bruveris and F.-X. Vialard. On completeness of groups of diffeomorphisms. *To appear in J. Europ. Math. Soc.*, pages 1–43, 2020.

[BW18] J. Birrell and J. Wehr. Langevin equations in the small-mass limit : Higher order approximations. *arXiv e-prints*, page arXiv :1809.01724, Sep 2018.

[Car01] E. Cartan. *Riemannian geometry in an orthogonal frame.* World Scientific, 2001.

[CFH18] D. Crisan, F. Flandoli, and D. Holm. Solution properties of a 3d stochastic euler fluid equation. *J. Nonlinear Sci.*, 242 :1–58, 2018.

[CFK+17] I. Chevyrev, P. K. Friz, A. Korepanov, I. Melbourne, and H. Zhang. Multiscale systems, homogenization, and rough paths. *ArXiv e-prints*, December 2017.

[CFK+19] I. Chevyrev, P.K. Friz, A. Korepanov, I. Melbourne, and H. Zhang. Deterministic homogenization for discrete-time fast-slow systems under optimal moment assumptions. *arXiv :1903.10418v1*, pages 1–24, 2019.

[CFM07] A.-B. Cruzeiro, F. Flandoli, and P. Malliavin. Brownian motion on volume preserving diffeomorphisms group and existence of global solutions of 2d stochastic euler equation. *J. Funct. Anal.*, 242 :304–326, 2007.

[Che18] I. Chevyrev. Random walks and Lévy processes as rough paths. *Probab. Theory Related Fields*, 170(3-4) :891–932, 2018.

[CHR18] A.-B. Cruzeiro, D. Holm, and T. Ratiu. Momentum maps and stochastic Clebsch action principles. *Comm. Math. Phys.*, 357(2) :873–912, 2018.

[CP03] M. Christensen and J. B. Pedersen. Diffusion in inhomogeneous and anisotropic media. *The Journal of Chemical Physics*, 119 :5171–5175, September 2003.

[CP11] K.A. Coulibaly-Pasquier. Brownian motion with respect to time-changing Riemannian metrics, applications to Ricci flow. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(2) :515–538, 2011.

[Cra91] M. Cranston. Gradient estimates on manifolds using coupling. *J. Funct. Anal.*, 99(1) :110–124, 1991.

[Cun17] Ch. Cuny. Invariance principles under the Maxwell-Woodroofe condition in Banach spaces. *Ann. Probab.*, 45(3) :1578–1611, 2017.

[CW16] T. Cass and M.P. Weidner. Tree algebras over topological vector spaces in rough path theory. *arXiv :1604.07352v2*, pages 1–25, 2016.

[Dav08] A.M. Davie. Differential equations driven by rough paths : An approach via discrete approximation. *AMRX, Appl. Math. Res. Express*, 2007 :40, 2008.

[DH18] D. Drivas and D. Holm. Circulation and energy theorem preserving stochastic fluids. *arXiv*, 1808.05308v1 :1–26, 2018.

[DM10] F. Delarue and S. Menozzi. Density estimates for a random noise propagating through a chain of differential equations. *J. Funct. Anal.*, 259(6) :1577–1630, 2010.

[Dro17] A. Drouot. Stochastic stability of Pollicott-Ruelle resonances. *Comm. Math. Phys.*, 356 :357–396, 2017.

[Dud66] R.M. Dudley. Lorentz-invariant Markov processes in relativistic phase space. *Ark. Mat*, 6 :241–268, 1966.

[Eel66] James Eells, Jr. A setting for global analysis. *Bull. Amer. Math. Soc.*, 72 :751–807, 1966.

[EM69] D.G. Ebin and J.E. Marsden. Groups of diffeomorphisms and the solution of the classical Euler equations for a perfect fluid. *Bull. Amer. Math. Soc.*, 75 :962–967, 1969.

[Fan02] Sh. Fang. Canonical brownian motion on the diffeomorphism group of the circle. *J. Funct. Anal.*, 196 :162–179, 2002.

[Fer70] X. Fernique. Intégrabilité des vecteurs gaussiens. *C. R. Acad. Sci. Paris*, 270(25) :1698–1699, 1970.

[FGL15] P.K. Friz, P. Gassiat, and T. Lyons. Physical Brownian motion in a magnetic field as a rough path. *Trans. Amer. Math. Soc.*, 367(11) :7939–7955, 2015.

[FH14] P.K. Friz and M. Hairer. *A course on rough paths*. Universitext. Springer, Cham, 2014. With an introduction to regularity structures.

[FLJ07] J. Franchi and Y. Le Jan. Relativistic diffusions and schwarzschild geometry. *Comm. Pure Appl. Math.*, 60(2) :187–251, 2007.

[FLJ11] J. Franchi and Y. Le Jan. Curvature diffusions in general relativity. *Comm. Math. Phys.*, pages 307–351, 2011.

[Fra19] J. Franchi. Small time equivalents for the density of a planar quadratic langevin equation. *Bull. Soc. Math. France*, 2019.

[FV10] P.K. Friz and N.B. Victoir. *Multidimensional stochastic processes as rough paths*, volume 120 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. Theory and applications.

[GBH17] F. Gay-Balmaz and D. Holm. Stochastic geometric models with non-stationary spatial correlations in Lagrangian fluid flows. *J. Nonlinear Sci.*, 28(3) :873–904, 2017.

[Gli11] Y.E. Gliklikh. *Global and stochastic analysis with applications to mathematical physics*. Theoretical and Mathematical Physics. Springer-Verlag London, Ltd., London, 2011.

[Gri04] D. Grieser. Notes on heat kernel asymptotics. Online version available on the author's website at http://www.staff.uni-oldenburg.de/daniel.grieser/wwwlehre/Schriebe/heat.pdf, 2004. Unpublished notes.

[Gro67] L. Gross. Abstract Wiener spaces. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. II : Contributions to Probability Theory, Part 1*, pages 31–42. Univ. California Press, Berkeley, Calif., 1967.

[Gro70] L. Gross. Abstract Wiener measure and infinite dimensional potential theory. In *Lectures in Modern Analysis and Applications, II*, pages 84–116. Lecture Notes in Mathematics, Vol. 140. Springer, Berlin, 1970.

[Hab18] K. Habermann. Small-time fluctuations for sub-riemannian diffusion loops. *Probability Theory and Related Fields*, 171(3) :617–652, Aug 2018.

[Hai09] M. Hairer. An introduction to stochastic PDEs. Lecture notes, 2009.

[HBR13] M. Heidernätsch, M.S. Bauer, and G. Radons. Characterizing n-dimensional anisotropic brownian motion by the distribution of diffusivities. *The Journal of Chemical Physics*, 139 18 :184105, 2013.

[HHV16] D.P. Herzog, S. Hottovy, and G. Volpe. The small-mass limit for Langevin dynamics with unbounded coefficients and positive friction. *J. Stat. Phys.*, 163(3) :659–673, 2016.

[HMR98] D.D. Holm, J.E. Marsden, and T.S. Ratiu. The Euler-Poincaré equations and semidirect products with applications to continuum theories. *Adv. Math.*, 137(1) :1–81, 1998.

[Hol15] D. Holm. Variational principles for stochastic fluid dynamics. *Proc. R. Soc. A*, 471 :20140963, 2015.

[Hör85] L. Hörmander. *The analysis of linear partial differential operators. III*, volume 274 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1985. Pseudodifferential operators.

[Hör90] L. Hörmander. *The analysis of linear partial differential operators. I*, volume 256 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1990. Distribution theory and Fourier analysis.

[Hsu02] E.P. Hsu. *Stochastic analysis on manifolds*, volume 38 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.

[IKN18a] S. Ishiwata, H. Kawabi, and R. Namba. Central limit theorems for non-symmetric random walks on nilpotent covering graphs : Part I. *ArXiv e-prints*, June 2018.

[IKN18b] S. Ishiwata, H. Kawabi, and R. Namba. Central limit theorems for non-symmetric random walks on nilpotent covering graphs : Part II. *ArXiv e-prints*, August 2018.

[Kam88] N. van Kampen. Diffusion in inhomogeneous media. *Journal of Physics and Chemistry of Solids*, 49 :673–677, 1988.

[Ken86] W.S. Kendall. Nonnegative Ricci curvature and the Brownian coupling property. *Stochastics*, 19(1-2) :111–129, 1986.

[KLMP13] B. Khesin, J. Lenells, G. Misiołek, and S.C. Preston. Geometry of diffeomorphism groups, complete integrability and geometric statistics. *Geom. Funct. Anal.*, 23(1) :334–366, 2013.

[KM16] D. Kelly and I. Melbourne. Smooth approximation of stochastic differential equations. *Ann. Probab.*, 44(1) :479–520, 2016.

[KM17] D. Kelly and I. Melbourne. Deterministic homogenization for fast-slow systems with chaotic noise. *J. Funct. Anal.*, 272(10) :4063–4102, 2017.

[Kol00] V.N. Kolokoltsov. *Semiclassical analysis for diffusions and stochastic processes*. Number 1724 in Lecture notes in Mathematics. Springer, 2000.

[Kol17] B. Kolev. Local well-posedness of the EPDiff equation : a survey. *J. Geom. Mech.*, 9(2) :167–189, 2017.

[Kun90] H. Kunita. *Stochastic flows*, volume 24 of *Probability Theory and Stochastic Modelling*. Cambridge Univ. Press, 1990.

[Kuw12] K. Kuwada. Convergence of time-inhomogeneous geodesic random walks and its application to coupling methods. *Ann. Probab.*, 40(5) :1945–1979, 2012.

[LCL07] T.J. Lyons, M. Caruana, and T. Lévy. *Differential equations driven by rough paths*, volume 1908 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007. Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004, With an introduction concerning the Summer School by Jean Picard.

[Li12] X.-M. Li. Effective Diffusions with Intertwined Structures. *arXiv e-prints*, page arXiv :1204.3250, Apr 2012.

[Li16a] X.-M. Li. Limits of random differential equations on manifolds. *Probab. Theory Related Fields*, 166(3-4) :659–712, 2016.

[Li16b] X.-M. Li. Random perturbation to the geodesic equation. *Ann. Probab.*, 44(1) :544–566, 2016.

[Li18] X.-M. Li. Homogenisation on homogeneous spaces. *J. Math. Soc. Japan*, 70(2) :519–572, 2018.

[Lia04] M. Liao. *Lévy processes in Lie groups*, volume 162 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2004.

[LS17] O. Lopusanschi and D. Simon. Area anomaly and generalized drift of iterated sums for hidden Markov walks. *ArXiv e-prints*, September 2017.

[LS18] O. Lopusanschi and D. Simon. Lévy area with a drift as a renormalization limit of Markov chains on periodic graphs. *Stochastic Process. Appl.*, 128(7) :2404–2426, 2018.

[LWL19] S.H. Lim, J. Wehr, and M. Lewenstein. Homogenization for Generalized Langevin Equations with Applications to Anomalous Diffusion. *arXiv e-prints*, page arXiv :1902.06496, Feb 2019.

[Lyo94] T.J. Lyons. Differential equations driven by rough signals. I. an extension of an inequality of L. C. Young. *Math. Res. Lett.*, 1(4) :451–464, 1994.

[Lyo98] T.J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14(2) :215–310, 1998.

[Mal99] P. Malliavin. The canonic diffusion above the diffeomorphism group of the circle. *C.R. Acad. Sci. Paris*, 329 :325–329, 1999.

[Mel93] R.B. Melrose. *The Atiyah-Patodi-Singer index theorem*, volume 4 of *Research Notes in Mathematics*. A K Peters, Ltd., Wellesley, MA, 1993.

[MJ00] K.V. Mardia and P.E. Jupp. *Directional statistics*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 2000.

[Pal68] R.S. Palais. *Foundations of global non-linear analysis*. W. A. Benjamin, Inc., New York-Amsterdam, 1968.

[Pao17] E. Paoli. Small time asymptotics on the diagonal for Hörmander's type hypoelliptic operators. *J. Dyn. Control Syst.*, 23(1) :111–143, 2017.

[Per18] P. Perruchaud. Homogenisation for anisotropic kinetic random motions. *arXiv e-prints*, page arXiv :1811.08415, Nov 2018.

[Pin76] M.A. Pinsky. Isotropic transport process on a Riemannian manifold. *Trans. Amer. Math. Soc.*, 218 :353–360, 1976.

[PMT+11] D. Pavlov, P. Mullen, Y. Tong, E. Kanso, J.E. Marsden, and M. Desbrun. Structure-preserving discretization of incompressible fluids. *Physica D : Nonlinear Phenomena*, 240(6) :443 – 458, 2011.

[PZ95] S. Peszat and J. Zabczyk. Strong Feller property and irreducibility for diffusions on Hilbert spaces. *Ann. Probab.*, 23(1) :157–172, 1995.

[Ros97] S. Rosenberg. *The Laplacian on a Riemannian manifold*, volume 31 of *Student texts*. London Mathematical Society, 1997.

[SE01] M. Serrano and P. Español. Thermodynamically consistent mesoscopic fluid particle model. *Phys. Rev. E*, 64 :046115, Sep 2001.

[SEZ05] M. Serrano, P. Español, and I. Zúñiga. Voronoi fluid particle model for Euler equations. *J. Stat. Phys.*, 121(1-2) :133–147, 2005.

[She16] Sh. Shen. Laplacien hypoelliptique, torsion analytique, et théorème de Cheeger-Müller. *J. Funct. Anal.*, 270(8) :2817–2999, 2016.

[Smo07] N.K. Smolentsev. Diffeomorphism group of compact manifolds. *Journal of Mathematical Sciences*, 146(6) :6213–6312, 2007.

[Sol95] M.R. Soloveitchik. Focker-Planck equation on a manifold. Effective diffusion and spectrum. *Potential Anal.*, 4(6) :571–593, 1995.

[Str93] D.W. Stroock. *Probability theory, an analytic view.* Cambridge University Press, Cambridge, 1993.

[WZ10] F.-Y. Wang and T.-Sh. Zhang. Gradient estimates for stochastic evolution equations with non-Lipschitz coefficients. *J. Math. Anal. Appl.*, 365(1) :1–11, 2010.

[Yos97] K. Yoshida. Riemannian curvature on the gorup of area-prserving diffeomorphisms (motions of fluid) of 2-sphere. *Physics D*, 100 :377–389, 1997.

[You36] L.Ch. Young. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.*, 67(1) :251–282, 1936.

[Zwo12] M. Zworski. *Semiclassical analysis*, volume 138 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2012.

UNIVERSITE BRETAGNE LOIRE / MATHSTIC

UNIVERSITÉ DE RENNES 1

**Titre :** Homogénéisation pour le mouvement brownien cinétique,
et quelques résultats sur son noyau

**Mot clés :** équations différentielles stochastiques ; géométrie différentielle stochastique ; chemins rugueux ; mécanique des fluides ; groupes de difféomorphismes ; noyau de la chaleur.

**Résumé :** Le mouvement brownien cinétique est une famille de processus stochastiques indexée par un paramètre de bruit, qui se veut une interpolation entre le flot géodésique et le mouvement brownien. Dans le cas d'une variété de dimension finie, il est connu que l'on peut donner un sens rigoureux à cette propriété d'homogénéisation. Ce travail de thèse construit ces diffusions et prouve un résultat de convergence dans l'exemple de dimension infinie de certaines variétés de difféomorphismes, issues de la mécanique des fluides (chapitre 3).

On se base sur des méthodes ergodiques et de chemins rugueux, qui, en dimension finie (chapitre 2), se montrent aussi efficaces pour une classe plus générale de diffusions cinétiques.

De manière indépendante, on propose deux pistes d'étude pour le noyau associé au mouvement brownien cinétique en dimension deux (chapitre 4). Plus précisément on s'intéresse à son comportement en temps petit, par des méthodes de parametrix adaptées à ce problème hypoelliptique particulier.

**Title:** Homogenisation for kinetic Brownian motion,
and some results about its kernel

**Keywords:** stochastic differential equations; stochastic analysis on manifolds; rough paths; fluid mechanics; diffeomorphism groups; heat kernels.

**Abstract:** The kinetic Brownian motion is a family of stochastic processes indexed by a noise parameter, which is expected to interpolate between the geodesic flow and the Brownian motion. For finite-dimensional manifolds, it is known that such a homogenisation property can be proved rigorously. This thesis provides a construction for these diffusions and a proof of a convergence result, in the infinite-dimensional example of certain manifolds of maps arising from fluid mechanics (chapter 3).

We work with ergodic methods and rough paths theory, which, in finite dimension (chapter 2), also prove effective for a general class of kinetic diffusions.

Independently, we offer two approaches for the study of the kernel associated to the two-dimensional kinetic Brownian motion (chapter 4). More precisely, we consider its small time asymptotics, using parametrix methods adapted to this particular hypoelliptic problem.