



HAL
open science

Contributions à l'étude mathématique et numérique de quelques modèles en simulation multi-échelle des matériaux

Frédéric Legoll

► **To cite this version:**

Frédéric Legoll. Contributions à l'étude mathématique et numérique de quelques modèles en simulation multi-échelle des matériaux. Analyse numérique [math.NA]. Université Pierre et Marie Curie - Paris VI, 2011. tel-00783334

HAL Id: tel-00783334

<https://theses.hal.science/tel-00783334>

Submitted on 31 Jan 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS VI PIERRE ET MARIE CURIE

HABILITATION À DIRIGER DES RECHERCHES

SPÉCIALITÉ : MATHÉMATIQUES APPLIQUÉES
ET APPLICATIONS DES MATHÉMATIQUES (CNU 26)

Frédéric LEGOLL

Sujet : *Contributions à l'étude mathématique et numérique de quelques modèles en simulation multi-échelle des matériaux*

soutenue le 17 octobre 2011 devant le jury composé de

Président : **Yves ACHDOU**

Rapporteurs : **Christian LUBICH**
Jean-Jacques MARIGO
Felix OTTO

Examineurs : **Michel BORNERT**
Claude LE BRIS

Coordinateur : **Yvon MADAY**

Remerciements

Il y a une dizaine d'années, je postulais pour un stage et rencontrais pour la première fois Eric Cancès, qui me mit rapidement en contact avec Claude Le Bris. Avec le recul aujourd'hui, je mesure la chance que j'ai eue de les rencontrer. Je leur dois beaucoup, et je souhaite ici les remercier pour leur soutien constant au cours de ces années, tant scientifique que humain.

Je souhaite aussi remercier Yvon Maday pour avoir accepté de tenir le rôle de coordinateur de mon habilitation. J'apprécie toujours autant son enthousiasme communicatif et lui suis reconnaissant de l'intérêt qu'il porte à mes travaux.

Mes remerciements vont également à Christian Lubich, Jean-Jacques Marigo et Felix Otto, qui ont accepté d'écrire un rapport sur mon habilitation, et Yves Achdou, qui a présidé le jury. Je leur suis très reconnaissant de leur intérêt pour mon travail, et de leurs suggestions, qui me permettent d'aller plus loin. Merci pour toutes les discussions que nous avons eues, avant, pendant et (espérons le) après mon habilitation.

J'exprime toute ma gratitude aux nombreuses personnes qui m'ont soutenu pour obtenir un poste à l'École des Ponts. Je remercie vivement Karam Sab, qui m'a accueilli au sein du Laboratoire Navier. Depuis mon arrivée, le laboratoire s'est restructuré, et une équipe Multiéchelle a été créée, dirigée par Michel Bornert. Je voudrais remercier Michel d'avoir fait partie de mon jury d'habilitation. Nous avons souvent des points de vue complémentaires sur nos thèmes d'intérêt commun, ce qui rend nos échanges au quotidien toujours enrichissants. Plus généralement, j'ai aussi une pensée pour tous mes collègues de Navier, que je remercie pour les discussions professionnelles ou amicales que nous avons pu avoir.

L'École des Ponts est un très bel endroit pour faire de la recherche, par la qualité de ses chercheurs et de ses étudiants. Je voudrais ici en remercier le Corps des Ponts, des Eaux et des Forêts.

J'ai aussi la chance de faire partie d'une équipe-projet de l'INRIA, que je souhaite remercier pour son aide précieuse.

Au quotidien, Catherine Baccaert, Marie-Françoise Kaspi, Nathalie Quelleu et Brigitte Vigo nous simplifient la tâche sur le plan administratif, tandis que Jacques Daniel est d'un secours toujours efficace sur le plan informatique. Merci à eux.

Depuis le début de ma thèse, j'ai eu la chance de pouvoir collaborer avec de nombreux chercheurs venus d'horizons divers. Je voudrais les remercier ici.

Depuis quelques années, je participe activement à l'encadrement de doctorants, ce qui est une expérience très enrichissante, bien sûr scientifiquement, mais aussi humainement. Je pense en particulier à Laurent Brochard, Amélie Deleurence et Laurent Monasse, avec qui j'ai eu le plaisir de longuement échanger. J'ai eu la chance de travailler activement avec Ronan Costaouec, Mouhamad Hammoud, Kimiya Minoukadeh et Florian Thomines. Je souhaite les remercier pour le travail effectué ensemble, ainsi que leurs directeurs de thèse respectifs pour m'avoir donné cette opportunité.

Enfin, je ne serais pas ce que je suis sans mon entourage, et je souhaite donc ici exprimer toute ma gratitude envers mes parents, mon frère, et surtout ma femme Nadine, mon fils Mathieu et bientôt mon second fils. Je sais pouvoir compter sur leur soutien sans faille, malgré une activité envahissante et des absences fréquentes.

Table des matières

Introduction	3
Liste des publications	5
Activités d'encadrement	7
1 Simulation moléculaire	9
1.1 Echantillonnage de la mesure de Boltzmann	10
1.1.1 Méthodes purement stochastiques	12
1.1.2 Chaînes de Markov	12
1.1.3 Un exemple de processus de Markov : la dynamique de Langevin amortie	14
1.1.4 Un exemple de méthode déterministe : la dynamique de Nosé-Hoover	15
1.2 Dynamique effective	20
1.2.1 Motivation	20
1.2.2 Construction	23
1.2.3 Estimation d'erreur	24
1.2.4 Résultats numériques	27
1.2.5 Extension	27
1.3 Problèmes Hamiltoniens	27
1.3.1 Convergence des moyennes temporelles	28
1.3.2 Problèmes hautement oscillants	28
1.3.3 Algorithmes parallèles en temps	36
1.4 Problèmes de conductivité thermique	43
2 Méthodes couplées discret-continu	47
2.1 Modèles à température nulle	49
2.2 Modèles à température finie	51
2.2.1 Le cas d'interactions à plus proches voisins	53
2.2.2 Extensions	57
3 Homogénéisation	63
3.1 Quelques résultats de base en homogénéisation	63
3.1.1 Le cas périodique	63
3.1.2 Homogénéisation stochastique	66
3.2 Approximation numérique et réduction de variance	68
3.2.1 Approche standard	68
3.2.2 Réduction de variance par la technique des variables antithétiques .	71
3.2.3 Un cas faiblement stochastique	76
3.3 Une variante de l'homogénéisation stochastique	79
3.3.1 Présentation du modèle	79
3.3.2 Approximation numérique	80
3.3.3 Perturbations de l'identité	81
3.3.4 Résidu dans le cas mono-dimensionnel	82
Conclusion et travaux en cours	85
Références	87

Introduction

La première partie de ce mémoire résume des travaux en simulation moléculaire. On s'intéresse à des systèmes de particules ponctuelles (représentant typiquement les noyaux des atomes d'un système moléculaire), qui interagissent via une énergie potentielle. Les degrés de liberté du système sont la position et l'impulsion de chaque particule. La complexité du problème vient du nombre de degrés de liberté en jeu : certains systèmes d'intérêt pratique comportent plusieurs centaines de milliers d'atomes, et les plus grands systèmes aujourd'hui simulés sont constitués de plusieurs dizaines de milliards d'atomes.

Une première question est l'échantillonnage de la mesure de Boltzmann-Gibbs associée à l'énergie potentielle du système. De nombreuses méthodes ont été proposées, basées sur des chaînes de Markov, des processus de Markov (solutions d'équations différentielles stochastiques), ou bien des équations différentielles ordinaires. L'objet de la Section 1.1 est d'en faire une brève revue, et de présenter nos résultats sur la non-ergodicité de certaines méthodes déterministes proposées dans la littérature (plus précisément, la méthode de Nosé-Hoover). L'utilisation de ces dernières pour échantillonner la mesure de Boltzmann-Gibbs est ainsi discutable.

Dans la Section 1.2, on s'intéresse à des questions de dynamique effective. Génériquement, l'énergie potentielle d'un système moléculaire possède de nombreux minima locaux, séparés par de hautes barrières. Cependant, il n'est souvent pas nécessaire de connaître tous les degrés de liberté du système pour déterminer dans quelle conformation le système se trouve : la connaissance d'une *fonction* de ces degrés de liberté suffit. Partant de la description microscopique du système (un vecteur X en grande dimension), on est donc amené à introduire une description macroscopique, fondée sur la connaissance d'une fonction de l'état microscopique, notée ici $\xi(X)$, supposée à valeur scalaire, et qui contient toute l'information qui nous intéresse. Nous proposons dans la Section 1.2 une méthode pour obtenir une dynamique approchant la dynamique $t \mapsto \xi(X_t)$ de la quantité d'intérêt, quand le système complet suit une dynamique $t \mapsto X_t$ régie par l'équation de Langevin amortie. Sous une hypothèse de séparation des échelles de temps, nous démontrons que la dynamique effective proposée est une bonne approximation de $t \mapsto \xi(X_t)$.

Les approches décrites ci-dessus concernent un système à température fixée, et donc non isolé. Nous nous sommes aussi intéressés au cas où le système est isolé. Son évolution est alors gouvernée par les équations de Newton, qui s'écrivent comme une dynamique Hamiltonienne. La simulation numérique de cette dynamique n'est pas simple. Il y a en effet une différence de plusieurs ordres de grandeur entre les temps caractéristiques des mouvements les plus rapides du système (qui se font à l'échelle de la femtoseconde), et les temps de simulation souhaités (qui sont de l'ordre de la micro-, voire de la milliseconde). Il est en pratique très difficile, sur des systèmes moléculaires de grande taille, d'atteindre de tels temps avec un schéma d'intégration usuel, pour lequel le pas de temps est contrôlé par les temps caractéristiques les plus rapides. Cette observation nous a conduit à développer des schémas numériques adaptés à des Hamiltoniens hautement oscillants (et pour lesquels le pas de temps n'est pas restreint par les mouvements les plus rapides), en suivant une démarche d'homogénéisation en temps (cf. la Section 1.3.2). Nous avons aussi proposé une adaptation au contexte Hamiltonien de l'algorithme *pararéel*, permettant d'obtenir la solution d'un problème d'évolution par des méthodes de calcul parallèle (cf. la Section 1.3.3).

Ces travaux en simulation moléculaire ont été effectués en collaboration avec E. Cancès, F. Castella, Ph. Chartier, X. Dai, B. Dickson, M. Dobson, E. Faou, P. Fleurat-Lessard, F. Galante, A. Iacobucci, S. Lahbabi, C. Le Bris, B. Leimkuhler, T. Lelièvre, M. Luskin, Y. Maday, M.-C. Marinica, K. Minoukadeh, R. Moeckel, R. Monneau, E. Noorizadeh, S. Olla, G. Stoltz, G. Turinici et F. Willaume.

La seconde partie de ce mémoire présente des travaux sur la dérivation de modèles à l'échelle du continuum à partir de modèles discrets, pour les solides, et sur le couplage de ces deux modèles, discret et continu. Le modèle fin est donc écrit à la même échelle d'espace que les modèles de la première partie, tandis que le modèle grossier est écrit à une échelle bien plus grande. Une première approche consiste à poser le problème sous forme variationnelle : la configuration recherchée est un minimiseur (parfois unique) d'une certaine énergie. Dans ce cas, nos résultats, en collaboration avec X. Blanc et C. Le Bris, portent sur des questions de couplage entre un modèle discret et un modèle de continuum, et ont été obtenus durant la thèse. Cette approche correspond à un système à température nulle. Nous nous sommes ensuite intéressés, avec X. Blanc, C. Le Bris et C. Patz, au cas de systèmes à température finie, pour lesquels les quantités d'intérêt macroscopiques sont des moyennes de fonctions dépendant de l'état microscopique du système (la position de tous les atomes) par rapport à la mesure de Boltzmann. Dans ce cas, nous nous sommes intéressés à l'obtention de modèles réduits, macroscopiques, où la température est un paramètre, en suivant des approches de type limite thermodynamique.

Dans la troisième partie du mémoire, on s'intéresse à des questions d'homogénéisation stochastique, pour des équations aux dérivées partielles elliptiques linéaires. Les matériaux sont donc modélisés à l'échelle du continuum (qui est l'échelle des modèles grossiers de la seconde partie). Le constat qui motive notre travail est le fait que, même dans les cas les plus simples sur le plan théorique, les méthodes numériques à ce jour disponibles en homogénéisation stochastique (par exemple pour calculer la matrice effective) conduisent à des calculs très lourds. Notre motivation ici est *pratique* : nous nous sommes intéressés à une équation très simple sur le plan théorique, et avons cherché à améliorer les méthodes numériques dans ce cas. Nous avons travaillé dans deux directions. La première consiste à réduire la variance des quantités calculées. En effet, bien que la matrice homogénéisée *exacte* soit, dans le cas considéré, déterministe, un artefact de la procédure d'approximation numérique est que cette matrice est en pratique approchée par une matrice *aléatoire*. Nous avons développé des méthodes pour réduire la variance de la matrice homogénéisée approchée, afin d'obtenir typiquement un intervalle de confiance plus petit pour les quantités homogénéisées, à coût de calcul fixe. Un second axe de travail est d'étudier le cas de problèmes *faiblement* stochastiques, en partant du constat que les matériaux hétérogènes, rarement périodiques, ne sont pas pour autant systématiquement fortement aléatoires. Le cas d'un matériau aléatoire pour lequel cet aléa n'est qu'une *petite perturbation* autour d'un modèle périodique est donc intéressant. Nous montrons que, dans ce cas, une approche spécifique peut être proposée, pour calculer, au moins aux premiers ordres, le comportement homogénéisé avec un coût bien plus faible que celui des approches générales en homogénéisation stochastique.

Ces travaux ont été réalisés en collaboration avec A. Anantharaman, X. Blanc, R. Costaouec, C. Le Bris et F. Thomines.

Liste des publications

Travaux effectués durant ma thèse

ARTICLES PUBLIÉS DANS DES REVUES INTERNATIONALES À COMITÉ DE LECTURE

- [FL1] *Designing reversible measure invariant algorithms with applications to molecular dynamics*, Journal of Chemical Physics, vol. 117 (23), 10452-10464, 2002 (avec R. Monneau).
- [FL2] *Numerical homogenization of nonlinear viscoplastic two-dimensional polycrystals*, Computational and Applied Mathematics, vol. 23 (2-3), 309-325, 2004.
- [FL3] *High-order averaging schemes with error bounds for thermodynamical properties calculations by molecular dynamics simulations*, Journal of Chemical Physics, vol. 121 (21), 10346-10355, 2004 (avec E. Cancès, F. Castella, Ph. Chartier, E. Faou, C. Le Bris et G. Turinici).
- [FL4] *Long-time averaging for integrable Hamiltonian dynamics*, Numerische Mathematik, vol. 100 (2), 211-232, 2005 (avec E. Cancès, F. Castella, Ph. Chartier, E. Faou, C. Le Bris et G. Turinici).
- [FL5] *Analysis of a prototypical multiscale method coupling atomistic and continuum mechanics*, Mathematical Modelling and Numerical Analysis, vol. 39 (4), 797-826, 2005 (avec X. Blanc et C. Le Bris).
- [FL6] *Analysis of a prototypical multiscale method coupling atomistic and continuum mechanics : the convex case*, Acta Mathematicae Applicatae Sinica, vol. 23 (2), 209-216, 2007 (avec X. Blanc et C. Le Bris).

Thèse

- [FL7] *Méthodes moléculaires et multi-échelles pour la simulation numérique des matériaux*, Thèse de l'Université Paris 6, août 2004.

Travaux postérieurs à ma thèse

ARTICLES PUBLIÉS DANS DES REVUES INTERNATIONALES À COMITÉ DE LECTURE

- [FL8] *Theoretical and numerical comparison of some sampling methods for molecular dynamics*, Mathematical Modelling and Numerical Analysis, vol. 41 (2), 351-389, 2007 (avec E. Cancès et G. Stoltz).
- [FL9] *Non-ergodicity of the Nosé-Hoover thermostatted harmonic oscillator*, Archives for Rational Mechanics and Analysis, vol. 184 (3), 449-463, 2007 (avec M. Luskin et R. Moeckel).
- [FL10] *Dérivation de schémas numériques symplectiques pour des systèmes hamiltoniens hautement oscillants (derivation of symplectic numerical schemes for highly oscillatory Hamiltonian systems)*, C. R. Acad. Sci. Paris, Série I, vol. 344 (4), 277-282, 2007 (avec C. Le Bris).
- [FL11] *Some remarks on sampling methods in molecular dynamics*, ESAIM Proceedings, vol. 22, 217-233, 2008 (avec T. Lelièvre et G. Stoltz).
- [FL12] *A temperature control technique for nonequilibrium molecular simulation*, Journal of Chemical Physics, vol. 128 (7), 074105, 2008 (avec B. Leimkuhler et E. Noorizadeh).

- [FL13] *Some improvements of the activation-relaxation technique method for finding transition pathways on potential energy surfaces*, Journal of Chemical Physics, vol. 130 (11), 114711, 2009 (avec E. Cancès, M.-C. Marinica, K. Minoukadeh et F. Willaime).
- [FL14] *Non-ergodicity of Nosé-Hoover dynamics*, Nonlinearity, vol. 22 (7), 1673-1694, 2009 (avec M. Luskin et R. Moeckel).
- [FL15] *Approximation numérique d'une classe de problèmes en homogénéisation stochastique (Numerical approximation of a class of problems in stochastic homogenization)*, C. R. Acad. Sci. Paris, Série I, vol. 348 (1-2), 99-103, 2010 (avec R. Costaeuec et C. Le Bris).
- [FL16] *Free energy calculations : An efficient adaptive biasing potential method*, Journal of Physical Chemistry B, vol. 114 (17), 5823-5830, 2010 (avec B. Dickson, T. Lelièvre, G. Stoltz et P. Fleurat-Lessard).
- [FL17] *Beyond multiscale and mutiphysics : multimaths for model coupling*, Networks and Heterogeneous Media, vol. 5 (3), 423-460, 2010 (avec X. Blanc, C. Le Bris et T. Lelièvre).
- [FL18] *Thermal conductivity of the Toda lattice with conservative noise*, Journal of Statistical Physics, vol. 140 (2), 336-348, 2010 (avec A. Iacobucci, S. Olla et G. Stoltz).
- [FL19] *Variance reduction in stochastic homogenization : proof of concept, using antithetic variables*, Boletin Soc. Esp. Mat. Apl., vol. 50, 9-27, 2010 (avec R. Costaeuec et C. Le Bris).
- [FL20] *Finite-temperature coarse-graining of one-dimensional models : mathematical analysis and computational approaches*, Journal of Nonlinear Science, vol. 20 (2), 241-275, 2010 (avec X. Blanc, C. Le Bris et C. Patz).
- [FL21] *Integrators for highly oscillatory Hamiltonian systems : an homogenization approach*, Discrete and Continuous Dynamical Systems - B, vol. 13 (2), 347-373, 2010 (avec C. Le Bris).
- [FL22] *Symplectic schemes for highly oscillatory Hamiltonian systems with varying fast frequencies (Intégrateurs symplectiques pour des systèmes Hamiltoniens hautement oscillants avec fréquences rapides variables)*, C. R. Acad. Sci. Paris, Série I, vol. 348 (17-18), 1033-1038, 2010 (avec M. Dobson et C. Le Bris).
- [FL23] *Effective dynamics using conditional expectations*, Nonlinearity, vol. 23 (9), 2131-2163, 2010 (avec T. Lelièvre).
- [FL24] *Variance reduction in stochastic homogenization using antithetic variables*, Markov Processes and Related Fields, à paraître (avec X. Blanc, R. Costaeuec et C. Le Bris).

ARTICLES SOUMIS

- [FL25] *Symplectic schemes for highly oscillatory Hamiltonian systems : the homogenization approach beyond the constant frequency case*, IMA Journal of Numerical Analysis, soumis, arXiv 1008.1030 (avec M. Dobson et C. Le Bris).
- [FL26] *Symmetric parareal algorithms for Hamiltonian systems*, Mathematical Modelling and Numerical Analysis, soumis, arXiv 1011.6222 (avec X. Dai, C. Le Bris et Y. Maday).
- [FL27] *A numerical strategy for coarse-graining two-dimensional atomistic models at finite temperature : the membrane case*, Computational Materials Science, soumis, arXiv 1109.5978 (avec X. Blanc).

- [FL28] *Asymptotic behaviour of Green functions of divergence form operators with periodic coefficients*, Applied Mathematics Research Express, soumis, arXiv 1110.4767 (avec A. Anantharaman et X. Blanc).
- [FL29] *Rate of convergence of a two-scale expansion for some "weakly" stochastic homogenization problems*, Asymptotic Analysis, soumis, arXiv 1110.5206 (avec C. Le Bris et F. Thomines).
- [FL30] *Multiscale Finite Element approach for "weakly" random problems and related issues*, Mathematical Modelling and Numerical Analysis, soumis, arXiv 1111.1524 (avec C. Le Bris et F. Thomines).
- [FL31] *Negative thermal conductivity of chains of rotors with mechanical forcing*, Physical Review E, soumis, arXiv 1107.1766 (avec A. Iacobucci, S. Olla et G. Stoltz).

CHAPITRES DE LIVRES

- [FL32] *Multiscale methods coupling atomistic and continuum mechanics : some examples of mathematical analysis*, dans Analytical and Numerical Aspects of Partial Differential Equations, édité par E. Emmrich et P. Wittbold, de Gruyter (2009), 193-245.
- [FL33] *Variance reduction in stochastic homogenization : the technique of antithetic variables*, dans Numerical Analysis and Multiscale Computations, édité par B. Engquist, O. Runborg et R. Tsai, Lect. Notes Comput. Sci. Eng., vol. 82, Springer, 2012 (avec X. Blanc, R. Costauque et C. Le Bris).
- [FL34] *Some remarks on free energy and coarse-graining*, dans Numerical Analysis and Multiscale Computations, édité par B. Engquist, O. Runborg et R. Tsai, Lect. Notes Comput. Sci. Eng., vol. 82, Springer, 2012 (avec T. Lelièvre).
- [FL35] *Introduction to numerical stochastic homogenization and the related computational challenges : some recent developments*, Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore, édité par W. Bao et Q. Du, volume 22, 2011 (avec A. Anantharaman, R. Costauque, C. Le Bris et F. Thomines).

ARTICLES EN PRÉPARATION

- [FL36] *Convergence of the residual process of a variant of stochastic homogenization in dimension one*, en préparation (avec F. Thomines).
- [FL37] *Effective dynamics for the Langevin equation*, en préparation (avec F. Galante et T. Lelièvre).
- [FL38] *Modified equations of parareal algorithms and application to Hamiltonian systems*, en préparation (avec X. Dai).
- [FL39] *Effective dynamics for a kinetic Monte Carlo model*, en préparation (avec S. Lahbabi et S. Olla).

Activités d'encadrement

Mon activité au sein du Laboratoire Navier m'a donné l'opportunité de participer à l'encadrement de plusieurs doctorants :

- Mouhamad Hammoud, thèse ayant pour titre *Modélisation et simulation numérique du couplage entre les milieux discret et continu* (soutenue le 17 décembre 2009 ; directeur de thèse : D. Duhamel) ;

- Kimiya Minoukadeh, thèse ayant pour titre *Méthodes numériques en dynamique moléculaire* (soutenue le 24 novembre 2010 ; directeur de thèse : E. Cancès) ;
- Ronan Costaouec, thèse ayant pour titre *Techniques numériques d'homogénéisation : application aux matériaux aléatoires* (depuis septembre 2008 ; directeur de thèse : C. Le Bris) ;
- Florian Thomines, thèse ayant pour titre *Techniques numériques de changement d'échelle : application à l'homogénéisation numérique des matériaux aléatoires et au couplage discret-continu* (depuis septembre 2009 ; directeur de thèse : C. Le Bris).

1 Simulation moléculaire

On s'intéresse dans cette première partie à des systèmes de N particules ponctuelles en dimension d , de positions $q_i \in \mathbb{R}^d$, de masses m_i et d'impulsions $p_i \in \mathbb{R}^d$. Ces particules, qui interagissent au travers d'un potentiel $V(q_1, \dots, q_N)$, représentent typiquement les atomes d'un système moléculaire. Toute la physique du système est contenue dans l'expression du potentiel V .

L'état du système est complètement caractérisé par la donnée de

$$(q, p) = (q_1, \dots, q_N, p_1, \dots, p_N) \in \mathbb{R}^{dN} \times \mathbb{R}^{dN}.$$

L'énergie totale du système (aussi appelée Hamiltonien), somme de l'énergie potentielle et de l'énergie cinétique, est

$$H(q, p) = V(q_1, \dots, q_N) + \sum_{i=1}^N \frac{p_i^T p_i}{2m_i}. \quad (1)$$

Lorsqu'on s'intéresse à des systèmes à l'échelle moléculaire, le nombre N de particules considéré est très grand (pour les praticiens, $N \geq 10^5$). La question n'est pas de suivre *individuellement* chaque particule (calculer sa trajectoire, ...), mais plutôt de calculer, à l'aide de ce modèle microscopique, très fin, des grandeurs macroscopiques.

On s'intéresse ici à des systèmes moléculaires avec un nombre N de particules connu et fixé, pouvant occuper un volume Ω connu et fixé, et dont l'énergie totale n'est connue qu'en moyenne. On dit alors qu'on travaille dans l'*ensemble canonique*. Dans toute la suite, on suppose que $\Omega = \mathbb{R}^d$, et que $\lim_{|q| \rightarrow \infty} V(q) = +\infty$ assez vite. Ainsi, la mesure (3)

ci-dessous est bien une mesure de probabilité. L'adaptation des résultats au cas où $\Omega = \mathbb{T}^d$, où $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ est le tore en dimension un, ne pose pas de difficultés.

On postule, suivant la mécanique statistique, que les grandeurs macroscopiques à l'équilibre sont données par

$$\int_{\mathbb{R}^{dN} \times \mathbb{R}^{dN}} \Phi(q, p) \mu(q, p) dqdp, \quad (2)$$

où Φ est une fonction de l'état microscopique du système (son choix est dicté par la grandeur macroscopique d'intérêt : pression dans un liquide, ...), et $\mu(q, p) dqdp$ est la mesure de Boltzmann dans l'espace des phases,

$$\mu(q, p) = Z^{-1} \exp(-\beta H(q, p)), \quad (3)$$

où $\beta > 0$ est un paramètre et Z est la constante de normalisation,

$$Z = \int_{\mathbb{R}^{dN} \times \mathbb{R}^{dN}} \exp(-\beta H(q, p)) dqdp.$$

Cette mesure est obtenue en maximisant une certaine fonctionnelle $\mathcal{H}(\psi)$ (appelée entropie) sur toutes les mesures de probabilité $\psi(q, p) dqdp$ satisfaisant la contrainte $\int H\psi = E_0$, où E_0 est l'énergie moyenne connue du système [13, Chap. 4]. Le paramètre β dans (3) est le multiplicateur de Lagrange associé à la contrainte d'énergie moyenne. Ainsi, β et E_0 sont reliés par

$$E_0 = \frac{\int_{\mathbb{R}^{dN} \times \mathbb{R}^{dN}} H(q, p) \exp(-\beta H(q, p)) dqdp}{\int_{\mathbb{R}^{dN} \times \mathbb{R}^{dN}} \exp(-\beta H(q, p)) dqdp}. \quad (4)$$

Dans la suite, on travaillera avec le paramètre β plutôt qu'avec E_0 . La température du système est définie (à une constante multiplicative près) par la moyenne de l'énergie cinétique par degré de liberté. Un calcul simple montre alors que

$$\int_{\mathbb{R}^{dN} \times \mathbb{R}^{dN}} \left(\frac{1}{dN} \sum_{i=1}^N \frac{p_i^T p_i}{2m_i} \right) \mu(q, p) dq dp = \frac{1}{2} \beta^{-1}. \quad (5)$$

Ainsi, β^{-1} peut aussi s'interpréter comme la température du système. Par conséquent, travailler dans l'ensemble canonique (où le nombre N de particules, le volume accessible et l'énergie moyenne E_0 sont fixés) revient à travailler avec la mesure (3), pour laquelle la température du système est connue et vaut β^{-1} , reliée à E_0 par (4). Les quantités (2) sont appelées moyennes dans l'ensemble *canonique*, ou bien à température constante. On prendra garde au fait que cette terminologie ne signifie pas qu'on se restreint à considérer

les configurations (q, p) telles que $\frac{1}{dN} \sum_{i=1}^N \frac{p_i^T p_i}{2m_i} = \frac{1}{2} \beta^{-1}$.

Notre objectif est de calculer (2). Sur le plan numérique, la difficulté centrale est que cette quantité est une intégrale en *très grande dimension* (car le nombre de particules est très grand), ce qui nécessite des méthodes de calcul adaptées.

Remarque 1 *La spécificité de la simulation moléculaire par rapport à d'autres domaines où apparaît le problème à n corps, comme la mécanique céleste par exemple, est le fait qu'on s'intéresse à des moyennes (sur des configurations, comme dans (2), ou bien sur des trajectoires, ...). Au contraire, en mécanique céleste, on s'intéresse à une trajectoire bien particulière, correspondant à une condition initiale particulière.* \diamond

Remarque 2 *Nous n'abordons pas ici la question de la détermination du potentiel V . C'est là tout un champ de recherches. Dans la suite, nous supposons disposer d'une expression analytique (ou au moins économiquement calculable) de V .* \diamond

Remarque 3 *Dans certaines applications, les particules ne représentent pas un unique atome, mais un groupe d'atomes ou de molécules (afin de diminuer le nombre de particules dans le système, et donc la complexité de la résolution numérique). La question de la détermination du potentiel d'interaction n'en est que plus difficile ! Dans d'autres applications (comme en physique de l'état solide), une particule est une lacune (i.e. une absence d'atome).* \diamond

1.1 Echantillonnage de la mesure de Boltzmann

Publications [FL8, FL9, FL14, FL13, FL16]

Le calcul de (2) revient à échantillonner la mesure de Boltzmann dans l'espace des phases. Plusieurs méthodes ont été proposées dans la littérature. Dans l'article de revue [FL8], nous avons comparé ces différentes méthodes, sur les plans théoriques et numériques. On reprend ici les principales conclusions de ce travail.

On remarque tout d'abord que la mesure étudiée est une mesure produit,

$$\mu(q, p) = \pi(q) Z_p^{-1} \exp \left(-\beta \sum_{i=1}^N \frac{p_i^T p_i}{2m_i} \right),$$

avec

$$\pi(q) = Z_q^{-1} \exp(-\beta V(q)), \quad Z_q = \int_{\mathbb{R}^{dN}} \exp(-\beta V(q)) dq, \quad (6)$$

et que l'échantillonnage de la mesure gaussienne en p ne pose pas de difficulté pratique. Le problème revient donc à échantillonner la mesure $\pi(q) dq$, et à calculer

$$\langle \Phi \rangle := \int_{\mathbb{R}^{dN}} \Phi(q) \pi(q) dq \quad (7)$$

pour une observable Φ ne dépendant que des positions des particules.

Remarque 4 Pour une observable $\tilde{\Phi}$ dépendant de q et p , on écrit

$$\int \tilde{\Phi}(q, p) \mu(q, p) dq dp = \int \Phi(q) \pi(q) dq$$

avec $\Phi(q) = Z_p^{-1} \int \tilde{\Phi}(q, p) \exp\left(-\beta \sum_{i=1}^N \frac{p_i^T p_i}{2m_i}\right) dp$. Calculer $\Phi(q)$ ne pose pas de difficulté,

et on est donc ramené au problème précédent, calculer $\int \Phi(q) \pi(q) dq$. \diamond

Pour calculer (7), les méthodes proposées dans la littérature peuvent être classées en quatre catégories :

1. les méthodes de type Monte-Carlo, qui consistent à tirer au hasard des configurations $\{q_m\}_{m \geq 1}$, indépendantes et identiquement distribuées suivant la mesure de Boltzmann $\pi(q) dq$. On calcule alors (7) via une loi des grands nombres,

$$\int_{\mathbb{R}^{dN}} \Phi(q) \pi(q) dq = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M \Phi(q_m). \quad (8)$$

2. les méthodes de type chaînes de Markov : l'idée est de générer une réalisation $\{q_m\}_{m \geq 1}$ d'une chaîne de Markov (les configurations successives sont donc maintenant *corréliées*), qui soit ergodique pour la mesure de Boltzmann $\pi(q) dq$ (cette mesure est alors l'unique mesure invariante de la chaîne). On obtient alors la convergence (8) sous un argument de loi forte des grands nombres pour les chaînes de Markov.

On rappelle qu'une mesure $\psi(q) dq$ est *invariante* par une chaîne de Markov si, lorsque l'itéré q_m de la chaîne est distribué suivant cette mesure, l'itéré suivant q_{m+1} est aussi distribué suivant la même mesure $\psi(q) dq$. La chaîne est *ergodique* pour $\psi(q) dq$ si la loi de q_m , lorsque $m \rightarrow \infty$, converge vers $\psi(q) dq$.

3. les méthodes de type processus de Markov : l'idée est d'introduire un processus $(q_t)_{t \geq 0}$ (*continu en temps*, par opposition au cas précédent), qui soit ergodique pour la mesure de Boltzmann. Les moyennes canoniques sont alors calculées via la relation

$$\int_{\mathbb{R}^{dN}} \Phi(q) \pi(q) dq = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \Phi(q_t) dt. \quad (9)$$

Typiquement, le processus $(q_t)_{t \geq 0}$ est la solution d'une équation différentielle stochastique. Se posent ensuite des questions relatives à la discrétisation numérique.

4. les méthodes déterministes : à l'image du cas précédent, une dynamique en temps continu est introduite. Très souvent, il s'agit d'une dynamique sur les variables (q, p, x) , où x est une variable supplémentaire de petite dimension. On calcule alors la moyenne temporelle de l'observable Φ le long de la trajectoire $(q(t), p(t), x(t))$. Sous une hypothèse d'ergodicité du système dynamique, on a

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \Phi(q(t)) dt = \int_{\mathbb{R}^{dN}} \Phi(q) \pi(q) dq. \quad (10)$$

Dans la suite, nous donnons brièvement un exemple de méthode, pour chacune de ces catégories, et rappelons les arguments pour obtenir la convergence des méthodes d'échantillonnage, au sens de (8) ou (9). On renvoie à l'article de revue [FL8] pour l'énoncé de résultats sur la *vitesse* de convergence, sur la convergence de la distribution de q_m vers la mesure d'équilibre lorsque $m \rightarrow \infty$, et sur la comparaison en pratique de ces différentes méthodes sur un cas concret. La présentation d'une méthode déterministe, la dynamique de Nosé-Hoover, à la Section 1.1.4, sera aussi l'occasion de décrire nos résultats démontrant, au moins dans un cas simple, la *non-ergodicité* de cette méthode, dans un certain régime de paramètres (cf. [FL9, FL14]). Par conséquent, on n'a pas la relation (10), et on ne peut pas utiliser cette dynamique pour calculer une moyenne canonique.

1.1.1 Méthodes purement stochastiques

Un exemple typique pour la première catégorie est la méthode d'échantillonnage par fonction d'importance, qui consiste à écrire

$$\langle \Phi \rangle = \int \Phi(q) \pi(q) dq = \int \left(\Phi(q) \frac{\pi(q)}{g(q)} \right) g(q) dq,$$

pour une mesure de probabilité g facile à échantillonner. Considérant des configurations $\{q_m\}_{m \geq 1}$, indépendantes et identiquement distribuées suivant $g(q) dq$, on obtient, en utilisant la loi forte des grands nombres, que

$$\langle \Phi \rangle = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M \Phi(q_m) \frac{\pi(q_m)}{g(q_m)}.$$

En utilisant le théorème de la limite centrale, on voit que, à M fixé grand, l'intervalle de confiance pour $\langle \Phi \rangle$ est relié à la variance de $\Phi \pi / g$, qui s'écrit

$$\text{Var} \left(\Phi \frac{\pi}{g} \right) = \int \left(\Phi \frac{\pi}{g} \right)^2 g - \left[\int \left(\Phi \frac{\pi}{g} \right) g \right]^2 = \int \Phi^2 \frac{\pi^2}{g} - \left[\int \Phi \pi \right]^2.$$

Pour obtenir l'approximation la plus précise possible de $\langle \Phi \rangle$, à coût calcul fixé, on souhaite choisir g qui minimise cette variance, ce qui conduit au choix

$$g = \frac{|\Phi| \pi}{\int |\Phi| \pi}.$$

On voit donc que, pour choisir g de manière optimale, il faut savoir calculer $\int |\Phi| \pi$, ce qui revient à savoir résoudre le problème initial, échantillonner π . En pratique, le choix de g résulte donc d'un compromis entre la nécessité de choisir une densité de probabilité qu'on sache échantillonner, et le souhait d'obtenir une bonne précision (i.e. un intervalle de confiance petit).

En pratique, pour les problèmes de dynamique moléculaire, ces méthodes sont très vite inefficaces quand la dimension augmente. Typiquement, pour la méthode décrite ci-dessus, il devient très difficile, en grande dimension, de trouver une fonction g facile à échantillonner et telle que l'intervalle de confiance obtenu soit petit.

1.1.2 Chaînes de Markov

Une classe générale de méthodes entrant dans cette seconde catégorie est celle des algorithmes de Metropolis-Hastings [115, 87]. On se donne une fonction $\mathcal{P}(q, q')$ définie sur $\mathbb{R}^{dN} \times \mathbb{R}^{dN}$ telle que, pour tout $q \in \mathbb{R}^{dN}$, la mesure $\mathcal{P}(q, q') dq'$ soit une mesure de probabilité facile à échantillonner. On considère alors la chaîne de Markov construite ainsi :

Algorithme 5 (Algorithme de Metropolis-Hastings) On se donne une configuration initiale q_0 . Pour $m \geq 0$,

1. on génère une configuration $\tilde{q} \in \mathbb{R}^{dN}$ distribuée suivant la mesure $\mathcal{P}(q_m, \tilde{q}) d\tilde{q}$;
2. soit

$$r = \min \left(1, \frac{\pi(\tilde{q})\mathcal{P}(\tilde{q}, q_m)}{\pi(q_m)\mathcal{P}(q_m, \tilde{q})} \right).$$

On tire un nombre aléatoire s distribué suivant la loi uniforme $\mathcal{U}[0, 1]$, et on définit q_{m+1} par

$$q_{m+1} = \tilde{q} \text{ si } s \leq r, \quad q_{m+1} = q_m \text{ sinon.}$$

La configuration \tilde{q} est ainsi acceptée avec la probabilité r .

On remarque que, pour évaluer r , il n'est pas nécessaire de connaître la constante de normalisation Z_q de π (cf. (6)), qui est en pratique difficilement calculable.

L'intérêt de cette classe d'algorithme est que, par construction, la mesure $\pi(q) dq$ est une mesure invariante de la chaîne de Markov ainsi construite, et ce quelle que soit la fonction de proposition \mathcal{P} . Un exemple consiste à choisir \tilde{q} en fonction de q_m suivant

$$\tilde{q} = q_m - \Delta t \nabla V(q_m) + \sqrt{2\Delta t \beta^{-1}} G_m, \quad (11)$$

où $\Delta t > 0$ est un paramètre numérique, et $G_m \in \mathbb{R}^{dN}$ est un vecteur dont toutes les composantes sont aléatoires, indépendantes les unes des autres, et distribuées suivant la loi normale centrée réduite. La motivation de ce choix apparaîtra clairement à la Section 1.1.3. L'algorithme ainsi obtenu est connu sous le nom de "Metropolis-adjusted Langevin algorithm" (MALA).

Pour montrer un résultat de type loi forte des grands nombres pour une chaîne de Markov, le théorème suivant est très utile :

Théorème 6 ([116, Théorème 17.1.7]) On considère une chaîne de Markov de noyau de transition $P(q, \cdot)$, vérifiant la propriété d'accessibilité suivante :

$$\forall q \in \mathbb{R}^{dN}, \quad \forall B \in \mathcal{B}(\mathbb{R}^{dN}), \quad \mu^{\text{Leb}}(B) > 0 \Rightarrow P(q, B) > 0, \quad (12)$$

où $\mathcal{B}(\mathbb{R}^{dN})$ est la tribu borélienne de \mathbb{R}^{dN} et μ^{Leb} est la mesure de Lebesgue sur \mathbb{R}^{dN} .

On suppose de plus qu'il existe une mesure $\pi(q) dq$ telle que

$$\pi(q) dq \text{ soit une mesure de probabilité laissée invariante par la chaîne.} \quad (13)$$

Sous les conditions (12) et (13), on a, pour toute fonction $\Phi \in L^1(\pi)$,

$$\lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M \Phi(q_m) = \int_{\mathbb{R}^{dN}} \Phi(q) \pi(q) dq \quad \text{presque sûrement,}$$

et ce pour π -presque toute condition initiale $q_0 \in \mathbb{R}^{dN}$.

On rappelle que, par définition du noyau de transition d'une chaîne de Markov, $P(q, B)$ est la probabilité que q' soit dans $B \in \mathcal{B}(\mathbb{R}^{dN})$, où q' est donnée par une itération de la chaîne de Markov partant de $q \in \mathbb{R}^{dN}$.

1.1.3 Un exemple de processus de Markov : la dynamique de Langevin amortie

Un exemple simple de méthode *continue* en temps est la dynamique de Langevin amortie,

$$dq_t = -\nabla V(q_t)dt + \sqrt{2\beta^{-1}} dW_t, \quad (14)$$

où $(W_t)_{t \geq 0}$ est un mouvement brownien en dimension dN . L'équation de Fokker-Planck associée à (14) est

$$\partial_t \psi = \operatorname{div}(\psi \nabla V) + \beta^{-1} \Delta \psi, \quad (15)$$

où, pour tout t , $\psi(t, q)$ est la densité de la variable aléatoire q_t . On vérifie aisément que $\pi(q) = Z_q^{-1} \exp(-\beta V(q))$ est une solution stationnaire de (15).

Pour établir un résultat de type (9), i.e. la convergence des moyennes temporelles le long d'une solution de (14) vers la moyenne canonique, on utilise des arguments très similaires à ceux du cas discret : propriété d'accessibilité et existence d'une mesure invariante. On renvoie à [FL8] pour plus de détails.

Un schéma très simple pour l'intégration numérique de (14) est le schéma d'Euler-Maruyama, qui est une généralisation du schéma d'Euler explicite pour les équations différentielles ordinaires : notant Δt le pas de temps, ce schéma s'écrit

$$q_{m+1} = q_m - \Delta t \nabla V(q_m) + \sqrt{2\Delta t \beta^{-1}} G_m, \quad (16)$$

où $(G_m)_{m \in \mathbb{N}}$ est une suite indépendante de vecteurs gaussiens standards en dimension dN .

On obtient ainsi une chaîne de Markov (en temps *discret*), dont l'ergodicité et les propriétés en temps long s'étudient via les outils évoqués à la Section 1.1.2. On peut espérer retrouver sur cette chaîne de Markov des propriétés vraies sur le processus (14), puisque (16) en est une approximation. C'est en effet le cas : sous de bonnes hypothèses, on montre que la chaîne de Markov formalisant (16) admet une unique mesure invariante $\pi^{\Delta t}(q) dq$, qu'elle est ergodique pour cette mesure [114, Théorème 7.3], et que cette mesure est proche, en un certain sens, de $\pi(q) dq$, l'unique mesure invariante de (14) (cf. [139]).

Ainsi, l'erreur commise après M pas de temps de taille Δt s'écrit comme la somme d'une erreur statistique et d'une erreur systématique (un biais) :

$$\frac{1}{M} \sum_{m=1}^M \Phi(q_m) - \int_{\mathbb{R}^{dN}} \Phi d\pi = \left[\frac{1}{M} \sum_{m=1}^M \Phi(q_m) - \int_{\mathbb{R}^{dN}} \Phi d\pi^{\Delta t} \right] + \left[\int_{\mathbb{R}^{dN}} \Phi d\pi^{\Delta t} - \int_{\mathbb{R}^{dN}} \Phi d\pi \right].$$

L'erreur statistique décroît typiquement comme $M^{-1/2}$, tandis que le biais, dans le cas de (16), décroît comme Δt . En général, le premier terme d'erreur domine le second.

Notons qu'il est possible de corriger le biais, en utilisant l'algorithme de Euler-Maruyama comme une fonction de proposition dans un algorithme de Metropolis-Hastings [131]. On obtient ainsi l'algorithme "Metropolis-adjusted Langevin algorithm" (MALA), déjà rencontré à la section précédente, et qui permet d'échantillonner sans biais la mesure $\pi(q) dq$.

Remarque 7 *La dynamique stochastique la plus utilisée par les physiciens est plutôt la dynamique de Langevin,*

$$\begin{cases} dq_t &= M^{-1} p_t dt, \\ dp_t &= -\nabla V(q_t) dt - \xi M^{-1} p_t dt + \sqrt{2\xi\beta^{-1}} dW_t, \end{cases} \quad (17)$$

où M est la matrice de masse (diagonale dans les cas simples), $\xi > 0$ est un paramètre, et $(W_t)_{t \geq 0}$ est un mouvement brownien en dimension dN . Lorsque $\xi = 0$, on retrouve

les équations de Newton. Dans la limite $\xi \rightarrow \infty$, on retrouve l'équation de Langevin amortie (14), après rescaling en temps (on retrouve aussi (14) dans la limite $M \rightarrow 0$, avec $\xi = 1$) [105, Section 2.2.4].

Pour tout $\xi > 0$, la mesure de Boltzmann dans l'espace des phases, $\mu(q,p)dqdp$, donnée par (3), est une mesure invariante de la dynamique. L'étude de la convergence des moyennes temporelles (ou du comportement en temps long de la densité de (q_t, p_t)) est rendue difficile par l'absence de bruit dans la première ligne de (17). L'équation de Fokker-Planck associée à (17) s'écrit $\partial_t \psi = \mathcal{L}^* \psi$, où l'opérateur \mathcal{L}^* n'est pas elliptique. \diamond

Remarque 8 En pratique, la convergence de la moyenne temporelle vers la moyenne d'espace, donnée par la relation (9), est très lente. Un résultat typique est que

$$\left| \int_{\mathbb{R}^{dN}} \Phi(q) \pi(q) dq - \frac{1}{T} \int_0^T \Phi(q_t) dt \right| \leq \frac{C}{\sqrt{T}}, \quad (18)$$

où q_t est la solution de (14) et C est une constante indépendante de T , mais en générale grande. Il faut donc des temps T de simulation très grands pour obtenir une précision raisonnable. Cette observation demeure si on remplace la dynamique (14) par la dynamique de Langevin (17). Ceci est dû au fait que, dans beaucoup de cas pratiques, l'énergie potentielle du système présente de nombreux minima locaux, séparés les uns des autres par de hautes barrières d'énergie. Ainsi, la dynamique du système passe beaucoup de temps au voisinage d'un minimum local avant de passer dans un autre puits. Le temps typique nécessaire pour avoir convergence dans (9) est donc de l'ordre du temps de résidence dans les puits (afin que, pendant le temps T , plusieurs puits aient pu être explorés), qui est de l'ordre de 10^{-6} , voire 10^{-3} seconde. Au contraire, le pas de temps d'intégration de la dynamique est lié aux temps caractéristiques les plus rapides, qui sont de l'ordre de 10^{-15} seconde, soit plusieurs ordres de grandeurs plus petits que le temps de résidence dans un puits.

Pour de tels systèmes, il est très difficile d'accéder à des temps de simulation T suffisamment grands pour que la borne dans (18) soit petite. Il existe une importante littérature visant à apporter des réponses à cette problématique. On renvoie à [FL13, FL16] pour nos contributions sur ce sujet. \diamond

1.1.4 Un exemple de méthode déterministe : la dynamique de Nosé-Hoover

Publications [FL9, FL14]

Toutes les méthodes déterministes proposées dans la littérature et visant à échantillonner l'ensemble canonique consistent à introduire, en plus des variables q et p (position et impulsion des particules), un petit nombre de variables supplémentaires x , et à écrire une dynamique sur cet ensemble étendu, s'inspirant des équations de Newton. Toutes ces dynamiques admettent une mesure invariante $\rho(q, p, x) dqdpdx$ dont la marginale, en (q, p) , est la mesure de Boltzmann :

$$\int \rho(q, p, x) dx = \mu(q, p),$$

avec μ définie par (3). De nombreuses dynamiques ont été proposées, vérifiant les (faibles) contraintes ci-dessus. Nous nous concentrons dans la suite sur l'une d'entre elles, la dynamique de Nosé-Hoover, qui est historiquement la première proposée. La plupart des autres dynamiques ont été construites sur ce modèle, et, sur le plan théorique, partagent les mêmes carences.

La dynamique de Nosé-Hoover [123, 90] s'écrit

$$\begin{aligned}\dot{q}_i &= \frac{p_i}{m_i}, \quad 1 \leq i \leq N, \\ \dot{p}_i &= -\nabla_{q_i} V(q) - \frac{\xi}{Q} p_i, \quad 1 \leq i \leq N, \\ \dot{\xi} &= \sum_{i=1}^N \frac{p_i^T p_i}{m_i} - dN\beta^{-1},\end{aligned}\tag{19}$$

où $(q, p, \xi) \in \mathbb{R}^{dN} \times \mathbb{R}^{dN} \times \mathbb{R}$. On a donc rajouté une unique variable scalaire ξ en plus de (q, p) . Dans ce système, $Q > 0$ est un paramètre de la méthode.

Au vu de la relation (5), on voit que le membre de droite de l'équation en ξ est égal à la différence entre (deux fois) l'énergie cinétique instantanée et sa valeur moyenne dans l'ensemble canonique. Cette différence contrôle l'évolution de ξ , qui peut être interprété comme un paramètre de friction (positif ou négatif) dans l'équation en p , ce qui a pour effet de diminuer ou d'augmenter l'énergie totale du système.

Un calcul simple montre que la mesure

$$d\rho_{\text{NH}}(q, p, \xi) = Z_{\text{NH}}^{-1} \exp\left(-\beta \left[H(q, p) + \frac{\xi^2}{2Q} \right]\right) dq dp d\xi,$$

où Z_{NH} est une constante de normalisation, est une mesure invariante de (19).

Remarque 9 *En plus de ξ , on introduit parfois la variable η donnée par*

$$\dot{\eta} = \frac{\xi}{Q}.\tag{20}$$

Un calcul simple montre alors que la dynamique (19)-(20) conserve la quantité

$$\tilde{H}_{\text{NH}} = \sum_{i=1}^N \frac{p_i^T p_i}{2m_i} + V(q) + \frac{\xi^2}{2Q} + dN\beta^{-1}\eta,\tag{21}$$

ainsi que la mesure $d\mu_{\text{NH}} = \exp(dN\eta) dq dp d\eta d\xi$. Le fait que la dynamique conserve (21) est ensuite souvent utilisé pour déterminer empiriquement un pas de temps d'intégration raisonnable pour (19)-(20). Cette idée est discutée dans [49]. \diamond

On rappelle que, par définition, le flot du système dynamique (19) est *ergodique* par rapport à la mesure $d\rho_{\text{NH}}$ si l'espace des phases, \mathbb{R}^{2dN+1} , ne peut pas être décomposé sous la forme $\mathbb{R}^{2dN+1} = A \cup B$, où les deux ensembles A et B , disjoints, sont invariants par la dynamique (19) et de mesure (par rapport à $d\rho_{\text{NH}}$) strictement positive [98].

Si le flot de (19) est ergodique par rapport à $d\rho_{\text{NH}}$, alors, pour tout $\Phi \in L^1(d\rho_{\text{NH}})$, le théorème ergodique de Birkhoff montre que, pour presque toute condition initiale,

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \Phi(q(t), p(t), \xi(t)) dt = \int_{\mathbb{R}^{2dN+1}} \Phi(q, p, \xi) d\rho_{\text{NH}},\tag{22}$$

où $(q(t), p(t), \xi(t))$ est une solution des équations de Nosé-Hoover (19). Dans le cas où $\Phi = \Phi(q, p)$ ne dépend pas de ξ , on voit immédiatement que (22) implique

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \Phi(q(t), p(t)) dt = \int_{\mathbb{R}^{2dN}} \Phi(q, p) d\mu,\tag{23}$$

et on retrouve ainsi les moyennes canoniques comme moyennes trajectorielles le long d'une solution des équations de Nosé-Hoover.

La dérivation ci-dessus a été faite sous l'hypothèse que la dynamique de Nosé-Hoover est ergodique pour la mesure $d\rho_{\text{NH}}$. Cependant, de nombreux résultats numériques sont en contradiction avec l'égalité (23) (même pour des temps T très longs), ce qui contredit l'hypothèse d'ergodicité. En particulier, les auteurs de [90, 113, 144] observent que, dans le cas d'un oscillateur harmonique mono-dimensionnel, i.e. lorsque

$$d = 1, N = 1, m_1 = 1, V(q) = \frac{1}{2}q^2$$

dans (19), il existe c et C tels que la solution de (19) vérifie

$$\forall t, \quad 0 < c \leq q^2(t) + p^2(t) \leq C, \quad (24)$$

et ce pour une grande plage de valeurs de Q et de β . Ceci est en contradiction avec (23).

Dans [FL9], nous avons *démontré* que, dans le cas de l'oscillateur harmonique mono-dimensionnel, pour des valeurs de Q assez grandes, la dynamique (19) n'est pas ergodique pour $d\rho_{\text{NH}}$, et que, de plus, il existe effectivement c et C tels que (24) soit vérifié pour tout t . Ce résultat a ensuite été étendu à une classe de systèmes plus grande dans [FL14], en suivant des arguments similaires. Ces travaux montrent que la dynamique (19) ne permet *pas* de calculer des moyennes canoniques, du moins si Q est choisi trop grand.

Pour contourner cette difficulté, plusieurs généralisations des équations de Nosé-Hoover ont été proposées (la méthode des chaînes de Nosé-Hoover [113, 144], qui consiste à coupler la variable ξ dans (19) avec un second thermostat, lui même possiblement couplé avec un autre thermostat, ... ; la dynamique de Nosé-Poincaré [28], qui est une généralisation *Hamiltonienne* de la dynamique de Nosé-Hoover ; ...). Dans tous les cas, on peut exhiber une mesure invariante de ces dynamiques, dont la marginale sur les variables (q, p) est la mesure de Boltzmann dans l'espace des phases. Cependant, l'exemple de la dynamique de Nosé-Hoover confirme l'intuition naturelle que cette propriété est loin de suffire pour obtenir l'ergodicité du système dynamique. Il est donc indispensable de garder la plus grande prudence lorsqu'on travaille avec ces méthodes déterministes.

Indépendamment des questions d'ergodicité, difficiles dans le cadre déterministe, ces méthodes soulèvent des questions intéressantes sur le plan numérique. Ce sont en effet des dynamiques qui préservent une mesure $d\rho$. Comme on l'a vu, cette préservation est importante pour espérer montrer une convergence du type (23). Il est donc intéressant de construire des algorithmes qui préservent aussi cette mesure. Dans [FL1], nous avons étudié cette question, pour un exemple d'équations généralisant les équations de Nosé-Hoover. Cette observation motive aussi le fait de construire des dynamiques Hamiltoniennes. En effet, il existe une littérature importante concernant la discrétisation de ce type d'équations, caractérisant les propriétés en temps long des trajectoires numériques, ... Nos contributions sur le sujet des équations Hamiltoniennes sont rassemblées dans la Section 1.3.

Nous concluons cette section en décrivant brièvement les arguments utilisés dans [FL9] pour montrer la non-ergodicité de (19), dans le cas de l'oscillateur harmonique mono-dimensionnel, et pour Q assez grand. L'argument central est d'utiliser la théorie KAM [136] pour montrer l'existence de tores invariants qui permettent de décomposer l'espace des phases en ensembles invariants de mesure positive. La projection de ces ensembles sur le plan (q, p) vérifie une inégalité du type (24).

Ecrivons les équations (19), pour un oscillateur harmonique monodimensionnel, en prenant $\beta = 1$. Dans les variables action-angle $(\hat{\tau}, \theta)$, reliées à (q, p) par

$$q = \sqrt{2\hat{\tau}} \cos \theta \quad \text{et} \quad p = -\sqrt{2\hat{\tau}} \sin \theta,$$

on obtient

$$\begin{aligned}\dot{\theta} &= 1 - \varepsilon \hat{\alpha} \sin \theta \cos \theta, \\ \dot{\hat{\tau}} &= -2\varepsilon \hat{\tau} \hat{\alpha} \sin^2 \theta, \\ \dot{\hat{\alpha}} &= \varepsilon(2\hat{\tau} \sin^2 \theta - 1),\end{aligned}\tag{25}$$

avec $\varepsilon = 1/\sqrt{Q}$ et $\hat{\alpha} = \varepsilon \xi$. Dans le régime $\varepsilon \ll 1$ (i.e. $Q \gg 1$), on voit que θ est une variable plus rapide que $\hat{\tau}$ ou $\hat{\alpha}$. Suivant la méthode de l'averaging [136], on introduit τ et α définis par

$$\hat{\tau} = \tau + \varepsilon \tau \alpha \sin \theta \cos \theta \quad \text{et} \quad \hat{\alpha} = \alpha - \varepsilon \tau \sin \theta \cos \theta,$$

et le système (25) se récrit comme

$$\begin{aligned}\dot{\theta} &= 1 - \varepsilon \alpha \sin \theta \cos \theta + O(\varepsilon^2), \\ \dot{\tau} &= -\varepsilon \tau \alpha + O(\varepsilon^2), \\ \dot{\alpha} &= \varepsilon(\tau - 1) + O(\varepsilon^2).\end{aligned}\tag{26}$$

On reconnaît dans l'équation en τ et α la moyenne par rapport à θ des termes de l'équation en $\hat{\tau}$ et $\hat{\alpha}$.

On s'intéresse alors à l'application de premier retour P_ε , dans le plan $\Sigma = \{(\theta, \tau, \alpha) : \theta = 0 \bmod 2\pi\}$, pour l'équation (26). Partant de la condition initiale $(0, \tau, \alpha)$, on suit ainsi le flot de (26) jusqu'au premier temps $t_{\text{retour}} > 0$ tel que $\theta(t_{\text{retour}}) = 0 \bmod 2\pi$. Par définition, on a alors $P_\varepsilon(\tau, \alpha) = (\tau(t_{\text{retour}}), \alpha(t_{\text{retour}}))$.

Au vu de (26), on voit que $t_{\text{retour}} = 2\pi + O(\varepsilon)$, et donc que $P_\varepsilon(\tau, \alpha) = Q_\varepsilon(\tau, \alpha) + O(\varepsilon^2)$, où Q_ε est le flot exact, pendant le temps 2π , du système

$$\dot{\tau} = -\varepsilon \tau \alpha, \quad \dot{\alpha} = \varepsilon(\tau - 1).\tag{27}$$

La suite de l'argument s'appuie sur le fait que la dynamique (27) admet comme invariant

$$G(\tau, \alpha) = \tau - \ln \tau + \frac{1}{2}\alpha^2 - 1.\tag{28}$$

Par conséquent, pour tout G_0 , la courbe

$$\Gamma_Q := \{(\tau, \alpha) \in \mathbb{R}^2; G(\tau, \alpha) = G_0\}$$

est une courbe invariante de l'application Q_ε (cf. la Figure 1).

En utilisant le théorème de la courbe invariante de Siegel et Moser [136, sections 32–34], on peut alors montrer que, pour ε assez petit, l'application P_ε , qui est une perturbation de l'application Q_ε , possède elle-aussi des courbes invariantes, qui sont des perturbations des courbes invariantes de Q_ε . Soit Γ_P une telle courbe invariante de P_ε : on sait qu'elle est proche d'une courbe invariante Γ_Q de Q_ε , et donc qu'il existe C et G_0 tels que

$$(\tau, \alpha) \in \Gamma_P \implies |G(\tau, \alpha) - G_0| \leq C\varepsilon.$$

Le long de Γ_P , on a donc

$$G(\tau, \alpha) = \tau - \ln \tau + \frac{1}{2}\alpha^2 - 1 \leq G_0 + C\varepsilon,$$

ce qui implique $\tau - \ln \tau \leq \overline{C}$. Comme $\tau = \hat{\tau} + O(\varepsilon)$ et $\hat{\tau} = (q^2 + p^2)/2$, on obtient l'existence de c et C tels que, le long de Γ_P , on ait

$$0 < c \leq q^2 + p^2 \leq C.$$

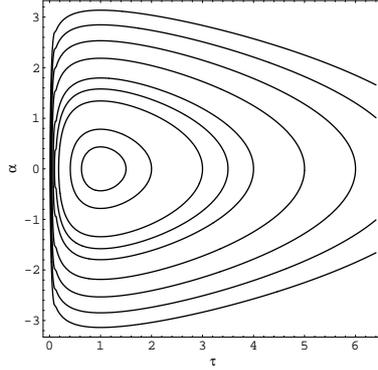


FIG. 1 – Courbes invariantes pour la dynamique (27), définies comme les courbes iso-valeurs de G défini par (28).

Par construction, Γ_P est une courbe invariante de l'application de premier retour P_ε associée à la dynamique de Nosé-Hoover, écrite sous la forme (26). Par conséquent, en considérant l'ensemble (pour tous les (τ, α) appartenant à une courbe invariante Γ_P) des trajectoires de (26) partant de la condition initiale $(0, \tau, \alpha)$ et arrivant au point $(2\pi, P_\varepsilon(\tau, \alpha))$, on forme un tube dans l'espace \mathbb{R}^3 (cf. la Figure 2), invariant par le flot des équations de Nosé-Hoover, et compris dans un cylindre

$$\{(q, p, \tau) \in \mathbb{R}^3; 0 < c \leq q^2 + p^2 \leq C\}.$$

Ceci confirme bien les observations numériques (certaines solutions des équations de Nosé-Hoover vérifient (24)), du moins dans le régime $Q \gg 1$.

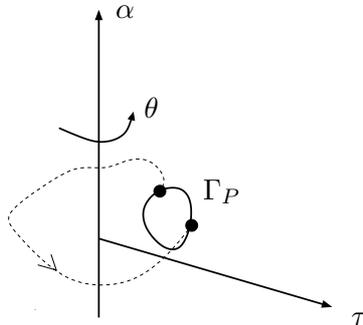


FIG. 2 – On suit le flot de (26) partant d'une condition initiale avec $\theta = 0$ (la trajectoire est représentée en pointillé), jusqu'à revenir sur le plan $\theta = 0 \pmod{2\pi}$. Si la condition initiale appartient à une courbe invariante Γ_P de P_ε (représentée en trait plein), alors le point final aussi. La réunion des trajectoires ainsi définies, lorsque la condition initiale décrit Γ_P , forme un tube dans l'espace \mathbb{R}^3 , que les trajectoires solutions de (26) ne croisent pas : elles sont soit à l'intérieur du tube, soit à l'extérieur.

On trace sur la Figure 3 les orbites de l'application de premier retour P_ε , pour deux valeurs de ε . On constate que, dans le cas $\varepsilon = 0.1$, l'application P_ε possède des courbes invariantes proches des courbes invariantes de Q_ε tracées sur la Figure 1. Dans le cas $\varepsilon = 1$, il semble que l'application P_ε possède encore plusieurs courbes invariantes. Dans ce cas, la dynamique de Nosé-Hoover n'est pas ergodique, même avec le choix $\varepsilon = 1$.

L'analyse ci-dessus, qui fait l'objet de l'article [FL9], est restreinte au cas de l'oscillateur harmonique mono-dimensionnel. Nous l'avons ensuite étendue au cas des systèmes

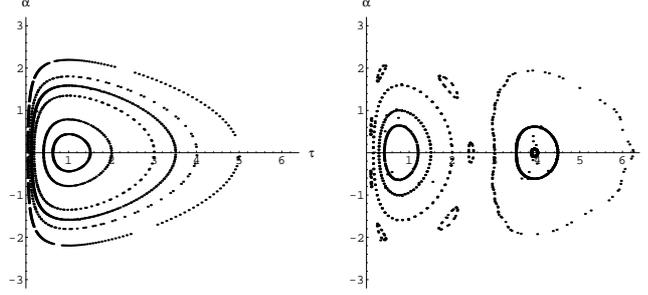


FIG. 3 – Orbites de l’application de premier retour P_ε de la dynamique (25) dans le plan $\theta = 0 \bmod 2\pi$, pour $\varepsilon = 0.1$ (à gauche) et $\varepsilon = 1$ (à droite).

Hamiltoniens complètement intégrables dans [FL14]. La preuve suit les mêmes arguments : on change de variables dans (19) pour écrire la dynamique dans les variables action-angle, on utilise à nouveau la méthode de l’averaging, et on identifie explicitement des invariants dans les équations de Nosé-Hoover moyennées.

1.2 Dynamique effective

Publications [FL23, FL34, FL37]

1.2.1 Motivation

En simulation moléculaire, il est très fréquent qu’on ne soit pas intéressé par toutes les coordonnées du système, mais seulement par quelques fonctions bien choisies de ces coordonnées. C’est typiquement le cas pour un alcane (ou une protéine), dont la conformation peut essentiellement être décrite par la connaissance de quelques angles dans le système. En effet, pour ce type de molécules, l’énergie potentielle possède plusieurs minima locaux, et la connaissance de quelques angles suffit pour caractériser le puits dans lequel le système se trouve.

Cette observation est en fait générique. Dans beaucoup de cas pratiques, l’énergie potentielle V du système présente de nombreux minima locaux, séparés les uns des autres par de hautes barrières d’énergie. Ainsi, le système, lorsqu’il suit la dynamique (14) (cette constatation est aussi vraie pour la dynamique de Langevin (17), ou une dynamique déterministe telle que (19)), passe beaucoup de temps au voisinage d’un minimum local de V avant de passer à un autre puits (cf. la courbe schématique de la Figure 4). Une quantité intéressante est alors le temps de résidence dans un bassin d’énergie potentielle (défini comme le temps moyen que le système passe dans ce bassin avant d’aller dans un autre bassin).

Dans la suite, on note $x \in \mathbb{R}^n$ l’état du système, et on suppose disposer d’une *coordonnée de réaction*, qui est une fonction

$$\xi : \mathbb{R}^n \rightarrow \mathbb{R},$$

contenant toute l’information d’intérêt. Par hypothèse, on suppose que ξ est telle que le système x est dans un bassin donné de V si et seulement si $\xi(x)$ est dans un certain intervalle. Il est donc suffisant de connaître l’évolution en temps de $\xi(X_t)$ pour connaître la suite des puits visités, le temps de résidence dans chaque puits, ...

On est donc bien dans le cas introduit ci-dessus où on ne s’intéresse en pratique qu’à une fonction (ici à valeur scalaire) de l’état du système. Pour certaines systèmes, choisir

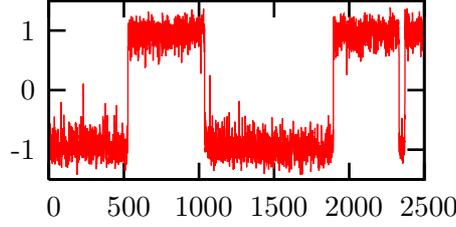


FIG. 4 – Trajectoire schématique $x(t)$ d'un système moléculaire métastable, avec deux configurations préférentielles, $x = \pm 1$.

cette fonction est facile. Pour d'autres, c'est beaucoup plus difficile, et cela nécessite une bonne connaissance intuitive du problème.

Remarque 10 *On a choisi ci-dessus une fonction ξ à valeur dans \mathbb{R} . Tous les résultats s'étendent au cas où ξ prend ses valeurs dans le tore mono-dimensionnel $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Par contre, il n'est pas clair que ces résultats s'étendent facilement au cas où ξ est à valeur vectorielle.* \diamond

A cette fonction ξ est naturellement associée une énergie effective A , appelée *énergie libre*, telle que

$$d(\xi \star \pi) = \exp(-\beta A(z)) dz,$$

où $\xi \star \pi$ est l'image par ξ de la mesure π définie par (6). Ainsi, par définition, pour toute fonction test $\Phi : \mathbb{R} \rightarrow \mathbb{R}$,

$$\int_{\mathbb{R}^n} \Phi(\xi(x)) Z^{-1} \exp(-\beta V(x)) dx = \int_{\mathbb{R}} \Phi(z) \exp(-\beta A(z)) dz.$$

Quand X est un variable aléatoire distribuée suivant la mesure de Boltzmann (6), la coordonnée de réaction $\xi(X)$ est distribuée suivant $\exp(-\beta A(z)) dz$. Obtenir une description réduite du système, avec l'objectif de calculer des moyennes thermodynamiques, revient donc à calculer l'énergie libre A . Plusieurs méthodes existent, on renvoie à la littérature (cf. par exemple [38, 105]). On se contente ici de rappeler les expressions de l'énergie libre A et de sa dérivée (cf. [42]), qui seront utiles pour la suite :

$$\begin{aligned} A(z) &= -\beta^{-1} \ln \left(\int_{\Sigma_z} Z^{-1} \exp(-\beta V) |\nabla \xi|^{-1} d\sigma_{\Sigma_z} \right), \\ A'(z) &= \int_{\Sigma_z} F d\pi_{\Sigma_z}, \end{aligned} \quad (29)$$

où F est donné par

$$F = \frac{\nabla V \cdot \nabla \xi}{|\nabla \xi|^2} - \beta^{-1} \operatorname{div} \left(\frac{\nabla \xi}{|\nabla \xi|^2} \right). \quad (30)$$

Dans ces expressions, Σ_z est la sous-variété

$$\Sigma_z = \{x \in \mathbb{R}^n; \xi(x) = z\}, \quad (31)$$

et π_{Σ_z} est la mesure de probabilité π (définie par (6)) conditionnée à la valeur z de la coordonnée de réaction :

$$d\pi_{\Sigma_z} = \frac{\exp(-\beta V) |\nabla \xi|^{-1} d\sigma_{\Sigma_z}}{\int_{\Sigma_z} \exp(-\beta V) |\nabla \xi|^{-1} d\sigma_{\Sigma_z}}, \quad (32)$$

où σ_{Σ_z} est la mesure de Lebesgue sur Σ_z induite par la mesure de Lebesgue dans \mathbb{R}^n . Par définition de σ_{Σ_z} , on a la propriété suivante, dite de la co-aire [67] : pour toute fonction régulière $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$\int_{\mathbb{R}^n} \Phi(x) |\nabla \xi(x)| dx = \int_{\mathbb{R}} \left[\int_{\Sigma_z} \Phi d\sigma_{\Sigma_z} \right] dz. \quad (33)$$

Remarque 11 *Considérons le cas $\xi(x_1, x_2) = x_2$, où $x = (x_1, x_2) \in \mathbb{R}^2$. L'énergie libre associée $A(x_2)$ vérifie*

$$\exp(-\beta A(x_2)) = Z^{-1} \int_{\mathbb{R}} \exp(-\beta V(x_1, x_2)) dx_1,$$

et on a

$$A'(x_2) = \frac{\int_{\mathbb{R}} \partial_{x_2} V(x_1, x_2) \exp(-\beta V(x_1, x_2)) dx_1}{\int_{\mathbb{R}} \exp(-\beta V(x_1, x_2)) dx_1}.$$

Dans ce cadre simple, la force F définie par (30) vaut $F(x_1, x_2) = \partial_{x_2} V(x_1, x_2)$, σ_{Σ_z} est la mesure de Lebesgue dans la variable $x_1 \in \mathbb{R}$, et la formule de la co-aire (33) est le théorème de Fubini. \diamond

On s'intéresse ici à l'obtention d'une description réduite du système, donnant accès à certaines propriétés *dynamiques*. Un exemple est le temps de résidence dans un bassin d'énergie potentielle, comme décrit ci-dessus. Calculer ce temps à partir du modèle complet consiste à calculer la trajectoire $t \mapsto X_t$ (simuler l'ensemble du système), puis analyser la projection $\xi(X_t)$ de cette trajectoire. Nous avons proposé dans [FL23] une dynamique mono-dimensionnelle qui approche, sous certaines hypothèses, le processus $t \mapsto \xi(X_t)$, et qui permet donc de calculer les temps de résidence en ne considérant qu'une dynamique en dimension un.

Pour rendre la question précise, il faut *choisir* une dynamique de référence sur X_t . Nous avons considéré la dynamique de Langevin amortie, déjà décrite à la Section 1.1.3, et qui s'écrit (cf. (14))

$$dX_t = -\nabla V(X_t) dt + \sqrt{2\beta^{-1}} dW_t. \quad (34)$$

On rappelle que cette dynamique est ergodique pour la mesure de Boltzmann (6). Dans ce cadre, on cherche à approcher (dans un sens à préciser) $\xi(X_t)$ par y_t solution de

$$dy_t = b(y_t) dt + \sqrt{2\beta^{-1}} \sigma(y_t) dB_t, \quad (35)$$

où B_t est un mouvement brownien mono-dimensionnel, et b et σ sont deux fonctions à valeurs scalaires. Ainsi, la dynamique (35) peut être considérée comme une dynamique *effective* pour la quantité d'intérêt $\xi(X_t)$. On demande bien sûr que (35) préserve les propriétés d'équilibre de $\xi(X_t)$, i.e. soit ergodique pour la mesure d'équilibre de $\xi(X)$, qui est $\exp(-\beta A(z)) dz$. Mais on souhaite obtenir plus que cela, et par exemple reproduire des propriétés dynamiques de $\xi(X_t)$ via y_t .

La difficulté de cette question vient du fait que, en général, le processus $t \rightarrow \xi(X_t)$ n'est pas un processus de Markov. Autrement dit, il ne suffit pas de connaître $\xi(X_{t_0})$ pour déterminer l'évolution de $\xi(X_t)$ pour $t \geq t_0$: en général, il faut connaître toute la configuration X_{t_0} . Une approche pour traiter ce problème est celle de Mori-Zwanzig, qui conduit à approcher $\xi(X_t)$ par un processus à mémoire, via une équation de Langevin généralisée (on renvoie à la littérature pour les détails, par exemple [48, 91, 39, 40, 86, 75]).

Dans la suite, nous allons faire une hypothèse de séparation des échelles de temps. On peut alors espérer, et c'est en effet le cas, que les effets de mémoire soient négligeables,

et qu'on puisse approcher $\xi(X_t)$ par un processus de Markov du type (35). Une telle hypothèse de séparation des échelles de temps peut être formalisée de plusieurs manières. Nous avons choisi de l'écrire dans le langage des inégalités de Sobolev logarithmiques, ce qui permet d'obtenir *in fine* une estimation d'erreur

- valable pour toute coordonnée de réaction ξ (et pas uniquement $\xi(x) = x_j$ pour $1 \leq j \leq n$, i.e. $\xi(x)$ est une des coordonnées cartésiennes du système),
- sans avoir besoin d'introduire *explicitement* un petit paramètre dans le potentiel V .

Soulignons aussi le fait que notre estimation d'erreur n'est pas un résultat asymptotique, valable dans la limite où un certain paramètre tend vers 0. On peut ensuite, *a posteriori*, chercher sous quelles conditions sur V et sur ξ la borne obtenue est intéressante (au sens l'erreur commise est petite). Ceci permet, au moins dans certains cas simples, de *guider* le choix de ξ .

1.2.2 Construction

Notre point de départ est d'observer que

$$d[\xi(X_t)] = (-\nabla V \cdot \nabla \xi + \beta^{-1} \Delta \xi)(X_t) dt + \sqrt{2\beta^{-1}} |\nabla \xi|(X_t) dB_t, \quad (36)$$

où X_t est la solution de (34), et B_t est le mouvement brownien mono-dimensionnel $dB_t = \frac{\nabla \xi}{|\nabla \xi|}(X_t) \cdot dW_t$. L'équation (36) n'est pas fermée, au sens où le membre de droite dépend de X_t , et pas seulement de $\xi(X_t)$.

Suivant Gyöngy [82], une procédure pour fermer le système est d'introduire les espérances conditionnelles des termes de dérive et de diffusion de (36). On définit ainsi

$$\begin{aligned} \tilde{b}(t, z) &= \mathbb{E} [(-\nabla V \cdot \nabla \xi + \beta^{-1} \Delta \xi)(X_t) \mid \xi(X_t) = z], \\ \tilde{\sigma}^2(t, z) &= \mathbb{E} [|\nabla \xi|^2(X_t) \mid \xi(X_t) = z], \end{aligned} \quad (37)$$

où les espérances sont calculées par rapport à la densité de probabilité de X_t , à l'instant t (on donne dans la Remarque 12 ci-dessous une expression explicite de \tilde{b} dans un cas simple). On introduit ensuite le processus \tilde{y}_t solution de

$$d\tilde{y}_t = \tilde{b}(t, \tilde{y}_t) dt + \sqrt{2\beta^{-1}} \tilde{\sigma}(t, \tilde{y}_t) dB_t. \quad (38)$$

Comme montré dans [82], cette procédure est exacte du point de vue des marginales en temps, au sens où, à chaque instant t , la loi de $\xi(X_t)$ est *égale* à la loi de \tilde{y}_t . Cependant, on remarque immédiatement que les fonctions \tilde{b} et $\tilde{\sigma}$ sont en pratique très difficiles à calculer. Elles nécessitent en effet la connaissance de la densité de probabilité $\psi(t, x)$ de X_t à chaque instant.

Remarque 12 *Considérons à nouveau le cas $\xi(x_1, x_2) = x_2$, où $x = (x_1, x_2) \in \mathbb{R}^2$. Alors (36) s'écrit*

$$dx_2 = -\partial_{x_2} V(x_1, x_2) dt + \sqrt{2\beta^{-1}} dB_t,$$

qui n'est effectivement pas fermée, puisque la dynamique de x_2 dépend de x_1 . On note $\psi(t, x_1, x_2)$ la densité de probabilité de X_t , solution de (34). Alors (37) s'écrit

$$\tilde{b}(t, z) = \frac{\int_{\mathbb{R}} (-\nabla V \cdot \nabla \xi + \beta^{-1} \Delta \xi) \psi(t, x_1, z) dx_1}{\int_{\mathbb{R}} \psi(t, x_1, z) dx_1} = -\frac{\int_{\mathbb{R}} \partial_{x_2} V(x_1, z) \psi(t, x_1, z) dx_1}{\int_{\mathbb{R}} \psi(t, x_1, z) dx_1}. \quad (39)$$

On voit bien que le calcul de \tilde{b} fait intervenir ψ . \diamond

Une simplification naturelle est d'utiliser le fait qu'on connaît, en temps long, la limite de ψ , qui n'est autre que la mesure de Boltzmann (6). On introduit donc

$$\begin{aligned} b(z) &= \mathbb{E}_\pi [(-\nabla V \cdot \nabla \xi + \beta^{-1} \Delta \xi)(X) \mid \xi(X) = z] \\ &= \int_{\Sigma_z} (-\nabla V \cdot \nabla \xi + \beta^{-1} \Delta \xi) d\pi_{\Sigma_z}, \end{aligned} \quad (40)$$

et

$$\sigma^2(z) = \mathbb{E}_\pi (|\nabla \xi|^2(X) \mid \xi(X) = z) = \int_{\Sigma_z} |\nabla \xi|^2 d\pi_{\Sigma_z}, \quad (41)$$

où \mathbb{E}_π est l'espérance par rapport à la mesure canonique (6), et π_{Σ_z} est définie par (32). Dans le cas simple de la Remarque 12, ceci revient à remplacer $\psi(t, x_1, z)$ par $\pi(x_1, z)$. Cette simplification est raisonnable si le temps caractéristique d'évolution de $\xi(X_t)$ est bien plus grand que le temps nécessaire à X_t pour échantillonner Σ_z défini par (31). Les hypothèses faites ci-dessous permettent de quantifier cette séparation des échelles de temps.

Dans l'esprit de (38), on introduit ensuite la dynamique

$$dy_t = b(y_t) dt + \sqrt{2\beta^{-1}} \sigma(y_t) dB_t, \quad (42)$$

qu'on appellera dans la suite dynamique effective, et qui présente le triple intérêt suivant :

- les fonctions b et σ définies par (40) et (41) sont *calculables* numériquement (à la différence de (37)), en utilisant par exemple les algorithmes proposés dans [42]. Le calcul de ces fonctions est en fait de la même complexité que le calcul de la dérivée de l'énergie libre, car ce sont des moyennes de certaines observables contre la mesure π_{Σ_z} , au même titre que A' (cf. l'expression (29)).
- la dynamique (42) est ergodique pour la mesure $d(\xi \star \pi)$: elle a donc les bonnes propriétés en temps long ;
- on peut estimer la distance entre les lois, à chaque instant t , des variables aléatoires y_t et $\xi(X_t)$, comme cela est présenté à la Section 1.2.3.

1.2.3 Estimation d'erreur

Nous donnons dans cette section une estimation, en terme d'entropie relative, entre les lois, à chaque instant t , de y_t , solution de la dynamique effective (42), et $\xi(X_t)$, où X_t est solution de (34).

On commence par rappeler que, si ν et η sont deux mesures de probabilité telles que ν est absolument continue par rapport à η , alors l'entropie relative de ν par rapport à η est définie par

$$H(\nu|\eta) = \int \ln \left(\frac{d\nu}{d\eta} \right) d\nu.$$

Par construction, on a toujours $H(\nu|\eta) \geq 0$. Dans le cas où ν et η ont une densité ψ_ν et ψ_η par rapport à la mesure de Lebesgue, alors on a l'inégalité de Csiszár-Kullback

$$\|\psi_\nu - \psi_\eta\|_{L^1} \leq \sqrt{2H(\nu|\eta)}. \quad (43)$$

On voit ainsi que $H(\nu|\eta) = 0$ implique que ν et η ont même densité. On aura besoin dans la suite de la définition suivante (cf. [4]) :

Definition 13 *Une mesure de probabilité η vérifie une inégalité de Sobolev logarithmique avec la constante $\rho > 0$ si, pour toute mesure de probabilité ν ,*

$$H(\nu|\eta) \leq \frac{1}{2\rho} I(\nu|\eta),$$

où l'information de Fisher $I(\nu|\eta)$ est définie par

$$I(\nu|\eta) = \int \left| \nabla \ln \left(\frac{d\nu}{d\eta} \right) \right|^2 d\nu.$$

Ces notions sont utiles pour démontrer la convergence en temps long de solutions d'équations de Fokker-Planck vers une mesure stationnaire. Ainsi, la densité de probabilité $\psi(t, x)$ de X_t solution de (34) vérifie l'équation de Fokker-Planck

$$\partial_t \psi = \operatorname{div} (\psi \nabla V) + \beta^{-1} \Delta \psi, \quad (44)$$

dont il est facile de voir que $\pi(x) = Z^{-1} \exp(-\beta V(x))$ est une solution stationnaire. On introduit alors l'entropie relative

$$\mathcal{E}(t) = H(\psi(t, \cdot) | \pi) = \int_{\mathbb{R}^n} \ln \left(\frac{\psi(t, x)}{\pi(x)} \right) \psi(t, x) dx.$$

Un calcul simple, utilisant (44), montre que

$$\frac{d\mathcal{E}}{dt} = -\beta^{-1} I(\psi(t, \cdot) | \pi). \quad (45)$$

Si π satisfait une inégalité de Sobolev logarithmique avec la constante $\rho > 0$, alors, pour tout $t \geq 0$,

$$H(\psi(t, \cdot) | \pi) \leq (2\rho)^{-1} I(\psi(t, \cdot) | \pi), \quad (46)$$

et on déduit de (45), (46) et de l'application du lemme de Gronwall que

$$\forall t \geq 0, \quad 0 \leq H(\psi(t, \cdot) | \pi) \leq H(\psi(0, \cdot) | \pi) \exp(-2\rho\beta^{-1}t). \quad (47)$$

On a donc la convergence exponentielle de l'entropie relative (et donc de $\|\psi(t, \cdot) - \pi\|_{L^1}$, via (43)) vers 0, avec un taux proportionnel à ρ . En particulier, plus ρ est grand, plus la convergence est rapide.

Dans [FL23], nous avons démontré le résultat suivant, dont la preuve s'inspire des travaux [81, 53] (on trouvera aussi dans [FL34] une preuve plus simple, dans le cas particulier de la Remarque 12, i.e. $\xi(x_1, x_2) = x_2$).

Proposition 14 ([FL23], Proposition 3.1) *On suppose que ξ est une fonction régulière à valeur scalaire, telle qu'il existe m et M tels que*

$$\text{pour tout } x \in \mathbb{R}^n, \quad 0 < m \leq |\nabla \xi(x)| \leq M < \infty. \quad (48)$$

On suppose que les mesures conditionnées π_{Σ_z} définies par (32) vérifient une inégalité de Sobolev logarithmique, avec une constante ρ uniforme en z : pour toute mesure de probabilité ν sur Σ_z , absolument continue par rapport à π_{Σ_z} , on a

$$H(\nu | \pi_{\Sigma_z}) \leq \frac{1}{2\rho} I(\nu | \pi_{\Sigma_z}). \quad (49)$$

On suppose aussi que le couplage entre la direction $\nabla \xi(x)$ et les directions orthogonales est borné, au sens

$$\kappa = \|\nabla_{\Sigma_z} F\|_{L^\infty} < \infty, \quad (50)$$

où F est définie par (30).

Enfin, on suppose que $|\nabla\xi|$ est proche d'une constante sur chaque sous-variété Σ_z :

$$\lambda = \left\| \frac{|\nabla\xi|^2 - \sigma^2 \circ \xi}{\sigma^2 \circ \xi} \right\|_{L^\infty} < \infty. \quad (51)$$

Soit ψ^ξ la densité de probabilité de $\xi(X_t)$, où X_t vérifie (34), et ϕ la densité de probabilité de la solution y_t de (42). On suppose qu'à $t = 0$, les distributions des conditions initiales de (34) et (42) sont consistantes l'une avec l'autre : $\psi^\xi(t = 0, \cdot) = \phi(t = 0, \cdot)$. Alors, pour tout $t \geq 0$,

$$E(t) \leq \frac{M^2}{4m^2} \left(\lambda^2 + \frac{m^2\beta^2\kappa^2}{\rho^2} \right) (H(\psi(0, \cdot)|\pi) - H(\psi(t, \cdot)|\pi)), \quad (52)$$

où $\pi(x) = Z^{-1} \exp(-\beta V(x))$, $\psi(t, \cdot)$ est la densité de probabilité de X_t , et $E(t)$ est l'entropie relative de ψ^ξ par rapport à ϕ :

$$E(t) = H(\psi^\xi(t, \cdot)|\phi(t, \cdot)) = \int_{\mathbb{R}} \ln \left(\frac{\psi^\xi(t, z)}{\phi(t, z)} \right) \psi^\xi(t, z) dz.$$

Remarque 15 La Proposition 14 donne une estimation uniforme en temps de l'entropie relative entre ψ^ξ et ϕ . Par ailleurs, on sait aussi que la dynamique effective (42) est ergodique pour $\exp(-\beta A(z)) dz$, qui est la mesure d'équilibre de $\xi(X_t)$, en temps long. Par conséquent, on s'attend à ce que ψ^ξ et ϕ , en temps long, aient la même limite. C'est effectivement le cas, et nous avons montré dans [FL23, Corollaire 3.1] que, sous des hypothèses supplémentaires assez faibles, la quantité $\|\psi^\xi(t, \cdot) - \phi(t, \cdot)\|_{L^1(\mathbb{R})}$ est bornée uniformément en temps, et décroît exponentiellement vite vers 0 quand $t \rightarrow \infty$. \diamond

L'estimation (52) n'est pas un résultat asymptotique. Bien sûr, elle est particulièrement intéressante lorsque le membre de droite est petit, ce qui nécessite en particulier que κ/ρ soit petit, ce que se produit lorsque :

- le paramètre ρ est grand ; ceci signifie que la mesure conditionnée π_{Σ_z} est facile à échantillonner (absence de métastabilité dans la sous-variété Σ_z). On se souvient en effet (cf. l'estimation (47)) qu'une mesure est d'autant plus rapide à échantillonner qu'elle vérifie une inégalité de Sobolev logarithmique avec une grande constante. Ceci implique en particulier qu'il n'y a pas de métastabilité dans la sous-variété Σ_z .
- le paramètre κ est petit ; illustrons ceci dans le cas simple où $x = (x_1, x_2) \in \mathbb{R}^2$, et $\xi(x) = x_2$. Alors $F = \nabla_{x_2} V$ et $\nabla_{\Sigma_z} F = \nabla_{x_1} F = \nabla_{x_1 x_2} V$. Demander que κ est petit revient à demander que $\nabla_{x_1 x_2} V$ est petit, et que V couple donc *faiblement* la direction x_2 de la coordonnée de réaction avec la direction orthogonale. Notons aussi que, dans le cas où $V(x) = \frac{1}{2} x^T H x$, où $H \in \mathbb{R}^{2 \times 2}$ est une matrice symétrique définie positive, on a $\kappa = 0$ si et seulement si X_1 et X_2 sont indépendantes, où $X = (X_1, X_2) \in \mathbb{R}^2$ est distribué suivant $Z^{-1} \exp(-\beta V(x)) dx$ (cf. aussi [104, 125]).

Le fait de dire que κ/ρ est petit est ainsi une façon de formaliser l'hypothèse de séparation des échelles de temps mentionnée dans la Section 1.2.1.

Remarquons enfin que l'estimation (52) peut être utilisée *a posteriori*, dans le cas d'un potentiel dépendant d'un petit paramètre ε . Dans certains cas, il est alors possible d'estimer comment ρ et κ dépendent de ε , et d'aboutir à des critères sur ξ pour que κ/ρ soit petit lorsque ε est petit. On renvoie à [FL23, Section 3.2] pour plus de détails.

1.2.4 Résultats numériques

La Proposition 14 donne une estimation de l'erreur commise sur les lois, à chaque instant t , entre la variable aléatoire exacte $\xi(X_t)$ et son approximation y_t . Nous n'avons à ce jour pas réussi à étendre ce résultat pour avoir une estimation de l'erreur sur des quantités dynamiques telles que les temps de résidence, qui ne peuvent pas être déduits de la simple connaissance des lois à chaque instant (il faudrait la loi du processus sur un intervalle de temps $[0, T]$).

Les résultats numériques donnés dans [FL23, FL34] montrent cependant que, si ξ est choisi telle que l'erreur dans (52) est petite, alors les temps de résidence sont aussi très bien reproduits par la dynamique effective (42).

Remarque 16 *La dynamique*

$$d\bar{y}_t = -A'(\bar{y}_t) dt + \sqrt{2\beta^{-1}} dB_t, \quad (53)$$

où A est l'énergie libre associée à la coordonnée de réaction ξ avec laquelle on travaille, est une dynamique mono-dimensionnelle, ergodique pour la mesure d'équilibre de $\xi(X)$, qui est $\exp(-\beta A(z)) dz$.

Dans [FL23], nous avons cherché à comprendre la relation entre la dynamique (42) que nous proposons et la dynamique (53), et ainsi à comprendre si l'énergie libre A a, ou n'a pas, une signification du point de vue de la dynamique. La motivation pour considérer (53) est que les praticiens déduisent souvent les temps de résidence dans les puits de la fonction $z \mapsto A(z)$. Si effectivement $\xi(X_t)$ suit une dynamique proche de (53), alors les temps de résidence au voisinage des minima locaux de A sont effectivement fonctions des barrières d'énergie libre, dans la limite des basses températures (on renvoie à la littérature sur les grandes déviations, et aussi à [99, 85]). \diamond

1.2.5 Extension

Nous avons choisi ci-dessus de considérer la dynamique de Langevin amortie (34) comme dynamique de référence, car son analyse mathématique est simple. Une autre dynamique fréquemment utilisée en simulation moléculaire est la dynamique de Langevin (17), qui a le mérite d'être plus proche d'une dynamique Hamiltonienne, et qui est ergodique pour la mesure de Boltzmann (3). Par contre, le bruit n'agit que sur p . L'équation de Fokker-Planck associée est hypo-elliptique, et prouver la convergence vers l'équilibre de sa solution est une question classique, mais plus difficile que dans le cas amorti. Dans un travail en cours [FL37], nous cherchons à étendre la *methodologie* présentée dans la Section 1.2.2 au cas de l'équation de Langevin. Il n'est pas clair que nous puissions obtenir sur cette dynamique un résultat du type de la Proposition 14.

1.3 Problèmes Hamiltoniens

Publications [FL10, FL21, FL22, FL25, FL26]

Dans les Sections 1.1 et 1.2 précédentes, nous nous sommes intéressés à la simulation de systèmes moléculaires dans l'ensemble canonique. Le point de départ est alors la mesure canonique (3), qu'il s'agit d'échantillonner, ce qui peut se faire via plusieurs dynamiques. La seule contrainte qu'on impose sur ces dynamiques est d'être ergodique par rapport à la mesure (3). On revient maintenant à l'étude de systèmes moléculaires *isolés*, dont l'évolution est régie par les équations de Newton, qu'on écrit ici comme la dynamique Hamiltonienne

$$\dot{q}(t) = \nabla_p H(q(t), p(t)), \quad \dot{p}(t) = -\nabla_q H(q(t), p(t)), \quad (54)$$

où H est donné par (1). Nous avons abordé trois thématiques dans ce domaine :

- l’étude et l’accélération de la convergence des moyennes temporelles le long des trajectoires de systèmes Hamiltoniens, y compris après discrétisation en temps de ces trajectoires.
- la construction de schémas symplectiques adaptés au cas *hautement oscillant*. Le système présente alors plusieurs échelles en temps, et le pas de temps admissible pour les algorithmes classiques est limité par les temps caractéristiques les plus rapides.
- la construction de schémas *parallèles* en temps. Le calcul de la trajectoire sur l’intervalle $[0, T]$ est fait de manière itérative, et une partie du calcul, la plus coûteuse, est réalisée en parallèle.

1.3.1 Convergence des moyennes temporelles

Pour aborder cette question, nous avons choisi un cadre très particulier, celui des systèmes Hamiltoniens *complètement intégrables*. Dans ce cas, un résultat classique du à Bohl, Sierpinski et Weyl (cf. [8] et les références incluses) montre, sous certaines hypothèses, qu’il existe une mesure $d\rho$ pour laquelle la dynamique (54) est ergodique, autrement dit

$$\lim_{T \rightarrow \infty} \langle \Phi \rangle(T) = \int \Phi d\rho, \quad (55)$$

où

$$\langle \Phi \rangle(T) := \frac{1}{T} \int_0^T \Phi(q(t), p(t)) dt$$

est la moyenne de l’observable Φ le long d’une solution $(q(t), p(t))$ de (54). Le taux de convergence dans (55) est en T^{-1} .

Nous avons tout d’abord démontré qu’il est possible d’accélérer la convergence en filtrant le signal temporel. Ainsi, pour tout $k \in \mathbb{N}^*$, il existe une fonction filtre permettant d’obtenir la convergence de la moyenne temporelle filtrée vers la moyenne d’ensemble avec un taux en T^{-k} (notons que cette approche a ensuite été réutilisée dans un cadre complètement différent, celui de l’homogénéisation des équations aux dérivées partielles, cf. [20]). Nous avons aussi montré la convergence des moyennes temporelles filtrées, calculées non pas le long de la solution exacte de (54), mais le long d’une trajectoire numérique, lorsque celle-ci est calculée par un schéma symplectique. L’erreur est alors bornée par $C(T^{-k} + (\Delta t)^r)$, où r est l’ordre du schéma, Δt le pas de temps, et C une constante indépendante de T et Δt .

Ces résultats, publiés dans [FL3, FL4], ont été décrits dans [FL7].

1.3.2 Problèmes hautement oscillants

Publications [FL10, FL21, FL22, FL25]

On s’intéresse ici au développement d’algorithmes numériques efficaces, dans le cas où la dynamique Hamiltonienne est hautement oscillante. Une motivation est que, pour de nombreux systèmes (en dynamique moléculaire mais aussi dans d’autres applications), la dynamique (54) fait apparaître plusieurs échelles de temps, indépendamment de l’échelle nécessaire pour (éventuellement) obtenir la convergence (55) des moyennes temporelles. Ainsi, dans une molécule, les liaisons entre atomes de carbone et d’hydrogène vibrent avec une période de l’ordre de 10^{-15} seconde, tandis que le temps caractéristique d’évolution de certains angles est beaucoup plus lent. Enfin, le temps nécessaire à l’exploration de l’espace des phases (et donc à la convergence des moyennes temporelles) est plutôt de l’ordre de

10^{-6} , voire 10^{-3} seconde. Se pose alors la question de construire des schémas numériques stables avec un pas de temps qui ne soit pas lié aux mouvements les plus rapides.

Notre objectif est de construire des schémas numériques *symplectiques*, en gardant dans la simulation *toutes les variables*, lentes et rapides (notons d'ailleurs qu'il n'est pas toujours évident d'isoler les unes des autres!), afin de calculer des *trajectoires* (et non pas des moyennes temporelles le long de celles-ci, par exemple). Une motivation est que les schémas symplectiques sont connus pour avoir de bonnes propriétés en temps long pour l'intégration de systèmes Hamiltoniens (on reviendra sur ce point à la fin de cette section). Dans le cas où on s'intéresse uniquement aux variables lentes, ou bien à des moyennes, d'autres approches ont été proposées dans la littérature.

L'approche que nous avons suivie consiste à associer à la dynamique Hamiltonienne son équation de Hamilton-Jacobi (il s'agit d'une reformulation équivalente, sous de bonnes hypothèses de régularité), à faire un Ansatz de type double-échelle dans cette équation pour obtenir une approximation de la solution, qui servira ensuite à construire un schéma numérique.

Quelques rappels sur le formalisme Hamilton-Jacobi On rappelle que, pour un Hamiltonien $H(q, p)$ donné, l'équation de Hamilton-Jacobi associée est

$$\partial_t S = H(q + \partial_P S, P), \quad S(0, q, P) = 0, \quad (56)$$

où S est une fonction du temps t , de q (qui représente la position initiale) et P (qui joue le rôle de l'impulsion finale). Il existe en fait plusieurs formulations de l'équation de Hamilton-Jacobi associée à un Hamiltonien H . Nous travaillerons essentiellement avec (56). Connaissant S solution de (56), on définit ensuite $(Q(t), P(t))$ pour tout temps t par

$$p = P(t) + \frac{\partial S}{\partial q}(t, q, P(t)), \quad Q(t) = q + \frac{\partial S}{\partial P}(t, q, P(t)), \quad (57)$$

où q et p sont des paramètres. Alors $(Q(t), P(t))$ est la solution de la dynamique Hamiltonienne

$$\dot{Q} = \nabla_p H(Q(t), P(t)), \quad \dot{P} = -\nabla_q H(Q(t), P(t)), \quad (58)$$

avec la condition initiale (q, p) .

Remarque 17 Dans le cas d'un oscillateur harmonique, $H(q, p) = (q^T q + p^T p)/2$, la solution de (56) est explicite :

$$S(t, q, P) = \frac{1}{2}(P^T P + q^T q) \tan t + P^T q \left(\frac{1}{\cos t} - 1 \right).$$

On remarque que S est 2π périodique de t . On utilisera cette observation ci-dessous. \diamond

Remarque 18 L'équation de Hamilton-Jacobi classiquement associée à un Hamiltonien $H(q, p)$ est

$$\partial_t S + H(Q, \partial_Q S) = 0, \quad (59)$$

d'inconnue $S(t, q, Q)$. Supposons connaître S . Soit alors (q, p) des paramètres, et soit $(Q(t), P(t))$ implicitement définis par

$$P(t) = \frac{\partial S}{\partial Q}(t, q, Q(t)), \quad p = -\frac{\partial S}{\partial q}(t, q, Q(t)).$$

Alors $(Q(t), P(t))$ est solution de la dynamique Hamiltonienne (58). Sous une hypothèse supplémentaire sur S à l'instant initial, on obtient que $(Q(0), P(0)) = (q, p)$, et donc $(Q(t), P(t))$ est la solution de (58) avec la condition initiale (q, p) [84, Théorème VI.5.6].

Dans le cas où $H(q, p) = V(q) + p^T p / 2$, on peut donner une interprétation physique de S . Soit $J(\phi)$ l'action associée à une trajectoire $\phi \in H^1(0, t)$ (pour simplifier les notations, on suppose $(q, p) \in \mathbb{R} \times \mathbb{R}$) :

$$J(\phi) = \int_0^t \left[\frac{1}{2} \dot{\phi}^2(s) - V(\phi(s)) \right] ds.$$

On considère le problème variationnel

$$\mathcal{S}(t, q, Q) = \inf \{ J(\phi), \phi \in H^1(0, t), \phi(0) = q, \phi(t) = Q \}. \quad (60)$$

Les fonctions ϕ points critiques de ce problème vérifient les équations de Newton. Supposons V régulier et borné supérieurement. Alors le problème (60) est atteint (cf. [FL8, Preuve du Lemme 1]) : il existe (au moins) un minimiseur ϕ^* , vérifiant donc $\mathcal{S}(t, q, Q) = J(\phi^*)$. Supposons que ϕ^* dépend régulièrement de t, q et Q . Alors on montre que la fonction $\mathcal{S}(t, q, Q)$ vérifie l'équation de Hamilton-Jacobi (59).

Ainsi, sous les hypothèses que nous avons faites, l'action minimale pour aller de q à Q dans le temps t est solution de l'équation de Hamilton-Jacobi (59). \diamond

On montre maintenant comment utiliser le formalisme Hamilton-Jacobi pour construire des schémas numériques. L'observation fondamentale est que, pour toute fonction $(q, P) \mapsto S(t, q, P)$ (on considère ici t comme un paramètre fixé), les équations (57) définissent une application

$$(q, p) \mapsto (Q(t), P(t))$$

qui est *symplectique*. Une méthode, initialement proposée dans [69], pour construire des schémas symplectiques est donc de chercher une approximation de la solution de (56) à un temps $t = \Delta t$ petit (on la note $\tilde{S}(\Delta t, q, P)$), et à l'insérer dans le changement de variables (57), ce qui définit implicitement $(Q_{\Delta t}, P_{\Delta t})$ par

$$p = P_{\Delta t} + \frac{\partial \tilde{S}}{\partial q}(\Delta t, q, P_{\Delta t}), \quad Q_{\Delta t} = q + \frac{\partial \tilde{S}}{\partial P}(\Delta t, q, P_{\Delta t}). \quad (61)$$

Le flot numérique est alors l'application $\Phi_{\Delta t}$ définie par $\Phi_{\Delta t}(q, p) = (Q_{\Delta t}, P_{\Delta t})$, et le schéma numérique associé est $\Phi_{\Delta t}(q_n, p_n) = (q_{n+1}, p_{n+1})$. Dans la mesure où (61) est résolu exactement, on obtient un schéma numérique qui est *par construction* symplectique.

A titre d'illustration, considérons le cas $H(q, p) = V(q) + \frac{1}{2} p^T p$. Soit S la solution de (56). On vérifie que $\partial_t S(0, q, P) = H(q, P)$. Une approximation naturelle est

$$S(\Delta t, q, P) \approx \tilde{S}(\Delta t, q, P) := S(0, q, P) + \Delta t \partial_t S(0, q, P) = \Delta t \left(V(q) + \frac{1}{2} P^T P \right).$$

Le schéma obtenu suivant (61) s'écrit

$$p = P_{\Delta t} + \Delta t \nabla V(q), \quad Q_{\Delta t} = q + \Delta t P_{\Delta t}.$$

On reconnaît le schéma d'Euler symplectique.

Le cas hautement oscillant : une approche de type homogénéisation On revient maintenant au cadre hautement oscillant. Nous nous sommes initialement intéressés, dans [FL10, FL21], au cas du Hamiltonien

$$H_\varepsilon(q_1, q_2, p_1, p_2) = \frac{p_1^T p_1}{2} + \frac{p_2^T p_2}{2} + V_0(q_1, q_2) + \Omega^2 \frac{q_2^T q_2}{2\varepsilon^2}, \quad (62)$$

où ε est un petit paramètre, $q = (q_1, q_2) \in \mathbb{R}^l \times \mathbb{R}^r$ et $p = (p_1, p_2) \in \mathbb{R}^l \times \mathbb{R}^r$, et V_0 est un terme d'énergie potentielle indépendant de ε , et borné inférieurement. On se donne des conditions initiales dépendant de ε telles que l'énergie soit bornée indépendamment de ε . Comme V_0 est borné inférieurement, on obtient ainsi que $p(t)$ et $q_2(t)/\varepsilon$ sont bornés par une constante indépendante de ε pour tout temps.

L'hypothèse importante dans (62) est la dépendance vis à vis des degrés de liberté rapides, qui sont considérés ici comme un oscillateur harmonique,

$$\frac{p_2^T p_2}{2} + \Omega^2 \frac{q_2^T q_2}{2\varepsilon^2}.$$

On a supposé aussi que Ω est une quantité scalaire et constante. Comme nous le montrons dans [FL22, FL25], notre approche permet aussi de traiter le cas où Ω dépend des variables lentes q_1 du système ($\Omega \in \mathbb{R}$ est remplacé par $\Omega(q_1) \in \mathbb{R}$ dans (62)), ainsi que le cas où Ω est une matrice symétrique constante (ce qui est équivalent, à une diagonalisation près, au cas où Ω est une matrice constante diagonale non sphérique). Un cas plus difficile est le cas où $\Omega(q_1)$ est une matrice dépendant de q_1 , et dont les *vecteurs propres* dépendent de q_1 . A ce jour, nous n'avons pas été en mesure d'étendre l'approche à ce cas, qui semble être le cas générique [84, Section XIV.3]. On renvoie aussi à l'article de revue [44], et à [102].

On revient maintenant au cas (62), que nous avons étudié dans [FL10, FL21]. Avant de pouvoir appliquer notre stratégie de développement à double échelle, quelques préliminaires algébriques sont nécessaires, qu'on décrit maintenant.

Dans le cas où V_0 ne dépend pas de q_2 , la trajectoire des variables rapides (q_2, p_2) est celle d'un oscillateur harmonique de pulsation $\Omega/\varepsilon \gg 1$. Il est intéressant pour la suite de changer de variables, afin de pré-conditionner le problème par ce mouvement rapide, analytiquement connu. Une idée similaire est utilisée pour construire d'autres schémas, comme par exemple les intégrateurs exponentiels, analysés dans [84, Chap. XIII], et les schémas proposés dans [79] et [93, 94].

On introduit donc le changement de variables (dépendant du temps) $(q_2, p_2) \mapsto (x_2, y_2) = \chi(t, q_2, p_2)$ défini par

$$\begin{aligned} q_2 &= \cos\left(\frac{\Omega t}{\varepsilon}\right) x_2 + \frac{\varepsilon}{\Omega} \sin\left(\frac{\Omega t}{\varepsilon}\right) y_2, \\ p_2 &= -\frac{\Omega}{\varepsilon} \sin\left(\frac{\Omega t}{\varepsilon}\right) x_2 + \cos\left(\frac{\Omega t}{\varepsilon}\right) y_2. \end{aligned}$$

La dynamique dans les variables (q_1, x_2, p_1, y_2) est Hamiltonienne pour

$$H_\varepsilon^{\text{pre}}(t, q_1, x_2, p_1, y_2) = \frac{p_1^T p_1}{2} + W_\varepsilon\left(\frac{\Omega t}{\varepsilon}, q_1, x_2, y_2\right), \quad (63)$$

avec

$$W_\varepsilon(\tau, q_1, x_2, y_2) = V_0\left[q_1, (\cos \tau)x_2 + \frac{\varepsilon}{\Omega}(\sin \tau)y_2\right].$$

Soit $\bar{S}_\varepsilon(t, q_1, x_2, P_1, Y_2)$ solution de l'équation de Hamilton-Jacobi associée à (63). La variable x_2 (comme la variable q_2) est d'ordre ε . Il est donc naturel de changer de variables,

et d'introduire

$$r_2 = \frac{\Omega}{\varepsilon} x_2 \quad \text{et} \quad S_\varepsilon(t, q_1, r_2, P_1, Y_2) = \bar{S}_\varepsilon\left(t, q_1, \frac{\varepsilon r_2}{\Omega}, P_1, Y_2\right).$$

On obtient alors que S_ε vérifie

$$\partial_t S_\varepsilon = \frac{P_1^T P_1}{2} + W_\varepsilon\left(\frac{\Omega t}{\varepsilon}, q_1 + \partial_{P_1} S_\varepsilon, \frac{\varepsilon r_2}{\Omega} + \partial_{Y_2} S_\varepsilon, Y_2\right), \quad (64)$$

avec la condition initiale $S_\varepsilon(0, q_1, r_2, P_1, Y_2) = 0$. C'est sur cette forme que nous allons maintenant utiliser un développement à deux échelles.

On cherche à calculer une approximation de $S_\varepsilon(\Delta t)$, pour un temps Δt petit, avec cependant $\Delta t \gg \varepsilon$. Comme Δt jouera ensuite le rôle du pas de temps (cf. (61)), on obtiendra ainsi un schéma dont le pas de temps est bien plus grand que celui admissible dans un schéma classique comme velocity-Verlet [84, Eq. (I.1.17)] (qui n'est stable que pour des pas de temps Δt de l'ordre de, ou plus petits que, ε). Notre idée est de faire l'Ansatz à double échelle suivant :

$$S_\varepsilon(t, q_1, r_2, P_1, Y_2) = S_0(t, \tau, q_1, r_2, P_1, Y_2) + \varepsilon S_1(t, \tau, q_1, r_2, P_1, Y_2) + \text{termes d'ordre plus élevé en } \varepsilon^k, k \geq 2, \quad (65)$$

où le temps rapide τ est défini par

$$\tau = \frac{t\Omega}{\varepsilon}. \quad (66)$$

Pour chaque k , les fonctions S_k dépendent de deux arguments de temps, le temps lent t et le temps rapide τ . On suppose que ces fonctions sont 2π périodiques en τ .

Cet Ansatz appelle plusieurs commentaires. Tout d'abord, dans le cas où V_0 est indépendant de q_2 , la solution exacte de (64) est effectivement consistante avec ce choix, puisqu'elle s'écrit

$$S_\varepsilon = S_{\text{lent}}(t, q_1, P_1) + \varepsilon S_{\text{rapide}}\left(\frac{t\Omega}{\varepsilon}, r_2, P_2\right),$$

où $S_{\text{lent}}(t, q_1, P_1)$ est la solution de l'équation de Hamilton-Jacobi

$$\partial_t S_{\text{lent}} = H_1(q_1 + \partial_{P_1} S_{\text{lent}}, P_1), \quad S_{\text{lent}}(0, q_1, P_1) = 0$$

associée au Hamiltonien $H_1(q_1, p_1) = \frac{p_1^T p_1}{2} + V_0(q_1)$, et

$$S_{\text{rapide}}(\tau, r_2, P_2) = \frac{1}{\Omega} \left[\left(\frac{P_2^T P_2 + r_2^T r_2}{2} \right) \tan \tau + P_2^T r_2 \left(\frac{1}{\cos \tau} - 1 \right) \right]$$

est effectivement 2π -périodique en τ et solution de l'équation de Hamilton-Jacobi associée à $H_2(r_2, p_2) = \frac{p_2^T p_2}{2} + \Omega^2 \frac{r_2^T r_2}{2}$.

Ensuite, dans l'Ansatz (65), l'hypothèse de périodicité vis à vis du temps rapide τ est liée au fait que la solution de l'équation de Hamilton-Jacobi associée à un oscillateur harmonique est périodique en temps (cf. la Remarque 17). Dans le cas où le potentiel rapide n'est pas harmonique en q_2 , faire une hypothèse raisonnable sur la dépendance au temps rapide est plus difficile.

L'idée est ensuite d'insérer cet Ansatz (65) dans (64), d'assimiler la première variable de W_ε au temps rapide τ , et d'identifier les puissances de ε . On obtient alors une hiérarchie d'équations, couplant les fonctions S_k les unes aux autres. Pour fermer cette hiérarchie, on utilise de façon cruciale l'hypothèse de périodicité par rapport à τ .

Remarque 19 *On remarquera bien sûr le parallèle avec l'analyse formelle pour l'homogénéisation du problème elliptique*

$$-div \left[A_{per} \left(\frac{x}{\varepsilon} \right) \nabla u^\varepsilon \right] = f \text{ dans } \mathcal{D}, \quad u^\varepsilon \in H_0^1(\mathcal{D}), \quad (67)$$

où A_{per} est une matrice \mathbb{Z}^d périodique. Pour trouver formellement le problème limite, une approche classique consiste à postuler que u^ε s'écrit sous la forme

$$u^\varepsilon(x) = u_0 \left(x, \frac{x}{\varepsilon} \right) + \varepsilon u_1 \left(x, \frac{x}{\varepsilon} \right) + \varepsilon^2 u_2 \left(x, \frac{x}{\varepsilon} \right) + \dots, \quad (68)$$

où, pour chaque k , la fonction u_k est supposée dépendre de deux arguments, une variable macroscopique x et une variable microscopique $\frac{x}{\varepsilon}$, et que $y \mapsto u_k(x, y)$ est \mathbb{Z}^d -périodique. On insère ensuite (68) dans (67) et on identifie en puissance de ε . Là encore, la périodicité (ou une hypothèse de structure similaire) est cruciale pour fermer la hiérarchie. On renvoie à la Section 3.1 pour une discussion plus complète des problèmes d'homogénéisation des équations aux dérivées partielles elliptiques. \diamond

Résultats numériques Nous avons mis en oeuvre cette stratégie dans [FL21]. On obtient ainsi un schéma numérique automatiquement symplectique. Dans le cas où V_0 est indépendant de q_2 , les variables lentes et rapides sont découplées. Notre schéma devient alors un schéma numérique standard (comme le schéma velocity-Verlet) sur les variables lentes, tandis que les variables rapides sont intégrées exactement. De plus, dans le cas général, ce schéma est consistant avec les quelques résultats théoriques connus dans la littérature sur l'équation limite satisfaite par les variables lentes, dans la limite $\varepsilon \rightarrow 0$ (cf. par exemple [31, 32]). Ainsi, en faisant formellement la limite $\varepsilon \rightarrow 0$ dans notre schéma, avec un pas de temps fixé, on obtient un schéma qui intègre bien la dynamique limite, qui préserve des quantités identifiées comme des invariants adiabatiques lorsque $\varepsilon > 0$, ...

L'intérêt des schémas obtenus est qu'on peut travailler avec un pas de temps $\Delta t \gg \varepsilon$, donc plus grand que les pas de temps admissibles par les schémas classiques. En fait, nos schémas étant obtenus via une approche de type homogénéisation, ils sont d'autant plus efficaces que ε est petit. Par ailleurs, un problème récurrent dans ce domaine est la présence de résonances : la plupart des schémas proposés dans la littérature se comportent très mal quand $\Delta t/(\varepsilon\pi)$ est un entier, i.e. le pas de temps Δt est un multiple de la demi-période rapide. Le problème n'est pas tant la présence de résonances, que leur "épaisseur" : si Δt est un pas résonant, est-il facile de trouver un autre pas de temps dans son voisinage qui ne soit pas résonant ? Pour certains schémas (comme Impulse, aussi connu sous le nom de Verlet-I/r-RESPA [80, 143]), les résonances sont très larges, au sens où, au delà d'un certain seuil (i.e. pour $\Delta t \geq \Delta t_0$), il devient très difficile de trouver un pas de temps non résonant. Les schémas construits suivant notre approche présentent aussi des résonances. On constate toutefois qu'elles sont très fines, que la dérive en énergie pour ces pas résonants est plus lente que pour les autres schémas, et que, pour des pas non-résonants, la précision obtenue (sur la conservation de l'énergie en temps long, sur la précision de la trajectoire dans un intervalle en temps borné, ...) est bien supérieure. On renvoie à [FL21] pour les résultats numériques détaillés sur un cas test typique de la littérature, une chaîne de ressorts de type Fermi-Pasta-Ulam (FPU) [84, Section XIII.2.1], pour laquelle le Hamiltonien est de la forme (62), avec $p_1 \in \mathbb{R}^3$, $p_2 \in \mathbb{R}^3$, et où V_0 est un potentiel polynomial.

De plus, avec nos schémas, nous simulons l'ensemble de la trajectoire, et non pas seulement les variables lentes. Ceci nous donne donc accès à beaucoup d'information. Illustrons ceci dans le cas de la chaîne de type FPU, qui est une succession de ressorts

harmoniques raides avec des ressorts non-harmoniques non-raides. On rappelle que, dans cet exemple, il y a 3 ressorts raides. Les variables rapides sont $(q_2, p_2) \in \mathbb{R}^3 \times \mathbb{R}^3$, où $q_{2,j} \in \mathbb{R}$ est la longueur du ressort raide j , et $p_{2,j}$ est le moment associé. L'énergie dans le ressort rapide j est

$$I_j(q, p) = \frac{p_{2,j}^2}{2} + \frac{q_{2,j}^2}{2\varepsilon^2}, \quad 1 \leq j \leq 3.$$

Introduisons la somme de ces énergies rapides :

$$I(q, p) = I_1(q, p) + I_2(q, p) + I_3(q, p). \quad (69)$$

Bien que les I_j ne dépendent que de variables rapides, leur évolution se fait sur un temps caractéristique lent. De plus, c'est un résultat classique que leur somme I est un invariant adiabatique de la trajectoire, au sens où il existe une constante C telle que

$$\forall t > 0, \quad |I(q(t), p(t)) - I(q(0), p(0))| \leq C\varepsilon.$$

Ainsi, I n'est pas un invariant de la dynamique Hamiltonienne, mais a une évolution très particulière.

Notre approche permet de suivre en temps l'évolution de ces quantités, le long de la trajectoire numérique (contrairement à une approche qui consiste à éliminer les variables rapides, pour aboutir à une dynamique uniquement sur les variables lentes). Comme la somme des I_j est quasiment conservée au cours du temps, il est intéressant d'observer les échanges d'énergie qui se produisent au cours du temps entre ces I_j . Nous avons observé que notre algorithme reproduit très bien ces échanges, comme on peut le voir sur la Figure 5.

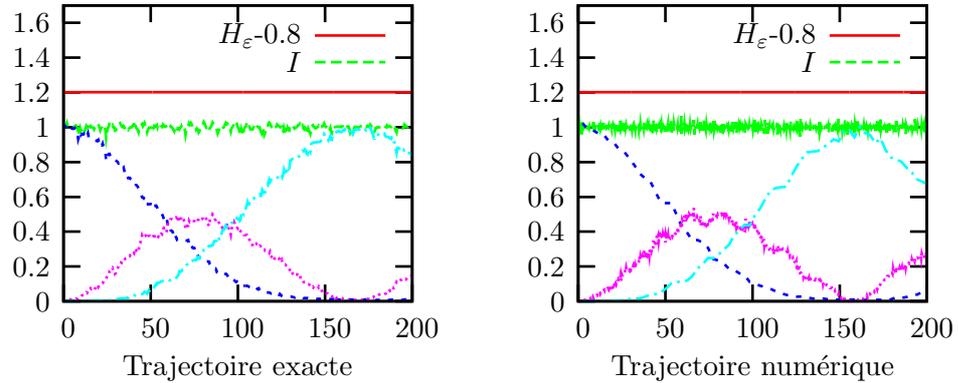


FIG. 5 – Conservation de l'énergie (on trace $H_\varepsilon - 0.8$), de l'invariant adiabatique I défini par (69) et échange entre les $\{I_j\}_{1 \leq j \leq 3}$, pour le cas d'une chaîne FPU avec $\varepsilon = 0.02$ (à gauche, la trajectoire exacte; à droite, la trajectoire numérique calculée avec notre algorithme avec le pas de temps $\Delta t = 0.03$). Résultats extraits de [FL21].

Une difficulté avec les schémas obtenus via l'approche Hamilton-Jacobi est qu'ils font très souvent intervenir les dérivées d'ordre deux et plus du potentiel (à l'opposé d'un schéma de type Verlet, qui ne fait appel qu'au gradient de V). A titre d'illustration, considérons le cas non-oscillant $H(q, p) = V(q) + \frac{1}{2}p^T p$. Soit S la solution de (56). On vérifie alors que

$$\partial_t S(0, q, P) = H(q, P), \quad \partial_{tt} S(0, q, P) = P^T \nabla V(q).$$

On a vu que l'approximation $S(\Delta t, q, P) \approx S(0, q, P) + \Delta t \partial_t S(0, q, P)$ conduit au schéma d'Euler symplectique, d'ordre un. Une idée naturelle pour obtenir un schéma d'ordre plus

élevé est de considérer une approximation d'ordre plus élevé de S . Considérons alors l'approximation d'ordre deux

$$\begin{aligned}\tilde{S}(\Delta t, q, P) &:= S(0, q, P) + \Delta t \partial_t S(0, q, P) + \frac{\Delta t^2}{2} \partial_{tt} S(0, q, P) \\ &= \Delta t \left(V(q) + \frac{1}{2} P^T P \right) + \frac{\Delta t^2}{2} P^T \nabla V(q).\end{aligned}$$

Le schéma obtenu suivant (61) s'écrit

$$p = P_{\Delta t} + \Delta t \nabla V(q) + \frac{\Delta t^2}{2} \nabla^2 V(q) P_{\Delta t}, \quad Q_{\Delta t} = q + \Delta t P_{\Delta t} + \frac{\Delta t^2}{2} \nabla V(q),$$

et on note qu'il fait intervenir la matrice hessienne de V .

De manière générale, le fait que des dérivées d'ordre élevé de V interviennent peut être limitant en pratique. Remplacer ces dérivées par des différences finies dans le schéma conduit à briser la symplecticité du schéma, et à perdre tout l'avantage de travailler dans le formalisme Hamilton-Jacobi. Nous avons montré dans [FL25] qu'il est possible de faire ce remplacement dans la *fonction génératrice* (i.e. dans l'approximation \tilde{S} de la solution de l'équation de Hamilton-Jacobi). On insère ensuite cette approximation, utilisant des différences finies plutôt que des dérivées d'ordre élevé, dans (61), qu'on résout exactement. Ainsi, la symplecticité est préservée.

Remarques Terminons par quelques remarques plus générales. Plusieurs algorithmes ont été proposés pour l'intégration numérique de la dynamique associée à (62) (on renvoie à [84, Chap. XIII et XIV], [79, 37, 66, 133, 92, 95, 103]). La spécificité de notre travail est d'utiliser le formalisme Hamilton-Jacobi (pour obtenir des schémas symplectiques), avec une approche de type double-échelle (pour tenir compte du caractère hautement oscillant du problème).

La motivation pour travailler avec des schémas symplectiques est qu'ils sont connus pour posséder de bonnes propriétés en temps long. Ils permettent d'obtenir une excellente conservation de l'énergie, au sens où il existe Δt_0 , Δt^* et C tels que, pour tout $\Delta t \leq \Delta t_0$, on a

$$|H(q_n, p_n) - H(q_0, p_0)| \leq C \Delta t^r$$

pour tout n tel que $n\Delta t \leq \exp(\Delta t^*/\Delta t)$ (on a noté r l'ordre du schéma). Ainsi, pour des temps exponentiellement longs, l'énergie reste proche de sa valeur initiale, cette proximité étant contrôlée par le pas de temps et l'ordre du schéma (il n'y a pas de dérive de l'énergie). Dans le cas de systèmes complètement intégrables, on montre de plus que l'écart entre la trajectoire réelle et la trajectoire numérique ne croît que linéairement en temps (avec un autre schéma, comme le schéma d'Euler, la croissance est exponentielle). La preuve de ces propriétés repose sur l'analyse rétrograde, qui consiste à montrer que la trajectoire numérique est (quasiment) égale à la trajectoire exacte d'un problème modifié. Si le schéma utilisé est symplectique, alors ce problème modifié est encore Hamiltonien, et hérite essentiellement de toute la structure présente dans le problème initial (cf. [84, Chap. IX et X], [132] et [16, 83, 128]).

Toutes ces propriétés sont cependant démontrées dans la limite de pas de temps Δt petits, ce qui signifie dans le cas présent $\Delta t \ll \varepsilon$. Or, c'est justement le régime dans lequel on ne souhaite *pas* travailler ici. Pour être efficace numériquement, on travaille avec $\Delta t \gg \varepsilon$. Par conséquent, les résultats de l'analyse rétrograde ne s'appliquent pas à notre cadre. Ainsi, nous observons d'excellentes propriétés numériques, mais nous ne disposons pas d'arguments d'analyse numérique les justifiant.

Remarque 20 *Pour intégrer une dynamique Hamiltonienne réversible en temps (ce qui est le cas de (54) dès que le Hamiltonien est de la forme (1)), une classe de schémas alternative aux schémas symplectiques est la classe des schémas symétriques. Il est en général plus facile de construire un schéma symétrique qu'un schéma symplectique. En utilisant à nouveau les outils de l'analyse rétrograde, on peut expliquer leur bon comportement, sous des hypothèses cependant plus fortes que dans le cas symplectique. Il existe aussi des contre-exemples, c'est-à-dire des schémas symétriques qui n'ont pas de bonnes propriétés en temps long (on renvoie par exemple à [68]). Enfin, de même que pour les schémas symplectiques, tous les résultats sont encore obtenus dans la limite de petits pas de temps.*

1.3.3 Algorithmes parallèles en temps

Publication [FL26]

Comme on l'a vu ci-dessus, les calculs de dynamique moléculaire nécessitent souvent la simulation d'une longue trajectoire, ce qui est très coûteux car le pas de temps admissible est la plupart du temps inférieur de plusieurs ordres de grandeur au temps total de simulation souhaité. Il apparaît donc intéressant de développer des algorithmes *parallèles* en temps, permettant ainsi d'accéder à de longues trajectoires.

Algorithme pararéel standard Dans [FL26], nous nous sommes intéressés à l'adaptation de l'algorithme pararéel, initialement introduit dans [110], au cas de problèmes Hamiltoniens. Commençons par rappeler la forme standard de l'algorithme pararéel, lorsqu'on intègre le problème

$$\dot{u} = f(u), \quad u(0) = u_0, \quad (70)$$

tel que présenté dans [9, 11]. Le calcul de la trajectoire sur l'intervalle $[0, T]$ est fait de manière itérative, en s'appuyant sur deux propagateurs (plusieurs variantes existent, on se restreint ici au cas le plus simple) :

- un propagateur grossier, qui consiste typiquement à utiliser un schéma à un pas, noté ψ , avec un pas de temps grossier δt_g . Utiliser le propagateur grossier pour avancer le système du temps $\Delta T = n_g \delta t_g$ consiste à faire n_g pas avec le pas de temps δt_g :

$$\mathcal{G}_{\Delta T} = (\psi_{\delta t_g})^{(n_g)}.$$

- un propagateur fin, qui consiste par exemple à utiliser le même schéma que le propagateur grossier, mais avec un petit pas de temps δt_f . Ainsi, utiliser le propagateur fin pour avancer le système du temps $\Delta T = n_f \delta t_f$ consiste à faire n_f pas de pas de temps δt_f :

$$\mathcal{F}_{\Delta T} = (\psi_{\delta t_f})^{(n_f)}.$$

Comme $\delta t_f < \delta t_g$, le propagateur fin a un coût plus élevé que le propagateur grossier. On cherche à calculer la solution de (70) sur l'intervalle $[0, T]$. Pour $N \in \mathbb{N}^*$, on pose $\Delta T = T/N$. Initialement, on définit $\{u_n^0\}_{0 \leq n \leq N}$ par

$$u_{n+1}^0 = \mathcal{G}_{\Delta T}(u_n^0), \quad u_0^0 = u_0,$$

ce qui revient simplement à utiliser le propagateur grossier $\mathcal{G}_{\Delta T}$ de manière séquentielle.

On suppose maintenant connaître une approximation $\{u_n^k\}_{0 \leq n \leq N}$ de la trajectoire, à une certaine itération k . L'algorithme pararéel consiste à définir l'itération suivante, $\{u_n^{k+1}\}_{0 \leq n \leq N}$, par

$$u_0^{k+1} = u_0 \text{ et, pour tout } 0 \leq n \leq N-1, \quad u_{n+1}^{k+1} = \mathcal{G}_{\Delta T}(u_n^{k+1}) + \mathcal{F}_{\Delta T}(u_n^k) - \mathcal{G}_{\Delta T}(u_n^k), \quad (71)$$

où $\mathcal{F}_{\Delta T}$ est le propagateur fin sur l'intervalle de temps ΔT .

Par construction, au début de l'itération $k + 1$, on connaît les valeurs de u_n^k , pour tout $0 \leq n \leq N$. On commence donc par calculer $\mathcal{F}_{\Delta T}(u_n^k) - \mathcal{G}_{\Delta T}(u_n^k)$ en *parallèle* sur chaque intervalle $[n\Delta T, (n+1)\Delta T]$. Il suffit ensuite de calculer $\mathcal{G}_{\Delta T}(u_n^{k+1})$ et de corriger ceci par la différence $\mathcal{F}_{\Delta T}(u_n^k) - \mathcal{G}_{\Delta T}(u_n^k)$. Cette étape est séquentielle, mais elle ne fait appel qu'au propagateur grossier. Le propagateur fin est systématiquement utilisé lors de calculs en *parallèle*, d'où l'intérêt de l'algorithme (71).

Remarque 21 *Par construction, le schéma (71) est tel que, pour $k \geq n$, on a $u_n^k = u_n^{\text{fin}}$, où $\{u_n^{\text{fin}}\}_{0 \leq n \leq N}$ est la trajectoire obtenue en utilisant le propagateur fin de manière séquentielle : $u_0^{\text{fin}} = u_0$ et, pour tout $0 \leq n \leq N - 1$,*

$$u_{n+1}^{\text{fin}} = \mathcal{F}_{\Delta T}(u_n^{\text{fin}}).$$

Lorsque $k \geq n$, le gain en terme de coût de calcul est nul par rapport à l'utilisation du propagateur fin de manière purement séquentielle. On s'intéressera donc exclusivement au cas $n > k$ dans la suite. \diamond

On s'intéresse ici au cas où l'équation étudiée est Hamiltonienne. Ainsi, dans (70), $u = (q, p) \in \mathbb{R}^d \times \mathbb{R}^d$ et la dynamique s'écrit

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}, \quad (72)$$

où le Hamiltonien $H(q, p)$ est une fonction régulière de q et p . Ce cadre a bien sûr des applications en simulation moléculaire, mais aussi dans d'autres domaines (mécanique céleste, élastodynamique, ...). On s'intéresse plus particulièrement au cas où H est de la forme (1). La dynamique (72) s'écrit alors

$$\dot{q} = M^{-1}p, \quad \dot{p} = -\nabla V(q). \quad (73)$$

Un schéma populaire pour intégrer cette dynamique est le schéma de velocity-Verlet, explicite et d'ordre 2, qui s'écrit

$$\begin{aligned} q_{n+1} &= q_n + M^{-1} \left(\delta t p_n - \frac{\delta t^2}{2} \nabla V(q_n) \right), \\ p_{n+1} &= p_n - \frac{\delta t}{2} (\nabla V(q_n) + \nabla V(q_{n+1})), \end{aligned} \quad (74)$$

où δt est le pas de temps.

Comme déjà mentionné ci-dessus, le flot d'un système Hamiltonien est symplectique, et il est connu que les schémas symplectiques, comme le schéma de velocity-Verlet, ont de bonnes propriétés en temps long. De plus, la dynamique Hamiltonienne (73) est réversible, au sens où $f \circ \rho = -\rho \circ f$, avec $\rho(q, p) = (q, -p)$ et où $f(q, p)$ est le membre de droite de la dynamique : $f(q, p) = (M^{-1}p, -\nabla V(q))$. En conséquence, le flot exact Ψ_t de (73) est ρ -réversible, au sens où

$$\forall t, \rho \circ \Psi_t = \Psi_t^{-1} \circ \rho. \quad (75)$$

Pour les dynamiques Hamiltoniennes *réversibles*, une classe de schémas intéressante est celle des schémas ρ -réversibles. On rappelle qu'un schéma à un pas défini par $u_{n+1} = \psi_{\delta t}(u_n)$ est ρ -réversible si $\psi_{\delta t}$ vérifie (75). Pour de tels schémas, appliqués à des problèmes Hamiltoniens réversibles, on peut démontrer d'excellentes propriétés en temps long (conservation de l'énergie, ...), à l'image des schémas symplectiques pour les problèmes Hamiltoniens généraux. Cependant, les conditions sous lesquelles ces résultats sont obtenus sont

plus restrictives que pour l'étude des schémas symplectiques : la dynamique Hamiltonienne est *réversible intégrable* (ceci implique en particulier que, si q et p sont de dimension d , alors la dynamique préserve d invariants), et le Hamiltonien satisfait une condition de *non-résonance* [84, Chap. XI].

Les schémas ρ -réversibles sont souvent construits comme des schémas symétriques. On rappelle qu'un schéma à un pas $u_{n+1} = \psi_{\delta t}(u_n)$ est *symétrique* (cf. [84, Définition V.1.4]) si, pour tout δt ,

$$\psi_{\delta t} \circ \psi_{-\delta t} = \text{Id}. \quad (76)$$

On montre alors qu'un schéma symétrique qui vérifie de plus la condition

$$\forall \delta t, \quad \rho \circ \psi_{\delta t} = \psi_{-\delta t} \circ \rho \quad (77)$$

est un schéma ρ -réversible au sens de (75) (cf. [84, Théorème V.1.5]). La condition (77) est très peu restrictive (elle est vérifiée par la plupart des schémas, y compris le schéma d'Euler explicite). Ainsi, en pratique, pour construire un schéma ρ -réversible, on commence par construire un schéma symétrique (vérifiant (76)), et on vérifie *a posteriori* que la relation (77) est satisfaite. Le schéma est alors ρ -réversible, et possède de bonnes propriétés en temps long. Le schéma velocity Verlet (74), qui est symplectique, est aussi symétrique et ρ -réversible.

Comme mentionné ci-dessus, les conditions sous lesquelles on peut démontrer de bonnes propriétés pour les schémas réversibles sont plus restrictives que dans le cas de schémas symplectiques. Cependant, l'intérêt des schémas symétriques vient du fait qu'il est souvent plus facile de construire un schéma symétrique qu'un schéma symplectique. Il est par exemple possible de "symétriser" un schéma : pour tout schéma à un pas $u_{n+1} = \psi_{\delta t}(u_n)$, le schéma

$$u_{n+1} = \psi_{\delta t/2} \circ (\psi_{-\delta t/2})^{-1}(u_n),$$

qu'on peut aussi écrire sous la forme

$$\psi_{-\delta t/2}(u_{n+1/2}) = u_n, \quad u_{n+1} = \psi_{\delta t/2}(u_{n+1/2}), \quad (78)$$

est symétrique (cf. [84, Chapitre V]). Il n'existe au contraire pas de méthode systématique pour rendre symplectique un schéma quelconque.

On revient maintenant à l'utilisation de l'algorithme pararéel (71) dans un contexte Hamiltonien. Il se trouve que, même si les propagateurs grossier $\mathcal{G}_{\Delta T}$ et fin $\mathcal{F}_{\Delta T}$ sont symplectiques (resp. symétriques), le schéma pararéel global, au sens l'application $u_0 = (q_0, p_0) \mapsto u_n^k = (q_n^k, p_n^k)$ (avec $n > k \geq 1$), n'est pas symplectique (resp. symétrique). Cette absence de structure a pour conséquence que, lorsqu'on intègre un problème Hamiltonien avec le schéma (71), l'énergie n'est pas conservée aux itérations pararéelles $k \geq 1$ (le comportement à l'itération $k = 0$ dépend exclusivement des propriétés de $\mathcal{G}_{\Delta T}$; ainsi, si $\mathcal{G}_{\Delta T}$ est symplectique, et le pas de temps δt_g assez petit, l'énergie ne dérive pas à $k = 0$). Par conséquent, en temps long, le schéma (71) n'a pas de bonnes propriétés lorsqu'on l'utilise pour intégrer un problème Hamiltonien.

Une variante symétrique de l'algorithme pararéel Dans [12], une variante *symplectique* de l'algorithme pararéel a été introduite. L'idée est de reconstruire la fonction génératrice $S : (q, p) \in \mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}$ associée à l'application symplectique $\mathcal{F}_{\Delta T} \circ \mathcal{G}_{\Delta T}^{-1}$ (la notion de fonction génératrice a été détaillée dans la Section 1.3.2). En pratique, on obtient une approximation de S en utilisant une interpolation dans $\mathbb{R}^d \times \mathbb{R}^d$. L'utilisation d'une telle approche en dimension grande ne paraît pas très simple. Ceci motive notre

objectif de construire une variante *symétrique* de l'algorithme pararéel (71), avec l'idée sous-jacente qu'une telle construction est plus simple que la construction d'un algorithme symplectique.

Pour ce faire, nous allons utiliser la formule de symétrisation (78). Il est auparavant utile de récrire le schéma pararéel standard (71) comme un schéma à un pas. Considérant les K premières itérations de l'algorithme, on définit

$$U_n := (u_n^0, u_n^1, \dots, u_n^K). \quad (79)$$

Alors le schéma (71) s'écrit

$$U_{n+1} = \Psi_{\Delta T}(U_n), \quad (80)$$

où l'application $\Psi_{\Delta T}$ est définie par

$$\Psi_{\Delta T}(U_n) = \begin{pmatrix} \mathcal{G}_{\Delta T}(u_n^0) \\ \mathcal{G}_{\Delta T}(u_n^1) + \mathcal{F}_{\Delta T}(u_n^0) - \mathcal{G}_{\Delta T}(u_n^0) \\ \dots \\ \mathcal{G}_{\Delta T}(u_n^K) + \mathcal{F}_{\Delta T}(u_n^{K-1}) - \mathcal{G}_{\Delta T}(u_n^{K-1}) \end{pmatrix}.$$

On peut maintenant appliquer la procédure de symétrisation (78) à $\Psi_{\Delta T}$, ce qui permet de définir le schéma

$$\begin{aligned} U_n &= \Psi_{-\Delta T/2}(U_{n+1/2}), \\ U_{n+1} &= \Psi_{\Delta T/2}(U_{n+1/2}). \end{aligned} \quad (81)$$

En utilisant la définition de l'application Ψ , on obtient la formulation équivalente suivante : à l'itération $k = 0$,

$$u_{n+1/2}^0 = \mathcal{G}_{-\Delta T/2}^{-1}(u_n^0), \quad u_{n+1}^0 = \mathcal{G}_{\Delta T/2}(u_{n+1/2}^0), \quad (82)$$

et, pour les itérations $k \geq 1$,

$$\begin{cases} u_{n+1/2}^{k+1} &= \mathcal{G}_{-\Delta T/2}^{-1} \left[u_{n+1/2}^{k+1} - \mathcal{F}_{-\Delta T/2}(u_{n+1/2}^k) + \mathcal{G}_{-\Delta T/2}(u_{n+1/2}^k) \right], \\ u_{n+1}^{k+1} &= \mathcal{G}_{\Delta T/2}(u_{n+1/2}^{k+1}) + \mathcal{F}_{\Delta T/2}(u_{n+1/2}^k) - \mathcal{G}_{\Delta T/2}(u_{n+1/2}^k). \end{cases} \quad (83)$$

Dans la suite, le schéma (82)-(83) est appelé *schéma pararéel symétrique*. Par construction, ce schéma est symétrique, lorsqu'on le considère comme un schéma sur U_n défini par (79). On peut aussi vérifier qu'il satisfait la condition (77). Par conséquent, comme détaillé ci-dessus, ce schéma est ρ -réversible.

Comme dans le cas standard, tous les calculs coûteux (ceux impliquant le propagateur fin) sont réalisés de manière indépendante sur chaque processeur, en parallèle. Sous l'hypothèse que la complexité des calculs en parallèle est supérieure à la complexité des calculs séquentiels (ce qui revient à supposer qu'avancer de l'intervalle ΔT par le propagateur fin est plus coûteux qu'avancer de l'intervalle T par le propagateur grossier), on observe que la complexité algorithmique de (82)-(83) est identique à celle du schéma standard (71).

La mise en oeuvre de l'algorithme (82)-(83) conduit malheureusement à des résultats numériques décevants, car très similaires à ceux obtenus avec l'algorithme standard (cf. la Figure 6) : à nouveau, aux itérations pararéelles $k \geq 1$, l'énergie n'est pas conservée. La raison est la suivante. A l'itération pararéelle K , on peut considérer que l'algorithme (71) est un schéma à un pas sur le système constitué par $1 + K$ copies identiques du système original (c'est l'idée qui a permis de reformuler (71) sous la forme (80)). A l'itération $k = 0$, au vu de (82), on voit qu'on intègre la première copie par un schéma symétrique. Si le système est *réversible intégrable*, alors son énergie est conservée le long de la trajectoire

numérique, et c'est effectivement le cas en pratique (cf. la Figure 6, courbe $k = 0$). Comme les copies n'interagissent pas les unes avec les autres, le système composé des $1 + K$ copies est à nouveau réversible intégrable, pour l'énergie $\sum_{k=0}^K H(q^k, p^k)$. Si le schéma symétrique (82)-(83) conserve l'énergie de ce système global, alors, par récurrence sur k , on obtient que l'énergie est préservée à chaque itération parallèle k . Ceci est contradictoire avec les résultats numériques ! Ce qui a été oublié dans le raisonnement ci-dessus, c'est le fait qu'on démontre la préservation de l'énergie par un schéma symétrique sous, en particulier, une *hypothèse de non-résonance*. Or, comme le système global est constitué de $1 + K$ copies du système original, il y a par construction des fréquences résonantes dans ce système. Par conséquent, les hypothèses permettant de prouver la conservation de l'énergie ne sont pas remplies, et les résultats numériques confirment que l'énergie n'est pas bien conservée. Il est donc nécessaire de corriger le schéma (82)-(83).

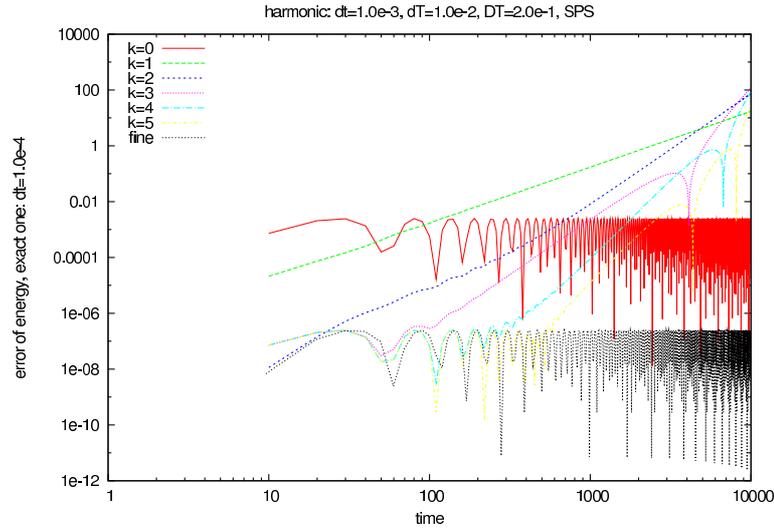


FIG. 6 – Erreur relative sur la conservation de l'énergie, $\frac{|H(q_n^k, p_n^k) - H(q_0, p_0)|}{|H(q_0, p_0)|}$, par l'algorithme (82)-(83), dans le cas de l'oscillateur harmonique (figure extraite de [FL26]).

Remarque 22 Une méthode pour supprimer les résonances est d'utiliser, à chaque itération k , un pas de temps $\delta t_g^{(k)}$ différent pour le propagateur grossier \mathcal{G} . Ceci supprime effectivement les résonances, et l'énergie ne diverge plus. Mais l'itération k est alors un moins bon préconditionneur pour l'itération $k + 1$, si bien que des problèmes de précision apparaissent. Il semble difficile de trouver des pas de temps $\delta t_g^{(k)}$ suffisamment différents pour que les systèmes à chaque itération k soient non-résonants, et suffisamment proches pour que l'itération k soit un bon préconditionneur pour l'itération $k + 1$. \diamond

Algorithme parallèle symétrique avec projection symétrique Afin d'améliorer l'algorithme (82)-(83), nous avons utilisé une procédure de projection, construite de façon à conserver le caractère symétrique de l'algorithme. On commence par décrire cette idée dans un contexte général. Soit

$$\mathcal{M} = \left\{ u = (q, p) \in \mathbb{R}^{2d}; H(q, p) = H_0 \right\}$$

la sous-variété des états d'énergie fixée, et soit $\Psi_{\Delta T}^{\text{sym}}$ un schéma symétrique. On considère l'algorithme suivant, pour définir une trajectoire numérique $\{u_n\}_{n \in \mathbb{N}}$ sur \mathcal{M} . On suppose

que $u_n \in \mathcal{M}$. On définit u_{n+1} par

$$\begin{cases} \tilde{u}_n &= u_n + \mu \nabla H(u_n), \\ \tilde{u}_{n+1} &= \Psi_{\Delta T}^{\text{sym}}(\tilde{u}_n), \\ u_{n+1} &= \tilde{u}_{n+1} + \mu \nabla H(u_{n+1}), \end{cases} \quad (84)$$

où μ est tel que $u_{n+1} \in \mathcal{M}$. On note $\pi_{\mathcal{M}}^{\text{sym}}$ cet algorithme :

$$u_{n+1} = \pi_{\mathcal{M}}^{\text{sym}}(u_n) \quad \text{si et seulement si (84) est satisfait.}$$

Il est important de remarquer que le *même* multiplicateur de Lagrange μ est utilisé à la première et dernière ligne de (84). Comme $\Psi_{\Delta T}^{\text{sym}}$ est symétrique, on obtient que $\pi_{\mathcal{M}}^{\text{sym}}$ est un schéma symétrique. On remarque aussi que, dans (84), μ et u_{n+1} sont définis de manière implicite. En pratique, ils sont calculés en utilisant un algorithme de type Newton.

Pour utiliser cette idée dans le contexte d'algorithmes de type pararéel, on revient au formalisme introduit précédemment. On écrit le schéma pararéel standard sous la forme d'un schéma à un pas (c'est la formulation (80)), sa version symétrisée est (81), qu'on note ici sous la forme $U_{n+1} = \Psi_{\Delta T}^{\text{sym}}(U_n)$, où l'application $\Psi_{\Delta T}^{\text{sym}}$ est symétrique au sens usuel. On introduit ensuite les énergies

$$H_k(U_n) := H(u_n^k), \quad 1 \leq k \leq K,$$

et on considère l'algorithme construit sur la base de l'application symétrique $\Psi_{\Delta T}^{\text{sym}}$ et sur la projection symétrique sur la sous-variété où les énergies H_k sont préservées. Suivant (84), cet algorithme s'écrit

$$\begin{aligned} \tilde{U}_n &= U_n + \sum_{k=1}^K \mu_k \nabla H_k(U_n), \\ \tilde{U}_{n+1} &= \Psi_{\Delta T}^{\text{sym}}(\tilde{U}_n), \\ U_{n+1} &= \tilde{U}_{n+1} + \sum_{k=1}^K \mu_k \nabla H_k(U_{n+1}), \end{aligned}$$

où les multiplicateurs de Lagrange μ_k sont tels que $H_k(U_{n+1}) = H_0$ pour tout $1 \leq k \leq K$. On obtient ainsi le schéma suivant, appelé *schéma pararéel symétrique avec projection symétrique* :

Algorithme 23 Soit u_0 la condition initiale.

1. *Initialisation* : on pose $u_0^0 = u_0$, et on calcule $\{u_{n+1/2}^0\}_{0 \leq n \leq N-1}$ et $\{u_{n+1}^0\}_{0 \leq n \leq N-1}$ de manière séquentielle par

$$u_{n+1/2}^0 = \mathcal{G}_{-\Delta T/2}^{-1}(u_n^0), \quad u_{n+1}^0 = \mathcal{G}_{\Delta T/2}(u_{n+1/2}^0).$$

On pose $\tilde{u}_{n+1/2}^0 = u_{n+1/2}^0$.

2. On suppose que, pour $k \geq 0$, la suite $\{\tilde{u}_{n+1/2}^k\}_{0 \leq n \leq N-1}$ est connue. On calcule alors $\{u_n^{k+1}\}_{0 \leq n \leq N}$ et $\{\tilde{u}_{n+1/2}^{k+1}\}_{0 \leq n \leq N-1}$ à l'itération $k+1$ par les étapes suivantes :

- (a) Pour tout $0 \leq n \leq N-1$, on calcule en parallèle $\mathcal{F}_{-\Delta T/2}(\tilde{u}_{n+1/2}^k)$, $\mathcal{F}_{\Delta T/2}(\tilde{u}_{n+1/2}^k)$, $\mathcal{G}_{-\Delta T/2}(\tilde{u}_{n+1/2}^k)$ et $\mathcal{G}_{\Delta T/2}(\tilde{u}_{n+1/2}^k)$;

(b) On pose $u_0^{k+1} = u_0$;

(c) Pour $0 \leq n \leq N - 1$, on calcule $\tilde{u}_{n+1/2}^{k+1}$ et u_{n+1}^{k+1} de manière séquentielle par

$$\begin{aligned}\tilde{u}_n^{k+1} &= u_n^{k+1} + \mu_{k+1} \nabla H(u_n^{k+1}) \\ \tilde{u}_{n+1/2}^{k+1} &= \mathcal{G}_{-\Delta T/2}^{-1} \left[\tilde{u}_n^{k+1} - \mathcal{F}_{-\Delta T/2}(\tilde{u}_{n+1/2}^k) + \mathcal{G}_{-\Delta T/2}(\tilde{u}_{n+1/2}^k) \right] \\ \tilde{u}_{n+1}^{k+1} &= \mathcal{G}_{\Delta T/2}(\tilde{u}_{n+1/2}^{k+1}) + \mathcal{F}_{\Delta T/2}(\tilde{u}_{n+1/2}^k) - \mathcal{G}_{\Delta T/2}(\tilde{u}_{n+1/2}^k) \\ u_{n+1}^{k+1} &= \tilde{u}_{n+1}^{k+1} + \mu_{k+1} \nabla H(u_{n+1}^{k+1})\end{aligned}$$

où μ_{k+1} est tel que $H(u_{n+1}^{k+1}) = H_0$.

A nouveau, tous les calculs coûteux (qui impliquent le propagateur fin) sont réalisés en parallèle. Sous la même hypothèse que ci-dessus, la complexité de l’Algorithme 23 est identique à celle de l’algorithme symétrique (82)-(83).

Nous avons mis en oeuvre cet algorithme sur deux cas tests, le problème de Kepler, qui correspond à l’énergie potentielle

$$V(q) = -\frac{1}{\|q\|}, \quad q \in \mathbb{R}^2,$$

et la simulation du système solaire extérieur (le Soleil et les 5 planètes à partir de Jupiter), qui correspond à l’énergie potentielle

$$V(q) = -\sum_{1 \leq i < j \leq 6} \frac{G m_i m_j}{\|q_i - q_j\|}, \quad q \in (\mathbb{R}^3)^6,$$

où G est la constante de gravitation et m_i la masse de la “particule” i .

Remarque 24 *Pour ces deux cas tests, la dynamique Hamiltonienne est réversible intégrable, et possède donc plusieurs invariants, en plus de l’énergie. On note que l’Algorithme 23 utilise une projection sur la sous-variété d’énergie constante, mais ne fait pas intervenir les autres invariants. Cet algorithme est donc générique, par opposition à un algorithme qui utiliserait une projection sur la sous-variété où tous les invariants de la dynamique – qui, en pratique, sont parfois difficilement identifiables – sont conservés. \diamond*

L’Algorithme 23 conduit à de bons résultats. Tout d’abord, il suffit de peu d’itérations pour résoudre le problème de projection sur la sous-variété d’énergie constante (c’est-à-dire la détermination de μ et u_{n+1} dans la formulation concise (84)) : en pratique, deux itérations suffisent. Ceci est en particulier du au fait qu’on dispose d’une bonne configuration pour initialiser l’algorithme de projection.

On observe aussi qu’après un petit nombre d’itérations pararélles, la trajectoire obtenue grâce à l’Algorithme 23 est d’une précision équivalente à celle de la trajectoire obtenue par l’utilisation purement séquentielle du propagateur fin. A titre d’exemple, pour le problème de Kepler avec les pas de temps considérés dans [FL26], 5 itérations suffisent pour obtenir une trajectoire précise¹ sur l’intervalle $[0, T]$ avec $T = 10^4$, alors qu’après 15 itérations avec l’algorithme standard (71), on obtient une trajectoire qui n’est précise que sur l’intervalle $[0, T]$ avec $T \approx 300$ (l’algorithme symétrique (82)-(83), qui n’utilise pas de projection, donne les mêmes résultats que l’algorithme standard (71)). Le moment

¹au sens : l’erreur entre cette trajectoire et une trajectoire de référence est identique à l’erreur entre la trajectoire obtenue par le propagateur fin utilisé séquentiellement et cette même trajectoire de référence.

cinétique $L = q \times p$, qui est un autre invariant de la dynamique exacte, est lui-aussi bien conservé. On renvoie à [FL26] pour plus de détails. Ces bons résultats sont aussi observés dans le cas du système solaire extérieur, qui est un problème plus difficile, car posé en dimension plus grande.

Nous avons mentionné ci-dessus que l'Algorithme symétrique 23, utilisant une projection symétrique, permet d'obtenir de meilleurs résultats que l'algorithme symétrique (82)-(83), qui n'utilise pas de projection. Nous avons aussi observé que cet Algorithme 23 donne de meilleurs résultats que l'algorithme pararéel standard couplé à une projection standard (non symétrique). Ainsi, les meilleurs résultats sont obtenus lorsqu'on utilise simultanément deux ingrédients, la *symétrisation* et la *projection symétrique*.

1.4 Problèmes de conductivité thermique

Publication [FL18]

On s'intéresse dans cette section à des problèmes de conductivité thermique, dans des chaînes d'atomes mono-dimensionnelles, avec interaction entre plus proches voisins. Commençons par poser le problème.

On se donne un Hamiltonien de la forme

$$H(q, p) = \sum_{i=1}^n \frac{p_i^2}{2} + \sum_{i=1}^n V(q_i - q_{i-1}), \quad (85)$$

où q_i représente la position courante de la particule i (on note $q = (q_1, \dots, q_n) \in \mathbb{R}^n$), p_i son impulsion, et V est le potentiel d'interaction de plus proches voisins. On supprime l'invariance par translation en fixant $q_0 = 0$. Pour simplifier la présentation, tous les atomes ont la même masse, qu'on a fixée à $m = 1$.

On considère ensuite la dynamique suivante :

- les particules qui ne sont pas aux extrémités de la chaîne suivent une dynamique Hamiltonienne : pour $2 \leq i \leq n - 1$,

$$\dot{q}_i = \partial_{p_i} H = p_i, \quad \dot{p}_i = -\partial_{q_i} H = V'(q_{i+1} - q_i) - V'(q_i - q_{i-1}), \quad (86)$$

où H est le Hamiltonien (85).

- les particules aux deux bords sont en contact avec un thermostat, la température T_n à droite étant différente de la température T_1 à gauche. Un modèle simple est l'équation de Langevin,

$$dq_1 = p_1 dt, \quad dp_1 = [V'(q_2 - q_1) - V'(q_1 - q_0) - \xi p_1] dt + \sqrt{2\xi T_1} dW_1 \quad (87)$$

et

$$dq_n = p_n dt, \quad dp_n = [-V'(q_n - q_{n-1}) - \xi p_n] dt + \sqrt{2\xi T_n} dW_n, \quad (88)$$

où $\xi > 0$ est un paramètre et $q_0 = 0$ dans (87).

Formellement, l'idée est d'imposer une différence de température aux deux extrémités du système, et de ne pas perturber la dynamique au sein du système. On s'attend à ce qu'un profil de température s'établisse dans la chaîne, dont les caractéristiques sont reliées aux propriétés de conductivité thermique du modèle.

Plus précisément, on définit l'énergie E_i de l'atome i par

$$E_i = \frac{p_i^2}{2} + \frac{1}{2} [V(q_i - q_{i-1}) + V(q_{i+1} - q_i)].$$

Un calcul simple montre que, pour tout $3 \leq i \leq n - 2$,

$$\frac{dE_i}{dt} = J_{i-1,i} - J_{i,i+1}, \quad (89)$$

avec

$$J_{i,i+1} = -\frac{1}{2}(p_{i+1} + p_i)V'(q_{i+1} - q_i).$$

Ainsi, pour tout atome n'interagissant pas avec un atome en contact avec un thermostat, la variation d'énergie est donnée par la relation (89), dans laquelle on peut interpréter $J_{i,i+1}$ comme le flux d'énergie de l'atome i à l'atome $i + 1$.

Supposons maintenant que la dynamique (86)-(87)-(88) est ergodique pour une mesure $\mu_{T_1, T_n}(q, p)dqdp$ (on renvoie à [36] et [129, Sec. 3] pour des travaux démontrant cette affirmation). Alors le courant moyen d'énergie entre les atomes i et $i + 1$ est

$$\mathcal{J}_{i,i+1} = \int J_{i,i+1} \mu_{T_1, T_n}(q, p)dqdp.$$

Le système étant homogène, on montre que $\mathcal{J}_{i,i+1}$ est en fait indépendant de i , et ne dépend que de la taille n du système, des températures imposées T_1 et T_n , et éventuellement de ξ . On le note $\mathcal{J}_n(T_1, T_n)$.

Remarque 25 Dans le cas $T_n = T_1$, la mesure $\exp(-H(q, p)/T_1)dqdp$ est une mesure stationnaire pour (86)-(87)-(88). Dans le cas $T_1 \neq T_n$, on ne sait pas, en général, exhiber une mesure invariante de la dynamique. Ce genre de situations est donc plus difficile que celle étudiée par exemple dans la Section 1.2 : dans les deux cas, on peut montrer que le problème admet une mesure invariante, mais ici, on ne sait pas l'écrire analytiquement.

Dans un modèle macroscopique (en dimension un pour simplifier), le flux d'énergie $\mathcal{J}(x)$ est relié au gradient de température locale par la loi de Fourier, $\mathcal{J}(x) = \kappa T'(x)$, où κ est la conductivité thermique du matériau. Dans le modèle microscopique ici considéré, on définit par analogie la conductivité thermique $\kappa_n(T_1, T_n)$ de la chaîne d'atomes par

$$\mathcal{J}_n(T_1, T_n) = \kappa_n(T_1, T_n) \frac{T_n - T_1}{n}. \quad (90)$$

Si la limite $\kappa := \lim_{n \rightarrow \infty} \lim_{T_n \rightarrow T_1} \kappa_n(T_1, T_n)$ existe, alors on dit que le matériau a une conductivité normale, et κ est sa conductivité [30].

Pour de nombreux modèles, la conductivité n'est pas *normale*. Par exemple, dans le cas où V est harmonique dans (85), on peut calculer analytiquement $\mathcal{J}_n(T_1, T_n)$, qui est constant avec n [130]. On déduit donc de (90) que la conductivité $\kappa_n(T_1, T_n)$ est proportionnelle à la taille n du système, et diverge dans la limite thermodynamique. Le même type de phénomène se produit dès que le Hamiltonien H est complètement intégrable [147, 142, 88].

La situation est très différente lorsque la dynamique interne (86) est perturbée par des termes aléatoires. En fonction du type de la perturbation, la conductivité est finie, ou bien diverge comme n^α pour un certain $0 < \alpha < 1$ (cf. par exemple [29, 18, 15]).

Dans [FL18], nous avons étudié numériquement le comportement asymptotique de la conductivité κ_n avec n , dans le cas où V est le potentiel de Toda,

$$V(r) = \frac{a}{b} \exp(-br) + ar + c,$$

où $a > 0$, $b > 0$ et c sont des paramètres du modèle. Ce choix est motivé par le fait que le Hamiltonien (85) est alors complètement intégrable. Nous avons introduit une perturbation

de la dynamique interne (86), qui consiste à échanger les moments de deux particules voisines à des temps aléatoires exponentiellement distribués (on note γ^{-1} le temps moyen entre deux échanges). On remarque que, lors d'un tel échange entre les particules i et $i + 1$, l'énergie locale $E_i + E_{i+1}$ est conservée, de même que l'impulsion locale $p_i + p_{i+1}$. Cette perturbation est donc considérée comme faible (par rapport à une perturbation dans laquelle l'énergie locale $E_i + E_{i+1}$ ou le moment local $p_i + p_{i+1}$ ne seraient pas préservés).

Lorsque $\gamma = 0$, la dynamique interne est (86), et la conductivité diverge comme n , le problème étant complètement intégrable. Lorsque $\gamma > 0$, nous avons constaté numériquement que la conductivité diverge encore, suivant la loi $\kappa_n(T_1, T_n) \sim n^\alpha$, pour un certain exposant $0 < \alpha \leq 1/2$. Le fait d'introduire une perturbation stochastique, aussi petite soit-elle, change donc immédiatement le taux de divergence de la conductivité.

Le cas que nous avons étudié numériquement appartient à une classe de problèmes pour lesquels l'existence (ou la non-existence) d'une valeur universelle pour l'exposant α fait l'objet d'un vif débat dans la littérature [120, 106]. Un intérêt de notre travail est que nous avons constaté, dans nos simulations, que l'exposant α *dépend* de γ , qui mesure l'intensité de la perturbation stochastique. Ceci semble donc être en contradiction avec l'existence d'une valeur universelle.

2 Méthodes couplées discret-continu

Cette seconde partie est consacrée à l'étude de quelques méthodes de couplage *discret-continu*, pour la simulation numérique des matériaux. L'idée est de coupler une description atomistique de la matière, utilisant comme degrés de liberté les positions des atomes (qui interagissent les uns avec les autres par des lois à l'échelle microscopique) et une description plus traditionnelle, via les concepts de la mécanique du continuum.

Pour illustrer notre propos, il est utile de brièvement rappeler quelques notions standard de mécanique du continuum, dans un cadre stationnaire en temps (on renvoie par exemple à [41] pour une présentation plus complète). Soit $\mathcal{D} \subset \mathbb{R}^d$ le domaine occupé par le système dans sa configuration de référence, et

$$\varphi : \mathcal{D} \rightarrow \mathbb{R}^d$$

la déformation du système : $\varphi(x)$ est la position dans la configuration déformée du point qui est à la position x dans la configuration de référence. On note $F = \nabla\varphi : \mathcal{D} \rightarrow \mathcal{M}_d$ le gradient de déformation (\mathcal{M}_d est l'espace des matrices de taille $d \times d$). L'équilibre du solide dans \mathcal{D} lorsqu'il est soumis à des forces de volume f s'écrit

$$-\operatorname{div} T = f \quad \text{dans } \mathcal{D}, \quad (91)$$

avec des conditions aux limites adéquates (typiquement, on impose sur le bord $\partial\mathcal{D}$ la position φ , ou bien les forces $T \cdot n$, où n est la normale sortante). Ici, T est le tenseur des contraintes (plus précisément, le premier tenseur des contraintes de Piola-Kirchhoff).

L'équation (91) est une équation universelle, qui exprime l'équilibre des forces. Pour fermer le système, il faut se donner une relation entre le tenseur des contraintes T et des grandeurs reliées à la cinématique du solide (telles que φ ou F). Cette relation, appelée *loi de comportement*, ou *équation constitutive*, dépend du matériau considéré. C'est elle qui contient la physique ou la mécanique du système. On se restreint ici à la modélisation de solides *hyperélastiques*, pour lesquels il existe une densité d'énergie élastique,

$$W : (x, F) \in \mathcal{D} \times \mathcal{M}_d \rightarrow \mathbb{R},$$

telle que la loi constitutive s'écrit

$$T(x, F) = \frac{\partial W}{\partial F}(x, F). \quad (92)$$

Ainsi, le tenseur des contraintes ne dépend de φ que via son gradient, et ce de manière locale.

Supposons pour simplifier qu'on munit (91)-(92) de conditions aux limites de type Dirichlet sur tout le bord $\partial\mathcal{D}$:

$$\forall x \in \partial\mathcal{D}, \quad \varphi(x) = \varphi_0(x). \quad (93)$$

On reconnaît alors dans (91)-(92)-(93) l'équation d'Euler-Lagrange associée au problème

$$\inf \{E_M(\varphi), \varphi \in \mathcal{A}_M\}, \quad (94)$$

où \mathcal{A}_M est l'ensemble des déformations *admissibles* (et qui vérifient donc en particulier les conditions aux limites), et $E_M(\varphi)$ est l'énergie macroscopique,

$$E_M(\varphi) = \int_{\mathcal{D}} W(x, \nabla\varphi(x)) dx - \int_{\mathcal{D}} f \varphi.$$

Remarque 26 *Pour tous les modèles réalistes, W est une fonction non-convexe de F (en raison de l'invariance du modèle par mouvement de corps rigide). Par conséquent, l'étude de problèmes variationnels du type (94) n'est pas simple, et fait par exemple appel à la notion de quasi-convexité, qui généralise celle de convexité. On renvoie à [14] et [41].* \diamond

Sur le plan de la modélisation, l'écriture d'une loi constitutive (par exemple du type (92)) n'est pas simple. Cela nécessite une très bonne compréhension du matériau, souvent acquise sur la base de résultats expérimentaux. Or, les matériaux sont soumis de plus en plus à des conditions de chargement extrêmes, en terme de température, pression, contraintes mécaniques (penser aux matériaux utilisés dans l'industrie nucléaire actuelle, et dans les réacteurs à fusion en projet), ... Le renouvellement fréquent des matériaux donne peu de temps pour parvenir à une compréhension empirique de leurs propriétés. Enfin, pour certaines sollicitations, la modélisation via un continuum n'est tout simplement pas adaptée. C'est par exemple le cas lorsqu'on s'intéresse à la propagation de fractures dans des matériaux cristallins (comme les métaux). La fracture avance car des liaisons atomiques cassent. La description intime du phénomène se fait donc à l'échelle atomistique.

Ceci motive l'apparition, depuis une quinzaine d'années, de modèles *discret-continu*, dont l'objectif est de compenser le manque d'information à l'échelle macroscopique (et la difficulté à en obtenir) par la mise en oeuvre de calculs à une échelle plus fine. L'idée est donc d'utiliser un modèle fin pour *calculer* la loi de comportement à l'échelle macroscopique. Par construction, la physique contenue à l'échelle microscopique, de même que les invariances du système, sont naturellement présentes à l'échelle macroscopique. On s'attend aussi à ce qu'il soit plus facile d'obtenir et de valider ces modèles fins, écrits à une échelle plus petite, et que le rôle des différents paramètres du modèle soit plus facile à appréhender.

Nous avons décrit ci-dessus le modèle macroscopique adopté, celui de la mécanique du continuum. Décrivons maintenant le modèle microscopique, qui sera ici le modèle de référence. Pour simplifier, on suppose que le modèle atomistique ne fait intervenir que des interactions de paire. On note N le nombre total d'atomes (il y a donc $N^{1/d}$ atomes par dimension), et $V : \mathbb{R}^d \rightarrow \mathbb{R}$ le potentiel interatomique. En l'absence de forces extérieures, l'énergie totale associée à la configuration $\{\varphi_i\}_{1 \leq i \leq N}$ est alors

$$E_\mu \left(\{\varphi_i\}_{1 \leq i \leq N} \right) = \frac{1}{2} \sum_{i=1}^N \sum_{j \neq i, j=1}^N V \left(\frac{\varphi_i - \varphi_j}{N^{-1/d}} \right). \quad (95)$$

On remarquera que, dans l'argument de V , on a mis à l'échelle toutes les distances, en utilisant le facteur multiplicatif $N^{-1/d}$. On voit ainsi que, si la distance d'équilibre de V (typiquement, V est à symétrie radiale, $V(x) = \mathcal{V}(|x|)$, et la distance d'équilibre de V est le minimiseur de la fonction $r \in \mathbb{R}^+ \mapsto \mathcal{V}(r)$) est de l'ordre de 1, alors la distance d'équilibre du modèle est d'ordre $N^{-1/d}$. Comme il y a $N^{1/d}$ particules par dimension, on obtient ainsi un système qui occupe à l'échelle macroscopique un volume fini.

Remarque 27 *On suppose dans toute la suite que V est régulier. Dans les cas pertinents, V est à symétrie radiale (par invariance du modèle par rotation), et $\lim_{|x| \rightarrow 0} V(x) = +\infty$, tandis que $\lim_{|x| \rightarrow +\infty} V(x) = 0$. A nouveau, comme dans le cas de la mécanique du continuum, ces hypothèses, essentielles du point de vue de la modélisation, rendent l'étude mathématique difficile.* \diamond

Remarque 28 *L'énergie potentielle E_μ donnée par (95) est exactement l'énergie potentielle de la Section 1, qui apparaît par exemple dans (6).* \diamond

Au bilan, la difficulté de travailler avec un modèle macroscopique tel que (94) est que, dans certaines situations, il est difficile de choisir une bonne densité W . Au contraire, on suppose que le modèle microscopique (95) contient toute la physique nécessaire. Le problème de ce modèle est bien sûr son coût : un échantillon macroscopique de matière contient de l'ordre de $N \sim 10^{23}$ atomes. Même si on travaille avec des échantillons très petits à l'échelle macroscopique, le nombre de degrés de liberté en jeu est encore immense.

Cependant, dans beaucoup de situations d'intérêt, il est très fréquent que la zone dans laquelle le modèle macroscopique est mis en défaut (et où il faut donc utiliser le modèle microscopique) est très localisée. Dans le cas de problèmes de propagation de fracture, il s'agit de la zone au voisinage de la pointe de fracture. Pour des problèmes de nano-indentation, la zone critique est située au voisinage de l'indenteur (cf. la Figure 7). Il est donc naturel d'essayer de développer une approche de type décomposition de domaine, où le domaine de calcul est écrit comme

$$\mathcal{D} = \mathcal{D}_{\text{reg}} \cup \mathcal{D}_{\text{sing}}, \quad (96)$$

où \mathcal{D}_{reg} est le domaine dans lequel on s'attend à une déformation *régulière* (cf. les Figures 8 et 9). L'idée est donc, dans \mathcal{D}_{reg} , d'utiliser un modèle macroscopique (tel que celui de mécanique du continuum rappelé ci-dessus), en tirant profit du fait qu'il conduit à une discrétisation abordable en terme de coût calcul. A l'opposé, dans $\mathcal{D}_{\text{sing}}$, on s'attend à une déformation irrégulière, que seul le modèle microscopique peut décrire avec précision.

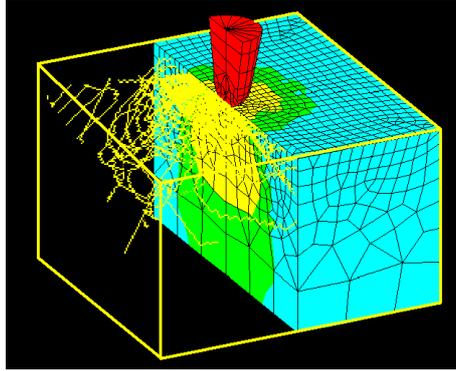


FIG. 7 – Une simulation numérique typique couplant un modèle discret avec un modèle de continuum, dans un cube de coté $2 \mu\text{m}$ (Image due à M. Fivel, INPG ; cf. aussi [70]).

2.1 Modèles à température nulle

On suppose pour commencer que le problème est posé comme un problème variationnel. Le modèle de référence est le modèle atomistique (95), et la solution de référence est le minimiseur de

$$\inf \left\{ E_\mu \left(\{\varphi_i\}_{1 \leq i \leq N} \right), \{\varphi_i\}_{1 \leq i \leq N} \in \mathcal{A}_\mu \right\}, \quad (97)$$

où \mathcal{A}_μ est l'ensemble des déformations microscopiques admissibles (pour simplifier la présentation, on suppose ici que les problèmes variationnels ont un unique minimiseur). En pratique, il est très difficile de résoudre ce problème, car le nombre de degrés de liberté est trop grand. L'approche discret-continu telle qu'esquissée ci-dessus consiste à passer par les étapes suivantes :

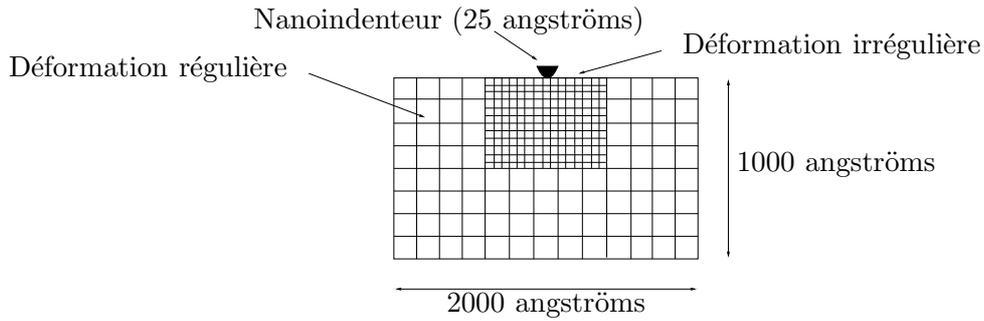


FIG. 8 – Représentation schématique d’une expérience de nanoindentation : près de l’indentateur, qui est très dur, on s’attend à une déformation irrégulière, dont la description nécessite un modèle fin. Plus loin, la déformation est régulière, et un modèle macroscopique, discrétisé sur un maillage grossier (ici, des quadrangles), permet déjà d’atteindre une bonne précision.

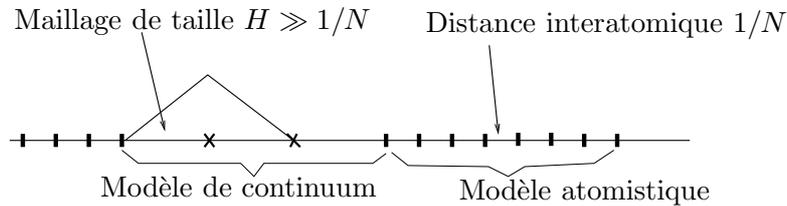


FIG. 9 – Partition du domaine \mathcal{D} (en dimension un) en un domaine *régulier* \mathcal{D}_{reg} , où on utilise un modèle de continuum, et un domaine *singulier* $\mathcal{D}_{\text{sing}}$, où on utilise un modèle atomistique.

1. dériver du modèle microscopique un modèle macroscopique, sous certaines hypothèses (typiquement de régularité de la déformation) ; on dispose ainsi de deux modèles, à deux échelles différentes, qui sont *consistants* l’un avec l’autre ;
2. proposer un modèle *couplé* utilisant les deux modèles, par exemple via une décomposition du domaine de calcul, et l’utilisation dans chaque zone du modèle le plus approprié. Les principales difficultés sont alors
 - (a) l’écriture de conditions de raccordement, entre un modèle discret et un modèle continu,
 - (b) le développement de stratégies adaptatives, où le choix des domaines \mathcal{D}_{reg} et $\mathcal{D}_{\text{sing}}$ dans (96) n’est pas fait *a priori*, mais est adapté par l’algorithme, un peu à la manière d’estimations *a posteriori*.

De nombreuses approches peuvent être utilisées pour dériver un modèle de continuum à partir d’un modèle atomistique. On renvoie ici à [21, 22] pour l’étude détaillée et rigoureuse du cas déterministe, à [24, 27, 25] pour le cas de réseaux aléatoires (cf. aussi les travaux reliés présentés dans [19]), et à l’article de revue [26].

Dans [FL5], nous avons proposé, dans un cadre mono-dimensionnel, un modèle couplé inspiré de la méthode QuasiContinuum (proposée initialement dans [138, 137], puis mise à jour dans [134]). En une dimension, les problèmes de raccordement de modèles sont considérablement simplifiés, ce qui nous a permis de nous concentrer sur les questions de choix de partition du domaine de calcul, et obtention d’estimations d’erreur, dans un cas où le modèle atomistique n’est *pas convexe* (cf. aussi [FL6] pour le traitement du cas convexe). Ces contributions ont été décrites dans [FL7]. On renvoie aussi, de manière plus générale, aux articles de revue [FL32, FL17, 26]. Notre principale conclusion est que, même

dans ce cadre mono-dimensionnel très simple, le couplage entre un modèle discret et un modèle continu peut conduire, dès lors que les modèles ne sont pas convexes, à des effets surprenants et non-désirés. Il est donc important de garder une grande prudence vis à vis de ces méthodes.

Notons enfin qu’il existe maintenant une littérature importante sur le couplage entre un modèle discret et continu, ou bien sur le couplage de modèles discrets dont l’un est la version “simplifiée” de l’autre (dans notre cas, cela consisterait à étudier le couplage entre le modèle atomistique d’une part et, d’autre part, la discrétisation du modèle de continuum). On renvoie à [5, 6, 7, 22, 54, 55, 56, 57, 58, 61, 62, 108, 109, 124, 1]. A notre connaissance, la plupart des études se concentre sur des problèmes mono-dimensionnels. De plus, le cas de potentiels non-convexes est rarement abordé.

2.2 Modèles à température finie

Publications [FL20, FL34, FL27]

Les travaux précédemment cités ont été effectués dans un cadre variationnel, ce qui correspond à une modélisation à température nulle. Nous nous sommes ensuite intéressés à l’extension de ces résultats quand le modèle de référence est à température finie. Nous nous sommes concentrés sur la question de la dérivation de modèles macroscopiques, partant d’un modèle atomistique très fin.

A température finie, l’objectif n’est pas de calculer une configuration particulière du système, mais de calculer des quantités du type

$$\langle A \rangle = \frac{\int_{\mathcal{D}^N} A(\varphi) \exp(-\beta E_\mu(\varphi)) d\varphi}{\int_{\mathcal{D}^N} \exp(-\beta E_\mu(\varphi)) d\varphi}, \quad (98)$$

où $\mathcal{D} \subset \mathbb{R}^d$ est le domaine accessible à chaque particule, $\varphi = (\varphi_1, \dots, \varphi_N)$ est le vecteur des positions des N atomes, A est une fonction dépendant de l’état du système (appelée aussi *observable*), E_μ est l’énergie atomistique (95), et β est un paramètre, représentant l’inverse de la température. La quantité (98) n’est autre que la moyenne canonique de l’observable A , déjà rencontrée dans la Section 1 (cf. les équations (6)-(7)), qui fait intervenir toutes les configurations du système (et pas seulement les minimiseurs globaux ou locaux de E_μ), avec des poids différents. On renvoie par exemple à [50] pour une discussion plus complète des questions de modélisation atomistique des solides.

Remarque 29 *Dans la limite de température basse, c’est-à-dire $\beta \rightarrow \infty$, les seules configurations qui comptent dans (98) sont les minimiseurs globaux de E_μ . On retrouve alors les problèmes variationnels de la Section 2.1. \diamond*

Une approche typique pour calculer (98) est d’écrire

$$\langle A \rangle = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T A(\varphi_t) dt \quad (99)$$

où φ_t est une solution de l’équation différentielle stochastique

$$d\varphi_t = -\nabla E_\mu(\varphi_t) dt + \sqrt{2\beta^{-1}} dB_t, \quad (100)$$

où $(B_t)_{t \geq 0}$ est un mouvement brownien en dimension dN . On renvoie à la Section 1.1.3 pour plus de détails. Le coût calcul est bien sûr croissant avec le nombre d’atomes N considérés.

Nous nous sommes intéressés au cas où l'observable A ne dépend pas de la position $\varphi \in \mathbb{R}^{dN}$ de tous les atomes, mais seulement de la position $\varphi_r \in \mathbb{R}^{dN_r}$ d'un *petit nombre* d'entre eux, avec

$$\varphi = (\varphi_1, \dots, \varphi_N) = (\varphi_r, \varphi_c), \quad \varphi_r \in \mathbb{R}^{dN_r}, \quad \varphi_c \in \mathbb{R}^{dN_c}, \quad N = N_r + N_c.$$

On cherche donc à calculer (98) pour une observable A qui ne dépend que de φ_r :

$$\langle A \rangle = Z^{-1} \int_{\mathcal{D}^N} A(\varphi_r) \exp(-\beta E_\mu(\varphi_r, \varphi_c)) d\varphi_r d\varphi_c. \quad (101)$$

Remarque 30 *On suppose ici que φ_r est donné. On ne discute pas son choix. Notons cependant que le fait qu'on distingue dans les atomes deux groupes, i.e. qu'on écrive $\varphi = (\varphi_r, \varphi_c)$ où φ_r et φ_c vont jouer des rôles différents, indique que les atomes ne sont pas indiscernables. Autrement dit, le système n'est pas invariant par permutation des atomes, par exemple parce que certains atomes, et pas d'autres, sont soumis à des forces importantes. \diamond*

Dans [FL20], nous avons proposé et analysé une méthode pour calculer la moyenne canonique d'une telle observable de façon plus efficace que via (99)-(100), dans un cadre mono-dimensionnel (cf. aussi [FL34] pour quelques résultats complémentaires). L'extension de l'approche à un cas bi-dimensionnel simple fait l'objet du travail [FL27] (cf. la Section 2.2.2). L'idée essentielle de notre approche est de travailler dans la limite thermodynamique où le nombre N_c de particules dont A ne dépend pas tend vers l'infini (on pourra lire une revue de l'état de l'art concernant les questions de limite thermodynamique en mécanique statistique, en lien avec la modélisation des solides à différentes échelles, dans [127]).

Dans cette limite, notre approche est *exacte*. De manière générale, elle est donc intéressante quand le nombre d'atomes éliminés est très grand devant le nombre d'atomes conservés : $N_c \gg N_r$. Notons de plus que, dans la limite $N_c \rightarrow \infty$, certains calculs deviennent en pratique possibles, alors qu'ils sont très coûteux pour N_c grand mais fini.

Remarque 31 *Cette problématique a aussi été abordée en adaptant la méthode Quasi-Continue, déjà évoquée ci-dessus, au cadre de la température finie (on renvoie à [60], et aussi aux travaux antérieurs [47, 107]). Le régime utilisé dans ces contributions est différent du nôtre : c'est un régime basse température. L'idée est essentiellement de prendre en compte les effets de température dans une approche perturbative par rapport au cas température nulle, en utilisant la méthode de Laplace pour calculer une approximation de (101) dans la limite $\beta \rightarrow +\infty$.*

Plus précisément, on introduit le potentiel effectif

$$E_{\text{CG}}(\varphi_r) := -\frac{1}{\beta} \ln \left[\int_{\mathcal{D}^{N_c}} \exp(-\beta E_\mu(\varphi_r, \varphi_c)) d\varphi_c \right], \quad (102)$$

qui n'est autre que l'énergie libre du système associée à la coordonnée de réaction $\xi(\varphi) = \xi(\varphi_r, \varphi_c) = \varphi_r$ (cf. la Section 1.2.1). On récrit ensuite (101) comme

$$\langle A \rangle = Z_r^{-1} \int_{\mathcal{D}^{N_r}} A(\varphi_r) \exp(-\beta E_{\text{CG}}(\varphi_r)) d\varphi_r, \quad (103)$$

avec $Z_r = \int_{\mathcal{D}^{N_r}} \exp(-\beta E_{\text{CG}}(\varphi_r)) du_r$. Sous l'hypothèse de pouvoir calculer de façon économique l'énergie libre (102), la moyenne canonique (103) est plus facile à calculer que (101), puisque c'est une intégrale en dimension $dN_r \ll dN$.

Pour approcher (102), une approche est de faire un développement de Taylor de E_μ par rapport aux positions φ_c des atomes qu'on élimine. Etant données les positions φ_r , on commence par déterminer des positions de "référence" $\overline{\varphi}_c(\varphi_r)$ pour les atomes éliminés, en s'appuyant sur une interpolation linéaire des positions φ_r . On écrit ensuite

$$\varphi_c = \overline{\varphi}_c(u_r) + \chi_c,$$

où χ_c est supposé petit. L'énergie est alors approchée par

$$E_\mu(\varphi_r, \varphi_c) = E_\mu(\varphi_r, \overline{\varphi}_c(\varphi_r) + \chi_c) \approx \tilde{E}_\mu(\varphi_r, \chi_c)$$

avec

$$\tilde{E}_\mu(\varphi_r, \chi_c) := E_\mu(\varphi_r, \overline{\varphi}_c(\varphi_r)) + \frac{\partial E_\mu}{\partial \varphi_c}(\varphi_r, \overline{\varphi}_c(\varphi_r)) \cdot \chi_c + \frac{1}{2} \chi_c \cdot \frac{\partial^2 E_\mu}{\partial \varphi_c^2} \cdot \chi_c.$$

On approche alors $E_{\text{CG}}(\varphi_r)$ par

$$E_{\text{HA}}(\varphi_r) = -\frac{1}{\beta} \ln \int_{\mathcal{D}^{N_c}} \exp(-\beta \tilde{E}_\mu(\varphi_r, \chi_c)) d\chi_c,$$

qui est analytiquement calculable. Cette méthode repose donc sur deux approximations, l'introduction d'une interpolation linéaire pour définir la position de référence $\overline{\varphi}_c(\varphi_r)$, et le développement limité de l'énergie autour de cette position de référence, dont il n'est pas facile de contrôler la validité quantitative. Notre approche, décrite ci-dessous, permet, dans certains cas, de valider ces approximations, ou d'en montrer les limites. \diamond

Avant de rentrer dans le détail, décrivons brièvement notre approche. Supposons pour simplifier que $N_r = 1$. Pour calculer $\langle A \rangle$ définie par (101), notre idée est de changer de variables, c'est-à-dire d'introduire $y = (y_1, \dots, y_N) = g(\varphi)$, et d'exprimer (101) sous la forme

$$\langle A \rangle = \int A \left(\frac{1}{N} \sum_{i=1}^N y_i \right) \nu(y) dy$$

pour une certaine mesure de probabilité $\nu(y)$ (cf. (108) ci-dessous pour un exemple concret).

On reconnaît ensuite $\langle A \rangle$ comme l'espérance $\mathbb{E} \left[A \left(\frac{1}{N} \sum_{i=1}^N Y_i \right) \right]$, où $Y = (Y_1, \dots, Y_N)$ est un vecteur de variables aléatoires distribuées suivant la mesure ν . La loi des grands nombres donne la limite de $\langle A \rangle$ quand $N \rightarrow +\infty$ (ce qui correspond ici à $N_c \rightarrow +\infty$, puisque $N_r = 1$).

On constate que l'outil théorique utilisé est très classique : c'est la loi forte des grands nombres (on utilisera aussi le théorème de la limite centrale, et des résultats de grandes déviations). L'originalité de ce travail est d'utiliser ces outils classiques pour bâtir une stratégie numérique efficace.

Dans la suite, on détaille le cas simple d'interactions à plus proches voisins dans la Section 2.2.1. On mentionnera ensuite plusieurs extensions de l'approche dans la Section 2.2.2.

2.2.1 Le cas d'interactions à plus proches voisins

On suppose ici que la dimension d'espace est $d = 1$, et que le domaine accessible aux particules est $\mathcal{D} = \mathbb{R}$. Le modèle le plus simple est celui d'une chaîne d'atomes avec interactions à plus proches voisins. Dans ce cas, l'énergie (95) s'écrit

$$E_\mu(\varphi_1, \dots, \varphi_N) = \sum_{i=1}^N V \left(\frac{\varphi_i - \varphi_{i-1}}{1/N} \right), \quad (104)$$

où $V : \mathbb{R} \rightarrow \mathbb{R}$ est le potentiel interatomique. Pour supprimer l'invariance par translation, on impose $\varphi_0 = 0$. Sans perte de généralité, on peut supposer que l'observable A ne dépend que des atomes 0 et N (cf. la Figure 10).



FIG. 10 – Cas d'une observable ne dépendant que des atomes d'indice 0 et N . Tous les atomes intermédiaires sont éliminés.

Dans ce cas, la moyenne (101) (qui dépend du nombre N d'atomes dans le système) s'écrit

$$\langle A \rangle_N = Z^{-1} \int_{\mathbb{R}^N} A(\varphi_N) \exp \left(-\beta \sum_{i=1}^N V \left(\frac{\varphi_i - \varphi_{i-1}}{1/N} \right) \right) d\varphi_1 \dots d\varphi_N. \quad (105)$$

Nous travaillons dans un régime de limite thermodynamique, c'est-à-dire ici dans la limite $N \rightarrow \infty$. Notre objectif est de calculer $\lim_{N \rightarrow \infty} \langle A \rangle_N$.

Remarque 32 Dans (105), les positions des atomes φ_i varient dans \mathbb{R} . On ne contraint pas les atomes à vérifier $\varphi_{i-1} \leq \varphi_i$, ce qui serait une contrainte très naturelle, exprimant le fait que les plus proches voisins demeurent plus proches voisins, quelle que soit la configuration du système considérée. Il se trouve en fait qu'on peut prendre en compte cette contrainte dans ce qui suit sans changer notre argument, en remplaçant V par

$$V_c(y) = V(y) \text{ si } y \geq 0, \quad V_c(y) = +\infty \text{ sinon.} \quad \diamond$$

Pour calculer la limite de $\langle A \rangle_N$, on commence par changer de variables dans (105), en introduisant

$$y_i := \frac{\varphi_i - \varphi_{i-1}}{1/N}, \quad i = 1, \dots, N, \quad (106)$$

ce qui implique en particulier que

$$\varphi_N = \frac{1}{N} \sum_{i=1}^N y_i. \quad (107)$$

Par conséquent, la moyenne (105) se réécrit

$$\langle A \rangle_N = Z^{-1} \int_{\mathbb{R}^N} A \left(\frac{1}{N} \sum_{i=1}^N y_i \right) \exp \left(-\beta \sum_{i=1}^N V(y_i) \right) dy_1 \dots dy_N, \quad (108)$$

avec maintenant $Z = \int_{\mathbb{R}^N} \exp \left(-\beta \sum_{i=1}^N V(y_i) \right) dy_1 \dots dy_N$. On peut maintenant réinterpréter (108) comme

$$\langle A \rangle_N = \mathbb{E} \left[A \left(\frac{1}{N} \sum_{i=1}^N Y_i \right) \right], \quad (109)$$

où les Y_i sont des variables aléatoires indépendantes et identiquement distribuées suivant la loi $z^{-1} \exp(-\beta V(y)) dy$, avec $z = \int_{\mathbb{R}} \exp(-\beta V(y)) dy$. En utilisant la loi forte des grands nombres, on déduit donc que

$$\lim_{N \rightarrow +\infty} \langle A \rangle_N = A(y^*), \quad (110)$$

avec

$$y^* = z^{-1} \int_{\mathbb{R}} y \exp(-\beta V(y)) dy. \quad (111)$$

Si A est assez régulière, on peut développer A au voisinage de y^* dans (109), ce qui donne le développement de $\langle A \rangle_N$ comme une série en puissance de $1/N$.

On cherche maintenant un résultat plus précis. Introduisons l'énergie libre E_{CG} associée au choix de la coordonnée de réaction $\xi(\varphi_1, \dots, \varphi_N) = \varphi_N$ (cf. la Section 1.2.1 pour quelques rappels sur l'énergie libre). Par définition,

$$E_{CG}(\varphi_N) = -\frac{1}{\beta} \ln \left[\int_{\mathbb{R}^{N-1}} \exp(-\beta E_{\mu}(\varphi_1, \dots, \varphi_N)) d\varphi_1 \dots d\varphi_{N-1} \right]. \quad (112)$$

Autrement dit, pour toute observable A ne dépendant que de φ_N , on a

$$\frac{\int_{\mathbb{R}} A(\varphi_N) \exp(-\beta E_{CG}(\varphi_N)) d\varphi_N}{\int_{\mathbb{R}} \exp(-\beta E_{CG}(\varphi_N)) d\varphi_N} = \frac{\int_{\mathbb{R}^N} A(\varphi_N) \exp(-\beta E_{\mu}(\varphi_1, \dots, \varphi_N)) d\varphi}{\int_{\mathbb{R}^N} \exp(-\beta E_{\mu}(\varphi_1, \dots, \varphi_N)) d\varphi}.$$

On montre facilement que E_{CG} croît proportionnellement à N (ce qui peut se vérifier explicitement dans le cas où V est un potentiel harmonique). Il est donc naturel d'introduire l'énergie libre par particule

$$F_N(x) := \frac{1}{N} E_{CG}(x). \quad (113)$$

Par définition, lorsque $\varphi \in \mathbb{R}^N$ est distribuée suivant $Z^{-1} \exp(-\beta E_{\mu}(\varphi)) d\varphi$, la variable aléatoire $\varphi_N \in \mathbb{R}$ est distribuée suivant $\tilde{z}^{-1} \exp(-\beta E_{CG}(\varphi_N)) d\varphi_N$ (où \tilde{z} est une constante de normalisation), autrement dit $\tilde{z}^{-1} \exp(-\beta N F_N(\varphi_N)) d\varphi_N$. En utilisant le changement de variable (106) et la relation (107), on voit aussi que $\tilde{z}^{-1} \exp(-\beta N F_N(x)) dx$ est la distribution de la variable $\frac{1}{N} \sum_{i=1}^N y_i$, lorsque les y_i sont des variables aléatoires indépendantes et identiquement distribuées suivant la loi $z^{-1} \exp(-\beta V(y)) dy$.

En utilisant la loi des grands nombres, on voit que $\frac{1}{N} \sum_{i=1}^N y_i$ est une variable aléatoire dont la distribution, quand $N \rightarrow \infty$, converge vers une masse de Dirac en y^* . Etudier le comportement de $F_N(x)$ quand $N \rightarrow \infty$ revient donc à comprendre à quelle vitesse cette convergence a lieu. Ceci revient aussi à comprendre à quelle vitesse la probabilité $\mathbb{P} \left(\frac{1}{N} \sum_{i=1}^N y_i \geq y^* + \varepsilon \right)$ tend vers 0 quand $N \rightarrow \infty$, pour tout $\varepsilon > 0$ fixé.

Dans [FL20, Théorème 2 et Corollaire 1], nous avons démontré le résultat de grandes déviations suivant (on renvoie à [63, 64, 65, 145, 51, 89] pour une présentation plus complète de la théorie des grandes déviations) :

Théorème 33 *On suppose que V est tel que*

$$\forall \xi \in \mathbb{R}, \quad \int_{\mathbb{R}} \exp(\xi y - \beta V(y)) dy < +\infty,$$

et $\exp(-\beta V) \in H^1(\mathbb{R} \setminus \{0\})$. Alors la limite de F_N est donnée par la transformée de Legendre

$$\lim_{N \rightarrow +\infty} \left(F_N(x) + \frac{1}{\beta} \ln \frac{z}{N} \right) = F_{\infty}(x), \quad (114)$$

avec

$$F_\infty(x) := \frac{1}{\beta} \sup_{\xi \in \mathbb{R}} \left(\xi x - \ln \left[z^{-1} \int_{\mathbb{R}} \exp(\xi y - \beta V(y)) dy \right] \right) \quad (115)$$

et $z = \int_{\mathbb{R}} \exp(-\beta V(y)) dy$. De plus, la convergence (114) a lieu aussi dans $L_{\text{loc}}^p(\mathbb{R})$ pour tout $p \in [1, +\infty[$. Ainsi, F'_N converge vers F'_∞ dans $W_{\text{loc}}^{-1,p}(\mathbb{R})$.

On remarque que l'hypothèse $\exp(-\beta V) \in H^1(\mathbb{R} \setminus \{0\})$ permet de prendre en compte le cas où V est continue par morceaux, avec une singularité en 0, ce qui est le cas du potentiel mentionné à la Remarque 32.

Remarque 34 Les résultats standards de grandes déviations ne sont pas formulés en terme de convergence (presque partout ou dans L_{loc}^p) de F_N , mais portent plutôt sur la limite quand $N \rightarrow \infty$ de $\frac{1}{N} \ln \mathbb{P} \left(\frac{1}{N} \sum_{i=1}^N y_i \in \mathcal{D} \right)$, où \mathcal{D} est (pour simplifier) un intervalle de \mathbb{R} , et les y_i sont des variables aléatoires i.i.d. (cf. par exemple [51, Section 2.2]). On rappelle que, par construction,

$$\mathbb{P} \left(\frac{1}{N} \sum_{i=1}^N y_i \in \mathcal{D} \right) = \tilde{z}^{-1} \int_{\mathcal{D}} \exp(-\beta N F_N(x)) dx,$$

où \tilde{z} est une constante de normalisation. Nous avons prouvé dans [FL20] un résultat sur F_N , qui est donc une sorte de version locale du théorème classique. La preuve repose en particulier sur l'utilisation d'une version élaborée du théorème de la limite centrale, donnée dans [111]. \diamond

L'observation suivante souligne l'intérêt de comprendre le comportement de F_N . On a considéré ci-dessus un matériau modélisé par l'énergie (104), et on a cherché à calculer la limite de la moyenne $\langle A \rangle_N$ d'une observable A ne dépendant que de $\varphi_0 = 0$ et φ_N . Dans le cas particulier où $A(\varphi_N) = \varphi_N$, la quantité $\langle A \rangle_N$ est exactement la longueur moyenne du matériau, lorsqu'il n'est soumis à aucune force de volume (cf. l'expression de l'énergie (104)). On s'intéresse maintenant à la situation inverse : plutôt que de prescrire une *force* (éventuellement nulle) et calculer une *longueur* moyenne, prescrivons une *longueur* et calculons une *force*. La force entre les deux derniers atomes est $V' \left(\frac{\varphi_N - \varphi_{N-1}}{1/N} \right)$. On en calcule la moyenne sur toutes les configurations, distribuées suivant la mesure canonique, sous la contrainte que la position du dernier atome est fixée : $\varphi_N = x$. La force moyenne est alors

$$\mathcal{T}_N(x) = \frac{\int_{\mathbb{R}^{N-1}} V' \left(\frac{x - \varphi_{N-1}}{1/N} \right) \exp(-\beta E_\mu(\varphi_1, \dots, \varphi_{N-1}, x)) d\varphi_1 \dots d\varphi_{N-1}}{\int_{\mathbb{R}^{N-1}} \exp(-\beta E_\mu(\varphi_1, \dots, \varphi_{N-1}, x)) d\varphi_1 \dots d\varphi_{N-1}}. \quad (116)$$

On souhaite calculer la limite de $\mathcal{T}_N(x)$ quand $N \rightarrow \infty$. On ne peut pas appliquer un résultat du type (110), car ici l'observable dépend de N . On remarque cependant que

$$\mathcal{T}_N(x) = F'_N(x),$$

où F_N est définie par (112)-(113). Le Théorème 33 rappelé ci-dessus donne la convergence de F'_N (et donc de la force $\mathcal{T}_N(x)$ pour une élongation x fixée) dans $W_{\text{loc}}^{-1,p}$.

Remarque 35 Par définition, lorsque $\varphi \in \mathbb{R}^N$ est distribuée suivant $Z^{-1} \exp(-\beta E_\mu(\varphi)) d\varphi$, la variable aléatoire $\varphi_N \in \mathbb{R}$ est distribuée suivant $\tilde{z}^{-1} \exp(-\beta N F_N(\varphi_N)) d\varphi_N$. On vient d'identifier la limite, en $N \rightarrow \infty$, de $F_N(x)$. Il est alors tentant d'approximer (105) par

$$\langle A \rangle_N = Z^{-1} \int_{\mathbb{R}} A(\varphi_N) \exp(-\beta N F_N(\varphi_N)) d\varphi_N \approx L_N,$$

avec

$$L_N = \frac{\int_{\mathbb{R}} A(\varphi_N) \exp(-\beta N F_\infty(\varphi_N)) d\varphi_N}{\int_{\mathbb{R}} \exp(-\beta N F_\infty(\varphi_N)) d\varphi_N}. \quad (117)$$

On a donc remplacé F_N par F_∞ dans le facteur exponentiel. Il se trouve en fait que cette stratégie n'est pas optimale. Ceci est dû au fait que $N F_\infty$ n'est pas une bonne approximation de $N F_N$, même si F_∞ est une bonne approximation de F_N . Dans [FL20], nous avons montré que

$$L_N = \langle A \rangle_N + \frac{C}{N} + o(N^{-1})$$

avec en général $C \neq 0$. Donc L_N n'est qu'une approximation d'ordre 0 de $\langle A \rangle_N$, au même titre que l'approximation $A(y^*)$ proposée ci-dessus (cf. les équations (110) et (111)). Cependant, il est beaucoup plus difficile de calculer L_N que $A(y^*)$. En effet, évaluer y^* revient à simplement évaluer une intégrale en dimension un, alors que le calcul de F_∞ , nécessaire pour construire L_N , est beaucoup plus coûteux. \diamond

Les résultats numériques obtenus dans [FL20] (cf. aussi [FL34]) montrent l'efficacité de la stratégie, que le potentiel V soit convexe ou pas, et qu'il soit du type de ceux mentionnés à la Remarque 32 ou pas. On notera bien sûr l'importante réduction en terme de coût calcul que la méthode présentée ici apporte. La convergence (110) permet de remplacer le calcul de $\langle A \rangle_N$, une intégrale N -dimensionnelle, par le calcul de y^* , qui est une intégrale en dimension un. De même, $F'_N(x)$ est une intégrale en dimension N , et est ici approchée par $F'_\infty(x)$, qui est issu d'un problème d'optimisation en dimension un.

Remarque 36 Nous avons ici suivi une approche de type limite thermodynamique. On a mentionné à la Remarque 31 qu'une autre approche, par exemple suivie par la Quasi-Continuum Method (QCM), consiste à étudier le régime basse température. Nous avons vérifié que, dans le régime basse température, nos résultats redonnent les résultats de la méthode QCM (cf. [FL20]). \diamond

2.2.2 Extensions

On décrit ici plusieurs extensions de l'approche proposée ci-dessus.

Le cas d'interactions à deux voisins On se place encore dans un cadre mono-dimensionnel, avec une chaîne d'atomes qui comporte maintenant des interactions à deux voisins. L'énergie s'écrit (cf. (104))

$$E_\mu(\varphi_1, \dots, \varphi_N) = \sum_{i=1}^N V_1 \left(\frac{\varphi_i - \varphi_{i-1}}{1/N} \right) + \sum_{i=1}^{N-1} V_2 \left(\frac{\varphi_{i+1} - \varphi_{i-1}}{1/N} \right).$$

Comme on va le voir, ce cas est mathématiquement beaucoup plus complexe que le précédent. Les outils théoriques sont les mêmes (loi forte des grands nombres, principe de grandes déviations), mais on les appliquera cette fois-ci à des chaînes de Markov, plutôt

qu'à des variables *indépendantes et identiquement distribuées*, comme c'était le cas dans la Section 2.2.1.

Comme ci-dessus, on s'intéresse pour commencer au calcul de moyennes d'observables. A nouveau, on supprime l'invariance par translation en posant $\varphi_0 = 0$, et on fait le changement de variables (106). En utilisant (107), on voit que la moyenne canonique d'une observable ne dépendant que de φ_N s'écrit

$$\begin{aligned}\langle A \rangle_N &= Z^{-1} \int_{\mathbb{R}^N} A(\varphi_N) \exp(-\beta E_\mu(\varphi_1, \dots, \varphi_N)) d\varphi_1 \dots d\varphi_N \\ &= Z^{-1} \int_{\mathbb{R}^N} A\left(\frac{1}{N} \sum_{i=1}^N y_i\right) e^{-\beta V_1(y_1)} f(y_1, y_2) \dots f(y_{N-1}, y_N) dy_1 \dots dy_N,\end{aligned}\quad (118)$$

avec

$$f(x, y) := \exp(-\beta V_2(x + y)) \exp(-\beta V_1(y)).$$

L'idée est alors d'écrire, comme dans (109), que

$$\langle A \rangle_N = \mathbb{E} \left[A \left(\frac{1}{N} \sum_{i=1}^N Y_i \right) \right],$$

où les Y_i sont une réalisation de la chaîne de Markov de noyau f . On rappelle que, dans le cas d'interactions à plus proches voisins, on a l'expression (109) où les Y_i sont des variables aléatoires *indépendantes et identiquement distribuées*. Ainsi, comme annoncé ci-dessus, passer du cas à plus proches voisins au cas à deux voisins revient à passer de variables i.i.d. à une chaîne de Markov. On espère alors obtenir un résultat de type (110) (respectivement (114)) par un argument de type loi forte des grands nombres (respectivement principe de grandes déviations).

Dans l'argument ci-dessus, nous avons implicitement supposé qu'il est possible de considérer f comme le noyau d'une chaîne de Markov. Or, en général,

$$\int_{\mathbb{R}} f(y_1, y_2) dy_2 = \int_{\mathbb{R}} \exp(-\beta V_2(y_1 + y_2)) \exp(-\beta V_1(y_2)) dy_2 \neq 1.$$

On ne peut donc pas considérer f comme un tel noyau. L'argument suivant permet cependant de revenir à un cas standard. On définit la fonction symétrique

$$\bar{f}(x, y) := \exp \left[-\beta V_2(x + y) - \frac{\beta}{2} V_1(x) - \frac{\beta}{2} V_1(y) \right],$$

et on considère le problème variationnel

$$\max_{\psi \in L^2(\mathbb{R})} \left\{ \int_{\mathbb{R}^2} \psi(y) \psi(z) \bar{f}(y, z) dy dz; \int_{\mathbb{R}} \psi^2(y) dy = 1 \right\}. \quad (119)$$

On peut montrer (cf. [59]) que ce problème admet un maximiseur (noté ici ψ_1), unique au signe près, et qu'on peut le choisir tel que $\psi_1 > 0$. L'équation d'Euler-Lagrange de (119) s'écrit

$$\lambda \psi_1(y) = \int_{\mathbb{R}} \bar{f}(y, z) \psi_1(z) dz$$

pour un certain λ . On définit alors

$$g(y, z) := \frac{\psi_1(z)}{\lambda \psi_1(y)} \bar{f}(y, z),$$

qui est telle que

$$\begin{cases} \int_{\mathbb{R}} g(y, z) dz = 1, \\ \int_{\mathbb{R}} \psi_1^2(y) g(y, z) dy = \psi_1^2(z). \end{cases} \quad (120)$$

Par ailleurs, la moyenne (118) s'écrit

$$\begin{aligned} \langle A \rangle_N &= Z_g^{-1} \int_{\mathbb{R}^N} A \left(\frac{1}{N} \sum_{i=1}^N y_i \right) \psi_1(y_1) e^{-\frac{\beta}{2} V_1(y_1)} \\ &\quad \times g(y_1, y_2) \dots g(y_{N-1}, y_N) \frac{e^{-\frac{\beta}{2} V_1(y_N)}}{\psi_1(y_N)} dy_1 \dots dy_N \end{aligned}$$

pour une certaine constante de normalisation Z_g . On voit donc que

$$\langle A \rangle_N = \frac{\mathbb{E} \left[A \left(\frac{1}{N} \sum_{i=1}^N Y_i \right) r(Y_N) \right]}{\mathbb{E} [r(Y_N)]}, \quad (121)$$

où (Y_1, \dots, Y_N) est une réalisation de la chaîne de Markov de noyau g , où Y_1 est distribuée suivant $\psi_1(y_1) \exp(-\beta V_1(y_1)/2)$, et où $r(y_N) = \frac{e^{-\frac{\beta}{2} V_1(y_N)}}{\psi_1(y_N)}$. La première ligne de (120) montre que g est maintenant correctement normalisé, tandis que la seconde ligne montre que ψ_1^2 est une mesure invariante de la chaîne de Markov.

Dans [FL20], en utilisant la loi forte des grands nombres pour les chaînes de Markov, nous avons identifié la limite de (121) quand $N \rightarrow \infty$, obtenant ainsi l'analogue (dans ce cas à deux voisins) du résultat (110). De la même façon, on peut aussi utiliser un principe de grandes déviations pour obtenir l'analogue du Théorème 33 (cf. [FL34]).

Remarque 37 *Le cas où l'observable A dépend de plus de deux atomes se traite de la même manière. Par ailleurs, le cas d'un modèle avec interactions à trois voisins se traite aussi d'une manière similaire : on introduit à nouveau une chaîne de Markov, cette fois-ci en dimension plus grande. \diamond*

Loi de comportement à température finie Dans la Section 2.2.1, nous avons considéré un modèle à plus proches voisins avec l'énergie (104), qui ne comprend aucun terme de force extérieure. En l'absence de force extérieure, la longueur moyenne de la chaîne comportant N atomes est exactement (105), avec $A(\varphi_N) = \varphi_N$ (on rappelle que $\varphi_0 = 0$). La relation (110) montre donc que, dans la limite thermodynamique, la longueur moyenne de la chaîne est y^* .

De manière similaire, on peut considérer une chaîne d'atomes dont l'extrémité droite est soumise à une force extérieure f . L'énergie, qui était donnée par (104), s'écrit maintenant

$$E_\mu^f(\varphi_1, \dots, \varphi_N) = \sum_{i=1}^N V \left(\frac{\varphi_i - \varphi_{i-1}}{1/N} \right) - N f \varphi_N.$$

Dans la limite thermodynamique, la longueur moyenne de la chaîne $y^*(f)$ est à nouveau la limite, quand $N \rightarrow \infty$, de

$$\langle A \rangle_N^f = Z^{-1} \int A(\varphi_N) \exp \left(-\beta E_\mu^f(\varphi_1, \dots, \varphi_N) \right) d\varphi_1 \dots d\varphi_N,$$

avec $A(\varphi_N) = \varphi_N$. En utilisant un argument similaire à celui de la Section 2.2.1, nous avons montré dans [FL34] que

$$y^*(f) = z_f^{-1} \int_{\mathbb{R}} y \exp(-\beta [V(y) - fy]) dy$$

avec $z_f = \int_{\mathbb{R}} \exp(-\beta [V(y) - fy]) dy$.

Ainsi, lorsqu'on impose une force extérieure f à l'extrémité de la chaîne, la longueur macroscopique est $y^*(f)$. Par ailleurs, lorsqu'on prescrit la longueur de la chaîne à la valeur x , la force moyenne dans la chaîne, à la limite thermodynamique, est $F'_\infty(x)$ (cf. le Théorème 33). Nous avons démontré dans [FL34] que ces deux relations étaient les réciproques l'une de l'autre : $y^* \circ F'_\infty = F'_\infty \circ y^* = \text{Id}$. Nous avons aussi généralisé ce résultat au cas d'interaction à deux voisins.

La relation $x \mapsto F'_\infty(x)$ peut ainsi être considérée comme la relation constitutive du matériau, à l'échelle macroscopique, à température finie. La force dans le système est $F'_\infty(x)$ lorsque son élongation est x .

Un cas bi-dimensionnel Dans le travail [FL27], nous cherchons à étendre la stratégie décrite dans la Section 2.2.1 à un cas bi-dimensionnel. Ce cas étant reconnu comme bien plus difficile que le cas mono-dimensionnel, nous nous sommes concentrés sur le modèle le plus simple possible. Les degrés de liberté sont les positions $\varphi_{i,j}$ d'un réseau d'atomes en deux dimensions, avec $1 \leq i, j \leq N$, et ces degrés de liberté sont *scalaires*, $\varphi_{i,j} \in \mathbb{R}$: chaque atome ne peut se déplacer que dans la direction verticale. Ce modèle représente d'une certaine façon une membrane (cf. la Figure 11). L'énergie a une expression très simple :

$$E_\mu(\{\varphi_{i,j}\}) = \frac{1}{2} \sum_{1 \leq i, j \leq N} \sum_{\substack{k, l \text{ tels que} \\ |i - k| + |j - l| = 1}} V\left(\frac{\varphi_{i,j} - \varphi_{k,l}}{1/N}\right). \quad (122)$$

Les interactions se font donc sur un réseau de maille carrée.

Pour supprimer l'invariance par translation, on fixe $\varphi_{0,0} = 0$, et on s'intéresse au calcul d'une moyenne d'observable ne dépendant que de $\varphi_{N,N}$, soit

$$\langle A \rangle_N = Z^{-1} \int A(\varphi_{N,N}) \exp(-\beta E_\mu(\{\varphi_{i,j}\})) \prod_{1 \leq i, j \leq N} d\varphi_{i,j}.$$

L'objectif du travail [FL27] est d'adapter la *méthodologie* élaborée dans le cadre mono-dimensionnel pour le calcul des moyennes (105) et (118). Le traitement rigoureux de cette question demeure cependant une question ouverte.

Comme on l'a vu dans la Section 2.2.1, une autre question intéressante est le calcul d'énergies libres du type (112) (à renormalisation près, cf. (113), le Théorème 33 donne la limite de cette énergie libre dans la limite thermodynamique). Dans le cadre du modèle (122) en dimension deux, une manière de formuler le problème est d'introduire

$$E_{\text{CG}}^{\Lambda_N}(u) = -\frac{1}{\beta} \ln \left[\int \exp(-\beta E_\mu(\{\varphi_{i,j}\})) \prod_{(i,j) \in \Lambda_N} d\varphi_{i,j} \prod_{(i,j) \in \partial \Lambda_N} \delta_{\varphi_{i,j} - u_{(i,j)}} \right], \quad (123)$$

où $\Lambda_N = \{(i, j) \in \mathbb{Z}^2; -N \leq i, j \leq N\}$, $\partial \Lambda_N = \{(i, j) \in \mathbb{Z}^2; |i| = N \text{ ou } |j| = N\}$ et $u \in \mathbb{R}^2$. Ainsi, on considère un système atomique dont les indices d'atomes sont dans $\Lambda_N \subset \mathbb{Z}^2$

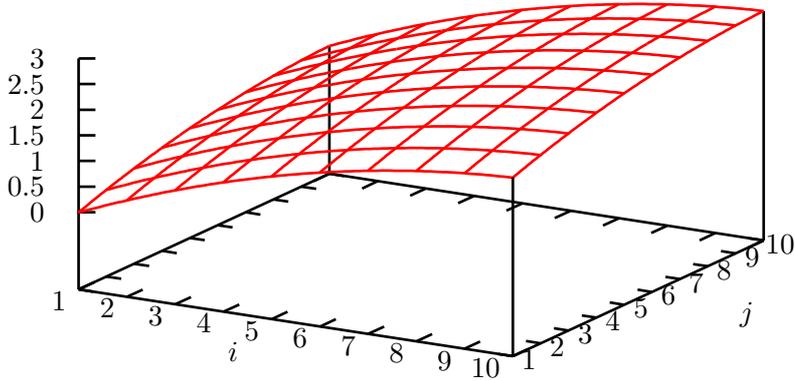


FIG. 11 – Réseau d’atomes en dimension deux; chaque atome ne peut se déplacer que dans la direction verticale.

(la limite thermodynamique correspondant à $N \rightarrow \infty$). La quantité (123) est l’énergie libre du système lorsqu’on prescrit la position des atomes au bord du domaine (ceux tels que $(i, j) \in \partial\Lambda_N$) à la valeur $u \cdot (i, j) = iu_1 + ju_2$, où $u = (u_1, u_2) \in \mathbb{R}^2$ (c’est le sens de la masse de Dirac dans (123)), et qu’on intègre sur les atomes intérieurs. Dans le langage de la Section 1.2.1, (123) est l’énergie libre du système associée à la coordonnée de réaction $\xi \left(\{\varphi_{i,j}\}_{(i,j) \in \Lambda_N} \right) = \{\varphi_{i,j}\}_{(i,j) \in \partial\Lambda_N}$. Remarquons enfin que la quantité (123) est l’analogue en dimension deux de la quantité (112), où $\Lambda_N = \mathbb{Z} \cap [0, N]$, et où la coordonnée de réaction est $\xi(\varphi_0, \dots, \varphi_N) = (\varphi_0, \varphi_N)$.

Remarque 38 Ici, le fait de travailler en dimension deux n’a pas d’importance particulière. Les mêmes résultats sont vrais en dimension $d \geq 2$. Il est par contre important que les variables soient scalaires : $\varphi_p \in \mathbb{R}$ pour tout $p \in \mathbb{Z}^d$. \diamond

Suivant (113), on introduit l’énergie libre par particule,

$$F^{\Lambda_N}(u) = \frac{1}{|\Lambda_N|} E_{\text{CG}}^{\Lambda_N}(u).$$

Le Théorème 33 donne l’expression de la limite de $F^{\Lambda_N}(u)$ quand $N \rightarrow \infty$, dans un cadre mono-dimensionnel ($\Lambda_N \subset \mathbb{Z}$). L’étude du cas de la dimension $d \geq 2$ a été fait dans [73], où les auteurs montrent que $F^{\Lambda_N}(u)$ converge vers un certain $F^\infty(u)$ quand $N \rightarrow \infty$, et démontrent que F^∞ est convexe et vérifie certaines bornes. Par contre, aucune formule permettant le *calcul en pratique* de F^∞ ne semble connue, par opposition au cas de la dimension un, où on dispose de (114)-(115). On renvoie à [52, 74, 118] pour des résultats reliés (cf. aussi [72] pour une revue sur le sujet), et, de manière plus générale, à la littérature sur les *gradient random fields*.

3 Homogénéisation

Cette dernière partie présente des problèmes en homogénéisation stochastique. L'objectif est de développer des méthodes numériques plus efficaces que celles actuellement disponibles, en étudiant éventuellement des classes de problèmes plus restreintes.

La motivation de ce travail vient de l'observation que, même dans les cas les plus simples sur le plan théorique, les formules dont on dispose pour évaluer en pratique les coefficients homogénéisés conduisent à des calculs très lourds. Nous nous sommes donc limités à des problèmes simples (équation aux dérivées partielles linéaire, sous forme divergence, avec une inconnue scalaire), pour lesquels il existe des résultats satisfaisants sur le plan théorique, afin de nous concentrer sur des *problèmes pratiques*. L'extension de certaines des méthodes décrites ci-dessous à des équations plus compliquées fait partie de nos objectifs, mais n'a pas à ce jour été étudiée.

On commence par présenter dans la Section 3.1 les principaux résultats théoriques dont nous aurons besoin dans la suite. Nos contributions sont décrites dans les Sections 3.2 et 3.3, et rassemblées dans l'article de revue [FL35].

3.1 Quelques résultats de base en homogénéisation

Avant d'aborder le cas stochastique, on rappelle quelques résultats classiques dans le cas périodique, qui permettront d'illustrer les difficultés de l'homogénéisation stochastique. On renvoie à la littérature (par exemple [17, 43, 96]) pour une présentation plus exhaustive.

3.1.1 Le cas périodique

On se donne pour toute la suite un ouvert borné régulier $\mathcal{D} \subset \mathbb{R}^d$ et une fonction $f \in L^2(\mathcal{D})$. On considère ici le problème

$$\begin{cases} -\operatorname{div} \left[A_{per} \left(\frac{x}{\varepsilon} \right) \nabla u^\varepsilon \right] = f & \text{dans } \mathcal{D}, \\ u^\varepsilon = 0 & \text{sur } \partial\mathcal{D}, \end{cases} \quad (124)$$

où la matrice $A_{per} \in (L^\infty(\mathbb{R}^d))^{d \times d}$ est symétrique, \mathbb{Z}^d -périodique, et telle qu'il existe $a_- > 0$ tel que

$$\forall \xi \in \mathbb{R}^d, \quad a_- \|\xi\|^2 \leq \xi^T A_{per}(x) \xi \quad \text{presque partout en } x \in \mathbb{R}^d.$$

On se concentre dans la suite sur le cas de matrices *symétriques*, mais les arguments peuvent être adaptés au cas général.

La fonction $u^\varepsilon \in H_0^1(\mathcal{D})$ est bien définie, et il existe une constante C , indépendante de ε , pour laquelle

$$\forall \varepsilon, \quad \|u^\varepsilon\|_{H^1(\mathcal{D})} \leq C.$$

En utilisant une méthode de discrétisation (par exemple la méthode des éléments finis, avec un pas de maillage h), on peut calculer une approximation numérique u_h^ε de u^ε , et c'est un résultat classique que, sous certaines hypothèses, u_h^ε converge vers u^ε quand $h \rightarrow 0$. Cependant, une analyse précise montre qu'il n'est pas possible de choisir h indépendant de ε tout en gardant une précision donnée. A précision donnée, il faut que h tende vers 0 quand ε tend vers 0, ce qui rend le calcul très cher lorsque ε est petit.

L'idée de l'homogénéisation est d'identifier la limite u^* de u^ε , et de calculer cette dernière, qui ne dépend pas de ε et ne nécessite donc pas un maillage dont la taille dépende de ε . Mieux, on s'attend à ce que cette méthode soit d'autant plus efficace que ε est petit,

car son coût est indépendant de ε (à l'opposé d'une méthode classique, dont le coût augmente quand ε devient petit), et sa précision est meilleure, la différence entre u^* et u^ε étant plus petite.

La théorie de l'homogénéisation fournit le résultat suivant. Lorsque ε tend vers 0, la solution u^ε de (124) converge (fortement dans $L^2(\mathcal{D})$ et faiblement dans $H_0^1(\mathcal{D})$) vers u^* solution de

$$\begin{cases} -\operatorname{div} [A^* \nabla u^*] = f & \text{dans } \mathcal{D}, \\ u^* = 0 & \text{sur } \partial \mathcal{D}. \end{cases} \quad (125)$$

Remarque 39 Une particularité du cadre dans lequel nous travaillons est que l'équation vérifiée par u^* est du même type que celle vérifiée par u^ε . Cette situation n'est pas générique en homogénéisation. Un contre-exemple bien connu est l'homogénéisation de l'équation

$$\partial_t u^\varepsilon + a \left(\frac{x}{\varepsilon} \right) u^\varepsilon = 0 \quad \text{pour } t > 0, \quad u^\varepsilon(t=0, x) = v(x), \quad x \in \mathbb{R},$$

où a est une fonction 1-périodique. Quand ε tend vers 0, u^ε converge vers u^* solution d'un problème en temps avec mémoire, du type

$$\partial_t u^* + a^* u^* = \int_0^t K(t-s) u^*(s, x) ds$$

pour un certain $a^* \in \mathbb{R}$ et une certaine fonction K . \diamond

La matrice homogénéisée A^* dans (125) est donnée par

$$[A^*]_{ij} = \int_Q e_i^T A_{per}(y) (e_j + \nabla w_{e_j}(y)) dy \quad (126)$$

où Q est le cube unité,

$$Q = [0, 1]^d,$$

et w_{e_i} est le correcteur associé au vecteur $p = e_i \in \mathbb{R}^d$, où $\{e_i\}_{1 \leq i \leq d}$ est la base canonique de \mathbb{R}^d . Pour un vecteur $p \in \mathbb{R}^d$ quelconque, le correcteur w_p est la solution (définie à une constante additive près) du problème du correcteur, qui s'écrit

$$\begin{cases} -\operatorname{div} (A_{per}(y) (p + \nabla w_p)) = 0 & \text{dans } \mathbb{R}^d, \\ w_p \text{ est } Q\text{-périodique.} \end{cases} \quad (127)$$

Comme mentionné ci-dessus, la convergence de u^ε vers u^* est forte dans $L^2(\mathcal{D})$, mais seulement faible dans $H_0^1(\mathcal{D})$. Les correcteurs peuvent aussi être utilisés pour construire une approximation de u^ε dans $H^1(\mathcal{D})$. On a en effet

$$u^\varepsilon - u^* - \varepsilon \sum_{i=1}^d w_{e_i} \left(\frac{\cdot}{\varepsilon} \right) \frac{\partial u^*}{\partial x_i} \xrightarrow{\varepsilon \rightarrow 0} 0 \quad \text{dans } H^1(\mathcal{D}) \text{ fort.} \quad (128)$$

Pour calculer une approximation de u^ε , on procède donc ainsi. Pour chaque $1 \leq i \leq d$, on calcule $w_{e_i} \in H^1(Q)$, fonction Q -périodique, en considérant la formulation variationnelle de (127), soit

$$\forall \varphi \in H^1(Q) \text{ avec } \varphi \text{ } Q\text{-périodique,} \quad \int_Q (\nabla \varphi)^T A_{per} (e_i + \nabla w_{e_i}) = 0. \quad (129)$$

En pratique, on peut utiliser une approximation de Galerkin de (129) via des éléments finis sur Q . Il est important de noter que le problème (129) est posé sur un domaine *borné* (ce qui ne sera *pas* le cas dans le cadre stochastique), et peut donc se résoudre facilement. Une fois les w_{e_i} déterminés, on calcule A^* par (126), et on résout le problème homogénéisé (125), où le petit paramètre ε n'apparaît pas. Le coût calcul est donc *indépendant* de ε .

La preuve du fait que u^ε converge vers u^* peut se faire suivant la *méthode des fonctions test oscillantes*, due à F. Murat et L. Tartar [117, 140]. Une autre approche consiste à utiliser la notion de *convergence à deux échelles*, introduite par G. Nguetseng et développée par G. Allaire [2, 121]. On renvoie à la littérature. On utilise de façon cruciale le résultat suivant :

Lemme 40 *Soit $b \in L^\infty(\mathbb{R}^d)$, où b est une fonction Q -périodique. Alors, pour toute fonction $\varphi \in L^1(\mathbb{R}^d)$, on a*

$$\lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^d} \varphi(x) b\left(\frac{x}{\varepsilon}\right) dx = \langle b \rangle \int_{\mathbb{R}^d} \varphi(x) dx \quad \text{avec} \quad \langle b \rangle = \int_Q b(x) dx.$$

Sur un plan plus formel, le résultat peut être obtenu en utilisant la technique du *développement à deux échelles*, qui consiste à *postuler* que la solution u^ε de (124) s'écrit sous la forme

$$u^\varepsilon(x) = u_0\left(x, \frac{x}{\varepsilon}\right) + \varepsilon u_1\left(x, \frac{x}{\varepsilon}\right) + \varepsilon^2 u_2\left(x, \frac{x}{\varepsilon}\right) + \dots, \quad (130)$$

où, pour chaque k , la fonction u_k est supposée dépendre de deux arguments, une variable macroscopique x et une variable microscopique x/ε . On suppose de plus que

$$\forall k \geq 0, \quad y \mapsto u_k(x, y) \text{ est } Q\text{-périodique.}$$

En insérant l'Ansatz (130) dans (124) et en identifiant les puissances de ε , on obtient une hiérarchie d'équations, qu'on ferme en utilisant l'hypothèse de périodicité de u_k par rapport à la deuxième variable, x/ε . On obtient ainsi les résultats (125), (127) et (126).

Remarque 41 *Dans ce cadre périodique, le fait d'obtenir un problème de cellule (127) posé sur un domaine borné est très lié au fait que l'équation (124) est l'équation d'Euler-Lagrange d'un problème d'optimisation convexe, qui s'écrit*

$$\inf \left\{ \int_{\mathcal{D}} W\left(\frac{x}{\varepsilon}, \nabla u\right) dx - \int_{\mathcal{D}} f(x)u(x) dx, \quad u \in H_0^1(\mathcal{D}) \right\}$$

avec

$$W(x, F) = \frac{1}{2} F^T A_{\text{per}}(x) F.$$

Un cas plus compliqué est le problème

$$\inf \left\{ \int_{\mathcal{D}} W\left(\frac{x}{\varepsilon}, \nabla u\right) dx - \int_{\mathcal{D}} f(x)u(x) dx, \quad u \in (H_0^1(\mathcal{D}))^3 \right\}$$

où $\mathcal{D} \subset \mathbb{R}^3$, u est à valeur vectorielle, W est Q -périodique de la première variable, et quasi-convexe en la seconde variable. On suppose aussi qu'il existe c et C tels que, pour tout $x \in \mathcal{D}$ et toute matrice $F \in \mathbb{R}^{3 \times 3}$,

$$c|F|^2 \leq W(x, F) \leq C(1 + |F|^2).$$

Alors la détermination du comportement homogénéisé passe par la résolution, pour tout $N \in \mathbb{N}^*$, de problèmes de cellule posés sur $(0, N)^3$, et le calcul de la limite d'une certaine quantité quand $N \rightarrow \infty$. On renvoie à [35] pour plus de détails, à [76] pour une analyse numérique de ce cas, et à [71] pour une récente généralisation de ce type de résultats.

On constate simplement ici que passer du cas convexe au cas quasi-convexe, même dans la situation périodique, accroît considérablement la difficulté du problème. \diamond

Remarque 42 *Le fait d'obtenir des formules explicites pour A^* , dans ce cadre linéaire, est relié au fait qu'on est dans un cadre périodique. Il existe aussi des résultats d'homogénéisation dans le cas où $u^\varepsilon \in H_0^1(\mathcal{D})$ est solution de*

$$-\operatorname{div}(A^\varepsilon \nabla u^\varepsilon) = f \quad \text{dans } \mathcal{D},$$

où A^ε une suite de matrices inversibles, uniformément bornées et coercives.

A nouveau, u^ε converge (à extraction près) vers $u^* \in H_0^1(\mathcal{D})$ solution de

$$-\operatorname{div}(A^* \nabla u^*) = f \quad \text{dans } \mathcal{D},$$

où la matrice homogénéisée A^* est indépendante de \mathcal{D} et de f . Toutefois, on ne dispose plus, en général, d'une expression explicite pour A^* . \diamond

3.1.2 Homogénéisation stochastique

Maintenant que nous avons rappelé les principaux résultats en homogénéisation périodique, passons au cas stochastique.

Théorème ergodique Comme on l'aura compris, le Lemme 40 est un résultat essentiel pour l'homogénéisation périodique. On présente ici le pendant de ce résultat dans le cadre stochastique, et plus précisément dans le cadre *stationnaire discret*, dont on commence par rappeler le formalisme.

On se donne un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. On note $\mathbb{E}(X) = \int_\Omega X(\omega) d\mathbb{P}(\omega)$ l'espérance d'une variable aléatoire $X \in L^1(\Omega, d\mathbb{P})$. On suppose disposer d'une action de groupe $(\tau_k)_{k \in \mathbb{Z}^d}$ sur Ω , telle que

– cette action préserve la mesure \mathbb{P} , au sens où

$$\forall k \in \mathbb{Z}^d, \quad \forall A \in \mathcal{F}, \quad \mathbb{P}(\tau_k A) = \mathbb{P}(A).$$

– cette action est ergodique, au sens où

$$\forall A \in \mathcal{F}, \quad \left(\forall k \in \mathbb{Z}^d, \quad \tau_k A = A \right) \implies (\mathbb{P}(A) = 0 \text{ ou } 1).$$

On renvoie à [27] pour des exemples explicites de telles actions de groupe.

On dit qu'une fonction $F \in L_{\text{loc}}^1(\mathbb{R}^d, L^1(\Omega))$ est *stationnaire* si

$$\forall k \in \mathbb{Z}^d, \quad F(x+k, \omega) = F(x, \tau_k \omega) \text{ presque partout en } x, \text{ presque sûrement.} \quad (131)$$

On reviendra sur l'interprétation de cette propriété dans la Remarque 45 ci-dessous.

Remarque 43 *On a ici choisi de travailler dans un cadre stationnaire discret. Un autre cadre possible est le cadre stationnaire continu, où on dispose d'une action de groupe $(\tau_y)_{y \in \mathbb{R}^d}$ sur Ω , et où (131) est remplacé par*

$$\forall y \in \mathbb{R}^d, \quad F(x+y, \omega) = F(x, \tau_y \omega) \text{ presque partout en } x, \text{ presque sûrement.} \quad (132)$$

Ce cadre n'est pas abordé dans ce mémoire. \diamond

Dans le cadre décrit ci-dessus, on a le théorème ergodique suivant [100, 135], qui joue le rôle ici du Lemme 40 dans le cadre périodique.

Théorème 44 Soit $F \in L^\infty(\mathbb{R}^d, L^1(\Omega))$ une fonction stationnaire au sens de (131). Pour $k = (k_1, k_2, \dots, k_d) \in \mathbb{Z}^d$, on pose $|k|_\infty = \sup_{1 \leq i \leq d} |k_i|$. Alors

$$\frac{1}{(2N+1)^d} \sum_{|k|_\infty \leq N} F(x, \tau_k \omega) \xrightarrow{N \rightarrow \infty} \mathbb{E}(F(x, \cdot)) \text{ dans } L^\infty(\mathbb{R}^d), \text{ presque sûrement.}$$

Ceci implique que

$$F\left(\frac{x}{\varepsilon}, \omega\right) \xrightarrow{\varepsilon \rightarrow 0} \mathbb{E}\left(\int_Q F(x, \cdot) dx\right) \text{ dans } L^\infty(\mathbb{R}^d), \text{ presque sûrement,}$$

où $Q = [0, 1]^d$ est le cube unité de \mathbb{R}^d .

Equation homogénéisée On revient maintenant au problème hautement oscillant

$$\begin{cases} -\operatorname{div} \left[A\left(\frac{x}{\varepsilon}, \omega\right) \nabla u^\varepsilon \right] = f & \text{dans } \mathcal{D}, \\ u^\varepsilon = 0 & \text{sur } \partial\mathcal{D}, \end{cases} \quad (133)$$

où la matrice A est stationnaire au sens de (131). De plus, comme dans le cas périodique, A est symétrique, bornée et elliptique, au sens où il existe $a_+ \geq a_- > 0$ tel que

$$\forall \xi \in \mathbb{R}^d, \quad a_- \|\xi\|^2 \leq \xi^T A(x, \omega) \xi \text{ presque partout en } x \in \mathbb{R}^d, \text{ presque sûrement,}$$

et $\|A(\cdot, \omega)\|_{L^\infty(\mathbb{R}^d)} \leq a_+$ presque sûrement.

Remarque 45 Comme τ_k préserve la loi de probabilité, la stationnarité de A implique que $\mathbb{E}(A(x, \cdot))$ est une fonction périodique. Ainsi, en espérance, les propriétés du matériau sont Q -périodiques. De plus, pour tout $k \in \mathbb{Z}^d$, et tout $x \in \mathbb{R}^d$, les variables aléatoires $A(x+k, \omega)$ et $A(x, \omega)$ ont la même loi. On dit ainsi que le matériau est statistiquement homogène.

Le fait que τ_k soit ergodique a pour conséquence qu'il est équivalent de considérer l'ensemble du matériau dans \mathbb{R}^d pour une réalisation de l'aléa, ou bien toutes les réalisations possibles du matériau dans la seule cellule Q . \diamond

Remarque 46 Si la matrice A est indépendante de ω , alors l'hypothèse de stationnarité discrète (131) implique que A est périodique. On retrouve ainsi le cas de l'homogénéisation périodique.

Remarquons que le cas périodique est aussi un cas particulier du cas stationnaire continu : il correspond au choix $\Omega = \mathbb{T}^d$, où \mathbb{T} est le tore en dimension un. Avec ce choix, une fonction F stationnaire au sens de (132) vérifie $F(x+k, \omega) = F(x, \omega)$ pour tout $k \in \mathbb{Z}^d$. \diamond

C'est un résultat classique (cf. [17, 96]) que la solution u^ε de (133) converge vers la solution u^* de (125), où la matrice homogénéisée est maintenant définie par

$$[A^*]_{ij} = \mathbb{E} \left(\int_Q e_i^T A(y, \cdot) (e_j + \nabla w_{e_j}(y, \cdot)) dy \right), \quad (134)$$

où, pour tout $p \in \mathbb{R}^d$, w_p est la solution (unique à une constante aléatoire additive près) de

$$\begin{cases} -\operatorname{div} [A(y, \omega) (p + \nabla w_p(y, \omega))] = 0 & \text{p.s. dans } \mathbb{R}^d, \\ \nabla w_p \text{ est stationnaire au sens de (131),} \\ \mathbb{E} \left(\int_Q \nabla w_p(y, \cdot) dy \right) = 0, \end{cases} \quad (135)$$

dans l'espace $\{w \in L^2_{\text{loc}}(\mathbb{R}^d, L^2(\Omega)), \nabla w \in L^2_{\text{unif}}(\mathbb{R}^d, L^2(\Omega))\}$. On rappelle que l'espace L^2_{unif} est l'espace des fonctions pour lesquelles la norme L^2 sur une boule de rayon unit e est born ee ind ependamment du centre de la boule.

A l'image du cas p eriodique, la convergence de u^ε vers u^* est forte dans $L^2(\mathcal{D})$ et faible dans $H^1(\mathcal{D})$, presque s urement.

Une diff erence essentielle entre les cas p eriodique et stochastique est que, dans le premier cas, le probl eme du correcteur (127) est pos e sur le domaine born e Q (ceci est li e au fait qu'on s'int eresse  a un probl eme elliptique lin eaire, cf. la Remarque 41). Son calcul ne pose donc pas de probl eme particulier. Au contraire, dans le cas stochastique, le probl eme du correcteur (135) est pos e dans l'espace \mathbb{R}^d tout entier. On remarquera en effet que la condition $\mathbb{E} \left(\int_Q \nabla w_p(y, \cdot) dy \right) = 0$ dans (135) est une condition globale, puisqu'elle s' ecrit aussi, ∇w_p  etant stationnaire, $\lim_{R \rightarrow +\infty} \frac{1}{|B_R|} \int_{B_R} \nabla w_p(y, \omega) dy = 0$ presque s urement, pour toute suite de boules B_R de rayon R .

Beaucoup de difficult es pratiques en homog enisation stochastique viennent du fait que le probl eme du correcteur est pos e sur l'espace entier. Il est alors courant d'introduire un probl eme du correcteur tronqu e (i.e. pos e sur un domaine grand mais fini, cf. (136) ci-dessous), ce qui permet de calculer une approximation de la matrice homog enis ee, dont il faut ensuite montrer la convergence vers la matrice homog enis ee exacte A^* lorsque le domaine tronqu e tend vers \mathbb{R}^d tout entier. On reviendra sur ces questions dans la Section 3.2.1. On retrouve en fait ici les difficult es mentionn ees dans la Remarque 41.

Remarque 47 *L'existence m eme d'une fonction w_p v erifiant (135) n'est pas imm ediate, ce probl eme  etant pos e dans \mathbb{R}^d tout entier. Une m ethode classique est de r egulariser l' equation, i.e. d'introduire $\eta > 0$, et de chercher $w_{p,\eta}(\cdot, \omega)$ stationnaire solution de*

$$-\text{div} [A(y, \omega) (p + \nabla w_{p,\eta}(y, \omega))] + \eta w_{p,\eta} = 0.$$

En utilisant des bornes a priori sur $\nabla w_{p,\eta}$, on passe ensuite  a la limite $\eta \rightarrow 0$. Notons qu'on n'a pas de borne sur $w_{p,\eta}$ lui-m eme. Ceci explique que,  a la limite $\eta \rightarrow 0$, on r ecup ere la stationnarit e de ∇w_p , mais pas celle de w_p . Le cas mono-dimensionnel montre que ce r esultat est optimal, au sens o u effectivement w_p solution de (135) n'est pas stationnaire.

3.2 Approximation num erique et r eduction de variance

Publications [FL19, FL33, FL24, FL28]

On s'int eresse ici au probl eme pratique de l'approximation num erique de la matrice homog enis ee, donn ee par (134)-(135). La difficult e vient du fait que le probl eme du correcteur est pos e sur \mathbb{R}^d tout entier.

3.2.1 Approche standard

Une m ethode standard pour approcher A^* est de proc eder comme suit. On introduit le probl eme du correcteur tronqu e

$$\begin{cases} -\text{div} [A(y, \omega) (p + \nabla w_p^N(y, \omega))] = 0 & \text{p.s. dans } \mathbb{R}^d, \\ w_p^N(\cdot, \omega) \text{ est } Q_N\text{-p eriodique,} \end{cases} \quad (136)$$

où $Q_N \subset \mathbb{R}^d$ est le cube de taille $(2N + 1)^d$ centré à l'origine. Ce problème est posé sur un domaine grand, mais fini. A partir du correcteur approché w_p^N , on calcule la matrice homogénéisée approchée

$$[A_N^*]_{ij}(\omega) = \frac{1}{|Q_N|} \int_{Q_N} e_i^T A(y, \omega) \left(e_j + \nabla w_{e_j}^N(y, \omega) \right) dy. \quad (137)$$

Il est important de noter que, bien que A^* soit un objet déterministe, son approximation A_N^* est une quantité aléatoire.

Dans [34, Théorème 1], les auteurs montrent que $A_N^*(\omega)$ converge vers A^* presque sûrement, dans la limite $N \rightarrow \infty$. L'obtention d'un *taux* de convergence est une question plus difficile. On renvoie à [34, Théorème 5], qui s'appuie sur les résultats antérieurs de [146]. Notons aussi que cette question, dans le cas de problèmes discrets (i.e. posés sur un *réseau*), a été étudié dans [78, 119].

Dans certains cas simples, comme celui décrit dans la Section 3.2.3 d'un problème faiblement stochastique, on peut montrer un résultat de type limite centrale (et donc convergence en $|Q_N|^{-1/2}$). Ainsi, typiquement, la convergence est très lente, et il est donc nécessaire de choisir un grand domaine Q_N pour espérer obtenir une bonne approximation de A^* , ce qui implique que résoudre (136) a un coût élevé.

Remarque 48 Dans [34], les auteurs étudient aussi deux autres approximations, qui consistent à choisir dans (136) des conditions aux limites de Dirichlet ou de Neumann, plutôt que des conditions aux limites périodiques. Ils montrent la convergence de $A_N^*(\omega)$ vers A^* indépendamment du choix des conditions aux limites. Par ailleurs, une comparaison numérique de ces différentes stratégies est décrite dans [97]. \diamond

On note que $A_N^*(\omega)$ diffère de A^* par deux sources d'erreur :

- une erreur *systématique*, liée au fait que $\mathbb{E}[A_N^*] \neq A^*$. On peut se convaincre de cette inégalité sur la base des expressions (146) et (147) ci-dessous (qui concernent un cas mono-dimensionnel simple) ;
- et une erreur *statistique*, liée au fait que $A_N^*(\omega)$ est un nombre aléatoire.

Dans la suite, on se concentre sur l'erreur statistique, avec l'idée que l'erreur systématique $\mathbb{E}([A_N^*]_{ij}) - [A^*]_{ij}$ (pour $1 \leq i, j \leq d$) est en pratique beaucoup plus petite que l'écart type $\sqrt{\text{Var}[A_N^*]_{ij}}$, qui donne une estimation grossière de l'erreur statistique. Cette idée est en accord avec des résultats numériques typiques (cf. par exemple la Figure 12 ci-dessous). Dans un cas mono-dimensionnel simple, on peut même quantifier ceci. En effet, dans le cas

$$a(x, \omega) = \sum_{k \in \mathbb{Z}} 1_{[k, k+1[}(x) a_k(\omega),$$

où les variables a_k sont i.i.d., on peut montrer que l'erreur systématique $\mathbb{E}(a_N^*) - a^*$ est d'ordre N^{-1} , tandis que l'écart type $\sqrt{\text{Var}(a_N^*)}$ est d'ordre $N^{-1/2}$. L'erreur statistique est donc bien plus grande que l'erreur systématique. On s'attache dans la suite à calculer le plus efficacement possible la matrice $\mathbb{E}[A_N^*]$.

Une méthode standard consiste, pour N fixé, à considérer $2M$ réalisations indépendantes et identiquement distribuées $\{A^m(x, \omega)\}_{1 \leq m \leq 2M}$ de la matrice aléatoire. Pour chacune de ces réalisations, on résout le problème du correcteur tronqué (136), et on calcule la matrice homogénéisée $A_N^{*,m}$ par (137). On définit ensuite, pour chaque coefficient de la

matrice, une moyenne et une variance empiriques par

$$\begin{aligned}\mu_{2M} \left([A_N^*]_{ij} \right) &= \frac{1}{2M} \sum_{m=1}^{2M} [A_N^{*,m}]_{ij}, \\ \sigma_{2M} \left([A_N^*]_{ij} \right) &= \frac{1}{2M-1} \sum_{m=1}^{2M} \left([A_N^{*,m}]_{ij} - \mu_{2M} \left([A_N^*]_{ij} \right) \right)^2.\end{aligned}$$

Comme les matrices $A_N^{*,m}$ sont indépendantes et identiquement distribuées, la loi forte des grands nombres s'applique :

$$\mu_{2M} \left([A_N^*]_{ij} \right) (\omega) \xrightarrow{M \rightarrow +\infty} \mathbb{E} \left([A_N^*]_{ij} \right) \text{ presque sûrement.}$$

De plus, en utilisant le théorème de la limite centrale, on obtient la convergence en loi

$$\sqrt{2M} \left(\mu_{2M} \left([A_N^*]_{ij} \right) - \mathbb{E} \left([A_N^*]_{ij} \right) \right) \xrightarrow{M \rightarrow +\infty} \mathcal{L} \sqrt{\text{Var} \left([A_N^*]_{ij} \right)} \mathcal{N}(0, 1),$$

où $\mathcal{N}(0, 1)$ est la loi Gaussienne centrée réduite. Il est alors classique de considérer que l'espérance $\mathbb{E} \left([A_N^*]_{ij} \right)$ est égale à la moyenne empirique $\mu_{2M} \left([A_N^*]_{ij} \right)$ avec une marge d'erreur de $1.96 \sqrt{\text{Var} \left([A_N^*]_{ij} \right) / M}$. La variance exacte, inconnue, est généralement approchée par la variance empirique. D'un point de vue pratique, on considère donc que l'espérance $\mathbb{E} \left([A_N^*]_{ij} \right)$ est dans l'intervalle de confiance

$$\left[\mu_{2M} \left([A_N^*]_{ij} \right) - 1.96 \frac{\sqrt{\sigma_{2M} \left([A_N^*]_{ij} \right)}}{\sqrt{2M}}, \mu_{2M} \left([A_N^*]_{ij} \right) + 1.96 \frac{\sqrt{\sigma_{2M} \left([A_N^*]_{ij} \right)}}{\sqrt{2M}} \right]. \quad (138)$$

La quantité calculée en pratique, $\mu_{2M} \left([A_N^*]_{ij} \right)$, est donc une bonne approximation de la valeur exacte $[A^*]_{ij}$ dans la limite $M \gg 1$ et $N \gg 1$.

A titre d'illustration, on donne sur la Figure 12 un résultat numérique en dimension deux, dans le cas où la matrice A s'écrit

$$A(x, \omega) = \sum_{k \in \mathbb{Z}^2} 1_{Q+k}(x) a_k(\omega) \text{Id}_2,$$

où les variables aléatoires a_k sont indépendantes et identiquement distribuées suivant une loi de Bernoulli :

$$\mathbb{P}(a_0 = \alpha) = \mathbb{P}(a_0 = \beta) = 1/2 \text{ avec } \alpha = 3, \beta = 20.$$

On constate, au moins sur ce cas, que pour $|Q_N| \geq 15000$, l'espérance $\mathbb{E} [A_N^*]_{11}$ est quasiment indépendante de N (l'erreur systématique est très petite), et que seule subsiste une erreur statistique.

L'approche qu'on vient de décrire est très chère en temps de calcul : pour un domaine Q_N suffisamment grand, il faut faire un grand nombre $2M$ de calculs pour obtenir une approximation raisonnable de A^* . Afin de réduire le coût calcul, deux approches peuvent être envisagées :

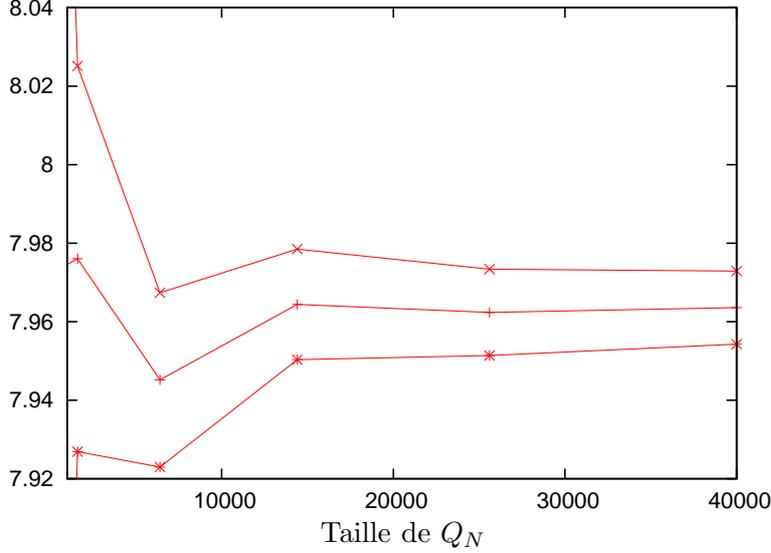


FIG. 12 – Intervalle de confiance (138) pour $[A_N^*]_{11}$ (calculé sur la base de $2M = 100$ réalisations), en fonction de la taille du domaine Q_N .

- proposer une approximation du problème du correcteur (135) alternative à (136), afin que l’erreur systématique $\mathbb{E}[A_N^*] - A^*$ converge plus vite vers 0 avec la taille du domaine de calcul. Il s’agit donc d’essayer d’améliorer le taux de convergence quand $N \rightarrow \infty$. On renvoie à [78] pour une approche en ce sens, dans le cas de problèmes stochastiques *discrets*;
- proposer une stratégie permettant de réduire la variance, et donc de diminuer l’erreur statistique. Typiquement, pour *tout* N (et pas uniquement dans le régime asymptotique $N \rightarrow \infty$), l’idée est de construire un intervalle de confiance pour $\mathbb{E}[A_N^*]$, dans l’esprit de (138), qui soit de taille plus petite, à coût calcul fixe.

Bien sûr, ces deux approches peuvent être combinées. Dans la suite, on se concentre sur la deuxième approche, de type réduction de variance.

Remarque 49 *Nous avons souligné à la Remarque 41 que, dans des cas périodiques non-linéaires, on ne peut pas toujours se ramener à un problème de cellule posée sur une unique cellule. Il est souvent nécessaire de considérer de grands domaines, et la question de l’accélération de la convergence se pose, comme ici dans le cadre stochastique. On renvoie à [20, 77] pour des approches en ce sens, dans le cas déterministe. \diamond*

3.2.2 Réduction de variance par la technique des variables antithétiques

Notre objectif est de calculer plus efficacement $\mathbb{E}[A_N^*]$. On commence par décrire ci-dessous une technique générale de réduction de variance, la technique des *variables antithétiques* [112, page 27]. On l’appliquera ensuite au cadre de l’homogénéisation.

La technique des variables antithétiques On se place dans un cadre très simple. On se donne une fonction $f : [0, 1] \mapsto \mathbb{R}$, et on souhaite calculer $\mathbb{E}(f(U))$, où U est une variable aléatoire uniformément distribuée dans $[0, 1]$. Une méthode standard consiste à introduire l’estimateur

$$I_{2M}(\omega) = \frac{1}{2M} \sum_{m=1}^{2M} f(U_m(\omega)),$$

où $\{U_m(\omega)\}_{1 \leq m \leq 2M}$ sont $2M$ réalisations indépendantes et identiquement distribuées de $U(\omega)$. D'après la loi forte des grands nombres, $I_{2M}(\omega)$ converge presque sûrement vers $\mathbb{E}(f(U))$. Cet estimateur nécessite $2M$ évaluations de la fonction f .

Une approche alternative (c'est la méthode des variables antithétiques) est de considérer l'estimateur

$$\tilde{I}_{2M}(\omega) = \frac{1}{M} \sum_{m=1}^M \frac{1}{2} \left(f(U_m(\omega)) + f(1 - U_m(\omega)) \right). \quad (139)$$

Comme $1 - U(\omega)$ a la même loi que $U(\omega)$, l'estimateur \tilde{I}_{2M} converge aussi presque sûrement vers $\mathbb{E}(f(U))$. Là encore, $2M$ évaluations de la fonction f sont nécessaires.

Ces deux estimateurs ont donc le même coût. On compare maintenant leur précision. En utilisant le théorème de la limite centrale, on peut construire deux intervalles de confiance pour $\mathbb{E}(f(U))$,

$$\left[I_{2M} - 1.96 \frac{\sqrt{\text{Var } f(U)}}{\sqrt{2M}}, I_{2M} + 1.96 \frac{\sqrt{\text{Var } f(U)}}{\sqrt{2M}} \right] \quad (140)$$

et

$$\left[\tilde{I}_{2M} - 1.96 \frac{\sqrt{\text{Var } g(U)}}{\sqrt{M}}, \tilde{I}_{2M} + 1.96 \frac{\sqrt{\text{Var } g(U)}}{\sqrt{M}} \right], \quad (141)$$

où $g(U) = \frac{1}{2} (f(U) + f(1 - U))$. Un calcul simple montre que l'intervalle de confiance (141) associé à \tilde{I}_{2M} est plus petit que l'intervalle de confiance (140) associé à I_{2M} si et seulement si $\text{Cov}(f(U), f(1 - U)) \leq 0$.

Le lemme suivant donne une condition suffisante pour que la technique des variables antithétique réduise la variance :

Lemme 50 *On suppose que $f : [0, 1] \mapsto \mathbb{R}$ est une fonction croissante, et que U est une variable aléatoire uniformément distribuée dans $[0, 1]$. Alors $\text{Cov}(f(U), f(1 - U)) \leq 0$.*

Ce résultat peut être généralisé au cas multi-dimensionnel, sous l'hypothèse que $f : [0, 1]^n \mapsto \mathbb{R}$ est une fonction croissante par rapport à chacun de ses arguments. On renvoie à [FL24].

Application L'idée est maintenant d'utiliser exactement la même méthode pour réduire la variance de $A_N^*(\omega)$, qui dépend du champ aléatoire $\{A(x, \omega)\}_{x \in Q_N}$. De la même manière que, dans l'estimateur (139), apparaissent simultanément les valeurs $f(U)$ et $f(1 - U)$, on va ici, à chaque fois qu'on effectue un calcul avec un champ aléatoire, effectuer le même calcul avec sa version antithétique.

Dans un souci de simplicité, on décrit l'approche dans le cas où le champ aléatoire est de la forme

$$A(x, \omega) = \sum_{k \in \mathbb{Z}^d} 1_{Q+k}(x) a_k(\omega) \text{Id}_d, \quad (142)$$

où les variables $a_k(\omega)$ sont indépendantes et uniformément distribuées dans $[\alpha, \beta]$, avec $0 < \alpha \leq \beta$. On introduit alors le champ antithétique

$$B(x, \omega) = \sum_{k \in \mathbb{Z}^d} 1_{Q+k}(x) b_k(\omega) \text{Id}_d, \quad (143)$$

où les variables aléatoires $b_k(\omega)$ sont les variables antithétiques de $a_k(\omega)$. Dans le cas considéré (a_k uniformément distribué dans $[\alpha, \beta]$), on a simplement $b_k(\omega) = \beta + \alpha -$

$a_k(\omega)$. Bien sûr, la notion de variables antithétiques n'est pas restreinte au cas de variables uniformément distribuées, on renvoie à [FL24] pour plus de détails.

On va maintenant comparer deux stratégies pour estimer $\mathbb{E}[A_N^*]$:

- comme à la Section 3.2.1, on se donne $2M$ réalisations indépendantes $A^m(x, \omega)$, auxquelles on associe les matrices homogénéisées $A_N^{*,m}(\omega)$. On construit ensuite l'estimateur

$$I_{2M} = \frac{1}{2M} \sum_{m=1}^{2M} A_N^{*,m}(\omega).$$

- on considère M réalisations indépendantes $A^m(x, \omega)$, et on construit les M champs antithétiques $B^m(x, \omega)$ associés (cf. (142) et (143)). Pour chacune des réalisations $A^m(x, \omega)$ (respectivement $B^m(x, \omega)$), on calcule le correcteur associé $w_p^{N,m}$ (respectivement $v_p^{N,m}$) ainsi que la matrice homogénéisée $A_N^{*,m}(\omega)$ (respectivement $B_N^{*,m}(\omega)$). On construit ensuite l'estimateur

$$\tilde{I}_{2M} = \frac{1}{M} \sum_{m=1}^M \frac{1}{2} (A_N^{*,m}(\omega) + B_N^{*,m}(\omega)).$$

Chacun des deux estimateurs, I_{2M} et \tilde{I}_{2M} , nécessite la résolution de $2M$ problèmes de correcteur. Ils sont donc de coût identique, et convergent tous les deux presque sûrement vers $\mathbb{E}[A_N^*]$.

En s'appuyant sur le théorème de la limite centrale, on considère en pratique que $\mathbb{E}[A_N^*]_{ij}$ appartient aux intervalles de confiance

$$\left[[I_{2M}]_{ij} - 1.96 \frac{\sqrt{\sigma_{2M}^{ij}}}{\sqrt{2M}}, [I_{2M}]_{ij} + 1.96 \frac{\sqrt{\sigma_{2M}^{ij}}}{\sqrt{2M}} \right] \quad (144)$$

et

$$\left[[\tilde{I}_{2M}]_{ij} - 1.96 \frac{\sqrt{\tilde{\sigma}_{2M}^{ij}}}{\sqrt{M}}, [\tilde{I}_{2M}]_{ij} + 1.96 \frac{\sqrt{\tilde{\sigma}_{2M}^{ij}}}{\sqrt{M}} \right], \quad (145)$$

où σ_{2M}^{ij} est la variance empirique de $[A_N^{*,m}]_{ij}$ (et de même pour $\tilde{\sigma}_{2M}^{ij}$). L'intervalle de confiance (145) est plus petit que l'intervalle de confiance (144) si et seulement si $\tilde{\sigma}_{2M}^{ij} \leq \sigma_{2M}^{ij}/2$. En confondant variance empirique et variance exacte, un calcul simple (analogue du calcul mené ci-dessus) montre que cette inégalité est équivalente au fait que $\text{Cov}([A_N^{*,m}]_{ij}, [B_N^{*,m}]_{ij}) \leq 0$.

A titre illustratif, on montre sur la Figure 13 un résultat numérique obtenu par cette méthode, dans le cas où

$$A(x, \omega) = \sum_{k \in \mathbb{Z}^2} 1_{Q+k}(x) a_k(\omega) \text{Id}_2,$$

où a_k sont des variables i.i.d., uniformément distribuées dans $[\alpha, \beta]$ (avec $\alpha = 3$ et $\beta = 20$). Dans ce cas précis, on observe que la taille de l'intervalle de confiance a été divisée par $\sqrt{6}$, pour un coût calcul identique (à précision donnée, le coût calcul est donc 6 fois plus petit).

Nous n'avons jusqu'ici considéré que des champs de la forme (142), qui sont isotropes, constant sur chaque cellule $Q+k$, et prennent des valeurs indépendantes d'une cellule $Q+k$ à une autre cellule $Q+k'$, $k' \neq k$. L'efficacité de la méthode est en fait avérée pour des

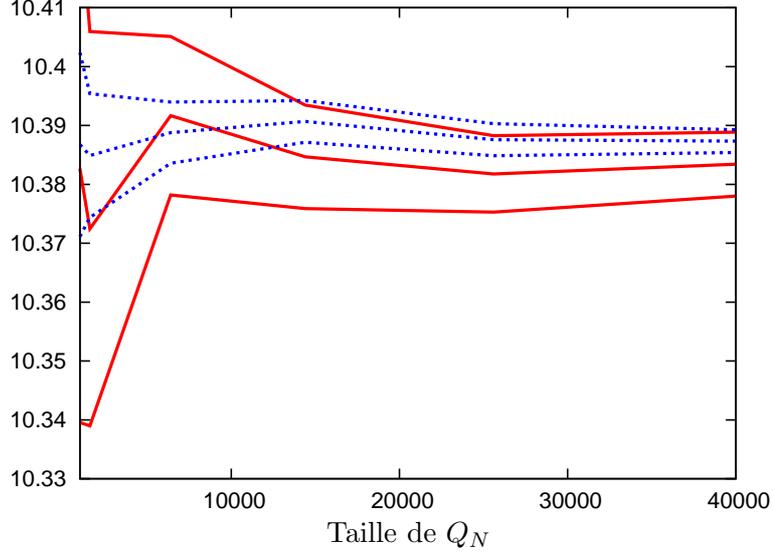


FIG. 13 – Intervalle de confiance pour $[A_N^*]_{11}$, (calculé sur la base de $2M = 100$ réalisations), en fonction de la taille du domaine Q_N (en rouge, estimateur (144); en bleu, estimateur (145)).

cas plus généraux. De la même manière, nous avons ici considéré comme quantité d'intérêt la matrice homogénéisée $\mathbb{E}[A_N^*]$. Cette méthode permet aussi de réduire la variance pour d'autres quantités. Tout ceci est documenté dans [FL19, FL33, 45].

On retiendra ici que, même si la réduction de variance n'est pas spectaculaire, elle est obtenue sans *aucun* coût calcul supplémentaire, et en utilisant une technique très générale. Il est plausible qu'une technique plus évoluée permette un gain encore plus important.

Eléments de preuve Pour conclure cette section, on donne quelques éléments de preuve, en commençant par le cas mono-dimensionnel. Le coefficient homogénéisé, donné par (134), vaut alors

$$a^* = \left[\mathbb{E} \left(\int_0^1 a^{-1}(x, \cdot) dx \right) \right]^{-1}. \quad (146)$$

Par ailleurs, l'approche numérique standard (résoudre (136) pour $Q_N =]0, N[$, évaluer (137)) conduit à la valeur

$$a_N^*(\omega) = \left[\frac{1}{N} \int_0^N a^{-1}(x, \omega) dx \right]^{-1}. \quad (147)$$

Suivant le cadre (142), supposons pour simplifier que

$$a(x, \omega) = \sum_{k \in \mathbb{Z}} 1_{[k-1, k[}(x) a_k(\omega),$$

où les variables a_k sont i.i.d. Alors

$$a_N^*(\omega) = \left[\frac{1}{N} \sum_{k=1}^N \frac{1}{a_k(\omega)} \right]^{-1} = f_N(a_1(\omega), a_2(\omega), \dots, a_N(\omega)),$$

où f_N est une fonction *croissante* de chacun de ses arguments. Or, on a vu que, dans le cas de fonctions croissantes, la méthode des variables antithétiques permet de réduire la

variance. On est donc ici exactement dans ce cadre, ce qui explique pourquoi la méthode permet de réduire la variance de $a_N^*(\omega)$.

Dans [FL24], nous avons établi un résultat en dimension quelconque, sous les hypothèses suivantes (on pourra vérifier que l'exemple (142) les satisfait) :

- On suppose que, pour tout N , il existe un entier n (éventuellement, mais pas nécessairement, $n = |Q_N|$) et une fonction \mathcal{A} , définie sur $Q_N \times \mathbb{R}^n$ et à valeur dans $\mathbb{R}^{d \times d}$, telle que la matrice $A(x, \omega)$ vérifie

$$\forall x \in Q_N, \quad A(x, \omega) = \mathcal{A}(x, X_1(\omega), \dots, X_n(\omega)) \quad \text{p.s.}, \quad (148)$$

où $\{X_k(\omega)\}_{1 \leq k \leq n}$ sont des variables aléatoires scalaires, indépendantes, et uniformément distribuées dans $[0, 1]$. En général, la fonction \mathcal{A} , comme le nombre n de variables aléatoires scalaires dans (148), dépendent de la taille de Q_N . On ne fait pas apparaître cette dépendance explicitement ici.

- On suppose que la fonction \mathcal{A} dans (148) est telle que, pour tout $x \in Q_N$ et pour tout $\xi \in \mathbb{R}^d$, l'application

$$(x_1, \dots, x_n) \in \mathbb{R}^n \mapsto \xi^T \mathcal{A}(x, x_1, \dots, x_n) \xi \quad (149)$$

est croissante de chacune de ses variables.

On a alors le résultat suivant :

Proposition 51 ([FL24], Proposition 1) *On fait les hypothèses (148)–(149). Soit $A_N^*(\omega)$ la matrice homogénéisée approchée, obtenue en travaillant sur le domaine $Q_N \subset \mathbb{R}^d$ (cf. (136)–(137)).*

On définit sur le domaine Q_N le champ

$$B(x, \omega) := \mathcal{A}(x, 1 - X_1(\omega), \dots, 1 - X_n(\omega)),$$

antithétique à $A(\cdot, \omega)$ défini par (148). On associe à ce champ le problème du correcteur (136) (où on a remplacé A par B), dont on note v_p^N la solution, et la matrice $B_N^(\omega)$, définie à partir de v_p^N suivant (137). On pose*

$$\tilde{A}_N^*(\omega) := \frac{1}{2} (A_N^*(\omega) + B_N^*(\omega)).$$

Alors $\mathbb{E}(\tilde{A}_N^) = \mathbb{E}(A_N^*)$, et, pour tout $\xi \in \mathbb{R}^d$,*

$$\text{Var}(\xi^T \tilde{A}_N^* \xi) \leq \frac{1}{2} \text{Var}(\xi^T A_N^* \xi). \quad (150)$$

Ainsi, \tilde{A}_N^ est une variable aléatoire dont l'espérance est identique à l'espérance de A_N^* , et dont la variance est plus petite que la moitié de celle de A_N^* , au sens de (150).*

Ainsi, pour tout $\xi \in \mathbb{R}^d$, on peut construire deux intervalles de confiance pour la quantité $\mathbb{E}(\xi^T A_N^* \xi)$ en résolvant $2M$ problèmes des correcteurs (comme on l'a fait avec $\mathbb{E}[A_N^*]_{ij}$ dans (144) et (145)). On déduit de (150) que l'intervalle de confiance bâti avec le second estimateur est plus petit que le premier, et donc que la méthode permet effectivement, à coût constant, d'obtenir un intervalle de confiance plus petit.

Nous ne détaillons pas ici la preuve de ce résultat, mais donnons simplement l'architecture de l'argument. Tout d'abord, on a vu que la technique des variables antithétiques permet de réduire la variance dans le calcul de $\mathbb{E}(f(X_1, \dots, X_n))$, où f est une fonction scalaire qui est croissante de chacun de ses arguments, et $X = (X_1, \dots, X_n)$ est un vecteur de

variables aléatoires indépendantes. Par ailleurs, on a fait l'hypothèse que la matrice $A(x, \omega)$ s'écrit comme une fonction croissante (au sens des matrices symétriques définies positives) des variables aléatoires indépendantes $X_k(\omega)$ (ce sont les hypothèses (148) et (149)). Enfin, on utilise le fait que l'homogénéisation *préserve l'ordre des matrices symétriques*. Par conséquent, $A_N^*(\omega)$ est aussi une fonction croissante des $X_k(\omega)$, ce qui implique que la technique des variables antithétiques permet de réduire la variance. De la même manière, on réduit aussi la variance de toute quantité qui est une fonction croissante de $A_N^*(\omega)$ (comme ses valeurs propres, ...). On renvoie à [FL24] pour plus de détails.

3.2.3 Un cas faiblement stochastique

Comme on l'a vu, les problèmes d'homogénéisation stochastique conduisent souvent à des calculs très lourds (résolution, pour un grand nombre de réalisations, du problème du correcteur (136) sur un grand domaine). Il est donc intéressant de chercher des cas particuliers du cas général, pour lesquels une approche spécifique peut être développée, qui conduise à des calculs moins chers. On s'intéresse dans la suite à des cas où le caractère aléatoire du matériau est une *petite perturbation* d'un comportement périodique. Ces cas sont motivés par le fait que, si très peu de matériaux hétérogènes sont rigoureusement périodiques, tous ne sont pas *fortement* stochastiques. Il existe de nombreux matériaux qui ne sont que *faiblement* aléatoires.

On suppose ici que la matrice A dans (133) est une *perturbation* d'une matrice périodique, au sens où

$$A(x, \omega) = A_{\text{per}}(x) + \eta A_1(x, \omega), \quad (151)$$

où A_{per} est Q -périodique, A_1 est stationnaire au sens de (131), et η est un *petit* paramètre déterministe. On suppose que A , A_{per} et A_1 sont trois matrices symétriques et bornées, et que A et A_{per} sont uniformément coercives. La relation (151) montre que A est une perturbation de A_{per} .

Remarque 52 *De nombreux cadres perturbatifs ont été proposés et étudiés dans la littérature. On renvoie à [101, FL35] pour une revue de différents cas possibles (cf. aussi [3]). Notons aussi qu'un autre cadre perturbatif, introduit dans [25], sera étudié dans la Section 3.3.3. \diamond*

En conséquence de (151), il est possible de développer la matrice homogénéisée A^* en une série en puissance de η (cette idée a déjà été suivie dans [25], sur un cas un peu différent). On montre alors que le correcteur ∇w_p , solution de (135), vérifie

$$\nabla w_p = \nabla w_p^0 + \eta \nabla w_p^1 + \dots, \quad (152)$$

où w_p^0 est la solution (unique à une constante additive près) du problème du correcteur périodique,

$$\begin{cases} -\text{div} [A_{\text{per}}(\nabla w_p^0 + p)] = 0 & \text{dans } \mathbb{R}^d, \\ w_p^0 \text{ est } Q\text{-périodique,} \end{cases} \quad (153)$$

et w_p^1 est la solution (unique à une constante aléatoire additive près) de

$$\begin{cases} -\text{div} [A_{\text{per}} \nabla w_p^1] = \text{div} [A_1 (\nabla w_p^0 + p)] & \text{p.s. sur } \mathbb{R}^d, \\ \nabla w_p^1 \text{ est stationnaire, } \int_Q \mathbb{E}(\nabla w_p^1) = 0. \end{cases} \quad (154)$$

De plus, la matrice homogénéisée (134) vérifie

$$A^* = A_{\text{per}}^* + \eta A_1^* + O(\eta^2), \quad (155)$$

où A_{per}^* est la matrice homogénéisée associée à A_{per} via (126)-(127), et

$$\begin{aligned} [A_1^*]_{ij} &= \int_Q \mathbb{E}(\nabla w_{e_i}^1)^T A_{\text{per}}(\nabla w_{e_j}^0 + e_j) + \int_Q (\nabla w_{e_i}^0 + e_i)^T A_{\text{per}} \mathbb{E}(\nabla w_{e_j}^1) \\ &+ \int_Q (\nabla w_{e_i}^0 + e_i)^T \mathbb{E}(A_1)(\nabla w_{e_j}^0 + e_j). \end{aligned} \quad (156)$$

Dans cet exemple très particulier, il n'y a en fait pas de problème de variance, et ce pour deux raisons. Tout d'abord, la fonction $\mathbb{E}(w_p^1)$, nécessaire pour évaluer (156), peut être directement calculée, sans passer par w_p^1 . On déduit en effet de (154), en prenant l'espérance, que $\bar{w}_p^1 := \mathbb{E}(w_p^1)$ vérifie

$$\begin{cases} -\text{div} [A_{\text{per}} \nabla \bar{w}_p^1] = \text{div} [\mathbb{E}(A_1) (\nabla w_p^0 + p)] & \text{sur } \mathbb{R}^d, \\ \bar{w}_p^1 \text{ est } Q\text{-périodique.} \end{cases} \quad (157)$$

Ce problème, déterministe et posé sur Q , est évidemment plus facile à résoudre que le problème (135) (ou sa version tronquée (136)). La détermination de A^* (au premier ordre en η) se fait donc simplement via la résolution des problèmes *déterministes* (153) et (157).

Par ailleurs, le correcteur *exact* w_p , qui est solution de (135), est en pratique approché par la solution w_p^N du problème tronqué (136), qui a elle-aussi un développement en puissance de η : à l'image de (152), on a

$$\nabla w_p^N = \nabla w_p^0 + \eta \nabla w_p^{1,N} + \dots, \quad (158)$$

où $w_p^{1,N}$ est une solution du problème *tronqué*

$$\begin{cases} -\text{div} [A_{\text{per}} \nabla w_p^{1,N}] = \text{div} [A_1 (\nabla w_p^0 + p)], \\ w_p^{1,N} \text{ est } Q_N\text{-périodique.} \end{cases}$$

De même, la matrice homogénéisée A^* est approchée en pratique par $A_N^*(\omega)$ définie par (137), qui vérifie

$$A_N^*(\omega) = A_{\text{per}}^* + \eta A_{1,N}^*(\omega) + O(\eta^2),$$

où

$$\begin{aligned} [A_{1,N}^*]_{ij}(\omega) &= \frac{1}{|Q_N|} \left[\int_{Q_N} \nabla w_{e_i}^{1,N}(x, \omega)^T A_{\text{per}}(x) (\nabla w_{e_j}^0(x) + e_j) dx \right. \\ &+ \int_{Q_N} (\nabla w_{e_i}^0(x) + e_i)^T A_1(x, \omega) (\nabla w_{e_j}^0(x) + e_j) dx \\ &\left. + \int_{Q_N} (\nabla w_{e_i}^0(x) + e_i)^T A_{\text{per}}(x) \nabla w_{e_j}^{1,N}(x, \omega) dx \right]. \end{aligned} \quad (159)$$

Il est alors facile de montrer, en utilisant (153), que le premier terme de (159) s'annule, ainsi que le troisième. On a donc

$$[A_{1,N}^*]_{ij}(\omega) = \frac{1}{|Q_N|} \int_{Q_N} (\nabla w_{e_i}^0(x) + e_i)^T A_1(x, \omega) (\nabla w_{e_j}^0(x) + e_j) dx.$$

Comme A_1 est stationnaire, une simple application du théorème ergodique montre que $A_{1,N}^*(\omega)$ converge presque sûrement vers A_1^* . En faisant des hypothèses supplémentaires sur A_1 , on peut aussi obtenir un taux de convergence, par un argument de type limite centrale. Enfin, d'un point de vue pratique, le calcul de $A_{1,N}^*$ ne pose pas de difficulté, car il ne nécessite pas la résolution de problèmes sur un grand domaine.

Ce cas particulier est en fait un très bon cadre pour comprendre plus en détail le lien entre la variance de $A_1(x, \omega)$ et celle de la matrice homogénéisée, mieux comprendre la méthode de réduction de variance exposée à la Section 3.2.2, ... On va donc utiliser cette méthode comme dans le cas général de la Section 3.2.2 (calcul, pour N fixé, d'un intervalle de confiance pour $\mathbb{E}(A_N^*)$ en utilisant plusieurs réalisations de $A(x, \omega)$, dans l'esprit de (144) et (145)), et utiliser les expressions analytiques par ailleurs disponibles pour mieux *quantifier* la qualité de l'approximation numérique, à la fois en terme de réduction de variance et en terme de taux de convergence avec N de la procédure d'approximation.

Dans la suite, on va s'intéresser aux quantités

$$\begin{aligned} \left[A_{1,N}^{*,\text{exact}} \right]_{ij}(\omega) = & \frac{1}{|Q_N|} \left[\int_{Q_N} \nabla w_{e_i}^1(x, \omega)^T A_{\text{per}}(x) (\nabla w_{e_j}^0(x) + e_j) dx \right. \\ & + \int_{Q_N} (\nabla w_{e_i}^0(x) + e_i)^T A_1(x, \omega) (\nabla w_{e_j}^0(x) + e_j) dx \\ & \left. + \int_{Q_N} (\nabla w_{e_i}^0(x) + e_i)^T A_{\text{per}}(x) \nabla w_{e_j}^1(x, \omega) dx \right], \quad (160) \end{aligned}$$

définies pour tout $N \in \mathbb{N}^*$ et $1 \leq i, j \leq d$. Les fonctions w_p^0 et w_p^1 sont les solutions respectives de (153) et de (154), et $Q_N = \cup_{|k|_\infty \leq N} (Q + k)$ est le cube de taille $(2N + 1)^d$ centré à l'origine. On remarque qu'on a ici utilisé le terme de premier ordre w_p^1 dans le développement (152) du correcteur *exact*. Si ce terme est remplacé par le terme de premier ordre $w_p^{1,N}$ dans le développement (158) du correcteur *approché* (solution du problème tronqué), alors on retrouve $A_{1,N}^*$ défini par (159).

Dans [FL24], nous avons démontré le résultat suivant :

Proposition 53 ([FL24], Proposition 2) *Soit A définie par (151), où A_{per} est périodique et Hölder continue. On suppose que A , A_{per} et A_1 sont symétriques et bornées, et que A et A_{per} sont uniformément coercives. On suppose de plus que A_1 est de la forme*

$$A_1(x, \omega) = \sum_{k \in \mathbb{Z}^d} 1_{Q+k}(x) X_k(\omega) Id_d,$$

où Id_d est la matrice identité et $(X_k)_{k \in \mathbb{Z}^d}$ est une suite de variables aléatoires scalaires i.i.d. et telles qu'il existe $0 < \alpha < \infty$ tel que $|X_k| \leq \alpha$ presque sûrement et pour tout k .

On considère la matrice $A_{1,N}^{*,\text{exact}}$ définie par (160). Il existe d^2 coefficients $C_N^{ij} > 0$, indépendants de $(X_k)_{k \in \mathbb{Z}^d}$, tels que

$$\text{Var} \left(\left[A_{1,N}^{*,\text{exact}} \right]_{ij} \right) = C_N^{ij} \text{Var}(X_0). \quad (161)$$

De plus, on a

$$C_N^{ij} \leq \frac{C}{|Q_N|}, \quad (162)$$

où C ne dépend pas de i, j et N , et dépend seulement de A_{per} .

La relation (161) montre que réduire la variance de X_0 (par exemple par la technique des variables antithétiques de la Proposition 51) réduit la variance de $A_{1,N}^{*,\text{exact}}$. On obtient aussi une estimation *quantitative* de la réduction de variance, ce dont nous ne disposons pas dans le cas général de la Section 3.2.2.

On peut de plus obtenir un résultat de type limite centrale sur $A_{1,N}^{*,\text{exact}}$ quand $N \rightarrow \infty$. On renvoie à [FL24] pour l'énoncé précis de ces résultats, leurs preuves, et des extensions. Notons ici que ces résultats font appel à des propriétés, exposées dans [FL28], sur le comportement asymptotique de la fonction de Green de l'opérateur $-\text{div} [A_{\text{per}} \nabla \cdot]$, i.e. la solution de

$$-\text{div}_x (A_{\text{per}}(x) \nabla_x G(x, y)) = \delta_y(x) \quad \text{dans } \mathbb{R}^d.$$

3.3 Une variante de l'homogénéisation stochastique

Publications [FL15, FL36]

Dans [FL15, FL36], nous nous sommes intéressés à un autre cadre d'homogénéisation stochastique, introduit et étudié dans [23, 25]. Cette variante, qui n'est pas un cas particulier du cadre étudié dans la Section 3.1.2, est motivée par la considération de matériaux dont la configuration courante, aléatoire, peut être reliée à une configuration périodique de référence par une déformation aléatoire.

Nous présentons tout d'abord ce modèle en toute généralité. On en discute ensuite une version faiblement stochastique.

3.3.1 Présentation du modèle

On se donne une matrice A_{per} , qu'on suppose \mathbb{Z}^d -périodique, bornée et elliptique, et on considère le problème

$$\begin{cases} -\text{div} \left[A_{\text{per}} \left(\Phi^{-1} \left(\frac{x}{\varepsilon}, \omega \right) \right) \nabla u^\varepsilon \right] = f & \text{dans } \mathcal{D}, \\ u^\varepsilon = 0 & \text{sur } \partial\mathcal{D}, \end{cases} \quad (163)$$

où, presque sûrement, $\Phi(\cdot, \omega)$ est un difféomorphisme de \mathbb{R}^d sur \mathbb{R}^d . On suppose de plus que

$$\text{EssInf}_{\omega \in \Omega, x \in \mathbb{R}^d} [\det(\nabla \Phi(x, \omega))] = \nu > 0, \quad \text{EssSup}_{\omega \in \Omega, x \in \mathbb{R}^d} (|\nabla \Phi(x, \omega)|) = M < \infty, \quad (164)$$

et que

$$\nabla \Phi(x, \omega) \quad \text{est stationnaire au sens de (131)}. \quad (165)$$

On voit donc que les propriétés du matériau sont périodiques, à la déformation aléatoire Φ près! D'autre part, le gradient de cette déformation est statistiquement homogène.

Ce problème d'homogénéisation a été étudié dans [23, 25], où le résultat suivant est démontré :

Théorème 54 *Soit Φ un difféomorphisme aléatoire vérifiant (164)-(165). Alors la solution $u^\varepsilon(x, \omega)$ de (163) converge presque sûrement vers $u^*(x)$, fortement dans $L^2(\mathcal{D})$ et faiblement dans $H^1(\mathcal{D})$, où u^* est la solution du problème homogénéisé*

$$\begin{cases} -\text{div} [A^* \nabla u^*] = f & \text{dans } \mathcal{D}, \\ u^* = 0 & \text{sur } \partial\mathcal{D}. \end{cases} \quad (166)$$

Dans (166), la matrice homogénéisée A^* est définie par

$$[A^*]_{ij} = \frac{\mathbb{E} \left(\int_{\Phi(Q, \cdot)} (e_i + \nabla w_{e_i}(y, \cdot))^T A_{per}(\Phi^{-1}(y, \cdot)) e_j dy \right)}{\det \left(\mathbb{E} \left(\int_Q \nabla \Phi(z, \cdot) dz \right) \right)}, \quad (167)$$

où, pour tout $p \in \mathbb{R}^d$, w_p est la solution (unique à une constante aléatoire additive près) dans $\{w \in L^2_{loc}(\mathbb{R}^d, L^2(\Omega)), \nabla w \in L^2_{unif}(\mathbb{R}^d, L^2(\Omega))\}$ du problème du correcteur

$$\left\{ \begin{array}{l} -\operatorname{div} [A_{per}(\Phi^{-1}(y, \omega)) (p + \nabla w_p)] = 0, \\ w_p(y, \omega) = \tilde{w}_p(\Phi^{-1}(y, \omega), \omega), \quad \nabla \tilde{w}_p \text{ est stationnaire au sens de (131),} \\ \mathbb{E} \left(\int_{\Phi(Q, \cdot)} \nabla w_p(y, \cdot) dy \right) = 0. \end{array} \right. \quad (168)$$

Remarquons que, dans le cas où Φ est l'identité, on retrouve la matrice homogénéisée du problème périodique.

3.3.2 Approximation numérique

En pratique, le problème du correcteur (168) est résolu sur un domaine tronqué, comme dans le cas du problème (135), qui est approché par (136). Pour une taille de maillage h fixée, on se donne donc une triangulation périodique $\mathcal{T}_h^{(Q)}$ de la cellule unité Q , qu'on étend par périodicité à une triangulation \mathcal{T}_h^N du domaine $Q_N \subset \mathbb{R}^d$, cube de taille $(2N+1)^d$ centré à l'origine. En suivant l'approche de [34], on considère l'approximation suivante de (168) :

$$\left\{ \begin{array}{l} \text{Trouver } \tilde{w}_p^{h,N}(\cdot, \omega) \in V_h^{\text{per}}(Q_N) \text{ tel que, pour tout } \tilde{v}_h \in V_h^{\text{per}}(Q_N), \\ \int_{Q_N} \det(\nabla \Phi) (\nabla \tilde{v}_h)^T (\nabla \Phi)^{-T} A_{per} \left(p + (\nabla \Phi)^{-1} \nabla \tilde{w}_p^{h,N}(\cdot, \omega) \right) = 0 \text{ p.s.,} \end{array} \right. \quad (169)$$

où $V_h^{\text{per}}(Q_N)$ est l'ensemble des fonctions définies sur \mathbb{R}^d , Q_N -périodiques, et dont la restriction à Q_N appartient aux fonctions éléments finis périodiques batis sur le maillage \mathcal{T}_h^N . On remarque que, dans (169), on a travaillé avec la fonction \tilde{w}_p plutôt que w_p , puisque l'hypothèse de stationnarité porte sur \tilde{w}_p . On définit ensuite l'approximation suivante de la matrice A^* :

$$(A_{h,N}^*)_{ij}(\omega) := \det \left(\frac{1}{|Q_N|} \int_{Q_N} \nabla \Phi \right)^{-1} \frac{1}{|Q_N|} \int_{Q_N} \det(\nabla \Phi) \left(e_i + (\nabla \Phi)^{-1} \nabla \tilde{w}_{e_i}^{h,N} \right)^T A_{per} e_j. \quad (170)$$

Comme dans la Section 3.1.2, la matrice homogénéisée exacte A^* est déterministe, tandis que son approximation $A_{h,N}^*$ est stochastique.

L'étude de la limite de $A_{h,N}^*$ lorsque h tend vers 0 ne pose pas de problème particulier, et se fait en utilisant des résultats classiques en éléments finis. On peut ainsi définir $A_N^*(\omega) = \lim_{h \rightarrow 0} A_{h,N}^*(\omega)$. L'étude de la limite de A_N^* quand N tend vers l'infini est une autre question, qui est très proche de l'étude de [34] dans le cadre standard de l'homogénéisation stochastique. Cette question, dans le cadre de la variante (163) présentée ici, fait l'objet du travail en cours [FL36].

3.3.3 Perturbations de l'identité

Dans le cadre présenté ci-dessus, il est possible d'introduire une version *faiblement* stochastique du modèle, dans l'esprit de la Section 3.2.3. On suppose ici que le difféomorphisme Φ est une perturbation de l'identité, au sens où

$$\Phi(x, \omega) = x + \eta \Psi(x, \omega) + O(\eta^2).$$

Alors, comme dans la Section 3.2.3, on peut développer la matrice homogénéisée (167) en puissance de η . Pour $\eta = 0$, le problème (163) est un problème d'homogénéisation périodique. Il est donc naturel d'introduire le correcteur périodique w_p^0 , solution de

$$-\operatorname{div} [A_{per} (p + \nabla w_p^0)] = 0, \quad w_p^0 \text{ est } Q\text{-périodique}, \quad (171)$$

et la matrice $[A_{per}^*]_{ij} = \int_Q e_i^T A_{per} (e_j + \nabla w_{e_j}^0)$. Les auteurs de [25] montrent que la matrice homogénéisée (167) admet le développement

$$A^* = A_{per}^* + \eta A_1^* + O(\eta^2),$$

où A_1^* est une matrice qui peut être calculée en connaissant la solution d'un problème auxiliaire, *déterministe* et posé sur le domaine borné Q . Ainsi, comme dans le cas de la Section 3.2.3 (cf. (155), (156) et (157)), il suffit de résoudre deux problèmes *déterministes*, posés sur le domaine borné Q , pour calculer la matrice homogénéisée au premier ordre en η .

Nous avons démontré dans [FL15] un résultat similaire, cette fois sur les versions *discrétisées* des problèmes (la preuve suit les mêmes étapes que la preuve du cas continu donnée dans [25]). L'objectif est d'obtenir un développement en puissance de η de la matrice $A_{h,N}^*$ définie par (170), en résolvant des problèmes plus simples que (169).

Il est naturel d'introduire la fonction $w_p^{0,h} \in V_h^{\text{per}}(Q)$ solution de

$$\forall v_h \in V_h^{\text{per}}(Q), \quad \int_Q (\nabla v_h)^T A_{per} (p + \nabla w_p^{0,h}) = 0,$$

qui est la formulation discrétisée du problème du correcteur périodique (171), ainsi que la matrice homogénéisée correspondante, $[A_{per,h}^*]_{ij} = \int_Q e_i^T A_{per} (e_j + \nabla w_{e_j}^{0,h})$. Dans [FL15] (cf. aussi [45, 46]), nous avons montré qu'il existe une constante $C(h, N, \omega)$ telle que, pour η assez petit,

$$\eta^{-2} |A_{h,N}^*(\omega) - A_{per,h}^* - \eta A_{1,h,N}^*(\omega)| \leq C(h, N, \omega),$$

où $A_{1,h,N}^*$ est la matrice définie par

$$\begin{aligned} [A_{1,h,N}^*]_{ij} = & -[A_{per,h}^*]_{ij} \frac{1}{|Q_N|} \int_{Q_N} \operatorname{div} \Psi + \frac{1}{|Q_N|} \int_{Q_N} (e_i + \nabla w_{e_i}^{0,h})^T A_{per} e_j \operatorname{div} \Psi \\ & + \frac{1}{|Q_N|} \int_{Q_N} (\nabla w_{e_i}^{1,h,N} - \nabla \Psi \nabla w_{e_i}^{0,h})^T A_{per} e_j, \end{aligned}$$

où $w_p^{1,h,N} \in V_h^{\text{per}}(Q_N)$ est la solution de

$$\begin{aligned} \forall v_h \in V_h^{\text{per}}(Q_N), \quad \int_{Q_N} (\nabla v_h)^T A_{per} \nabla w_p^{1,h,N} = \\ \int_{Q_N} (\nabla v_h)^T \left[A_{per} \nabla \Psi \nabla w_p^{0,h} + (\nabla \Psi^T - (\operatorname{div} \Psi) \operatorname{Id}) A_{per} (p + \nabla w_p^{0,h}) \right]. \quad (172) \end{aligned}$$

La résolution de (172) est aussi coûteuse que la résolution de (169), car c'est un problème posé sur le grand domaine Q_N . On observe cependant que l'espérance de $w_p^{1,h,N}$ est une fonction indépendante de N , solution d'un problème *déterministe*, posé sur le petit domaine Q . Par conséquent, la détermination de $\mathbb{E}(w_p^{1,h,N})$, et par conséquent de $\mathbb{E}(A^{1,h,N})$, est facile.

Introduisons les deux estimateurs d'erreur suivants :

$$\begin{aligned} e^{h,N}(\omega) &:= \eta^{-2} (A_{h,N}^*(\omega) - A_{per,h}^* - \eta A_{1,h,N}^*(\omega)), \\ \bar{e}^{h,N}(\omega) &:= \eta^{-2} (A_{h,N}^*(\omega) - A_{per,h}^* - \eta \mathbb{E}(A_{1,h,N}^*)). \end{aligned}$$

On montre dans [FL15] que la variance de $\bar{e}^{h,N}$ diverge quand $\eta \rightarrow 0$. Par conséquent, pour l'*analyse* de l'approche, il est nécessaire de considérer l'estimateur $e^{h,N}$, ce qui permet de vérifier que $A_{per,h}^* + \eta A_{1,h,N}^*$ est une approximation précise de $A_{h,N}^*$, au second ordre en η . En *pratique*, on utilise bien sûr $A_{per,h}^* + \eta \mathbb{E}(A_{1,h,N}^*)$, bien plus facile à calculer, pour approcher $A_{h,N}^*$.

3.3.4 Résidu dans le cas mono-dimensionnel

Dans tous les problèmes d'homogénéisation décrits ci-dessus, on a vu que u^ε converge vers u^* fortement dans $L^2(\mathcal{D})$, et faiblement dans $H^1(\mathcal{D})$. On sait que, dans le cas périodique, il est possible d'améliorer la qualité de l'approximation (et ainsi obtenir des convergences dans $H^1(\mathcal{D})$), en considérant $u^* - \varepsilon \sum_{i=1}^d w_{e_i} \left(\frac{\cdot}{\varepsilon} \right) \frac{\partial u^*}{\partial x_i}$ plutôt que u^* (cf. (128)).

Dans le cadre de l'homogénéisation stochastique présentée à la Section 3.1.2, ce résultat est encore vrai [126, Théorème 3]. On peut se convaincre de ceci dans le cas mono-dimensionnel, où tous les calculs peuvent être faits analytiquement. Le problème (133) s'écrit alors (on a choisi $\mathcal{D} =]0, 1[$)

$$\begin{cases} -\frac{d}{dx} \left[a \left(\frac{x}{\varepsilon}, \omega \right) \frac{d}{dx} u^\varepsilon \right] = f \text{ dans }]0, 1[, \\ u^\varepsilon(0, \omega) = u^\varepsilon(1, \omega) = 0. \end{cases}$$

Le coefficient homogénéisé a^* est donné par (146). On introduit la fonction

$$v^\varepsilon(x, \omega) = u^\varepsilon(x, \omega) - \left[u^*(x) + \varepsilon \frac{du^*}{dx}(x) w \left(\frac{x}{\varepsilon}, \omega \right) \right],$$

où u^* est la solution du problème homogénéisé (125), et w est le correcteur, solution de (135). Cette fonction converge vers 0 dans $L^2(0, 1)$ presque sûrement. De plus, un calcul simple montre que

$$\frac{dv^\varepsilon}{dx}(x, \omega) = \left(c_\varepsilon - \int_0^1 F(t) dt \right) a^{-1} \left(\frac{x}{\varepsilon}, \omega \right) - \varepsilon \frac{d^2 u^*}{dx^2}(x) w \left(\frac{x}{\varepsilon}, \omega \right) \quad (173)$$

avec

$$F(t) = \int_0^t f(s) ds \quad \text{et} \quad c_\varepsilon = \frac{\int_0^1 F(t) a^{-1} \left(\frac{t}{\varepsilon}, \omega \right) dt}{\int_0^1 a^{-1} \left(\frac{t}{\varepsilon}, \omega \right) dt}.$$

Grace au théorème ergodique, c_ε converge presque sûrement vers $\int_0^1 F(t) dt$, et donc $\frac{dv^\varepsilon}{dx}$ converge presque sûrement vers 0, dans $L^\infty(0, 1)$ et donc dans $L^2(0, 1)$. Ainsi, le correcteur

permet à nouveau d'améliorer la convergence, puisqu'on obtient que v^ε converge vers 0 dans H^1 fort.

Pour préciser la vitesse de convergence de $\frac{dv^\varepsilon}{dx}$ vers 0, il est nécessaire de préciser la vitesse de convergence dans le théorème ergodique, ce qui nécessite des hypothèses supplémentaires. Comme a (et donc a^{-1}) est stationnaire, on sait que les variables aléatoires $X_k(\omega) = \int_k^{k+1} a^{-1}(t, \omega) dt$ sont identiquement distribuées. Faisons l'hypothèse supplémentaire qu'elles sont indépendantes. Alors, en utilisant le théorème de la limite centrale, on obtient que

$$\varepsilon^{-1/2} \left(c_\varepsilon - \int_0^1 F(t) dt \right) \text{ converge en loi vers une Gaussienne.} \quad (174)$$

Bien sûr, l'hypothèse d'indépendance des X_k peut être affaiblie.

On déduit donc de (173) et (174) que, pour tout x dans $]0, 1[$, $\varepsilon^{-1/2} \frac{dv^\varepsilon}{dx}(x, \omega)$ "ressemble", quand $\varepsilon \rightarrow 0$, à une variable aléatoire gaussienne (indépendante de x) multipliée par la fonction hautement oscillante $a^{-1}\left(\frac{x}{\varepsilon}, \omega\right)$. Ce résultat est différent de celui obtenu dans le cas périodique, où on peut montrer que $\left\| \varepsilon^{-1} \frac{dv^\varepsilon}{dx} \right\|_{L^2(0,1)} \leq C$ pour une constante C indépendante de ε . Ainsi, dans le cas périodique, on a une convergence au taux ε , alors que le taux est $\sqrt{\varepsilon}$ dans le cas stochastique.

On peut aussi utiliser (174) pour obtenir des estimations sur la fonction $v^\varepsilon(x, \omega)$ elle-même, plutôt que sa dérivée. On renvoie à [10, 33] pour un résultat détaillé sur ce cas, dans le cas mono-dimensionnel.

Une question intéressante est la généralisation de ce type de résultats au cas de la variante de l'homogénéisation stochastique présentée dans la Section 3.3.1, où intervient à la fois une matrice *périodique* et un difféomorphisme *aléatoire*. Cette question fait l'objet du travail en cours [FL36].

Conclusion et travaux en cours

Une grande partie de mon activité de recherche a pour but de développer des modèles et des méthodes numériques, pour la simulation des matériaux à plusieurs échelles, en espace et aussi en temps. Les travaux effectués jusqu'à présent ont fait apparaître de nombreuses questions intéressantes, qui font l'objet de travaux en cours, et que nous avons parfois déjà évoquées dans ce mémoire. Nous en détaillons quelques-unes ici.

Sur la thématique simulation moléculaire, nous poursuivons nos travaux dans plusieurs directions. Nous avons présenté dans la Section 1.2 une stratégie pour construire une dynamique effective, approchant la dynamique $t \mapsto \xi(X_t) \in \mathbb{R}$ d'une quantité d'intérêt, lorsque la dynamique $t \mapsto X_t \in \mathbb{R}^N$ du système complet est donnée. Dans les exemples que nous avons traités, la précision de cette dynamique repose sur une séparation des échelles de temps entre composantes rapides du mouvement et composantes lentes. Dans le cas de systèmes homogènes comportant un grand nombre d'atomes (dont l'exemple le plus simple est une chaîne mono-dimensionnelle d'atomes avec interactions à plus proches voisins), il est plausible de pouvoir trouver une coordonnée de réaction $\xi(X)$ définie comme une moyenne sur le système, et qui de ce fait évolue plus lentement que les positions individuelles $X_i \in \mathbb{R}$, $1 \leq i \leq N$, de chaque particule. Dans ce cas, la séparation des échelles de temps ne vient pas de l'énergie potentielle utilisée (le système est homogène, il n'y a pas de petit paramètre dans l'énergie potentielle), mais d'une limite thermodynamique. La dynamique effective associée à ce choix de ξ est ainsi une dynamique à l'échelle macroscopique, partant d'une dynamique à l'échelle microscopique. La Proposition 14 permet de quantifier la précision de cette dynamique. L'idée est donc de choisir ξ de telle façon à ce que les constantes m , M , λ , κ et ρ , qui dépendent *a priori* de la longueur N de la chaîne, soient telles que le membre de droite de l'estimation (52) converge vers 0 dans la limite thermodynamique $N \rightarrow \infty$.

Une autre direction que nous explorons fait suite au travail [FL26], présenté dans la Section 1.3.3, dans lequel nous proposons une adaptation de l'algorithme pararéel au cas de problèmes Hamiltoniens. Comme décrit dans cette section, la simple symétrisation de l'algorithme ne suffit pas pour obtenir de bons résultats, et une procédure de projection s'avère utile. Dans le travail [FL38], nous nous intéressons aux équations modifiées correspondant aux différents schémas pararéels construits, afin de mieux comprendre leurs propriétés en temps long.

D'autres travaux sont en cours, d'une part sur des questions de conductivité thermique dans des chaînes d'atomes mono-dimensionnelles soumises à un gradient thermique et à des forces mécaniques extérieures [FL31], et d'autre part sur l'obtention de modèles effectifs sur la base de modèles de Monte-Carlo cinétique comportant un grand nombre de puits, et présentant une séparation d'échelles de temps (certaines transitions étant beaucoup plus fréquentes que d'autres) [FL39].

Dans la thématique homogénéisation stochastique, plusieurs directions de travail sont aussi à l'étude. Tout d'abord, sur le thème de la réduction de variance, nous avons pour l'instant utilisé une technique très générique, celle des variables antithétiques. Nous souhaitons à l'avenir utiliser des approches plus spécifiques, ce qui pourrait permettre d'obtenir des gains plus importants. Par ailleurs, nous nous sommes concentrés pour l'instant sur le problème homogénéisé, et donc sur le calcul de la matrice homogénéisée. Cette approche est intéressante quand la longueur caractéristique des hétérogénéités est infiniment plus petite que la dimension du matériau considéré. Il est cependant des cas où on souhaite traiter directement le problème elliptique originel (ici, le problème (133)), et non pas sa limite quand $\varepsilon \rightarrow 0$. Pour ce faire, on peut par exemple utiliser la *Multiscale Finite*

Element Method (MsFEM). L'idée de cette méthode est de faire une approximation de Galerkin du problème étudié, non pas en utilisant des fonctions de forme génériques (comme les fonctions de forme de la méthode des éléments finis), mais en utilisant des fonctions de forme qui sont issues d'un calcul numérique, et qui encodent certaines propriétés du problème étudié. Dans le cas simple du problème déterministe (124), ces fonctions de forme $\phi_i^\varepsilon \in H_0^1(\mathcal{D})$ sont typiquement définies élément par élément. On se donne ainsi une triangulation du domaine \mathcal{D} , et, dans chaque élément \mathbf{K} , on considère le problème

$$-\operatorname{div} \left(A_{per} \left(\frac{x}{\varepsilon} \right) \nabla \phi_{i,\mathbf{K}}^\varepsilon \right) = 0 \text{ dans } \mathbf{K}, \quad \phi_{i,\mathbf{K}}^\varepsilon = \phi_i^0|_{\mathbf{K}} \text{ sur } \partial\mathbf{K}, \quad (175)$$

où les ϕ_i^0 sont les fonctions de forme \mathbb{P}_1 associées à la triangulation de \mathcal{D} . On définit ensuite ϕ_i^ε sur \mathcal{D} par $\phi_i^\varepsilon|_{\mathbf{K}} := \phi_{i,\mathbf{K}}^\varepsilon$ pour tout \mathbf{K} . Les problèmes (175), indexés par \mathbf{K} , sont indépendants les uns des autres et peuvent donc être résolus en parallèle. On réalise ensuite une approximation de Galerkin du problème originel en utilisant ces fonctions de forme ϕ_i^ε , qui sont mieux adaptées au problème que les fonctions \mathbb{P}_1 usuelles, car elles encodent en particulier les oscillations de la matrice A_{per} .

Dans un cadre stochastique, cette méthode conduit à des calculs très coûteux, puisque, pour chaque réalisation de $A(\cdot, \omega)$, il est nécessaire de calculer à nouveau une base de fonctions de forme oscillantes (via la résolution de (175)), qui est ensuite utilisée pour faire une approximation de Galerkin de (133). Dans le travail [FL29, FL30] (cf. aussi [141]), nous étudions l'adaptation de cette méthode à un cadre *faiblement* stochastique, afin de construire une méthode qui, lorsque la perturbation aléatoire est petite, permet d'obtenir une précision identique à la méthode MsFEM originelle, pour un coût de calcul plus faible. L'idée sous-jacente est de définir les fonctions de forme par un problème du type (175), en utilisant la partie *déterministe* de $A(\cdot, \omega)$, et en négligeant donc, à cette étape, la petite perturbation aléatoire.

De façon plus générale, notre souhait est d'étendre l'applicabilité des stratégies jusqu'ici développées, de façon à pouvoir les tester, en partenariat avec les praticiens du domaine, sur des cas plus réalistes.

Références

- [1] A. Abdulle, P. Lin, A. Shapeev, *Homogenization-based analysis of QuasiContinuum method for complex crystals*, arXiv preprint 1006.0378.
- [2] G. Allaire, *Homogenization and two-scale convergence*, SIAM J. Math. Anal., 23(6) :1482–1518, 1992.
- [3] A. Anantharaman, *Analyse mathématique de quelques modèles en calcul de structures électroniques et homogénéisation*, Thèse de l'Université Paris Est, Ecole des Ponts ParisTech, novembre 2010.
- [4] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, G. Scheffer, *Sur les inégalités de Sobolev logarithmiques*, Société Mathématique de France, 2000.
- [5] M. Anitescu, D. Negrut, P. Zapol, A. El-Azab, *A note on the regularity of reduced models obtained by nonlocal quasi-continuum-like approach*, Mathematical Programming, 118(2) :207–236, 2009.
- [6] M. Arndt, M. Griebel, *Derivation of higher order gradient continuum models from atomistic models for crystalline solids*, SIAM J. Multiscale Model. Simul., 4(2) :531–562, 2005.
- [7] M. Arndt, M. Luskin, *Error estimation and atomistic-continuum adaptivity for the quasicontinuum approximation of a Frenkel-Kantorova model*, SIAM J. Multiscale Model. Simul., 7(1) :147–170, 2008.
- [8] V.I. Arnold, *Mathematical methods of classical mechanics*, Graduate Texts in Mathematics, vol. 60, Springer-Verlag, Berlin, 1978.
- [9] L. Baffico, S. Bernard, Y. Maday, G. Turinici, G. Zérah, *Parallel-in-time molecular dynamics simulations*, Phys. Rev. E., 66 :057701, 2002.
- [10] G. Bal, J. Garnier, S. Motsch, V. Perrier, *Random integrals and correctors in homogenization*, Asymptotic Analysis, 59(1-2) :1–26, 2008.
- [11] G. Bal, Y. Maday, *A parareal time discretization for nonlinear PDE's with application to the pricing of an American put*, in *Recent developments in domain decomposition methods*, L.F. Pavarino, A. Toselli eds., Lect. Notes Comput. Sci. Eng., vol. 23, pp. 189–202, Springer Verlag, 2002.
- [12] G. Bal, Q. Wu, *Symplectic parareal*, in *Domain Decomposition Methods in Science and Engineering XVII*, U. Langer, M. Discacciati, D.E. Keyes, O.B. Widlund, W. Zulehner eds., Lect. Notes Comput. Sci. Eng., vol. 60, pp. 401–408, Springer, 2008.
- [13] R. Balian, *From microphysics to macrophysics : methods and applications of statistical physics*, Theoretical and mathematical physics series, Springer, 2007.
- [14] J.M. Ball, *Some open problems in elasticity*, in *Geometry, Mechanics, and Dynamics*, Springer, 2002, pp. 3–59.
- [15] G. Basile, C. Bernardin, S. Olla, *Momentum conserving model with anomalous thermal conductivity in low dimensional systems*, Phys. Rev. Lett., 96 :204303, 2006.
- [16] G. Benettin, A. Giorgilli, *On the Hamiltonian interpolation of near to the identity symplectic mappings with application to symplectic integration algorithms*, J. Stat. Phys., 74 :1117–1143, 1994.
- [17] A. Bensoussan, J.-L. Lions, G. Papanicolaou, *Asymptotic analysis for periodic structures*, Studies in Mathematics and its Applications, vol. 5. North-Holland Publishing Co., Amsterdam-New York, 1978.

- [18] C. Bernardin, S. Olla, *Fourier's law for a microscopic model of heat conduction*, J. Stat. Phys., 121 :271–289, 2005.
- [19] X. Blanc, C. Le Bris, *Définition d'énergies d'interfaces à partir de modèles atomiques [A definition of interface energies using atomistic models]*, C. R. Acad. Sci. Paris, Série I, 340 :535–540, 2005.
- [20] X. Blanc, C. Le Bris, *Improving on computation of homogenized coefficients in the periodic and quasi-periodic settings*, Netw. Heterog. Media, 5(1) :1–29, 2010.
- [21] X. Blanc, C. Le Bris, P.-L. Lions, *Convergence de modèles moléculaires vers des modèles de mécanique des milieux continus [From molecular models to continuum mechanics]*, C. R. Acad. Sci. Paris, Série I, 332 :949–956, 2001.
- [22] X. Blanc, C. Le Bris, P.-L. Lions, *From molecular models to continuum mechanics*, Arch. Rat. Mech. Anal., 164(4) :341–381, 2002.
- [23] X. Blanc, C. Le Bris, P.-L. Lions, *Une variante de la théorie de l'homogénéisation stochastique des opérateurs elliptiques [A variant of stochastic homogenization theory for elliptic operators]*, C. R. Acad. Sci. Paris, Série I, 343 :717–724, 2006.
- [24] X. Blanc, C. Le Bris, P.-L. Lions, *Du discret au continu pour des modèles de réseaux aléatoires d'atomes [Discrete to continuum limit for some models of stochastic lattices of atoms]*, C. R. Acad. Sci. Paris, Série I, 342(8) :627–633, 2006.
- [25] X. Blanc, C. Le Bris, P.-L. Lions, *Stochastic homogenization and random lattices*, Journal de Mathématiques Pures et Appliquées, 88 :34–63, 2007.
- [26] X. Blanc, C. Le Bris, P.-L. Lions, *Atomistic to continuum limits for computational materials science*, Math. Mod. Num. Anal. (M2AN), 41 :391–426, 2007.
- [27] X. Blanc, C. Le Bris, P.-L. Lions, *On the energy of some microscopic stochastic lattices*, Arch. Rat. Mech. Anal., 184 :303–339, 2007.
- [28] S.D. Bond, B.J. Leimkuhler, B.B. Laird, *The Nosé-Poincaré method for constant temperature molecular dynamics*, J. Comput. Phys., 151 :114–134, 1999.
- [29] F. Bonetto, J.L. Lebowitz, J. Lukkarinen, S. Olla, *Heat conduction and entropy production in anharmonic crystals with self-consistent stochastic reservoirs*, J. Stat. Phys., 134 :1097–1119, 2009.
- [30] F. Bonetto, J.L. Lebowitz, L. Rey-Bellet, *Fourier's law : a challenge for theorists*, in *Mathematical Physics 2000*, A. Fokas, A. Grigoryan, T. Kibble, B. Zegarlinsky eds., Imperial College Press, pp. 128-151, 2000.
- [31] F. Bornemann, *Homogenization in time of singularly perturbed mechanical systems*, Lect. Notes Math., vol. 1687, Springer, Berlin, 1998.
- [32] F. Bornemann, Ch. Schütte, *Homogenization of Hamiltonian systems with a strong constraining potential*, Physica D, 102 :57–77, 1997.
- [33] A. Bourgeat, A. Piatnitski, *Estimates in probability of the residual between the random and the homogenized solutions of one-dimensional second-order operator*, Asymptotic Analysis, 21(3-4) :303–315, 1999.
- [34] A. Bourgeat, A. Piatnitski, *Approximation of effective coefficients in stochastic homogenization*, Ann. I. H. Poincaré - PR, 40(2) :153–165, 2004.
- [35] A. Braides, A. Defranceschi, *Homogenization of multiple integrals*, Oxford University Press, 1998.
- [36] P. Carmona, *Existence and uniqueness of an invariant measure for a chain of oscillators in contact with two heat baths*, Stochastic Processes and their Applications, 117(8) :1076–1092, 2007.

- [37] F. Castella, Ph. Chartier, E. Faou, *An averaging technique for highly-oscillatory Hamiltonian problems*, SIAM J. Numer. Anal., 47 :2808–2837, 2009.
- [38] C. Chipot, A. Pohorille eds., *Free energy calculations*, Springer Series in Chemical Physics, vol. 86, Springer, 2007.
- [39] A.J. Chorin, O.H. Hald, R. Kupferman, *Optimal prediction and the Mori-Zwanzig representation of irreversible processes*, Proceedings of the National Academy of Sciences, 97(8) :2968–2973, 2000.
- [40] A.J. Chorin, O.H. Hald, R. Kupferman, *Optimal prediction with memory*, Physica D, 166(3-4) :239–257, 2002.
- [41] Ph.G. Ciarlet, *Mathematical elasticity*, vol. 1, North Holland, 1993.
- [42] G. Ciccotti, T. Lelièvre, E. Vanden-Eijnden, *Projection of diffusions on submanifolds : application to mean force computation*, Comm. Pure and Applied Math., 61(3) :371–408, 2008.
- [43] D. Cioranescu, P. Donato, *An introduction to homogenization*, Oxford Lecture Series in Mathematics and its Applications, vol. 17. The Clarendon Press, Oxford University Press, New York, 1999.
- [44] D. Cohen, T. Jahnke, K. Lorenz, Ch. Lubich, *Numerical integrators for highly oscillatory Hamiltonian systems : a review*, in *Analysis, Modeling and Simulation of Multiscale Problems*, Mathematics and Statistics, A. Mielke ed., pp. 553–576, Springer, 2006.
- [45] R. Costaouec, Thèse de l’Université Paris Est, Ecole des Ponts ParisTech, en préparation.
- [46] R. Costaouec, *Asymptotic expansion of the homogenized matrix in two weakly stochastic homogenization settings*, Applied Mathematics Research Express, à paraître, <http://amrx.oxfordjournals.org/content/early/2011/10/28/amrx.abr011.short?rss=1>
- [47] S. Curtarolo, G. Ceder, *Dynamics of an inhomogeneously coarse grained multiscale system*, Phys. Rev. Lett., 88(25) :255504, 2002.
- [48] E. Darve, J. Solomon, A. Kia, *Computing generalized Langevin equations and generalized Fokker-Planck equations*, Proceedings of the National Academy of Sciences, 106(27) :10884–10889, 2009.
- [49] R.L. Davidchack, *Discretization errors in molecular dynamics simulations with deterministic and stochastic thermostats*, J. Comput. Phys., 229(24) :9323–9346, 2010.
- [50] P. Deák, Th. Frauenheim, M.R. Pederson eds., *Computer simulation of materials at atomic level*, Wiley, 2000.
- [51] A. Dembo, O. Zeitouni, *Large deviations techniques and applications*, Jones and Bartlett Publishers, Boston, MA, 1993.
- [52] J.-D. Deuschel, G. Giacomin, D. Ioffe, *Large deviations and concentration properties for $\nabla\varphi$ interface models*, Probab. Theory Relat. Fields, 117 :49–111, 2000.
- [53] D. Dizdar, *Towards an optimal rate of convergence in the hydrodynamic limit for Kawasaki dynamics*, Ph.D. thesis, Bonn University, 2007.
- [54] M. Dobson, M. Luskin, *Analysis of a force-based quasicontinuum approximation*, Math. Model. Numer. Anal. (M2AN), 42(1) :113–139, 2008.
- [55] M. Dobson, M. Luskin, *An analysis of the effect of ghost force oscillation on quasicontinuum error*, Math. Model. Numer. Anal. (M2AN), 43(3) :591–604, 2009.
- [56] M. Dobson, M. Luskin, *An optimal order error analysis of the one-dimensional quasicontinuum approximation*, SIAM J. Numer. Anal., 47(4) :2455–2475, 2009.

- [57] M. Dobson, M. Luskin, C. Ortner, *Stability, instability, and error of the force-based quasicontinuum approximation*, Arch. Rat. Mech. Anal., 197(1) :179–202, 2010.
- [58] M. Dobson, M. Luskin, C. Ortner, *Accuracy of quasicontinuum approximations near instabilities*, J. Mech. Phys. Solids., 58(10) :1741–1757, 2010.
- [59] N. Dunford, J.T. Schwartz, *Linear operators. Volume 2 : Spectral theory : self adjoint operators in Hilbert space*, Wiley, New York, 1963.
- [60] L.M. Dupuy, E.B. Tadmor, R.E. Miller, R. Phillips, *Finite temperature Quasicontinuum : molecular dynamics without all the atoms*, Phys. Rev. Lett., 95 :060202, 2005.
- [61] W. E, P.B. Ming, *Analysis of multiscale methods*, J. Comp. Math., 22(2) :210–219, 2004.
- [62] W. E, P.B. Ming, *Cauchy-Born rule and stability of crystals : static problems*, Arch. Rat. Mech. Anal., 183(2) :241–297, 2007.
- [63] R.S. Ellis, *Entropy, large deviations, and statistical mechanics*, volume 271 of Grundlehren der Mathematischen Wissenschaften. Springer-Verlag, New York, 1985.
- [64] R.S. Ellis, *Large deviations and statistical mechanics*, in *Particle systems, random media and large deviations*, (Brunswick, Maine, 1984), volume 41 of Contemp. Math., pp. 101–123. Amer. Math. Soc., Providence, RI, 1985.
- [65] R.S. Ellis, *An overview of the theory of large deviations and applications to statistical mechanics*, Scand. Actuar. J., 1 :97–142, 1995. Harald Cramer Symposium (Stockholm, 1993).
- [66] B. Engquist, R. Tsai, *Heterogeneous multiscale methods for stiff ordinary differential equations*, Math. Comput., 74 :1707–1742, 2005.
- [67] L.C. Evans, R.F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, 1992, Studies in Advanced Mathematics.
- [68] E. Faou, E. Hairer, T.-L. Pham, *Energy conservation with non-symplectic methods : examples and counter-examples*, BIT, 44 :699–709, 2004.
- [69] K. Feng, *Difference schemes for Hamiltonian formalism and symplectic geometry*, J. Comp. Math., 4 :279–289, 1986.
- [70] M.C. Fivel, C.F. Robertson, G.R. Canova, L. Boulanger, *Three-dimensional modeling of indent-induced plastic zone at a mesoscale*, Acta. Mater., 46 :6183–6194, 1998.
- [71] I. Fonseca, S. Krömer, *Multiple integrals under differential constraints : two-scale convergence and homogenization*, Indiana Univ. Math. J., 59(2) :427–457, 2010.
- [72] T. Funaki, *Stochastic interface models*, in *Lectures on Probability Theory and Statistics, Ecole d’été de probabilités de Saint-Flour XXXIII - 2003*, J. Picard ed., Lect. Notes Math., vol. 1869, pp. 103–274, Springer, 2005.
- [73] T. Funaki, H. Spohn, *Motion by mean curvature from the Ginzburg-Landau $\nabla\phi$ interface model*, Comm. Math. Phys., 185 :1–36, 1997.
- [74] G. Giacomin, S. Olla, H. Spohn, *Equilibrium fluctuations for $\nabla\phi$ interface model*, The Annals of Probability, 29(3) :1138–1172, 2001.
- [75] D. Givon, R. Kupferman, A.M. Stuart, *Extracting macroscopic dynamics : model problems and algorithms*, Nonlinearity, 17(6) :55–127, 2004.
- [76] A. Gloria, *A direct approach to numerical homogenization in finite elasticity*, Netw. Heterog. Media, 1(1) :109–141, 2006.

- [77] A. Gloria, *Reduction of the resonance error. Part 1 : Approximation of homogenized coefficients*, Mathematical Models and Methods in Applied Sciences, 21(8) :1601–1630, 2011.
- [78] A. Gloria, F. Otto, *An optimal error estimate in stochastic homogenization of discrete elliptic equations*, preprint.
<http://hal.archives-ouvertes.fr/inria-00457020/en/>
- [79] V. Grimm, M. Hochbruck, *Error analysis of exponential integrators for oscillatory second-order differential equations*, J. Phys. A, 39 :5495–5507, 2006.
- [80] H. Grubmüller, H. Heller, A. Windemuth, K. Schulten, *Generalized Verlet algorithm for efficient molecular dynamics simulations with long range interaction*, Mol. Sim., 6 :121–142, 1991.
- [81] N. Grunewald, F. Otto, C. Villani, M.G. Westdickenberg, *A two-scale approach to logarithmic Sobolev inequalities and the hydrodynamic limit*, Ann. Inst. H. Poincaré Probab. Statist., 45(2) :302–351, 2009.
- [82] I. Gyöngy, *Mimicking the one-dimensional marginal distributions of processes having an Itô differential*, Probab. Theory Relat. Fields, 71 :501–516, 1986.
- [83] E. Hairer, Ch. Lubich, *The life-span of backward error analysis for numerical integrators*, Numer. Math., 76 :441–462, 1997
(erratum disponible à <http://www.unige.ch/math/folks/hairer>).
- [84] E. Hairer, Ch. Lubich, G. Wanner, *Geometric Numerical Integration, Structure-Preserving Algorithms For Ordinary Differential Equations*, Springer Series in Computational Mathematics, vol. 31, 2nd edition. Springer-Verlag, Berlin, 2006.
- [85] P. Hänggi, P. Talkner, M. Borkovec, *Reaction-rate theory : fifty years after Kramers*, Reviews of Modern Physics, 62(2) :251–342, 1990.
- [86] C. Hartmann, *Model reduction in classical molecular dynamics*, Ph.D. thesis, Freie Universität Berlin, 2007.
- [87] W.K. Hastings, *Monte Carlo sampling methods using Markov chains and their applications*, Biometrika, 57(1) :97–109, 1970.
- [88] T. Hatano, *Heat conduction in the diatomic Toda lattice revisited*, Phys. Rev. E, 59(1) :R1–R4, 1999.
- [89] F. den Hollander, *Large deviations*, Fields Institute Monographs, vol. 14. American Mathematical Society, Providence, RI, 2000.
- [90] W.G. Hoover, *Canonical dynamics : Equilibrium phase-space distributions*, Phys. Rev. A, 31(3) :1695–1697, 1985.
- [91] I. Horenko, C. Hartmann, C. Schütte, F. Noe, *Data-based parameter estimation of generalized multidimensional Langevin processes*, Phys. Rev. E, 76 :016706, 2007.
- [92] J.A. Izaguirre, S. Reich, R.D. Skeel, *Longer time steps for Molecular Dynamics*, J. Chem. Phys., 110 :9853–9864, 1999.
- [93] T. Jahnke, *Long-time-step integrators for almost-adiabatic quantum dynamics*, SIAM J. Sci. Comput., 25 :2145–2164, 2004.
- [94] T. Jahnke, Ch. Lubich, *Numerical integrators for quantum dynamics close to the adiabatic limit*, Numer. Math., 94 :289–314, 2003.
- [95] Z. Jia, B. Leimkuhler, *Geometric integrators for multiple time-scale simulation*, J. Phys. A, 39 :5379–5403, 2006.
- [96] V.V. Jikov, S.M. Kozlov, O.A. Oleinik, *Homogenization of differential operators and integral functionals*, Springer-Verlag, 1994.

- [97] T. Kanit, S. Forest, I. Galliet, V. Mounoury, D. Jeulin, *Determination of the size of the representative volume element for random composites : statistical and numerical approach*, Int. J. Solids and Structures, 40 :3647–3679, 2003.
- [98] A. Khinchin, *Mathematical Foundations of Statistical Mechanics*, Dover, New York, 1949.
- [99] H.A. Kramers, *Brownian motion in a field of force and the diffusion model of chemical reactions*, Physica, 7(4) :284–304, 1940.
- [100] U. Krengel, *Ergodic theorems*, de Gruyter Studies in Mathematics, vol. 6, de Gruyter, 1985.
- [101] C. Le Bris, *Some numerical approaches for “weakly” random homogenization*, Numerical mathematics and advanced applications, Proceedings of ENUMATH 2009, G. Kreiss, P. Lötstedt, A. Malqvist, M. Neytcheva eds., Lect. Notes Comput. Sci. Eng., pp. 29–45, Springer, 2010.
- [102] C. Le Bris, F. Legoll, *Integrators for highly oscillatory Hamiltonian systems : an homogenization approach [Schémas numériques pour systèmes hamiltoniens hautement oscillants : une approche par l’homogénéisation]*, Rapport de recherche INRIA RR-6252, 2007. <http://hal.inria.fr/inria-00165293/en>
- [103] B.J. Leimkuhler, S. Reich, *A reversible averaging integrator for multiple time-scale dynamics*, J. Comput. Phys., 171 :95–114, 2001.
- [104] T. Lelièvre, *A general two-scale criteria for logarithmic Sobolev inequalities*, J. Funct. Anal., 256(7) :2211–2221, 2009.
- [105] T. Lelièvre, M. Rousset, G. Stoltz, *Free energy computations : A mathematical perspective*, Imperial College Press, 2010.
- [106] S. Lepri, R. Livi, A. Politi, *Thermal conduction in classical low-dimensional lattices*, Phys. Rep., 377 :1–80, 2003.
- [107] R. LeSar, R. Najafabadi, D.J. Srolovitz, *Finite-temperature defect properties from free-energy minimization*, Phys. Rev. Lett., 63 :624–627, 1989.
- [108] P. Lin, *Theoretical and numerical analysis of the quasi-continuum approximation of a material particle model*, Math. Comput., 72 :657–675, 2003.
- [109] P. Lin, *Convergence analysis of a quasi-continuum approximation for a two-dimensional material*, SIAM J. Numer. Anal., 45(1) :313–332, 2007.
- [110] J.-L. Lions, Y. Maday, G. Turinici, *A parareal in time discretization of PDE’s*, C. R. Acad. Sci. Paris, Série I, 332 :661–668, 2001.
- [111] P.-L. Lions, G. Toscani, *A strengthened central limit theorem for smooth densities*, J. Funct. Anal., 128 :148–176, 1995.
- [112] J.S. Liu, *Monte-Carlo strategies in scientific computing*, Springer Series in Statistics, 2001.
- [113] G.J. Martyna, M.L. Klein, M.E. Tuckerman, *Nosé-Hoover chains : The canonical ensemble via continuous dynamics*, J. Chem. Phys., 97(4) :2635–2643, 1992.
- [114] J.C. Mattingly, A.M. Stuart, D.J. Higham, *Ergodicity for SDEs and approximations : Locally Lipschitz vector fields and degenerate noise*, Stochastic Processes and their Applications, 101 :185–232, 2002.
- [115] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, *Equations of state calculations by fast computing machines*, J. Chem. Phys., 21 :1087–1091, 1953.
- [116] S.P. Meyn, R.L. Tweedie, *Markov Chains and Stochastic Stability*, Springer, 1993.

- [117] F. Murat, *Compacité par compensation*, Ann. Scuola Norm. Sup. Pisa. Cl. Sci., 5(4) :485–507, 1978.
- [118] A. Naddaf, T. Spencer, *On homogenization and scaling limit of some gradient perturbations of a massless free field*, Comm. Math. Phys., 183 :55–84, 1997.
- [119] A. Naddaf, T. Spencer, *Estimates on the variance of some homogenization problems*, preprint.
<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.35.1978>.
- [120] O. Narayan, S. Ramaswamy, *Anomalous heat conduction in one-dimensional momentum-conserving systems*, Phys. Rev. Lett., 89 :200601, 2002.
- [121] G. Nguetseng, *A general convergence result for a functional related to the theory of homogenization*, SIAM J. Math. Anal., 20(3) :608–623, 1989.
- [122] S. Nosé, *A Molecular Dynamics method for simulations in the canonical ensemble*, Mol. Phys., 52 :255–268, 1984.
- [123] S. Nosé, *A unified formulation of the constant temperature Molecular Dynamics method*, J. Chem. Phys., 81 :511–519, 1985.
- [124] C. Ortner, E. Süli, *Analysis of a quasicontinuum method in one dimension*, Math. Model. Numer. Anal. (M2AN), 42(1) :57–91, 2008.
- [125] F. Otto, M.G. Reznikoff, *A new criterion for the logarithmic Sobolev inequality and two applications*, J. Func. Anal., 243 :121–157, 2007.
- [126] G. Papanicolaou, S.R.S. Varadhan, *Boundary value problems with rapidly oscillating random coefficients*, in *Proc. Colloq. on Random Fields : Rigorous Results in Statistical Mechanics and Quantum Field Theory*, J. Fritz, J.L. Lebaritz, D. Szasz eds., vol. 10 of Colloquia Mathematica Societ. Janos Bolyai, pp. 835–873, 1979.
- [127] E. Presutti, *Scaling limits in statistical mechanics and microstructures in continuum mechanics*, Springer-Verlag, 2008.
- [128] S. Reich, *Backward error analysis for numerical integrators*, SIAM J. Numer. Anal., 36 :1549–1570, 1999.
- [129] L. Rey-Bellet, *Open classical systems*, in *Quantum Open Systems II. The Markovian approach*, Lect. Notes Math., vol. 1881, pp. 41–78, Springer, 2006.
- [130] Z. Rieder, J.L. Lebowitz, E. Lieb, *Properties of a harmonic crystal in a stationary nonequilibrium state*, J. Math. Phys., 8(5) :1073–1078, 1967.
- [131] G.O. Roberts, R.L. Tweedie, *Exponential convergence of Langevin diffusions and their discrete approximations*, Bernoulli, 2 :341–364, 1996.
- [132] J.M. Sanz-Serna, M.P. Calvo, *Numerical Hamiltonian problems*, Chapman & Hall, 1994.
- [133] R. Sharp, R. Tsai, B. Engquist, *Multiple time scale numerical methods for the inverted pendulum problem*, in *Lect. Notes Comput. Sci. Eng.*, vol. 44, pp. 241–261, Springer, 2005.
- [134] V.B. Shenoy, R. Miller, E.B. Tadmor, D. Rodney, R. Phillips, M. Ortiz, *An adaptive finite element approach to atomic-scale mechanics - the quasicontinuum method*, J. Mech. Phys. Solids, 47 :611–642, 1999.
- [135] A.N. Shiryaev, *Probability*, Graduate Texts in Mathematics, vol. 95, Springer, 1984.
- [136] C. Siegel, J. Moser, *Lecture on Celestial Mechanics*, Springer, New York, 1971.
- [137] E.B. Tadmor, M. Ortiz, R. Phillips, *Quasicontinuum analysis of defects in solids*, Phil. Mag. A, 73 :1529–1563, 1996.

- [138] E.B. Tadmor, R. Phillips, *Mixed atomistic and continuum models of deformation in solids*, Langmuir, 12 :4529–4534, 1996.
- [139] D. Talay, L. Tubaro, *Expansion of the global error for numerical schemes solving stochastic differential equations*, Stochastic Analysis and Applications, 8(4) :94–120, 1990.
- [140] L. Tartar, *Compensated compactness and applications to partial differential equations*, Nonlinear analysis and mechanics : Heriot-Watt Symposium, vol. IV, pp. 136–212, Res. Notes in Math., vol. 39, Pitman, Boston, Mass.-London, 1979.
- [141] F. Thomines, Thèse de l'Université Paris Est, Ecole des Ponts ParisTech, en préparation.
- [142] M. Toda, *Solitons and heat conduction*, Physica Scripta, 20 :424–430, 1979.
- [143] M.E. Tuckermann, B.J. Berne, G.J. Martyna, *Reversible multiple time scale molecular dynamics*, J. Chem. Phys., 97 :1990–2001, 1992.
- [144] M.E. Tuckerman, G. Martyna, *Understanding modern molecular dynamics : Techniques and applications*, J. Phys. Chem. B, 104 :159–178, 2000.
- [145] S.R.S. Varadhan, *Large deviations and applications*, SIAM, Philadelphia, 1984.
- [146] V.V. Yurinskii, *Averaging of symmetric diffusion in random medium*, Sibirskii Mat. Zh., 27(4) :167–180, 1986.
- [147] X. Zotos, *Ballistic transport in classical and quantum integrable systems*, Journal of low temperature physics, 126(3-4) :1185–1194, 2002.