

Simulation moléculaire : modèles, calcul d'énergie libre et aspects dynamiques

1- Introduction

Tony Lelièvre

Ecole des Ponts ParisTech et INRIA



Quiberon, Septembre 2020

Introduction



Mon activité scientifique

Mon activité de recherche :

- 1/3 : analyse mathématique de modèles issus de la physique
- 1/3 : analyse mathématique d'algorithmes utilisés en physique computationnelle
- 1/3 : développement de nouveaux algorithmes, implémentation dans des codes et tests (avec des collègues physiciens ou des ingénieurs en entreprise)

Applications: mécanique des fluides, physique statistique.

Outils mathématiques: analyse des équations aux dérivées partielles et analyse stochastique.

Algorithmes: éléments finis pour des problèmes à surface libre, algorithmes stochastiques pour l'échantillonnage (événements rares, échantillonnage sur des sous-variétés, accélération de dynamique), réduction de dimension (tensor method, reduced basis, auto-encoders)

Bibliographie: DM

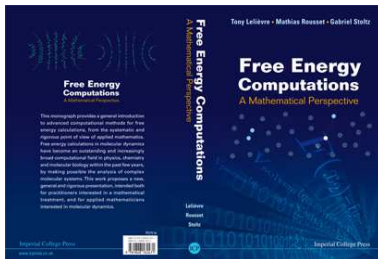
Ouvrages de référence en dynamique moléculaire:

- M. Allen and D. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, 1989.
- D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, Oxford University Press, 2001.
- E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, Springer-Verlag, 2006.
- B. Leimkuhler, and S. Reich, *Simulating Hamiltonian Dynamics*, Cambridge University Press, 2005.
- D. Rapaport, *The Art of Molecular Dynamics Simulation*, Cambridge University Press, 2004.
- M. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation*, Oxford University Press, 2010.

Bibliographie: ce cours

Quelques références sur ce que je vais raconter:

- TL and G. Stoltz, *Partial differential equations and stochastic methods in molecular dynamics*, Acta Numerica, **25**, 681-880, 2016.
- TL, *Mathematical foundations of Accelerated Molecular Dynamics methods*, Handbook of Materials Modeling Vol.1, W. Andreoni S. Yip Ed., 2018.
- TL, M. Rousset and G. Stoltz, *Free energy computations: a mathematical perspective*, Imperial College Press, 2010.

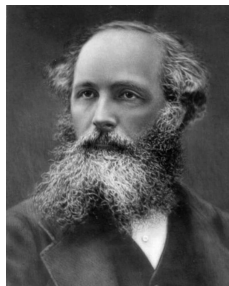


Physique statistique computationnelle...

...le fruit de deux révolutions

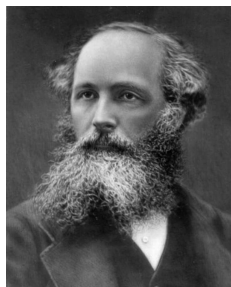
Une première révolution

Physique statistique et vision atomiste de la matière



Une première révolution

Physique statistique et vision atomiste de la matière



Ludwig Boltzmann (1844-1906) James Clerk Maxwell (1831-1879)

- A une échelle microscopique, la matière est constituée d'*atomes* qui interagissent.
- De cette interaction naît, à notre échelle macroscopique, les propriétés de la matière.

... qui pose des problèmes mathématiques toujours d'actualité

Comment expliquer le passage d'une échelle microscopique à notre échelle macroscopique ?

Problème 1: Des équations de Newton, aux équations de Boltzmann:

$$\left\{ \begin{array}{l} \frac{dx(t)}{dt} = v(t) \\ M \frac{dv(t)}{dt} = -\nabla V(x(t)) \end{array} \right. \longrightarrow \partial_t f(t, x, v) + v \cdot \nabla_x f(t, x, v) = Q(f, f)$$

O.E. Lanford (1975),

I. Gallagher, L. Saint-Raymond et B. Texier (2013)

... qui pose des problèmes mathématiques encore d'actualité

Problème 2:

Equations de Boltzmann, régularité, irréversibilité:

$$\frac{d}{dt} \int f \ln f \, dx \, dv \leq 0$$

L. Desvillettes et C. Villani (2005), ...



Problème 3:

Des équations de Boltzmann aux équations de la mécanique des fluides

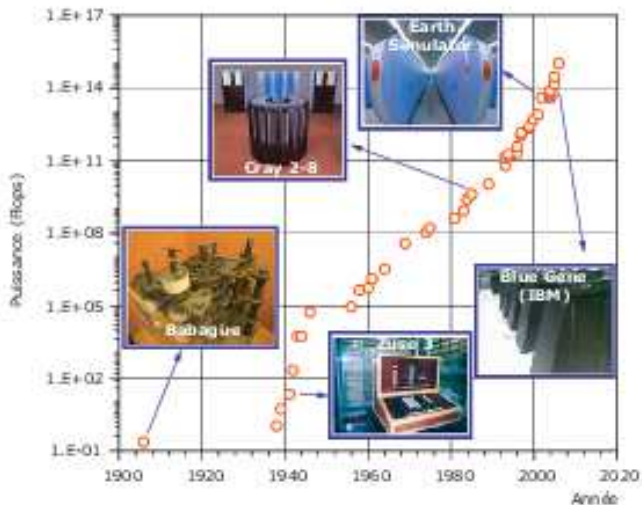
$$\partial_t f(t, x, v) + v \cdot \nabla_x f(t, x, v) = Q(f, f)$$

$$\longrightarrow \begin{cases} \partial_t u + u \cdot \nabla u - \nu \Delta u + \nabla p = 0 \\ \operatorname{div} u = 0 \end{cases}$$

F. Golse et L. Saint-Raymond (2004), ...



Une deuxième révolution ...



La puissance des ordinateurs a été multipliée par 10^{15} au cours des 70 dernières années !

... qui permet de simuler sur ordinateur la matière à l'échelle microscopique

- **1950**: Apparition de la dynamique moléculaire et des méthodes de Monte Carlo.
- **Un microscope numérique**: Explorer la matière à l'échelle atomique ; Comprendre le lien entre les échelles microscopiques et les propriétés macroscopiques ; Simuler de nouveaux matériaux, de nouvelles molécules ; Interpréter les résultats d'expériences.
- **Applications**: biologie, chimie, science des matériaux.

... qui permet de simuler sur ordinateur la matière à l'échelle microscopique

Une science aujourd'hui reconnue:

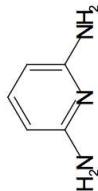
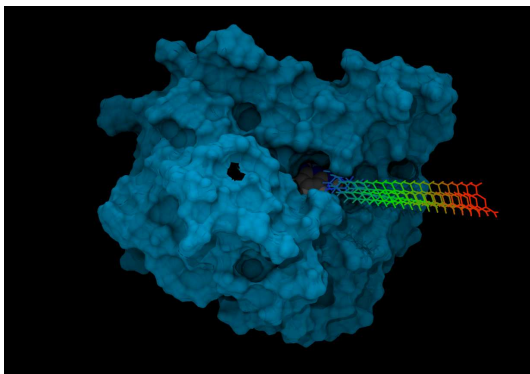
- Centre Européen de Calcul Atomique et Moléculaire :
"Nowadays, computer simulations are considered the third way of doing science, beyond theory and experiment."
- Prix Nobel de chimie 2013: Arieh Warshel, Martin Karplus et Michael Levitt. *"Aujourd'hui l'ordinateur est un outil tout aussi important pour les chimistes que l'éprouvette."*
- 1/4 du temps CPU dans le monde est consacré à des calculs à l'échelle moléculaire.

Quelques exemples

Des applications en biologie

- Quelles sont les conformations typiques d'une protéine ?
- Comment la protéine passe-t'elle d'une conformation à une autre ?
- Comment calculer l'affinité d'un ligand pour un site cible ?
- Quels sont les mécanismes moléculaires qui régulent la cellule (canaux ioniques, réplication de l'ADN, etc...) ?
- ...

Des applications en biologie



Détachement d'un ligand d'une protéine (Diaminopyridine-HSP90, Courtesy of SANOFI)

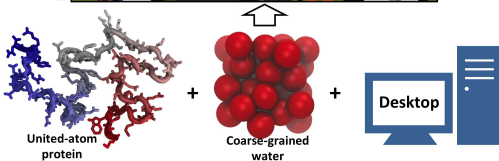
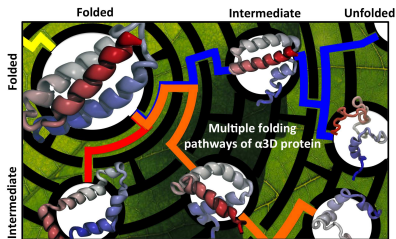
Nombre d'atomes: 40 000

Pas de temps de la simulation: 10^{-15} s

Temps de dissociation: 0.5 s

Des applications en biologie

Sera-t-on un jour capable de déterminer la structure tri-dimensionnelle d'une protéine à partir de sa séquence d'acides aminés ?

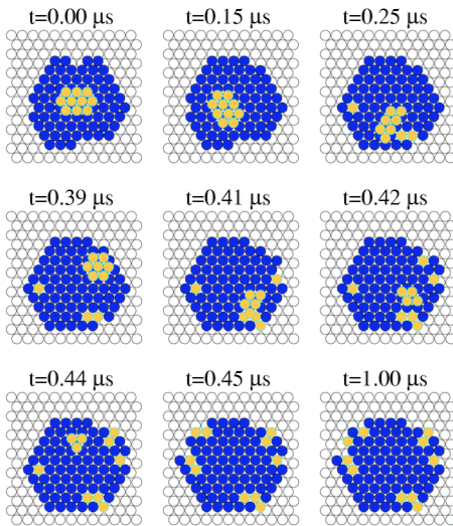


Everybody can fold protein (Courtesy of K. Schulten group) \longrightarrow une simulation...

Des applications en science des matériaux

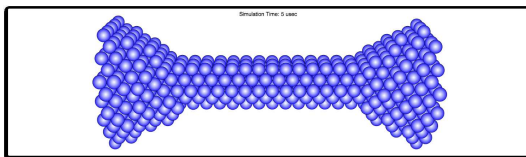
- Comment se propage les défauts dans les matériaux (dislocation, matériaux sous irradiation, ...) ?
- Quels sont les propriétés macroscopiques d'un nouveau polymère ?
- Comment améliorer les propriétés d'un catalyseur ?
- ...

Des applications en science des matériaux



Diffusion d'*adatoms* sur une surface (Courtesy of A. Voter, Los Alamos National Laboratory)

Des applications en science des matériaux



Expérience d'extension sur un nanofil d'argent à température $T=300\text{K}$. Calculs effectués au Los Alamos National Laboratory sur 120 000 processeurs. —→ une simulation...

(Courtesy of A. Voter, Los Alamos National Laboratory)

Des applications en science des matériaux

Les propriétés des nanoclusters métalliques sont très sensibles à la forme et la taille. Les nanoclusters oscillent entre différentes conformations.

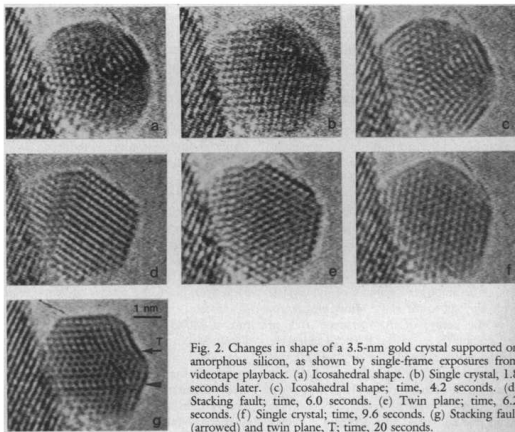
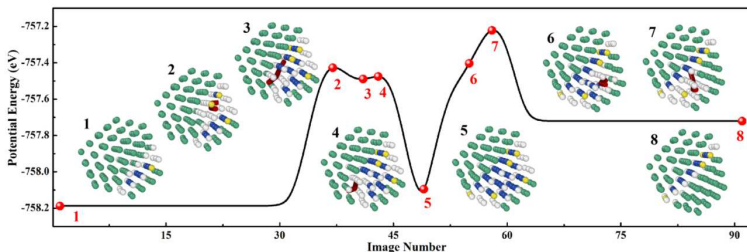


Fig. 2. Changes in shape of a 3.5-nm gold crystal supported on amorphous silicon, as shown by single-frame exposures from videotape playback. (a) Icosahedral shape. (b) Single crystal, 1.8 seconds later. (c) Icosahedral shape; time, 4.2 seconds. (d) Stacking fault; time, 6.0 seconds. (e) Twin plane; time, 6.2 seconds. (f) Single crystal; time, 9.6 seconds. (g) Stacking fault (arrowed) and twin plane, T; time, 20 seconds.

Smith *et al.*, Science, vol. 233, 1986

Des applications en science des matériaux

En utilisant des [méthodes de dynamique accélérée](#), D. Perez, A.F. Voter et collaborateurs peuvent atteindre $100\mu\text{s}$.



FCC to twin: Faulting and twinning occurs through repeated nucleation and propagation of $1/6\langle 211 \rangle$ partials.

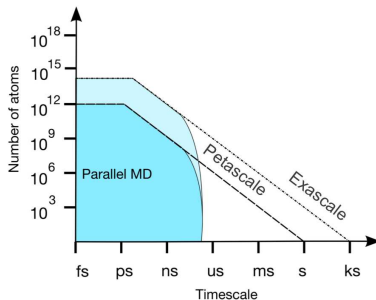
R. Huang *et al.*, J Chem. Phys., vol 147, 2017

Les principaux défis

Les principaux défis en physique statistique computationnelle :

- Améliorer les modèles (champs de force, modèles gros-grains): polarisabilité, modèles d'eau, réactions chimiques.
- Améliorer les méthodes d'échantillonnage, atteindre des temps macroscopiques (propriétés thermodynamiques et dynamiques).
- Interaction données expérimentales et simulations : approches bayésiennes, apprentissage machine.

On sait paralléliser efficacement les calculs en espace, mais il reste une barrière temporelle: les calculs directs de DM sur des systèmes réalistes atteignent au maximum $1 \mu\text{s}$.



Courtesy of Danny Perez (LANL)

Les principaux défis

A quoi peuvent servir les développements abstraits, et les mathématiques en particulier ?

- Faire un lien propre entre les modèles à différentes échelles (espace et temps): coarse-graining, couplage, certification.
- Développer de nouveaux algorithmes (efficacité, robustesse, analyse d'erreur)...
- ...en particulier sur des architectures parallèles (parallélisation en espace et en temps).
- Méthodes d'assimilation de données modernes (apprentissage machine).

... mais faible implication de la communauté: nombre d'articles sur MathSciNet: fluid 95543, Navier Stokes 24026, Molecular dynamics 3166, ... ???

Modèle et simulation :

Les briques fondamentales d'une simulation de dynamique moléculaire

Le principal ingrédient de modélisation...

L'essentiel de la modélisation est contenue dans une **fonction potentielle** V qui à une configuration $(\mathbf{q}_1, \dots, \mathbf{q}_N) = \mathbf{q} \in \mathbb{R}^{3N}$ associe une énergie $V(\mathbf{q}_1, \dots, \mathbf{q}_N)$. La fonction $-\nabla V$ s'appelle le **champ de force**.

Idéalement, V est obtenu par un calcul de structure électronique (dynamique de Born-Oppenheimer ou Ab Initio Molecular Dynamics) \rightarrow possible seulement pour certains systèmes.

Dans l'immense majorité des cas, on utilise un **champ de force empirique** dont on veut qu'il soit rapide et précis.

Exemples: Stillinger-Weber (C, Si), Tersoff (C, Si), modified EAM, Bond-order, Buckingham (oxides), ReaxFF (Charge Transfer), TIP3P ou TIP4P (eau), etc.

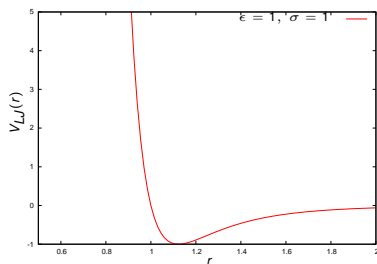
Le principal ingrédient de modélisation...

Pour un champ de force empirique, typiquement, V est la somme de potentiels modélisant l'interaction entre deux, trois ou quatre particules:

$$V = \sum_{i < j} V_1(\mathbf{q}_i, \mathbf{q}_j) + \sum_{i < j < k} V_2(\mathbf{q}_i, \mathbf{q}_j, \mathbf{q}_k) + \sum_{i < j < k < l} V_3(\mathbf{q}_i, \mathbf{q}_j, \mathbf{q}_k, \mathbf{q}_l).$$

Par exemple,

$V_1(\mathbf{q}_i, \mathbf{q}_j) = V_{LJ}(|\mathbf{q}_i - \mathbf{q}_j|)$ où
 $V_{LJ}(r) = 4\epsilon \left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right)$ est
 le potentiel de Lennard-Jones.



En pratique, choisir le bon potentiel empirique pour un problème donné relève plus de l'art que de la science.

Dans toute la suite, on suppose que le champ de force est donné.

Les équations du mouvement: à énergie fixée

Pour un système isolé, on utilise les équations de Newton:

$$\begin{cases} d\mathbf{Q}(t) = M^{-1} \mathbf{P}(t) dt \\ d\mathbf{P}(t) = -\nabla V(\mathbf{Q}(t)) dt \end{cases}$$

où M est le tenseur de masse, $(\mathbf{Q}(t), \mathbf{P}(t)) \in \mathbb{R}^{3N} \times \mathbb{R}^{3N}$.

Propriété fondamentale: l'énergie est conservée:

$$\frac{d}{dt} H(\mathbf{Q}(t), \mathbf{P}(t)) = 0$$

où

$$H(\mathbf{q}, \mathbf{p}) = \frac{\mathbf{p}^T M^{-1} \mathbf{p}}{2} + V(\mathbf{q})$$

est le hamiltonien du système.

Les équations ont une structure particulière (symplectique)

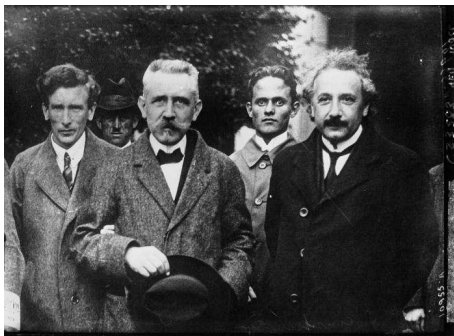
$$\begin{cases} d\mathbf{Q}(t) = \nabla_{\mathbf{p}} H(\mathbf{Q}(t), \mathbf{P}(t)) dt \\ d\mathbf{P}(t) = -\nabla_{\mathbf{q}} H(\mathbf{Q}(t), \mathbf{P}(t)) dt \end{cases}$$

Les équations du mouvement: à température fixée

Pour un système plongé dans un environnement qui fixe la température, il faut ajouter un thermostat aux équations de Newton. Ce sont les équations de Langevin:

$$\begin{cases} d\mathbf{Q}(t) = M^{-1}\mathbf{P}(t) dt \\ d\mathbf{P}(t) = -\nabla V(\mathbf{Q}(t)) dt - \gamma M^{-1}\mathbf{P}(t) dt + \sqrt{2\gamma\beta^{-1}}d\mathbf{W}(t) \end{cases}$$

avec $\gamma > 0$ le paramètre de friction et $\beta^{-1} = k_B T$.



Les équations du mouvement: température fixée

$$\begin{cases} d\mathbf{Q}(t) = M^{-1}\mathbf{P}(t) dt, \\ d\mathbf{P}(t) = -\nabla V(\mathbf{Q}(t)) dt - \gamma M^{-1}\mathbf{P}(t) dt + \sqrt{2\gamma\beta^{-1}} d\mathbf{W}(t), \end{cases}$$

Pour un potentiel régulier et confinant, cette dynamique est ergodique par rapport à la mesure de Boltzmann-Gibbs:

$$Z^{-1} \exp(-\beta H(\mathbf{q}, \mathbf{p})) d\mathbf{q} d\mathbf{p}$$

où $Z = \iint_{\mathbb{R}^{6N}} \exp(-\beta H(\mathbf{q}, \mathbf{p})) d\mathbf{q} d\mathbf{p} < \infty$.

On a donc:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \frac{\mathbf{P}(t)^T M^{-1} \mathbf{P}(t)}{2} dt = \frac{3N}{2} k_B T$$

et, pour toute fonction $\varphi : \mathbb{R}^{3N} \rightarrow \mathbb{R}$,

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \varphi(\mathbf{Q}(t)) dt = \int_{\mathbb{R}^{3N}} \varphi(\mathbf{q}) Z^{-1} \exp(-\beta V(\mathbf{q})) d\mathbf{q}.$$

Les équations du mouvement: discrétisation

Pour simuler une trajectoire, on choisit un pas de temps $h > 0$, et on discrétise les équations du mouvement aux temps $t_n = n\Delta t$ ($n \in \mathbb{N}$).

Exemple 1: Equations de Newton

$$\begin{cases} d\mathbf{Q}(t) = M^{-1} \mathbf{P}(t) dt \\ d\mathbf{P}(t) = -\nabla V(\mathbf{Q}(t)) dt \end{cases} \longrightarrow \begin{cases} \mathbf{Q}(t_{n+1}) - \mathbf{Q}(t_n) = \int_{t_n}^{t_{n+1}} M^{-1} \mathbf{P}(t) dt \\ \mathbf{P}(t_{n+1}) - \mathbf{P}(t_n) = - \int_{t_n}^{t_{n+1}} \nabla V(\mathbf{Q}(t)) dt \end{cases}$$

... discrétisation (schéma d'Euler explicite) ...

$$\begin{cases} \mathbf{Q}^{n+1} - \mathbf{Q}^n = M^{-1} \mathbf{P}^n \Delta t \\ \mathbf{P}^{n+1} - \mathbf{P}^n = -\nabla V(\mathbf{Q}^n) \Delta t \end{cases}$$

avec bien sûr $(\mathbf{Q}^0, \mathbf{P}^0) = (\mathbf{Q}(0), \mathbf{P}(0))$.

On espère que $(\mathbf{Q}^n, \mathbf{P}^n)$ est une bonne approximation de $(\mathbf{Q}(t_n), \mathbf{P}(t_n))$.

Les équations du mouvement: discrétisation

Exemple 2: Equations de Langevin

$$\begin{cases} d\mathbf{Q}(t) = M^{-1} \mathbf{P}(t) dt \\ d\mathbf{P}(t) = -\nabla V(\mathbf{Q}(t)) dt - \gamma M^{-1} \mathbf{P}(t) dt + \sqrt{2\gamma\beta^{-1}} d\mathbf{W}(t) \end{cases}$$

... discrétisation (schéma d'Euler Maruyama) ...

$$\begin{cases} \mathbf{Q}^{n+1} - \mathbf{Q}^n = M^{-1} \mathbf{P}^n \Delta t \\ \mathbf{P}^{n+1} - \mathbf{P}^n = -\nabla V(\mathbf{Q}^n) \Delta t - \gamma M^{-1} \mathbf{P}^n \Delta t + \sqrt{2\gamma\beta^{-1} \Delta t} \mathbf{G}^n \end{cases}$$

où $\mathbf{G}^n = (G_1^n, \dots, G_{3N}^n)$ avec G_i^n gaussiennes centrées réduites indépendantes (et $(\mathbf{q}^0, \mathbf{p}^0) = (\mathbf{q}(0), \mathbf{p}(0))$).

Remarque: L'implémentation pratique sur ordinateur nécessite de savoir simuler des suites de nombres aléatoires indépendants (méthode de Monte Carlo).

Quelques codes gratuits bien distribués

Ab initio MD and materials:

- ABINIT (Ab Initio MD): mainly electronic structure calculation.
- CP2K (Ab Initio MD): solid state, liquid, biological systems.
- CPMD (Ab Initio MD): Car Parinello Molecular Dynamics
- GROMACS: protein, lipids.
- LAMMPS (with force fields): materials science (ReaxFF, EAM, ...)
- VASP (Ab Initio MD)

Simulation of biomolecules:

- AMBER (with force field): MD of biomolecules.
- CHARMM (with force field): DNA, RNA, lipids.
- CP2K (Ab Initio MD): solid state, liquid, biological systems.
- CPMD (Ab Initio MD): Car Parinello Molecular Dynamics
- GROMACS: protein, lipids.
- NAMD and VMD (visualisation): a parallel code for biomolecules.
- OpenMM (high performances on GPU).
- Tinker: partly developed in France (J.P. Piquemal).

Discrétisation en temps: Le schéma de Verlet et généralisation à Langevin

Discretisation d'un système hamiltonien

On cherche une “bonne” discrétisation des équations de Newton:

$$\begin{cases} d\mathbf{Q}(t) = \nabla_{\mathbf{p}} H(\mathbf{Q}(t), \mathbf{P}(t)) dt = M^{-1} \mathbf{P}(t) dt \\ d\mathbf{P}(t) = -\nabla_{\mathbf{q}} H(\mathbf{Q}(t), \mathbf{P}(t)) dt = -\nabla V(\mathbf{Q}(t)) dt \end{cases}$$

où $H(\mathbf{q}, \mathbf{p}) = \frac{\mathbf{p}^T M^{-1} \mathbf{p}}{2} + V(\mathbf{q})$.

Schéma d'Euler explicite:

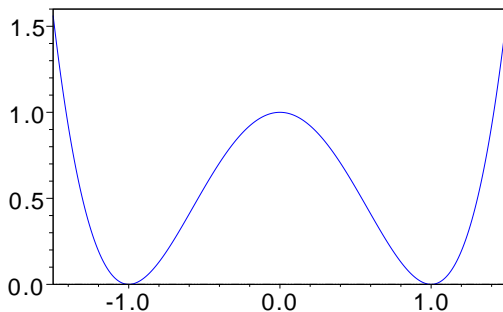
$$\begin{cases} \mathbf{Q}^{n+1} - \mathbf{Q}^n = M^{-1} \mathbf{P}^n \Delta t \\ \mathbf{P}^{n+1} - \mathbf{P}^n = -\nabla V(\mathbf{Q}^n) \Delta t \end{cases}$$

Une simulation en dimension 1

Potentiel double puits: $V(q) = (q^2 - 1)^2$, $V'(q) = -2q(q^2 - 1)$:

$$\begin{cases} dQ(t) = P(t) dt \\ dP(t) = -V'(Q(t)) dt \end{cases}$$

Condition initiale: $(Q(0), P(0)) = (-1.45, 0)$.



Une simulation en dimension 1

Quelques simulations...

- On trace $t_n \mapsto H(Q^n, P^n)$.
- On trace également $(Q^n, V(Q^n))_{n \geq 1}$.

Une simulation en dimension 1

Quelques simulations...

- On trace $t_n \mapsto H(Q^n, P^n)$.
- On trace également $(Q^n, V(Q^n))_{n \geq 1}$.

Le schéma de Verlet:

$$\begin{cases} \mathbf{P}^{n+1/2} - \mathbf{P}^n = -\frac{1}{2} \nabla V(\mathbf{Q}^n) \Delta t \\ \mathbf{Q}^{n+1} - \mathbf{Q}^n = M^{-1} \mathbf{P}^{n+1/2} \Delta t \\ \mathbf{P}^{n+1} - \mathbf{P}^{n+1/2} = -\frac{1}{2} \nabla V(\mathbf{Q}^{n+1}) \Delta t \end{cases}$$

Pourquoi le schéma de Verlet permet-il de conserver l'énergie sur des temps plus longs ?

→ Analyse rétrograde de l'erreur en temps (E. Hairer, C. Lubich, J.M. Sanz Serna,... 1990-).

Schéma de Verlet

$$\begin{cases} \mathbf{P}^{n+1/2} - \mathbf{P}^n = -\frac{1}{2} \nabla V(\mathbf{Q}^n) h \\ \mathbf{Q}^{n+1} - \mathbf{Q}^n = M^{-1} \mathbf{P}^{n+1/2} h \\ \mathbf{P}^{n+1} - \mathbf{P}^{n+1/2} = -\frac{1}{2} \nabla V(\mathbf{Q}^{n+1}) h \end{cases}$$

Théorème

Pour le schéma de Verlet, pour tout $k \geq 1$, il existe un fonction $\tilde{H}(\mathbf{q}, \mathbf{p})$ tel que la solution $(\tilde{\mathbf{Q}}(t), \tilde{\mathbf{P}}(t))$ du problème: $\forall t \in [t_n, t_{n+1}]$,

$$\begin{cases} d\tilde{\mathbf{Q}}(t) = \nabla_{\mathbf{p}} \tilde{H}(\tilde{\mathbf{Q}}(t), \tilde{\mathbf{P}}(t)) dt \\ d\tilde{\mathbf{P}}(t) = -\nabla_{\mathbf{q}} \tilde{H}(\tilde{\mathbf{Q}}(t), \tilde{\mathbf{P}}(t)) dt \end{cases} \quad (\tilde{\mathbf{Q}}(t_n), \tilde{\mathbf{P}}(t_n)) = (\mathbf{Q}^n, \mathbf{P}^n)$$

satisfait:

$$\|\mathbf{Q}^{n+1} - \tilde{\mathbf{Q}}(t_{n+1})\| + \|\mathbf{P}^{n+1} - \tilde{\mathbf{P}}(t_{n+1})\| \leq C_k h^{2k+2}.$$

De plus, $\tilde{H}(\mathbf{q}, \mathbf{p}) = \frac{\mathbf{p}^T M^{-1} \mathbf{p}}{2} + V(\mathbf{q}) + h^2 H_2(\mathbf{q}, \mathbf{p}) + \dots + h^{2k} H_{2k}(\mathbf{q}, \mathbf{p})$.

Schéma de Verlet

Résultat du théorème précédent:

$$\|\mathbf{Q}^{n+1} - \tilde{\mathbf{Q}}(t_{n+1})\| + \|\mathbf{P}^{n+1} - \tilde{\mathbf{P}}(t_{n+1})\| \leq C_k h^{2k+2}$$

et le schéma de Verlet est tel que *l'équation modifiée* est hamiltonienne !

Conséquences:

- Puisque $\tilde{H}(\tilde{\mathbf{Q}}(t_{n+1}), \tilde{\mathbf{P}}(t_{n+1})) = \tilde{H}(\tilde{\mathbf{Q}}(t_n), \tilde{\mathbf{P}}(t_n))$, on a $|\tilde{H}(\mathbf{Q}^{n+1}, \mathbf{P}^{n+1}) - \tilde{H}(\mathbf{Q}^n, \mathbf{P}^n)| = \mathcal{O}(h^{2k+2})$
- $\tilde{H}(\mathbf{Q}^n, \mathbf{P}^n) = \tilde{H}(\mathbf{Q}^0, \mathbf{P}^0) + \mathcal{O}(nh^{2k+2})$
- $H(\mathbf{Q}^n, \mathbf{P}^n) = H(\mathbf{Q}^0, \mathbf{P}^0) + \mathcal{O}(h^2) + \mathcal{O}(nh^{2k+2})$
- L'énergie est conservée à $\mathcal{O}(h^2)$ sur des temps de l'ordre $nh \leq C_k h^{1-2k}$

Remarque: On peut construire une équation modifiée pour de nombreux intégrateurs. La spécificité du schéma de Verlet (et de manière plus générale des intégrateurs *symplectiques*) est que l'équation modifiée est hamiltonienne.

Schéma de Verlet

Remarques:

- Sous des hypothèses d'analyticité, on montre même que l'erreur sur l'énergie est exponentiellement petite sur des temps exponentiellement longs.
- Ce type de résultat est valable dès qu'on utilise un intégrateur *symplectique*.
- Des difficultés apparaissent en pratique pour des problèmes raides (intégration de systèmes hautement oscillants).

L'algorithme de Verlet est

- Peu coûteux: un seul appel à la force par pas de temps ;
- D'ordre deux en temps ;
- Symplectique ;
- Réversible en temps.

Choix du pas de temps: la contrainte de stabilité est liée aux modes de vibration les plus rapides. Exemples: 2-6 fs pour des métaux ; 0.5-1 fs pour des systèmes avec des atomes légers (molécules biologiques dans un solvant).

Généralisation à Langevin

$$\begin{cases} d\mathbf{Q}(t) = M^{-1} \mathbf{P}(t) dt, \\ d\mathbf{P}(t) = -\nabla V(\mathbf{Q}(t)) dt - \gamma M^{-1} \mathbf{P}(t) dt + \sqrt{2\gamma\beta^{-1}} d\mathbf{W}(t), \end{cases}$$

Algorithme de splitting: 1/2 pas de temps **thermostat** (midpoint Euler) puis 1 pas de temps hamiltonien (Verlet) puis 1/2 pas de temps **thermostat** (midpoint Euler).

$$\left\{ \begin{array}{l} \mathbf{P}^{n+1/4} = \mathbf{P}^n - \frac{\Delta t}{4} \gamma M^{-1} (\mathbf{P}^n + \mathbf{P}^{n+1/4}) + \sqrt{\frac{\Delta t}{2}} \sigma G^n, \\ \mathbf{P}^{n+1/2} = \mathbf{P}^{n+1/4} - \frac{\Delta t}{2} \nabla V(\mathbf{Q}^n), \\ \mathbf{Q}^{n+1} = \mathbf{Q}^n + \Delta t M^{-1} \mathbf{P}^{n+1/2}, \\ \mathbf{P}^{n+3/4} = \mathbf{P}^{n+1/2} - \frac{\Delta t}{2} \nabla V(\mathbf{Q}^{n+1}), \\ \mathbf{P}^{n+1} = \mathbf{P}^{n+3/4} - \frac{\Delta t}{4} \gamma M^{-1} (\mathbf{P}^{n+3/4} + \mathbf{P}^{n+1}) + \sqrt{\frac{\Delta t}{2}} \sigma G^{n+1/2}, \end{array} \right.$$

où $\sigma = \sqrt{2\gamma\beta^{-1}}$.

Généralisation à Langevin

Remarques:

- Pour déboguer le calcul des forces: Vérifier la conservation de l'énergie avec $\gamma = 0$ (dynamique hamiltonienne).
- Ne pas utiliser des thermostats déterministes (Nosé-Hoover, Nosé-Poincaré, Berendsen, ...)
- Attention au choix du générateur de nombres aléatoires, et de la graine (*seed*).

Nous allons maintenant parler essentiellement de dynamique thermostatée, et de comment simuler des systèmes moléculaires sur des temps longs.