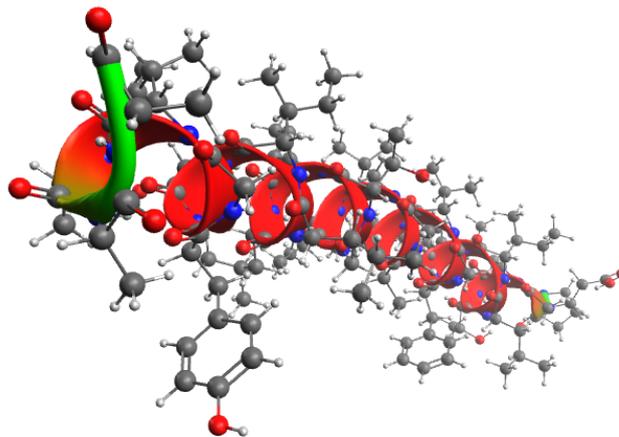**TECHNISCHE UNIVERSITÄT KAISERSLAUTERN**

Department of Mathematics

BACHELOR THESIS

Molecular Dynamical Simulation for Polymers

Steffen Plunder

supervised by

Dr. Wolfgang BOCK

1st August 2016

MOLECULAR DYNAMICAL SIMULATION FOR POLYMERS

STEFFEN PLUNDER

Contents

This bachelor thesis gives an introduction to molecular dynamic simulations, on both theoretical and numerical sides. It builds a connection between mathematical tools of numerical and stochastic analysis and physical models. First we build up the needed background in the theory of SDEs in Chapter 1 and recap the physical model in Chapter 2. The theory for symplectic numerical schemes and common implementation techniques are content of Chapter 3. A short discussion of numerical methods for the stochastic Langevin equation follows in Chapter 4. Our main focus relies on the preservation of geometric invariants. Simultaneous to writing this thesis a C++ plugin for the open source application *Avogadro* was build. At the end of Chapter 4 we will test these techniques on the simulation of a complex molecule.

Notation 0.1. Within this thesis we will use the following Notations.

N		number of particles
d		space dimension
m_i		mass of the ith particle
$\mathrm{D}f$		Jacobian of f
$\mathrm{D}^2 f$		Hessian matrix of f
$\mathbf{X}$	$= (\mathbf{x}_1, \ldots, \mathbf{x}_N) \in (\mathbb{R}^d)^N$	position coordinates
$\mathbf{V}$	$= (\mathbf{v}_1, \ldots, \mathbf{v}_N) \in (\mathbb{R}^d)^N$	velocity coordinates
$\mathbf{Q}$	$= (\mathbf{q}_1, \ldots, \mathbf{q}_N) \in (\mathbb{R}^d)^N$	position coordinates
$\mathbf{P}$	$= (\mathbf{p}_1, \ldots, \mathbf{p}_N) \in (\mathbb{R}^d)^N$	inertia coordinates
$\mathcal{T}$	$= \sum_i \frac{1}{2m_i} \langle \mathbf{p}_i, \mathbf{p}_i \rangle$	kinetic energy
$\mathcal{V}$		potential energy
$\mathcal{H}$	$= \mathcal{T} + \mathcal{V}$	Hamiltonian function
γ		friction coefficient
$(\Omega, \mathcal{A}, \mathcal{P})$		a fixed probability space
$L^2(\Omega)$		the associated Hilbert space on Ω
$\mathcal{B}(U)$		Borel-σ-algebra on U
$\mathrm{E}(\,\cdot\,)$		expectation value
$\mathrm{Var}(\,\cdot\,)$		variance
$\sigma(X)$		by X generated σ-algebra
$\mathbb{1}_A$		the indicator function of A
W	$= \{W_t\}_{t \in T}$	Brownian motion on Ω

1. BASIC CONCEPTS OF STOCHASTIC DIFFERENTIAL EQUATIONS

This chapter gives an introduction to stochastic differential equations (SDEs). For a more consistent introduction to SDEs we refere to textbooks in this field like [Øks03] or [KP92]. For a quick overview we

recommend [Eva12] or [KK13, Chapter 2]. A more general introduction to Brownian motion and stochastic integration can be found in [KS12].

As the solutions of SDEs will be given by stochastic processes, we will start with some basic definitions and an brief introduction to Brownian motion, which describes the motion of a particle which chooses its position randomly at every time. Brownian motion is a classical example for a stochastic process and plays an important role in stochastic differential equations.

1.1. **Stochastic Processes.**

Definition 1.1 (Stochastic process, [Øks03, Definition 2.1.4]). A *stochastic process* is collection of random variables

$$\{X_t : \Omega \to \mathbb{R}^n\}_{t \in T}$$

parametrized by some index set T and defined for a probability space $(\Omega, \mathcal{A}, \mathcal{P})$ and assuming values in $\mathbb{R}^n$ equipped with the Borel-σ-algebra as usual.

In this thesis, the parameter space T will always be $[0, \infty)$ or $[0, t_{end}]$. There are various interpretations of a stochastic process, therefore we fix some notations.

Notation 1.2. It is common to interpret a stochastic process as a mapping

$$X : T \times \Omega \to \mathbb{R}^n : (t, \omega) \mapsto X(t, \omega) := X_t(\omega).$$

If we fix $\omega \in \Omega$, we get one *path* of a stochastic process

$$X^\omega : T \to \mathbb{R}^n : t \mapsto X(t, \omega).$$

Somehow we could view ω as an individual "particle" or "experiment" which follows the path X^ω.

A modification of a stochastic processes on a set of measure zero has still much in common with the original process. This leads to the next definition.

Definition 1.3 (Versions of a stochastic process). Let X and $\widetilde{X}$ be stochastic processes on the probability space assuming values in $\mathbb{R}^n$. Then $\widetilde{X}$ is called a *version* of X if the are for almost all times equal, i.e. $\mathcal{P}(X(t, \cdot) = \widetilde{X}(t, \cdot)) = 1$ for all $t \in T$.
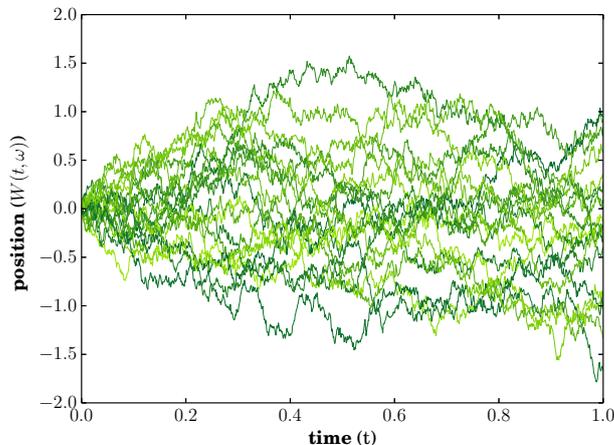
FIGURE 1. Paths of a Brownian motion.

1.2. **Brownian motion.** The name of this stochastic process refers to Robert Brown, how observed in the year 1826 that pollen grains in a fluid follow a continuous, irregular movement. Einstein [Ein05] and Smoluchowski independently explained these movements as a result of collisions with liquid molecules. For example, a colloidal particle of radius 10^{-6} m in a liquid, is subject to approximately 10^{20} molecular collissions each second, each of which changes its velocity by an amount of the order 10^{-8} m/s. One important consequence of these theoretical ideas was due to Perrin, who estimated the Avogadro number $N_A \approx 7 \cdot 10^{23}$, which was a major evidence for the atomic model.

One the mathematical side, Norbert Wiener proposed several different methods to construct a Brownian motion. In honour to his work the stochastic process of a Brownian motion is also called Wiener process.

Definition 1.4 (Brownian motion)**.** A real-valued stochastic process $\{W_t\}_{t \in [0,\infty)}$ is called *Brownian motion* (or *Wiener process*) if

   (i) $W_0 = 0$ almost surely,

   (ii) $W_t - W_s \sim \mathcal{N}(0, t - s)$ for all $t, s \in [0, \infty)$ with $t > s$,

   (iii) For all times $0 < t_0 < t_1 < \cdots < t_k$ the so called increments $W_{t_1} - W_{t_0}, \ldots, W_{t_k} - W_{t_{k-1}}$ are independent.

   (iv) The sample paths W^ω are continuous for almost every ω.

The existence of such a process is not trivial, proofs are given for example in [Øks03, Section 2.2] or [Eva12, Section 3.3]. There are different ways to define a Brownian motion, some authors neglect the last condition and instead prove that a continuous Brownian motion exists.

We just note that a Brownian motion with continuous paths exists and denote it by W.

Remark 1.5 (Properties of a Brownian motion , [Eva12, Lemma 3.2.3]). Let us summarize some basic properties of a Brownian motion. For $s, t \in [0, \infty)$ holds

- $\mathrm{E}\,(W_t) = 0$

- $\mathrm{Var}\,(W_t) = t$

- $\mathrm{Cov}\,(W_t, W_s) = \mathrm{E}\,(W_t W_s) = \min\{t, s\}$.

Another remarkable property of the Brownian motion is the following scaling invariance: If $W(t, \cdot)$ is a Brownian motion then for all $\lambda > 0$ the process

$$X(t, \omega) := \frac{1}{\lambda} W(\lambda^2 t, \omega)$$

is also a Brownian motion. This property is one of the reasons why we will later introduce a stochastic integral to define stochastic differential equations.

Before getting too enthusiastic, we also have to note that the paths of a Brownian motion are only continuous, but almost surely nowhere differentiable.

Remark 1.6 (Properties of sample paths, [Eva12, Theorem 3.4.2]). For each $\frac{1}{2} < \gamma \leq 1$ and almost every ω, the sample path W^ω is nowhere Hölder continuous with exponent γ, i.e. for each time s there exists no constant K such that

$$\left| W^\omega(t) - W^\omega(s) \right| \leq K |t - s|^\gamma \quad \text{for all } t \in [0, T]$$

holds. In particular, for almost every ω, the sample path W^ω is nowhere differentiable and is of infinite variation on each subinterval.

Remark 1.7 (Numerical simulation of a Brownian motion). For a simple simulation of a Brownian motion one could use the Euler-Maruyama scheme

$$W_{t+h} \approx W_t + \sqrt{h} \Delta_h,$$

where $\Delta_h \sim \mathcal{N}(0, h)$ are normal distributed random values. This is directly motivated by $W_t - W_s \sim \mathcal{N}(0, t - s)$ from Definition 1.4 (ii).

1.3. **Filtration.** A plain stochastic process can behave quite strange. In particular, the position of a particle can depend not only on some initial data give for one time, they can also depend on the previous path. This makes analysis more complicated and is not necessary for our application. Filtrations represent the information available at certain times.

Definition 1.8 (Filtration, [Øks03, Definition 3.2.2]). A *filtration* (on $(\Omega, \mathcal{A})$) is a family $\mathcal{M} = \{\mathcal{M}_t\}_{t \geq 0}$ of σ-algebras $\mathcal{M}_t \subseteq \mathcal{A}$ such that

$$0 \leq s < t \quad \Rightarrow \quad \mathcal{M}_s \subseteq \mathcal{M}_t$$

(i.e. $\{\mathcal{M}_t\}$ is increasing). A n-dimensional stochastic process $\{M_t\}_{t \geq 0}$ on $(\Omega, \mathcal{A}, \mathcal{P})$ is called a *martingale* with respect to a filtration $\{\mathcal{M}_t\}_{t \geq 0}$ (and with resprect to P) if

(i) M_t is $\mathcal{M}_t$-measurable for all t,

(ii) $\mathrm{E}\left(|M_t|\right) < \infty$ for all t and

(iii) $\mathrm{E}\left(M_s \mid \mathcal{M}_t\right) = M_t$ for all $s \geq t$

Definition 1.9 (Adapted processes). A stochastic process X is *adapted* to the filtration $\mathcal{M}$ if, for each $t \geq 0$, X_t is an $\mathcal{M}_t$-measurable random variable.

Like random variables generate a σ-algebra, each stochastic process also has a "naturally" associated filtration.

Definition 1.10 (Natural Filtration). Let X be a stochastic process. Then the Filtration $\mathcal{F}$ defined by

$$\mathcal{F}_t = \sigma\left(\{X_s \mid 0 \leq s \leq t\}\right)$$

the *natural filtration* of X.

Clearly X is $\mathcal{F}$ adapted, i.e. we have for each process at least one filtration.

1.4. **Conditional expectation.** We still need to introduce another important concept, namely the conditional expectation of a stochastic process. A simple example is given by the case of a piecewise constant random variable

$$Y = \sum_{j=1}^{n} a_j \mathbb{1}_{A_i}.$$

For another random variable X, the conditional expectation $\mathrm{E}\left(X \mid Y\right)$ then could been seen as the best approximation of X on the generated σ-algebra $\sigma(Y)$.

$$\mathrm{E}\left(X \mid Y\right) = \begin{cases} \frac{1}{\mathcal{P}(A_1)} \int_{A_1} X \, \mathrm{d}\mathcal{P}, & \text{on } A_1 \\ \vdots \\ \frac{1}{\mathcal{P}(A_n)} \int_{A_n} X \, \mathrm{d}\mathcal{P}, & \text{on } A_n \end{cases}$$

We see that only the generated σ-algebra matters, thus the definition is given for σ-subalgebras. The definition now is rather abstract, we will

not give any further explanations here and refer to [Eva12] instead. We continue with the defining properties of the conditional expectation.

Definition 1.11 (Conditional expectation, [Eva12, Definition 2.6])**.** Let $(\Omega, \mathcal{A}, \mathcal{P})$ be a probability space and suppose $\mathcal{V}$ is a σ-subalgebra of $\mathcal{A}$. If $X : \Omega \to \mathbb{R}$. is an integrable random variable, we define the *conditional expectation*

$$\mathrm{E}\left(X \mid \mathcal{V}\right)$$

to be a random variable on Ω such that

   (i) $\mathrm{E}\left(X \mid \mathcal{V}\right)$ is $\mathcal{V}$-measurable and

   (ii) $\int_A X \, \mathrm{d}\mathcal{P} = \int_A \mathrm{E}\left(X \mid \mathcal{V}\right) \mathrm{d}\mathcal{P}$ for all $A \in \mathcal{V}$.

Remark 1.12 (Existence and Uniqueness of the conditional expectation, [Øks03, Appendix B]). First of all we shortly ensure existence and uniqueness. We the define the measure

$$\mu(A) = \int_A X \, \mathrm{d}\mathcal{P}, \text{ for } A \in \mathcal{V}$$

on $\mathcal{V}$ which is absolute continuous w.r.t. $\mathcal{P}|_{\mathcal{V}}$, by the Radon-Nikodym Theorem there exists a $\mathcal{V}$-measurable function F on Ω such that

$$\mu(A) = \int_A F \, \mathrm{d}\mathcal{P}, \text{ for all } A \in \mathcal{V}.$$

Thus $\mathrm{E}\left(X \mid \mathcal{V}\right) := F$ does the job and according to the Radon-Nikodym Theorem this function is unique almost surely w.r.t. the measure $\mathcal{P}|_{\mathcal{V}}$.

Lemma 1.13 (Properties of the conditional expectation, [Øks03, Appendix B])**.** *Suppose $X, Y : \Omega \to \mathbb{R}^n$ are integrable random variables on $(\Omega, \mathcal{A}, \mathcal{P})$ and let $a, b \in \mathbb{R}$. Then*

   (i) $\mathrm{E}\left(aX + bY \mid \mathcal{V}\right) = a\,\mathrm{E}\left(X \mid \mathcal{V}\right) + b\,\mathrm{E}\left(Y \mid \mathcal{V}\right)$ *(linearity)*

   (ii) $\mathrm{E}\left(\mathrm{E}\left(X \mid \mathcal{V}\right)\right) = \mathrm{E}\left(X\right)$

   (iii) $\mathrm{E}\left(X \mid \mathcal{V}\right) = X$ *if X is $\mathcal{V}$-measurable*

   (iv) $\mathrm{E}\left(X \mid \mathcal{V}\right) = \mathrm{E}\left(X\right)$ *if X is independent of $\mathcal{V}$*

   (v) $\mathrm{E}\left(Y \cdot X \mid \mathcal{V}\right) = Y\,\mathrm{E}\left(X \mid \mathcal{V}\right)$ *if Y is $\mathcal{V}$-measurable,*
     where $\cdot$ denotes the usual inner product on $\mathbb{R}^n$.

1.5. **The Itô integral.** We want add a stochastic term to an ordinary differential equation. Like the Brownian motion, the paths of solutions also will not necessarily be differentiable. Therefore we chose an integral representation to set up a stochastic differential equation

$$X(t, \omega) = X(0, \omega) + \int_0^t a(X(s, \omega), s) ds + \int_0^t b(X(s, \omega), s) \, \mathrm{d}W_s,$$

where we still need to define the last integral. Later we will use the differential notation to denote the equation above as

$$dX_t = a(X_t, t)\, dt + b(X_t, t)\, dW_t.$$

The definition of a stochastic integral is somehow involved and there exist other useful stochastic integrals as well. We choose the integral preferred by mathematicians, the Itô integral, which has some analytical advantages. In physics one usually considers the Stratonovich Integral, which satisfies the usual chain rule. In this section we will concentrate on the definition of a stochastic integral in the Itô sense.

1.6. **Construction of the Itô integral.** We will first define the Itô integral for simple processes and extend this definition onto the class of square integrable functions. Here we follow [KP92, chapter 3.2].

Definition 1.14 (Progressively measurable processes, [KP92, Section 3.2])**.** Suppose we have a probability space $(\Omega, \mathcal{A}, \mathcal{P})$, a Wiener process W and an increasing family $(\mathcal{A}_t)_{t \geq 0}$ of σ-subalgebras of $\mathcal{A}$ such that W_t is $\mathcal{A}_t$-measurable and almost surely hold

$$\mathrm{E}\left(W_t \mid \mathcal{A}_0\right) = 0 \text{ and } \mathrm{E}\left(W_t - W_s \mid \mathcal{A}_s\right) = 0,$$

for all $0 \leq s \leq t$. For $0 < T < \infty$ we define a class L_T^2 of functions $f : [0, T] \times \Omega \to \mathbb{R}$ satisfying

(i) f is jointly $\mathcal{B}([0, T]) \times \mathcal{A}$-measurable,

(ii) $\int_0^T \mathrm{E}\left(f(t, \cdot)^2\right) dt < \infty$,

(iii) $\mathrm{E}\left(f(t, \cdot)^2\right) < \infty$ for each $0 \leq t \leq T$,

(iv) $f(t, \cdot)$ is $\mathcal{A}_t$-measurable for each $0 \leq t \leq T$.

In addition we consider two functions in L_T^2 to be identical if they are equal for all (t, ω) except possibly on a subset of $dt \times \mathcal{P}$-measure zero. The scalar product in this space is defined by

$$\langle f, g \rangle_{2,T} := \int_0^T \mathrm{E}\left(f(t, \cdot)g(t, \cdot)\right) dt.$$

Remark 1.15. The space $\left(L_T^2, \langle \cdot, \cdot \rangle_{2,T}\right)$ is a complete Hilbert space. We denote the norm on this space as $\|\cdot\|_{2,T}$ is has the form

$$\|f\|_{2,T} = \int_0^T \mathrm{E}\left(f^2(t, \cdot)\right) dt.$$

Definition 1.16 (Space of step functions)**.** A function $f \in L_T^2$ that can be represented as

$$f(t, \omega) = \sum_{i=0}^n f_j(\omega) \cdot \mathbb{1}_{[t_i, t_{i+1}]}(t).$$

for a partition $0 = t_0 < \cdots < t_{n+1} = T$ is called a *step function*. We denote the subspace of step functions in L_T^2 by S_T^2

Lemma 1.17. S_T^2 *is dense in* L_T^2.

Sketch of the proof. The main idea is to apply the dominated convergence theorem twice. First to show that the set of mean square continuous functions lies dense in L_T^2 and then to approximate these functions by equidistant step functions. See [KP92, Lemma 3.2.1] or [KS12, Section 3.2, Proposition 2.8] for details. $\square$

We now want to define the Itô integral on S_T^2 and extend the definition onto L_T^2.

Definition 1.18 (Itô integral for step functions, [KP92, Page 84])**.** For $f \in S_T^2$ we define the *Itô integral for step functions* by

$$\mathrm{I}\left(f\right)\left(\omega\right) := \sum_{i=1}^{n} f_j(\omega) \cdot \left(W_{t_{j+1}}(\omega) - W_{t_j}(\omega)\right).$$

Later we will use the notation

$$\int_0^T f(t, \omega)\,\mathrm{d}W_t := \mathrm{I}\left(f\right)\left(\omega\right).$$

To extend the linear functional $\mathrm{I}(\cdot)$ onto L_T^2 we have to check that for any convergent sequence of step functions, the integral also converges against a unique random variable. Therefore we need to prove the Itô isometry for step functions.

Theorem 1.19 (Basic properties of the Itô integral for step function, [Eva12] 4.2.3)**.** *For all constants* $a, b \in \mathbb{R}$ *and for all* $G, H \in S_T^2$, *we have:*

   (i) $\mathrm{I}\left(f\right)$ *is* $\mathcal{A}_T$*-measurable.*

   (ii) $\mathrm{I}\left(aG + bH\right) = a\,\mathrm{I}\left(G\right) + b\,\mathrm{I}\left(H\right)$ *almost surely*

   (iii) $\mathrm{E}\left(\mathrm{I}\left(G\right)\right) = 0$

   (iv) $\mathrm{E}\left(\mathrm{I}\left(f\right)^2\right) = \int_0^T \mathrm{E}\left(f(s, \cdot)^2\right)\mathrm{d}s$

Rewriting (iv) gives the so called Itô isometry $\left\|\mathrm{I}\left(f\right)\right\|_{L^2(\Omega)}^2 = \|f\|_{2,T}^2$.

Proof. **(i)** Let $f \in S_T^2$ be a step function corresponding to a partition $0 = t_1 < t_2 < \cdots < t_{n+1} = T$.
By Definition 1.14 (iv) we have for each $j = 1, 2, \ldots, n$ that $f_j = f(t_j, \cdot)$ is $\mathcal{A}_{t_j}$-measurable and again by Definition 1.14 we also know that the increments $W_{t_{j+1}} - W_{t_j}$ are $\mathcal{A}_{t_{j+1}}$-measurable.
The product $f_j\left(W_{t_{j+1}} - W_{t_j}\right)$ is also $\mathcal{A}_T$-measurable, since we have an

increasing filtration, i.e. $\mathcal{A}_{t_j} \subset \mathcal{A}_{t_{j+1}} \subset \mathcal{A}_T$. As a sum of measurable function, $I(f)$ is also $\mathcal{A}_T$ measurable.

(ii) The linearity follows by algebraic operations.

(iii) By the Cauchy-Schwarz inequality we have

$$
E\left(|I(f)|\right) \leq \sum_{j=1}^{n} \sqrt{E\left(f_j^2\right)} \cdot \sqrt{E\left(\left(W_{t_{j+1}} - W_{t_j}\right)^2\right)} < \infty,
$$

since f_j is mean-square integrable by Definition 1.14 (iii) and the increment of a Brownian motion is normal distributed with variance $|t - s|$ by Definition 1.4. Hence is also mean-square integrable.

Using Lemma 1.13 (iii) and (v) we get

$$
E\left(I(f)\right) = \sum_{j+1}^{n} E\left(f_j \cdot (W_{t_{j+1}} - W_{t_j})\right)
$$

$$
= \sum_{j+1}^{n} E\left(f_j \cdot E\left(W_{t_{j+1}} - W_{t_j} \middle| \mathcal{A}_j\right)\right)
$$

$$
= \sum_{j+1}^{n} E\left(f_j\right) \underbrace{E\left(W_{t_{j+1}} - W_{t_j} \middle| \mathcal{A}_{t_j}\right)}_{=0,\text{ by Definition 1.14 (iv).}} = 0.
$$

(iv) We next compute the variance. Here we use that $f_i f_j \cdot \left(W_{t_{j+1}} - W_{t_j}\right)$ are $\mathcal{A}_{t_j}$-measurable for any $i < j$. We get

$$
E\left(I(f)^2\right) = \sum_{j=1}^{n} E\left(f_j^2 (W_{t_{j+1}} - W_{t_j})^2\right)
$$

$$
+ 2\sum_{j=1}^{n} \sum_{i=j+1}^{n} E\left(f_j f_i (W_{t_{j+1}} - W_{t_j})(W_{t_{i+1}} - W_{t_i})\right)
$$

$$
= \sum_{j=1}^{n} E\left(f_j^2\right) \underbrace{E\left((W_{t_{j+1}} - W_{t_j})^2 \middle| \mathcal{A}_{t_j}\right)}_{=t_{j+1} - t_j,\text{ by Definition 1.4.}}
$$

$$
+ 2\sum_{j=1}^{n} \sum_{i=j+1}^{n} E\left(f_j f_i\right) \underbrace{E\left(W_{t_{j+1}} - W_{t_j} \middle| \mathcal{A}_{t_j}\right)}_{=0,\text{ by Definition 1.14.}}
$$

$$
= \sum_{j=1}^{n} E\left(f_j^2\right)(t_{j+1} - t_j)
$$

$$
= \int_0^T E\left(f(t, \cdot)^2\right) dt.
$$

In the last equation we used the definition of the Riemann integral for the non-random step function $s \mapsto E\left(f(s, \cdot)^2\right)$. $\qquad \square$

Remark 1.20 (Convergence of the integrals of step functions). Since S_T^2 lies dense in L_T^2 we have for every $f \in L_T^2$ a Cauchy sequence $f^{(n)} \in S_T^2$ of step functions, such that $f^{(n)} \xrightarrow{L_T^2} f$.

By linearity of $I(\cdot)$ and the Itô isometry from Theorem 1.22 we get

$$
\left\| I\left(f^{(n)} \right) - I\left(f^{(n+m)} \right) \right\|_{L_T^2}^2 = E\left(\left| I\left(f^{(n)} \right) - I\left(f^{(n+m)} \right) \right|^2 \right)
$$

$$
= E\left(\left| I\left(f^{(n)} - f^{(n+m)} \right) \right|^2 \right)
$$

$$
= \left\| f^{(n)} - f^{(n+m)} \right\|_{2,T}^2 .
$$

The inequality $\|a + b\|^2 \le 2\|a\|^2 + 2\|b\|^2$ finally gives

$$
\left\| f^{(n)} - f^{(n+m)} \right\|_{2,T}^2 \le 2 \left\| f^{(n)} - f \right\|_{2,T}^2 + 2 \left\| f - f^{(n+m)} \right\|_{2,T}^2
$$

$$
\to 0, \text{ as } n, m \to \infty.
$$

Hence the sequence $I\left(f^{(n)} \right)$ is a Cauchy sequence in the Banach space $L^2(\Omega)$. We call the limit $I(f)$ the Itô integral of f. We shortly check that this definition is independent of the choice of the approximating sequence $f^{(n)}$. Let $\widetilde{f}^{(n)}$ be another approximating sequence, using the Itô isometry again gives

$$
\left\| I\left(f^{(n)} \right) - I\left(\widetilde{f}^{(n+m)} \right) \right\|_{\mathcal{L}^2}^2 = \left\| f^{(n)} - \widetilde{f}^{(n+m)} \right\|_{2,T}^2
$$

$$
\le 2 \left\| f^{(n)} - f \right\|_{2,T}^2 + 2 \left\| f - \widetilde{f}^{(n+m)} \right\|_{2,T}^2
$$

$$
\to 0, \text{ as } n, m \to \infty.
$$

We can now finally define the Itô integral on L_T^2.

Definition 1.21 (Itô integral). For $f \in L_T^2$, let $(f^n)_{n \in \mathbb{N}}$ be a sequence of step function in S_T^2 with $f^n \xrightarrow{L_T^2} f$. Then we define the *Itô integral* over f as the limit

$$
\int_0^T f(s) \, dW_s := \lim_{n \to \infty} I(f^n).
$$

Due to the construction we can restate Theorem 1.22 on L_T^2.

Theorem 1.22 (Basic properties of the Itô integral, [KP92] Theorem 3.2.3). *The properties (i)-(iv) from Theorem 1.19 also hold on L_T^2 instead of S_T^2.*

Remark 1.23 (Integrating over measurable sets). As usual we define for a measurable set $A \in \mathcal{A}$ the integral over A as

$$\int_A f(t, w) \, \mathrm{d}W_t := \int_0^T \mathbb{1}_A f(t, w) \, \mathrm{d}W_t,$$

where $\mathbb{1}_A$ denotes the indicator function. For $A = [t_0, t_1]$ we write $\int_{t_0}^{t_1} f(s, \omega) \, \mathrm{d}W_s(\omega)$ instead.

Theorem 1.24 (Martingale property , [KP92] , Theorem 3.2.5). *For $t_0 \leq s \leq t, \leq T$ and $f \in L_T^2$ the integrals*

$$Z_t(\omega) := \int_{t_0}^t f(s, \omega) \, \mathrm{d}W_s(\omega)$$

fulfil the martingale property

$$\mathrm{E}\left(Z_t - Z_s \mid \mathcal{A}_s\right) = 0.$$

Sketch. The basic idea is to check the claim for step functions by induction. The fundamental step is to calculate

$$\mathrm{E}\left(f_j^{(n)} \cdot \left(W_{t_{j+1}^{(n)}} - W_{t_j^{(n)}}\right) \middle| \mathcal{A}_{t_j}^{(n)}\right) = f_j^{(n)} \, \mathrm{E}\left(W_{t_{j+1}^{(n)}} - W_{t_j^{(n)}} \middle| \mathcal{A}_{t_j}^{(n)}\right) = 0.$$

$\square$

Theorem 1.25 (Continuous sample paths, [Øks03, Theorem 3.2.5]). *Let $f \in L_T^2$. We define a stochastic process Z as*

$$Z_s := \int_0^s f(t, \omega) \, \mathrm{d}W_t(\omega)$$

for $s \in [t_0, T]$. Then there exists a version of Z, which is almost surely continuous w.r.t. s.

1.7. **Stochastic differential equations.** Now we can finally define stochastic differential equations. Since even the Brownian motion is almost surely nowhere differentiable, the integral representation for differential equations is the leading model to define stochastic differential equations.

Definition 1.26 (Stochastic differential, [Eva12, Definition 4.3.1]). Suppose that X is a real-valued stochastic process satisfying

$$X(r, \omega) = X(s, \omega) + \int_s^r f(t, \omega) \, \mathrm{d}t + \int_s^r g(t, \omega) \, \mathrm{d}W_t(\omega),$$

for some $f \in L_T^1$, $g \in L_T^2$ and all times $0 \leq s \leq r \leq T$. We say that X has the *stochastic differential*

$$\mathrm{d}X = f \, \mathrm{d}t + g \, \mathrm{d}W_t$$

for $0 \leq t \leq T$.

We now come to the first large difference between stochastic calculus in comparison to usual integral rules. Namely the chain and product rules for Itô integrals.

Theorem 1.27 (Itô's product rule, [Eva12, Theorem 4.3.3]). *Suppose*

$$\begin{cases} \mathrm{d}X_1 = f_1 \, \mathrm{d}t + g_1 \, \mathrm{d}W_t \\ \mathrm{d}X_2 = f_2 \, \mathrm{d}t + g_2 \, \mathrm{d}W_t \end{cases}, \qquad on \ 0 \le t \le T,$$

for $f_1, f_2 \in \mathcal{L}^1_T, g_1, g_2 \in \mathcal{L}^2_T$. Then $X_1 X_2$ has the stochastic differential

$$\mathrm{d}(X_1 X_2) = X_1 \, \mathrm{d}X_2 + X_2 \, \mathrm{d}X_1 + g_1 g_2 \, \mathrm{d}t.$$

Theorem 1.28 (Itô's chain rule , [Eva12, Theorem 4.3.1]). *Suppose that X has a stochastic differential*

$$\mathrm{d}X = f \, \mathrm{d}t + g \, \mathrm{d}W_t$$

for some $f \in L^1_T$, $g \in L^2_T$. Assume that $u : \mathbb{R} \times [0, T] \to \mathbb{R}$ is continuous and that its partical derivatives $u_t = \frac{\partial u}{\partial t}, u_x = \frac{\partial u}{\partial x}, u_{xx} = \frac{\partial^2 u}{\partial x^2}$ exist and are continuous.
Then $Y(t, \omega) = u(X(t, \omega), t)$ has the stochastic differential

$$\mathrm{d}u(X, t) = u_t \, \mathrm{d}t + u_x \, \mathrm{d}X + \frac{1}{2} u_{xx} g^2 \, \mathrm{d}t$$

$$= (u_t + u_x f + \frac{1}{2} u_{xx} g^2) \, \mathrm{d}t + u_x g \, \mathrm{d}W_t.$$

Proofs could be found in [Eva12].

1.8. **Taylor approximation.** An application of Itô's rule gives us

Remark 1.29 (Towards an Taylor expansion, [KP92, Page 163]).
Assume X has the stochastic differential

$$\mathrm{d}X = f \, \mathrm{d}t + g \, \mathrm{d}W_t.$$

Applying the Itô formula on $u(x) = f(x)$ and $u(x) = g(x)$ for $f, g \in \mathcal{C}^2(\mathbb{R})$ yields in

$$f(X_t) = f(X_0) + \int_0^t f(X_s) \frac{\partial f}{\partial x}(X_s) + \frac{1}{2} g(X_s) \frac{\partial^2 f}{\partial x^2}(X_s) \, \mathrm{d}s$$

$$+ \int_0^t g(X_s) \frac{\partial f}{\partial x}(X_s) \, \mathrm{d}W_s$$

and

$$g(X_t) = g(X_0) + \int_0^t f(X_s) \frac{\partial g}{\partial x}(X_s) + \frac{1}{2} g(X_s) \frac{\partial^2 g}{\partial x^2}(X_s) \, \mathrm{d}s$$

$$+ \int_0^t g(X_s) \frac{\partial g}{\partial x}(X_s) \, \mathrm{d}W_s.$$

Inserting these equations into the integral equation of X we get

$$X_t = X_0 + f(X_0) \int_0^t \mathrm{d}s + g(X_0) \int_0^t \mathrm{d}W_s + R.$$

This is a Taylor like extension of X_t. Of cause one could proceed to apply the Itô formula again and again to obtain higher order extension. For us this rather simple extension is enough. It leads to the Euler-Maruyama scheme. Collecting the remaining terms gives

$$R = \int_0^t \int_0^s L^0 f(X_z) \, \mathrm{d}z \, \mathrm{d}s + \int_0^t \int_0^s L^1 f(X_z) \, \mathrm{d}W_z \, \mathrm{d}s$$
$$+ \int_0^t \int_0^s L^0 g(X_z) \, \mathrm{d}z \, \mathrm{d}W_s + \int_0^t \int_0^s L^1 g(X_z) \, \mathrm{d}W_z \, \mathrm{d}W_s.$$

2. The physical background of molecular dynamics

Our goal is to describe the motion of atoms and molecules. In practice we will use a classical model, which describes the interaction between particles by given potentials and Newtons laws. The influence of smaller particles is modelled by a stochastic term in the equations of motion. For an argumentation why this model does not contradict quantum mechanical basics we refer to [GKZ10] or [FS01].

2.1. **Potentials.** For each model we have to chose the forces $\mathbf{F}_i$ acting on the i-th particle. It is common to use potentials which depend on the positions of the particles to describe the forces. The force applied to each particle then has the form $\mathbf{F}_i = -\nabla_{\mathbf{q}_i} V$, where V is a given potential function. Usually the potential is separated into two parts $\mathcal{V} = \mathcal{V}_{\text{non-bonded}} + \mathcal{V}_{\text{bonded}}$, where $\mathcal{V}_{\text{bonded}}$ is the potential caused by bonds between atoms and $\mathcal{V}_{\text{non-bonded}}$ describes forces acting between all pairs of atoms.

2.1.1. *The non-bonded potentials.* Typical examples for the non-bonded potentials are the Lennard-Jones potential or the Coulomb potential. They act between all particles in contrast to bonded potentials. Let r be the distance between two particles, q_i be the electric charge of a particle and ε, r_m are particle depended parameters. Then the potentials have to form:

Lennard-Jones $\quad \mathcal{V}(r) = \varepsilon \left(\left(\frac{r_m}{r} \right)^{12} - \left(\frac{r_m}{r} \right)^6 \right),$
$\qquad$**Coulomb**$\qquad \mathcal{V}(r) = \frac{1}{4\pi\epsilon} \frac{q_1 q_2}{r}.$

We note that the sum of these potentials are bounded from below, even if $q_1 q_2 < 0$.

2.1.2. *The bonded potentials.* In view towards molecular simulations we have to consider a second type of potentials. These potentials also include the interaction between three and four particles. If two atoms are bounded the distribution probabilities of their electrons change and at this level quantum mechanical effects are relevant. To overcome this, models were developed which try to approximate the quantum mechanical behaviour. These models include interaction between two, three and four atoms. They are often called *force fields*. One has to consider that currently different models of force fields exist, not only for atomic interaction, also for the interaction of small groups of atoms. They all have different strengths and weaknesses.

We summarize the quantitative behaviour of typical potentials. Let r be the distance between two atoms, θ should denote the angle between two atoms w.r.t. to a third atom and ω denotes a complex torsion angle, then typical potentials are

Bond $\qquad\qquad\qquad \mathcal{V}(r) = c \cdot (r - r_{min})^2$
Angle $\qquad\qquad\qquad \mathcal{V}(\theta) = c \cdot (cos(\theta) - 1)$
Torsion $\quad \mathcal{V}(\omega) = c_1(1 - \cos(\omega)) + c_2(1 - \cos(2\omega)) + c_3(1 - \cos(3\omega)).$

In this thesis we threat these potential models as a black box. We just assume that the resulting potential is smooth for $\mathbf{q}_i \neq \mathbf{q}_j$ and bounded from below. Simulation results in this thesis are based on the *universal force field* [RCC$^+$92].

2.2. **Newton.** To model just a molecular system no complicated physics are required, only Newtons equation

$$\frac{\mathrm{d}^2 \mathbf{q}_i}{\mathrm{d}t^2} = \frac{\mathbf{F}_i(\mathbf{q}_1, \ldots, \mathbf{q}_N)}{m_i} = \frac{\mathbf{F}_i(\mathbf{Q})}{m_i}.$$

The force will be given as the gradient of a potential function depending only on the positions $F_i = -\nabla_{\mathbf{q}_i} \mathcal{V}(\mathbf{q}_1, \ldots, \mathbf{q}_N)$. Let $\mathbf{p}_i$ denote the inertia of the ith particle, then we can rewrite Newtons equation as a Hamilton system. The Hamiltonian of the system is given by

$$\mathcal{H} = \mathcal{V} + \mathcal{T}, \quad \text{with } \mathcal{T}(\mathbf{Q}) = \sum_{i=1}^N \frac{1}{2m_i} \|\mathbf{q}_i\|_2^2.$$

The Hamilton system

$$\frac{\mathrm{d}\mathbf{q}_i}{\mathrm{d}t} = \nabla_{\mathbf{p}_i} \mathcal{H}(\mathbf{Q}, \mathbf{P}), \qquad \frac{\mathrm{d}\mathbf{p}_i}{\mathrm{d}t} = -\nabla_{\mathbf{q}_i} \mathcal{H}(\mathbf{Q}, \mathbf{P})$$

is equivalent to Newtons law. Since the Hamilton function is time independent, it is just the total energy of the system.

Mathematical Interlude I. Since we are mathematicians, we need to ensure the existence and uniqueness of solutions. Some of the potentials have singularities. The reason why we still get a unique solution is simple: If a solution would approach a singular point of the potential, the energy will become infinitely large, but that is a contradiction to the conservation of energy.

Definition 2.1 (Assumptions on the Hamiltonian system). Let

$$D = \{(\mathbf{Q}, \mathbf{P}) \in \mathbb{R}^{2dN} \mid \mathbf{q}_i \neq \mathbf{q}_j \text{ for } i \neq j\}$$

be the set of non overlapping states. Obviously this set is open.

The potential function $\mathcal{V} : D \to \mathbb{R}$ should have the fulfil

- $\mathcal{V} \in \mathcal{C}^2(D)$,

- $\mathcal{V}$ is bounded form below,

- $\mathcal{V}$ depends only on the positions $\mathbf{Q}$,

- for each sequence $\mathbf{y}_k$, with

$$\mathbf{y}_k \to y \in \partial D \text{ as } k \to \infty,$$

the potential function has a blow-up, i.e.

$$\mathcal{V}(\mathbf{y}_k) \to \infty, \text{ as } k \to \infty.$$

Then $\mathcal{H} : D \to \mathbb{R} : (\mathbf{Q}, \mathbf{P}) \mapsto \mathcal{V}(\mathbf{Q}) + \mathcal{T}(\mathbf{P})$ is a separable Hamilton function. To shorten the notation we may write $\mathbf{y}(t) = (\mathbf{Q}(t), \mathbf{P}(t))$ instead and define

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix},$$

the Hamilton system then transforms to

$$\dot{\mathbf{y}} = -J^1 \nabla \mathcal{H}(\mathbf{y}).$$

Lemma 2.2 (Existence and Uniqueness in the deterministic case). *The solutions of the system defined in Definition 2.1 exists and are unique.*

Proof. We fix some initial states $\mathbf{y}_0 \in D$. Since the Hamiltonian $\mathcal{H}$ is in $\mathcal{C}^2(D)$, the right hand side of the Hamilton system is continuous differentiable, hence locally Lipschitz continuous.

By Picard-Lindelöf a local solution of the initial problem exists, we call it $\mathbf{y}(t) = (\mathbf{Q}(t), \mathbf{P}(t))$. By a Corollary of the Picard-Lindelöf Theorem one of the following three cases holds true.

(i) *Global solution case:* The solution exists for all times $t > 0$.

(ii) *Boundary case:* For some finite time t_+ the solution exists and

$$\lim_{t \to t_+} \mathbf{y}(t) \in \partial M$$

(iii) *Blow-up case:* For some finite time t_+ the solution exists and

$$\lim_{t \to t_+} \left\| \mathbf{y}(t) \right\|_2 = \infty$$

We prove that we are in the first case by contradiction to the conservation of energy.

Conservation of energy:

(1)
$$\begin{aligned}
\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t} &= \sum_{i=1}^{dN} \frac{\partial \mathcal{H}}{\partial q_i} \frac{\mathrm{d}q_i}{\mathrm{d}t} + \sum_{i=1}^{dN} \frac{\partial \mathcal{H}}{\partial p_i} \frac{\mathrm{d}p_i}{\mathrm{d}t} + \frac{\partial \mathcal{H}}{\partial t} \\
&= \sum_{i=1}^{dN} -\frac{\mathrm{d}p_i}{\mathrm{d}t} \frac{\mathrm{d}q_i}{\mathrm{d}t} + \sum_{i=1}^{dN} \frac{\mathrm{d}q_i}{\mathrm{d}t}^T \frac{\mathrm{d}p_i}{\mathrm{d}t} + \frac{\partial \mathcal{H}}{\partial t} = \frac{\partial \mathcal{H}}{\partial t} = 0
\end{aligned}$$

Assume: The boundary case holds true. In this case we can assume that the distance between two particles becomes zero $\lim \mathbf{q}_i(t) - \mathbf{q}_j(t) = \mathbf{0}$ Then also $\mathcal{H} \to \infty$ in contradiction to the conservation of energy (1).

Assume: The blow-up case holds true.

1. Case: $\left\| \mathbf{p}_i(t) \right\| \to \infty$, *as* $t \to t_+ < \infty$.
Since $\mathcal{T} = \sum_{k=1}^{N} \frac{\left\| \mathbf{p}_k \right\|^2}{m_i}$, we have a blow-up in the kinetic energy as well $T(\mathbf{Q}(t), \mathbf{P}(t)) \to \infty$. But since the potential $\mathcal{V}$ is bounded from below we have again a contradiction to the conservation of $\mathcal{H} = \mathcal{T} + \mathcal{V}$ (1).

2. Case: $\left\| \mathbf{q}_i(t) \right\| \to \infty$, *as* $t \to t_+ < \infty$.
By definition of the kinetic energy $\mathcal{T}$ we have the relation $\dot{\mathbf{q}}_i(t) = \frac{\mathbf{p}_i}{m_i}$, hence if the positions increases in finite time towards infinity, then also the inertia does and we are in case 1 again.

In conclusion we are in the global solution case and the claim holds. $\quad\square$

Integration methods of the deterministic Hamilton system are discussed in §3. Before we want to introduce another fundamental equation for molecular dynamics simulations.

2.3. **The Langevin equation.** For $\gamma \in (0, 1)$ and $\sigma \in \mathbb{R}^{dN \times dN}$, the Langevin equation

$$\mathrm{d}\mathbf{x}_i = \mathbf{v}_i \, \mathrm{d}t,$$

$$\mathrm{d}\mathbf{v}_i = \frac{1}{m_i} \mathbf{F}_i \, \mathrm{d}t - \gamma \mathbf{v}_i \, \mathrm{d}t + \sigma \, \mathrm{d}\mathbf{W}_t$$

models the movement of particles in a fluid of constant temperature. We use a different notation to underline that due to the friction this system cannot be written as a Hamiltonian system.

The motivation to include additional terms into our model is, that due to the high number of fluid molecules, we are simply not able to capture the initial conditions. Besides these deterministic models would be by far to large for many practical computations. Instead we consider an ensemble of systems, i.e. we fix only the temperature of the initial states. An experiment ω then represents a particular initial state. Before we give some details on the fixed temperature ensemble we want to present a heuristic approach towards the Langevin equation.

The friction term. Each particle is surrounded by many small fluid molecules. The faster the particles moves, the more fluid molecules are in the path of the particle. Therefore the more collisions are expected. To model these collisions the Langevin equations includes an additional friction term $-\gamma \mathbf{v}_i$,. The coefficient γ can be chosen according to Stokes formula as $\gamma = 6\pi\eta R$, where R denotes radius of the particles and η represents the viscosity of the fluid [LL87, §60].

The diffusion term. The last term represents the influence of thermal fluctuation. As a consequence of the equipartition law

$$\frac{1}{2}m\,\mathrm{E}\left(v^2\right) = \frac{1}{2}kT,$$

the thermal velocity of the particles is given by $\mathrm{E}\left(v^2\right) = \sqrt{\frac{kT}{m}}$. In the case of no interaction between the particles we would assume that

(2) $$\mathrm{Var}\left(\|\mathbf{V}\|\right) = \frac{kT}{m}$$

holds. In the next section we compute the variance of the Langevin equation. This will lead to the relation

$$\sigma_{ii} = \frac{2\gamma kT}{m_i}$$

Mathematical Interlude II. We shortly analyse the case of no intermolecular interactions.

Example 2.3 (Ornstein-Uhlenbeck process , [Eva12], Example 5.6)**.** For simplicity we consider the one dimensional case

$$\begin{cases} \mathrm{d}X = V\,\mathrm{d}t, \\ \mathrm{d}V = -\gamma V\,\mathrm{d}t + \sigma\,\mathrm{d}W, \\ X(0) = X_0, \quad V(0) = V_0. \end{cases}$$

The velocity process. We start with the behaviour of the velocity process. The solution is given by

$$V(t) = e^{-\gamma t}V_0 + \sigma \int_0^t e^{-\gamma(t-s)}\,\mathrm{d}W_s.$$

We use the basic properties in Theorem 1.22 to calculate the expected velocity

$$\mathrm{E}\left(V(t)\right) = e^{-\gamma t}\,\mathrm{E}\left(V_0\right)$$

and

$$\mathrm{E}\left(V(t)^2\right) = \mathrm{E}\left(e^{-2\gamma t}V_0^2 + 2\sigma e^{-\gamma t}V_0 \int_0^t e^{-\gamma(t-s)}\,\mathrm{d}W_s\right.$$

$$\left. + \sigma^2 \left(\int_0^t e^{-\gamma(t-s)}\,\mathrm{d}W_s\right)^2\right)$$

$$= e^{-2\gamma t}\,\mathrm{E}(V_0^2) + 2\sigma e^{-\gamma t}\,\mathrm{E}\left(V_0\right)\mathrm{E}\left(\int_0^t e^{-\gamma(t-s)}\,\mathrm{d}W_s\right)$$

$$+ \sigma^2 \int_0^t e^{-2\gamma(t-s)}\,\mathrm{d}s$$

$$= e^{-2\gamma t}\,\mathrm{E}\left(V_0^2\right) + \frac{\sigma^2}{2\gamma}(1 - e^{-2\gamma t}).$$

By $\mathrm{Var}\left(V(t)\right) = \mathrm{E}\left(V^2(t)\right) - \mathrm{E}\left(V(t)\right)^2$ we get

$$\mathrm{Var}\left(V(t)\right) = e^{-2\gamma t}\,\mathrm{Var}\left(V_0\right) + \frac{\sigma^2}{2\gamma}(1 - e^{-2\gamma t}).$$

Assuming a finite variance of V_0 we get

$$\begin{cases} \mathrm{E}\left(V(t)\right) \to 0 \\ \mathrm{Var}\left(V(t)\right) \to \frac{\sigma^2}{2\gamma} \end{cases}, \text{ as } t \to \infty.$$

In particular we see a balance between friction and random fluctuation. If we insert this result into (2) we get $\frac{\sigma^2}{2\gamma} = \frac{kT}{m}$, hence $\sigma = \frac{2\gamma kT}{m}$.
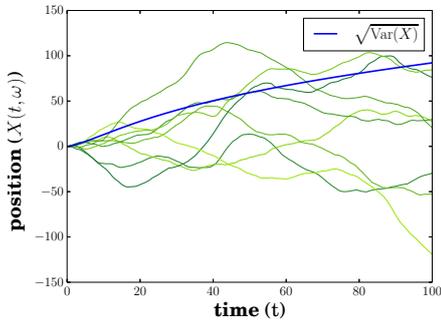
FIGURE 2. Position paths of an Ornstein-Uhlenbeck process.
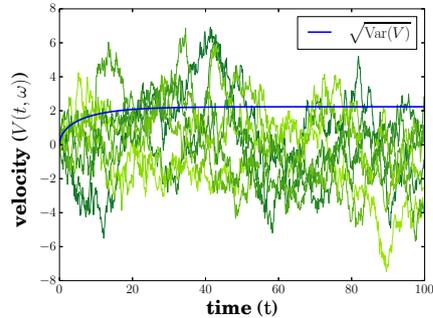
FIGURE 3. Velocities of an Ornstein-Uhlenbeck process.

The position process. Knowing the velocity process, we can analyse the positions. The process is given by

$$X(t) = X_0 + \int_0^t V(s)ds.$$

The expected value is given by

$$\mathrm{E}\left(X(t)\right) = \mathrm{E}\left(X_0\right) + \mathrm{E}\left(V_0\right)e^{-\gamma t}.$$

A lengthy calculations also gives the value of the variance as

$$\mathrm{Var}\left(X(t)\right) = \mathrm{Var}\left(X_0\right) + \frac{\sigma^2}{\gamma^2}t + \frac{\sigma^2}{2\gamma^3}\left(-3 + 4e^{-\gamma t} - e^{-2\gamma t}\right).$$

Some numerical simulations are presented in Figure 2 and Figure 3.
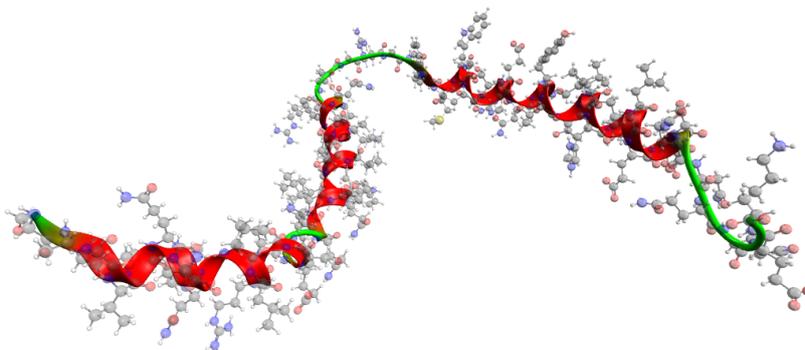
FIGURE 4. A polymer with the label '2mk3' [RLG$^+$14] taken from the RSCB-Database . The red helix represents the secondary structure of this molecule.

2.4. **Polymers.** Polymers are a special class of molecules. These are molecules composed out of many monomers, simple groups of atoms. A more complex example is given in Figure 4. To simulate polymers an usual approach is to model each monomer of the chain as one particle and introduce intermolecular potentials or fixed constrains between the particles. In this thesis we concentrate on a molecular simulation of polymers, i.e. we simulate the molecule without any further simplifications of the model.

3. NUMERICAL METHODS FOR MOLECULAR DYNAMICS

The goal of molecular dynamical simulations is to solve the associated deterministic Hamilton system. The main difficulties are the high dimension of the phase space, the long computation time of the forces and the presence of fast oscillations. Therefore lower order methods are often preferred, since these only require one force evaluation per

step. Our main candidate is the Störmer-Verlet scheme. Later we discuss splitting methods and implementation techniques to calculate the forces. The numerical simulations are computed in high dimensions and with the molecular model of the *Universal Force Field* [RCC$^+$92].

Since we simulate a Hamilton system, a geometric integration approach is reasonable. The book of Hairer, Lubich and Wanner on geometric integration [HLW06] is a fantastic resource for both theory and application of this field. A more specific view on molecular dynamical simulation is presented in [GKZ10].

In general, a good choice is given by the Störmer-Verlet scheme. We will discuss alternatives and point out the reasons why Störmer-Verlet perform so well.

Notation. We only consider one-step methods. The numerical flow Φ_h maps an initial state $\mathbf{x}_0$ to the numerical approximation of the system after time h, i.e. $\Phi_h(x_0) = x_1$. In contrast, the flow φ_t maps an initial state $\mathbf{x}_0$ to the exact solution at time t, i.e. $\varphi_t(x_0) = x(t)$.

3.1. **Nyström methods.** To handle differential equations of the type $\ddot{x} = f(x)$ we apply two Runge-Kutta methods to each component. The resulting methods are called Nyström methods. We first introduce partitioned Runge-Kutta methods, which are applied to slightly more general systems of the form

$$\dot{\mathbf{x}} = g(t, \mathbf{x}, \mathbf{v})$$
$$\dot{\mathbf{v}} = h(t, \mathbf{x}, \mathbf{v}).$$

Definition 3.1 (Partitioned Runge-Kutta Methods). For given coefficients $a_{ij}, \hat{a}_{ij}, b_i$ and $\hat{b}_i$ a partitioned Runge-Kutta method is given by

$$\mathbf{k}_i = g(\mathbf{x}_0 + h\sum_{j=0}^{i-1} a_{ij}\mathbf{k}_j, \mathbf{v}_0 + h\sum_{j=0}^{i-1} \hat{a}_{ij}\mathbf{l}_j)$$

$$\mathbf{l}_i = h(\mathbf{x}_0 + h\sum_{j=0}^{i-1} a_{ij}\mathbf{k}_j, \mathbf{v}_0 + h\sum_{j=0}^{i-1} \hat{a}_{ij}\mathbf{l}_j)$$

$$\mathbf{x}_1 = \mathbf{x}_0 + \sum_{i=0}^{s} b_i\mathbf{k}_i, \quad \mathbf{v}_1 = \mathbf{v}_0 + \sum_{i=0}^{s} \hat{b}_i\mathbf{l}_i.$$

It is convenient to write the coefficients into so called *Butcher Arrays*. In these Arrays the coefficients are arranged as illustrated below.

In the case of partitioned methods we use two arrays to write down the coefficients.

$$
\begin{array}{c|ccc}
c_1 & a_{11} & a_{12} & a_{13} \\
c_2 & a_{21} & a_{22} & a_{23} \\
c_3 & a_{31} & a_{32} & a_{33} \\
\hline
 & b_1 & b_2 & b_3
\end{array}
\qquad
\begin{array}{c|ccc}
\hat{c}_1 & \hat{a}_{11} & \hat{a}_{12} & \hat{a}_{13} \\
\hat{c}_2 & \hat{a}_{21} & \hat{a}_{22} & \hat{a}_{23} \\
\hat{c}_3 & \hat{a}_{31} & \hat{a}_{32} & \hat{a}_{33} \\
\hline
 & \hat{b}_1 & \hat{b}_2 & \hat{b}_3
\end{array}
$$

Since we consider an autonomous ODE system, the c_i coefficients are not important for us. Next we consider a simple looking scheme with various important properties.

Method 3.2 (Störmer-Verlet)**.** An order 2 method is given by the coefficients:

$$
\begin{array}{c|cc}
0 & 0 & \\
1 & 1/2 & 1/2 \\
\hline
 & 1/2 & 1/2
\end{array}
\qquad
\begin{array}{c|cc}
1/2 & 1/2 & \\
1/2 & 1/2 & 0 \\
\hline
 & 1/2 & 1/2
\end{array}
$$

Fitting this method into the framework of partitioned Runge-Kutta methods gives us access to general order conditions. Next we specialize these methods to our second order system of the form $\ddot{\mathbf{x}} = f(t, \mathbf{x}, \dot{\mathbf{x}})$, hence we apply a partitioned Runge-Kutta method to $\dot{\mathbf{x}} = \mathbf{v}, \dot{\mathbf{v}} = f(t, \mathbf{x}, \mathbf{v})$. Due to the simplicity of the first equations some implicit relations cancel. The resulting methods are called Nyström methods. We here restrict the definition to autonomous systems.

Method 3.3 (Autonomous Nyström Methods)**.** Let $c_i, \bar{b}_{ij}, \bar{a}_{ij}$ and $\hat{b}_{ij}, \hat{a}_{ij}$ be real coefficients. A *Nyström methods* for the solution of $\ddot{\mathbf{x}} = f(\mathbf{x}, \dot{\mathbf{x}})$ is given by

$$
\mathbf{l}_i = f(\mathbf{x}_0 + c_i h \mathbf{v}_0 + h^2 \sum_{j=1}^{s} \bar{a}_{ij} \mathbf{l}_j, \mathbf{v}_0 + h \sum_{j=1}^{s} \hat{a}_{ij} \mathbf{l}_j)
$$

$$
\mathbf{x}_1 = \mathbf{x}_0 + h \mathbf{v}_0 + h^2 \sum_{i=1}^{s} \bar{b}_i \mathbf{l}_i, \quad \mathbf{v}_1 = \mathbf{v}_0 + h \sum_{i=1}^{s} \hat{b}_i \mathbf{l}_i.
$$

To transform partitioned Runge-Kutta methods, the additional coefficients can be calculated as

$$
\bar{a}_{ij} = \sum_{k=1}^{s} a_{ik} \hat{a}_{kj}, \quad \bar{b}_i = \sum_{k=1}^{s} b_k \hat{a}_{ki}.
$$

We finaly get an explicit method if we apply the Störmer-Verlet method 3.2 on a second order system.

Method 3.4 (Störmer-Verlet for second order systems)**.**

$$\mathbf{l}_1 = f(\mathbf{x}_0), \qquad\qquad \mathbf{l}_2 = f(\mathbf{x}_0 + h\mathbf{v}_0 + \frac{1}{2}h^2 f(\mathbf{x}_0)) = f(\mathbf{x}_1),$$

$$\mathbf{x}_1 = \mathbf{x}_0 + h\mathbf{v}_0 + \frac{1}{2}h^2 f(\mathbf{x}_0), \qquad\qquad \mathbf{v}_1 = \mathbf{v}_0 + \frac{1}{2}h(f(\mathbf{x}_0) + f(\mathbf{x}_1)).$$

3.2. **Symplectic Integration of Hamilton systems.** The flow of a Hamilton system has some important invariants. For example, the Liouville Theorem implies that volumes in phase space are preserved. Since these invariants are fundamental to each system, we aim at respecting them also numerically. The condition of preserving certain properties is not equivalent to a small local or global error. We will see that many popular numerical methods fail at this point.

Example 3.5 (Hamiltonian oscillator)**.** As a first example we plot the numerical flow and the total energy of three numerical integrators applied to an one-dimensional oscillator

$$\begin{cases} \dot{q} = p \\ \dot{p} = -q. \end{cases}$$

Explicit Euler	**Implicit Euler**	**Symplectic Euler**
$p_{i+1} = p_i - hq_i$	$p_{i+1} = {}^1/_{1+h^2}(p_i - hq_i)$	$p_{i+1} = p_i - hq_i$
$q_{i+1} = q_i + hp_i$	$q_{i+1} = {}^1/_{1+h^2}(q_i + hp_i)$	$q_{i+1} = q_i + hp_{i+1}$
Stepsize: 0.05	*Stepsize: 0.05*	*Stepsize: 0.05*
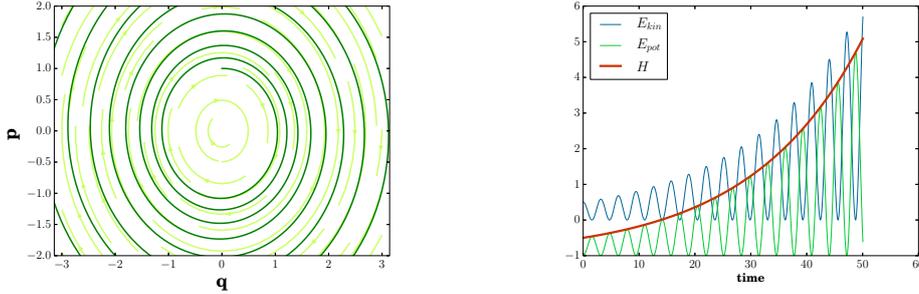
FIGURE 5. The explicit Euler method results in a dilation of the phase space. The energy increases.

All integrators are of first order, the implicit Euler method is even transformed to an explicit method, but only the symplectic variant does nearly preserve the volume in phase space and the total energy. Especially the long time behaviour is better predicted by the symplectic method. Choosing different time steps or higher order Runge-Kutta method does not change the principal behaviour.
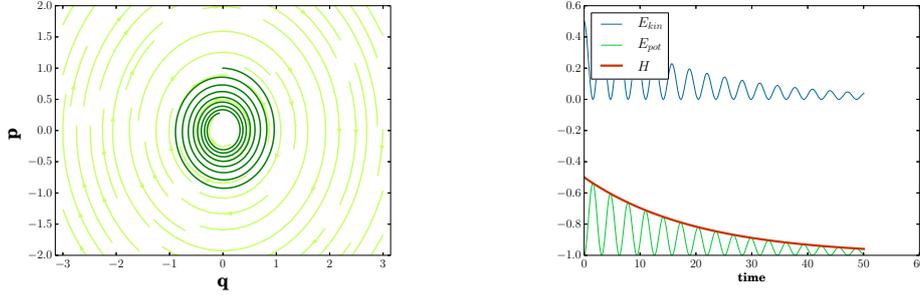
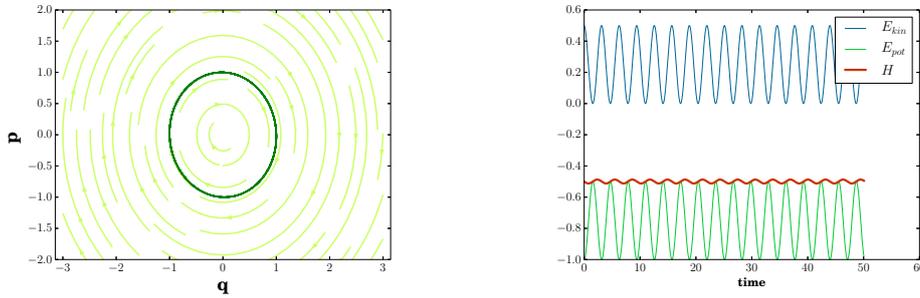FIGURE 6. The implicit Euler method has a reversed behaviour.

FIGURE 7. Finally the symplectic Euler does a good job
in both plots.

After this first observations, we start with the abstract definition of a
symplectic mapping. In a two dimensional phase space a symplectic
map will be exactly a volume preserving diffeomorphism. In higher
dimensions the *sum of projected orientated areas* will be preserved.
There exists a more general concept of symplectic manifolds, we here
only consider the compact manifold D with coordinates arranged like
$(\mathbf{Q}, \mathbf{P})$. Together with the following differential 2-form

$$w^2 = d\mathbf{P} \wedge d\mathbf{Q} = \sum_{i=1}^{N} d\mathbf{p}_i \wedge d\mathbf{q}_i,$$

the pair $(\mathbb{R}^{2dN}, w^2)$ is called a symplectic manifold. A map is called
symplectic if it preserves this structure. We are interested in explicit
formulas. Let $\boldsymbol{\xi} = (\boldsymbol{\xi}^q, \boldsymbol{\xi}^p), \boldsymbol{\eta} = (\boldsymbol{\eta}^q, \boldsymbol{\eta}^p) \in \mathrm{T}_{(Q,P)} M$ be two tangent
vectors at $(\mathbf{Q}, \mathbf{P})$, then the explicit formula for w is given by

$$w(\boldsymbol{\xi}, \boldsymbol{\eta})\big|_{(\mathbf{Q}, \mathbf{P})} = \sum_{j=1}^{dN} \det \begin{pmatrix} \xi_j^q & \eta_j^q \\ \xi_j^p & \eta_j^p \end{pmatrix} = \sum_{j=1}^{dN} \xi_j^q \eta_j^p - \xi_j^p \eta_j^q.$$

In matrix notation we get

$$w(\boldsymbol{\xi}, \boldsymbol{\eta})\big|_{(\mathbf{Q}, \mathbf{P})} = \boldsymbol{\xi}^T J \boldsymbol{\eta}, \quad \text{where } J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.$$

Since w does not depend on the position in phase space $(\mathbf{Q}, \mathbf{P})$, we can omit the position in our notation and just write w instead of $w\big|_{(\mathbf{Q},\mathbf{P})}$. Therefore the value of w will be invariant under a differentiable map if

$$w(\boldsymbol{\xi}, \boldsymbol{\eta}) = w(\ \mathrm{D}g(\mathbf{Q}, \mathbf{P})\boldsymbol{\xi},\ \mathrm{D}g(\mathbf{Q}, \mathbf{P})\boldsymbol{\eta})$$

holds.

Definition 3.6 (Symplectic Mapping)**.** Let $g : U \to \mathbb{R}^{2dN}$ be a differentiable map, we call it *symplectic*, if

$$\mathrm{D}g(\mathbf{Q}, \mathbf{P})^T \circ J \circ \mathrm{D}g(\mathbf{Q}, \mathbf{P}) = J$$

holds.

Let φ_t be the flow of a Hamilton system. Using the matrix J from above we can rewrite the Hamilton system as

$$\dot{\mathbf{y}} = J^{-1} \nabla \mathcal{H}(\mathbf{y})$$

Theorem 3.7 (Poincare 1899, [HLW06, Theorem VI.2.4])**.** *Let $\mathcal{H}(\mathbf{Q}, \mathbf{P})$ be a twice continuously differentiable function on $U \subset \mathbb{R}^{dN}$ open. Then, for each fixed time t, the flow φ_t is a symplectic transformation wherever it is defined.*

Proof. The derivative of $\varphi_t(\mathbf{Q}_0, \mathbf{P}_0)$ with respect to the initial value solves the variational problem

$$\dot{\psi}(t) = -J\,\mathrm{D}^2\,\mathcal{H}(\psi(t)), \quad \psi(0) = I$$

Due to the assumptions, a unique solution exists. We get $\mathrm{D}\varphi_t = \psi$.

For $t = 0$ we have $\mathrm{D}\varphi_t^T J \,\mathrm{D}\varphi_t = J$ by the initial conditions of the variational problem. Now we will prove that the left term is constant. In the last expression we will use $(-J)^T = J$.

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{D}\varphi_t^T J \,\mathrm{D}\varphi_t\right) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\psi(t)^T J \psi(t)\right) \\
&= \dot{\psi}(t)^T J \psi(t) \qquad\qquad + \psi(t)^T J \dot{\psi}(t) \\
&= \left(-J\,\mathrm{D}^2\,\mathcal{H}(\Phi(t))\right)^T J\psi(t) + \psi(t)^T J\left(-J\,\mathrm{D}^2\,\mathcal{H}(\psi(t))\right) \\
&= \mathrm{D}^2\,\mathcal{H}(\psi(t)) J J \psi(t) \qquad - \psi(t)^T J J\,\mathrm{D}^2\,\mathcal{H}(\psi(t)) \\
&= 0
\end{aligned}
$$

In the last expression we used $(-J)^T = J$. In conclusion the expression is constant which gives the claim. $\qquad\square$

Having seen that the Hamilton flow is symplectic, we are now could ask the same question for any numerical integrator.

Definition 3.8 (Symplectic scheme)**.** We call a numerical integrator symplectic if for any step size $h > 0$ the numerical flow Φ_h is symplectic.

Example 3.9 (Symplecticity of the symplectic Euler and Störmer-
-Verlet schemes)**.** The symplectic Euler method and the Störmer-Verlet
scheme are both symplectic.

Before we compare various methods we will turn our attention towards
backward analysis of symplectic methods. Symplectic methods will
not conserve the total energy, but conserve the energy of a modified
Hamilton equation. We will see that the modified Hamilton function
could be written in terms of the original Hamiltonian.

3.3. **Backward error analysis for symplectic methods.** Our first
experiments already pointed out that symplectic methods also preserve
the Hamiltonian of a system approximately. We will prove that sym-
plectic methods in fact solve a modified Hamiltonian system, therefore
preserve the modified Hamiltonian.

Here we again follow Hairer et al [HLW06, Section IX.3].

Backward error analysis. Instead of considering the error between nu-
merical solutions and the exact solutions, in backward analysis we con-
sider the difference between the exact solutions of the original equation
and the exact solutions of a modified equation. Let $\mathbf{y}_n$ be the numerical
solution, for a step size h then the modified equation is defined as

$$\begin{cases} \dot{\widetilde{\mathbf{y}}} & = f(\widetilde{\mathbf{y}}) + h f_2(\widetilde{\mathbf{y}}) + h^2 f_3(\widetilde{\mathbf{y}}) + \dots \\ \dot{\widetilde{\mathbf{y}}}(nh) & = \mathbf{y}_n \text{ for all } j = 0, 1, \dots. \end{cases}$$

The solution should coincides with the numerical solution, but we first
need to find the missing functions f_j. We cannot expect the series
$f + h f_2 + h^2 f_3 + \dots$ to converge, instead the series is considered as a
asymptotic series and we truncate at some order. The next theorems
will show that a global modified equation for symplectic methods can
be found and is Hamiltonian again.

Theorem 3.10 (Order of the modified equation)**.** *Suppose that the
method* $\mathbf{y}_{n+1} = \Phi_h(\mathbf{y}_n)$ *is of order p, i.e.*

$$\Phi_h(\mathbf{y}) = \phi_h(\mathbf{y}) + h^{p+1} \delta_{p+1}(\mathbf{y}) + \mathcal{O}(h^{p+2}),$$

where $\phi_t(\mathbf{y})$ *denotes the exact flow of* $\dot{\mathbf{y}} = f(\mathbf{y})$, *and* $h^{p+1} \delta_{p+1}(\mathbf{y})$ *the
leading term of the local truncation error. The modified equation then
satisfies*

$$\dot{\widetilde{\mathbf{y}}} = f(\widetilde{\mathbf{y}}) + h^p f_{p+1}(\mathbf{y}) + h^{p+1} f_{p+2}(\widetilde{\mathbf{y}}) + \dots, \qquad \widetilde{\mathbf{y}}(0) = \mathbf{y}_0$$

with $f_{p+1} = \delta_{p+1}(\widetilde{\mathbf{y}})$.

Proof. First we need to find candidates for the functions f_j, therefore we write down a Taylor-expansion of the solution of the modified equation

$$\widetilde{\mathbf{y}}(t+h) = \mathbf{y} + h\left(f(\mathbf{y}) + hf_2(\mathbf{y}) + h^2 f_3(\mathbf{y}) + \dots\right)$$

$$(3) \qquad + \frac{h^2}{2!}\left(f'(\mathbf{y}) + hf_2'(\mathbf{y}) + \dots\right)\left(f(\mathbf{y}) + hf_2(\mathbf{y}) + \dots\right) + \dots.$$

We also expand the numerical flow as a Taylor-series, where d_j denotes some function normally composed by terms in $f, f', f'', \dots$

$$\Phi_h(\mathbf{y}) = \mathbf{y} + hd_1(\mathbf{y}) + h^2 d_2(\mathbf{y}) + h^3 d_3(\mathbf{y}) + \dots.$$

Next we compare the coefficients of both expansions and simply choose the functions f_j for $j = 2, \dots, p$ such that the coefficients coincide in both extensions coincide. For example the first defining equations are

$$f_2(\mathbf{y}) = d_2(\mathbf{y}) - \frac{1}{2!}f'(\mathbf{y})f(\mathbf{y})$$

$$f_3(\mathbf{y}) = d_3(\mathbf{y}) - \underbrace{\frac{1}{3!}(f''(\mathbf{y})f(\mathbf{y})f(\mathbf{y}) + f'(\mathbf{y})f'(\mathbf{y})f(\mathbf{y}))}_{=\frac{1}{3!}\frac{\partial^3 \varphi_h(\mathbf{y})}{\partial h^3}|_{h=0}}$$

$$- \underbrace{\frac{1}{2!}(f'(\mathbf{y})f_2(\mathbf{y}) + f_2'(\mathbf{y})f(\mathbf{y}))}_{:=R_3(f,f_2)(\mathbf{y})}.$$

We note that the second terms on the right hand side are equal to the Taylor expansion of the exact solution. We define $R_j(f, f_2, \dots, f_{j-1})$ as the additional terms on the left hand side, which depend only on $f, f_2, \dots, f_{j-1}$. We note that $R_j(f, 0, 0, 0, \dots) = 0$.

Since Φ_h is an order p method we get for $j = 1, 2, \dots, p$

$$d_j(\mathbf{y}) = \frac{1}{j!}\frac{\partial^j \Phi_h(\mathbf{y})}{\partial h^j}\bigg|_{h=0} = \frac{1}{j!}\frac{\partial^j \varphi_h(\mathbf{y})}{\partial h^j}\bigg|_{h=0}$$

To shorten notation we denote the additional terms in the expansion of $\widetilde{\mathbf{y}}$ in (3) by $R_j(f, f_2, \dots, f_{j-1})$. We get

$$f_j = d_j - \frac{1}{j!}\frac{\partial^j \varphi_h(\mathbf{y})}{\partial h^j}\bigg|_{h=0} - R_j(f, \dots, f_{j-1})(\mathbf{y}).$$

For $j = 2$ the rest term $R_2(f)$ is zero, therefore $f_2 = 0$.
Now by induction we get that all f_j vanish for $j = 2, \dots, p$, since $R_j(f, 0, 0, 0, \dots) = 0$ for all $j = 2, \dots, p$. $\square$

For the next theorem on local existence of a Hamiltionian modified equation we first need the following lemma. Note that these theorems are taken from [HLW06], we only sometimes restricted the domain is additionally restricted to D, which is mentioned in the last paragraph of Section IX.3.1 by the authors therein.

Lemma 3.11 (Integrability Lemma, [HLW06] VI.2.7). *Let $D \subset \mathbb{R}^n$ be open and $f : D \to \mathbb{R}^n$ be continuously differentiable, and assume that the Jacobian $\mathrm{D}f(\mathbf{y})$ is symmetric for all $\mathbf{y} \in D$. Then, for every $\mathbf{y}_0 \in D$ there exists a neighbourhood and a function $H(\mathbf{y})$ such that*

$$f(\mathbf{y}) = \nabla \mathcal{H}(\mathbf{y})$$

on this neighbourhood. In other words, the differential from $f_1(\mathbf{y})\,\mathrm{d}\mathbf{y}_1 + \ldots + f_n(\mathbf{y})\,\mathrm{d}\mathbf{y}_n = \mathrm{d}\mathcal{H}$ is a total differential.

Proof. Assume $\mathbf{y}_0 = 0$, by shifting the domain D. Consider a ball around y_0 which is contained in D. On this ball we define

$$H(y) = \int_0^1 \mathbf{y}^T f(t\mathbf{y})\,\mathrm{d}t + \mathrm{const.}$$

Due to the symmetry assumption $\frac{\partial f_i}{\partial y_k} = \frac{\partial f_k}{\partial y_i}$ we have

$$\frac{\partial H}{\partial y_k}(y) = \int_0^1 \left(f_k(t\mathbf{y}) + \mathbf{y}^T \frac{\partial f}{\partial y_k}(t\mathbf{y})t \right) \mathrm{d}t = \int_0^1 \frac{\mathrm{d}t f_k(t\mathbf{y})}{\mathrm{d}t}\,\mathrm{d}t = f_k(\mathbf{y}),$$

which proves the statement. $\qquad\square$

Theorem 3.12 (Existence of a Local Modified Hamiltonian, [HLW06] Theorem IX.3.1). *If a symplectic method $\Phi_h(y)$ is applied to a Hamiltonian system with a smooth Hamiltonian $H : \mathbb{R}^n \to \mathbb{R}$, the the modified equation is also Hamiltonian. More precisely, there exists smooth function $H_j : \mathbb{R}^n \to \mathbb{R}$ for $j = 2, 3, \ldots$, such that*

$$f_j(\mathbf{y}) = J^{-1}\nabla H_j(\mathbf{y}).$$

Proof. The proof is done by induction again. For $j = 1$ the claim is true since, we have $f_1 = f$ in the standard modified equation, hence $f_1(\mathbf{y}) = J^{-1}\nabla H(\mathbf{y})$

Now assume that $f_j = J^{-1}\nabla H_j$ for $j = 1, 2, \ldots, r$.
We consider the truncated modified equation

$$\dot{\tilde{\mathbf{y}}} = J^{-1}\nabla(\mathcal{H}(\mathbf{y}) + h\mathcal{H}_2(\mathbf{y}) + \ldots h^{r-1}\mathcal{H}_r(\mathbf{y})).$$

Let $\varphi_{r,t}$ denote the flow of this truncated system, then numerical flow satisfies

$$\Phi_h(\mathbf{y}_0) = \varphi_{r,t}(\mathbf{y}_0) + h^{r+1}f_{r+1}(\mathbf{y}_0) + \mathcal{O}(h^{r+2}),$$

and also

$$\Phi_h'(y_0) = \varphi_{r,t}'(y_0) + h^{r+1}f_{r+1}'(y_0) + \mathcal{O}(h^{r+2}).$$

We now use the identity $\varphi_{r,h}'(y_0) = I + \mathcal{O}(h)$. Since the method Φ_h is symplectic we have

$$J = \Phi_h'(y_0)^T J \Phi_h'(y_0) = J + h^{r+1}\left(f_{r+1}'(y_0)^T J f_{r+1}'(y_0) \right) + \mathcal{O}(h^{r+2}).$$

Consequently, the matrix $Jf'_{r+1}(y)$ is symmetric and the existence of $H_{r+1}(y)$ satisfying $f_{r+1}(y) = J^{-1}\nabla H_{r+1}(y)$ follows from the Integrability Lemma 3.11. □

Example 3.13 (Modified Hamiltonian of the Störmer-Verlet scheme)**.** Since the proof is in principle constructive, it is possible to calculate the modified Hamiltonians for a given scheme. Explicit formulas for the modified Hamiltonian, derived by the theory of Butcher trees, are available [HLW06, Theorem IX.9.8]. The resulting modified Hamiltonian is given by

$$\widetilde{\mathcal{H}}(\mathbf{Q},\mathbf{P}) = \mathcal{H}(\mathbf{Q},\mathbf{P}) + h^2\Big(-\frac{1}{24}(M^{-1}\mathbf{P})^T\frac{\partial^2 \mathcal{V}}{\partial q^2}(\mathbf{Q},\mathbf{P})M^{-1}\mathbf{P}$$
$$+\frac{1}{12}(\nabla_Q\mathcal{V}(\mathbf{Q}))^T M^{-1}\nabla_Q\mathcal{V}(\mathbf{Q})\Big) + \dots$$

3.4. **Time step control.** Changing the time step size during the simulation will still remain the symplecticity of the numerical method. In practice time step control can lead to a significant change of the modified equation and therefore the global error in energy can become even worst. The next example illustrates the effect.

Example 3.14 (Bad time steps)**.** We again consider the pendulum system. To illustrate the effect of changing time step sizes we compare a the following two time stepping strategies

$$h_1(t) = 1, \qquad h_2(t) = \begin{cases} 1, \text{ if } t\,\mathrm{mod}\,2 > 0.5 \\ 0.73, \quad \text{else.} \end{cases}$$

On a first view the second choice seems to be more accurate, but in fact simulation results show that the contrary is true.
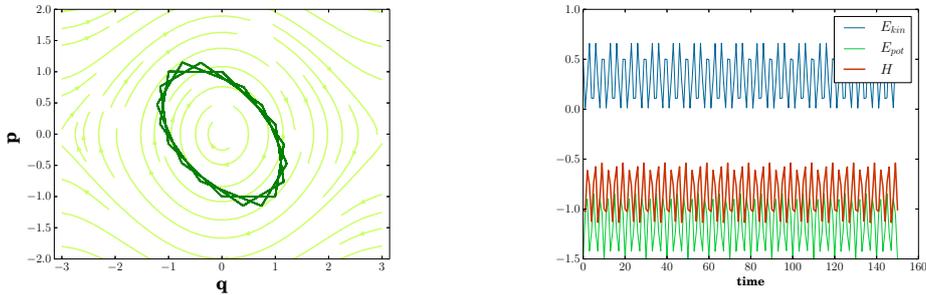
FIGURE 8. Due to the large step size, the solution is not correct, but the energy is still stable.

It is possible to develop stable adaptive energy conserving methods, but special care is necessary.
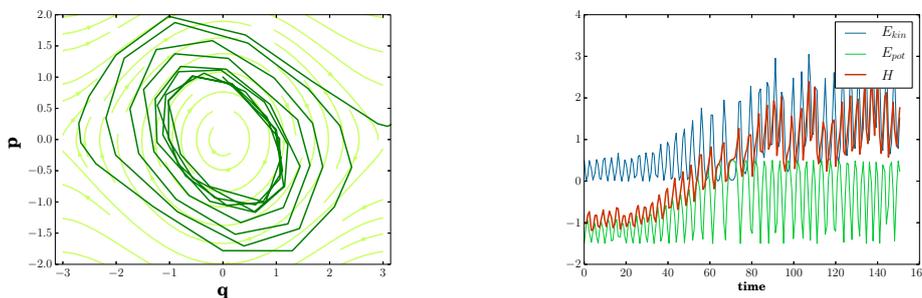
FIGURE 9. The second strategy h_2 fails totally at this example.

3.5. **Higher order schemes.** There exists higher order symplectic Nyström methods. Since the frequencies of a molecular system are usually high compared to the time integration step, higher order schemes are no real benefit, since the need more force computations. We do not elaborate on this, but refer to [GKZ10] and [HLW06].

3.6. **The Impuls method / RESPA-method.** In our final Hamilton system for a molecule oscillations with different frequencies arise. Since our time steps have to be smaller than the highest frequency we are rather limited in choosing a step size. We could avoid this by using a splitting method. This method is called the Impuls method [HLW06]. We will split the force into two forces $f = f^{[fast]} + f^{[slow]}$. Let $\Phi_h^{[fast]}, \Phi_h^{[slow]}$ denote the numerical flow function of the Störmer-Verlet scheme 3.2 applied to the fast force and the slow force respectively.

Method 3.15 (Impuls method)**.** For a fixed number $n > 0$ the Impuls method is composed as

$$\Phi_h^{Impuls} = \Phi_{h/2}^{[slow]} \circ \left(\Phi_{h/n}^{[fast]} \right)^n \circ \Phi_{h/2}^{[slow]}.$$

In our example of a molecule simulation, we could set

$$f^{[fast]} = -\nabla V_{\text{bonded}}, \quad f^{[slow]} = -\nabla V_{\text{non-bonded}}.$$

3.7. **Implementation.** Since the runtime of any molecular dynamic simulation will be governed by the speed of the force calculation, we will give some details on how to implement a force field effectively and which simplifications are common. Main references for this section are [GKZ10] and [All04].

3.7.1. *Pair interaction, Cut-off and Linked Cell method.* One of the main issues is the Lennard-Jones potential, since it requires to compute the distances between all pairs of atoms. Newtons third law implies $F_{ij} = -F_{ji}$, thus only one gradient calculation for each pair is required. If we additionally modify the potential such that $-\nabla V_{\text{L-J}}(r) = 0$ for $r >$

$r_\text{cut-off}$, the runtime can be optimized further. The modified potential function is then

$$V_\text{L-J,Cut} = \begin{cases} V_\text{L-J}(r), & \text{if } r < r_\text{cut-off}, \\ V_\text{L-J}(r_\text{cut-off}), & \text{if } r > r_\text{cut-off}. \end{cases}$$

To deicide effectively for which pair the gradient must be computed the method of *linked cells* is reasonable. We decompose the domain into small cubes, then only the interactions between particles of neighbouring cells are taken into account (see Figure 10). In [GKZ10] this method is called the *linked cell method*, since linked list structures are useful to store the particles of each cell in. At each frame (or less of-
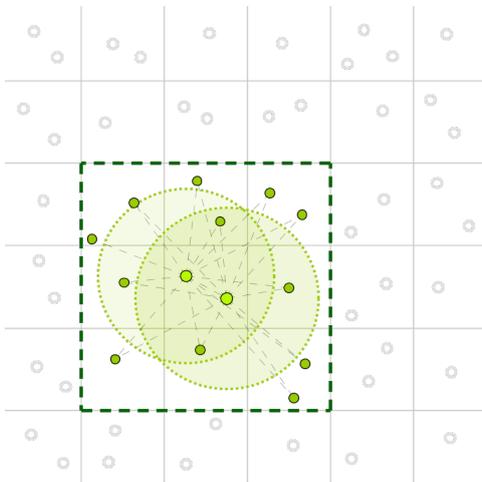
FIGURE 10. The cut-off balls are included in the neighbourhood of the current cell.

ten) one builds up the linked list structure, which takes $\mathcal{O}(N)$ time. If the particles are uniformly distributed the runtime has in the three-dimensional case the form

$$c_1 \cdot N + c_2 \cdot \left(8 \cdot \frac{N}{N_{cells}}\right)^2 \cdot N_{cells} + c_3,$$

where c_1, c_2, c_3 are implementation depended coefficients.

To obtain an energy conserving scheme the edges of the cells have to be larger than the cut-off radius $r_\text{cut-off}$ to ensure the continuity of $\mathcal{H}$.

3.8. **Numerical results for the deterministic model.** Now we present simulation result of full molecule (see Figure 11) to ensure that the theory of symplectic integration has practical results.
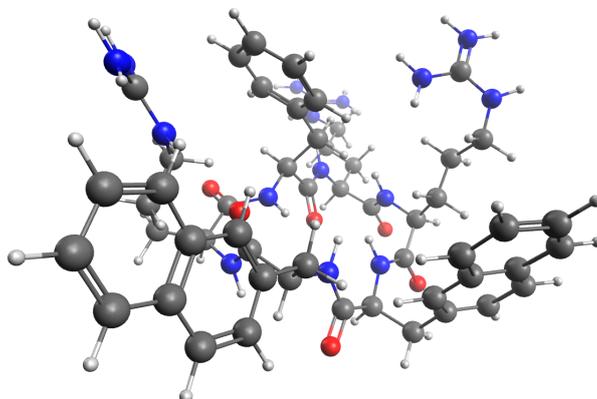
FIGURE 11.   A randomly chosen molecule with the PDB ID "*1skl*" [AWS⁺05] from the RCSB online database. The molecule contains 144 atoms and 149 bonds.

explicit Euler
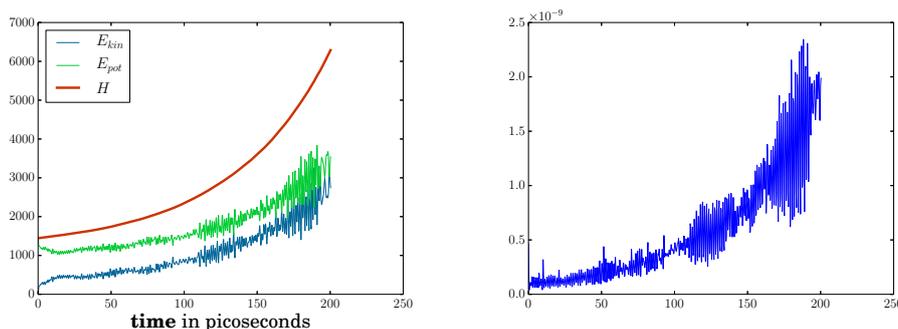
FIGURE 12. Both the energy and the local errors increase exponentially.   Therefore even small time steps will not lead to a stable behaviour for long time scales. In the simulation high frequencies arise and the atom distances are critical small.

To show the particular strength we repeat these simulations using a larger time step. Here we only compare both euler methods.

4. Quasi-symplectic methods for Langevin Equations

In this section we present the generalisation of symplectic methods for numerical solutions of SDEs. The concept of quasi symplectic methods has been taken from a publication of Milstein et al [MT03].

The Euler-scheme has a stochastic counterpart, namely the Euler-Maruyama-scheme. We also obtain a Verlet like integrator using operator splitting.

Symplectic Euler
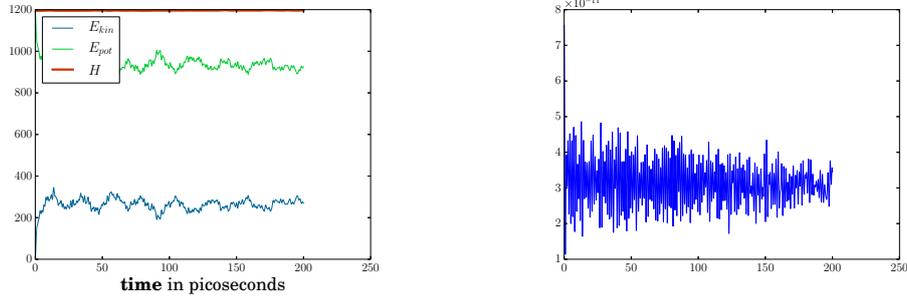
FIGURE 13. The symplecitc Euler method does nearly preserve the energy and has a stable behaviour.

Verlet scheme
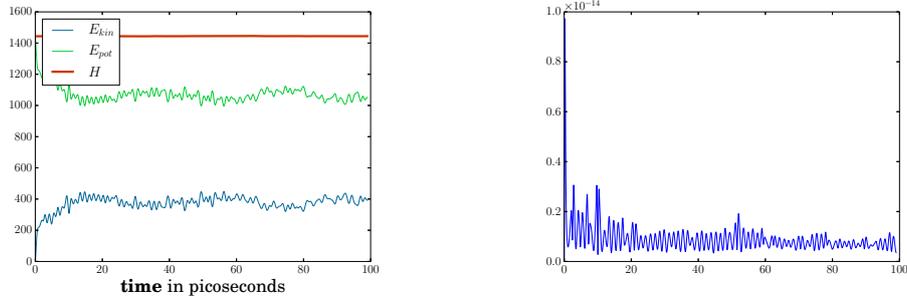
FIGURE 14. Finally the Verlet method could archive a smaller local error with the same number of force calculations.
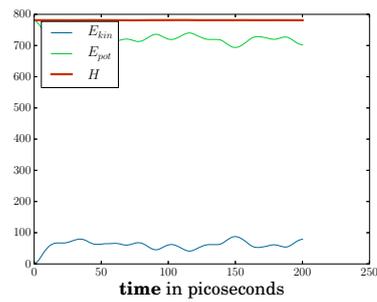
Explicit Euler, $h = 10^{-2}$ *Symplectic Euler, $h = 10^{-1}$*
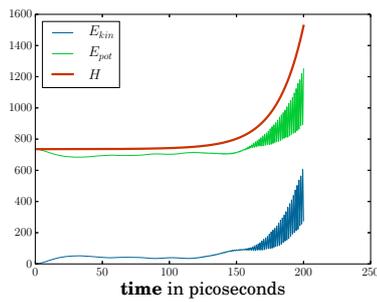
FIGURE 15. A step size of $h = 10^{-2}$ ns leads to an explosion of the explicit euler approximation, whereas the symplectic method still remains stable with a larger step size of $h = 10^{-1}$ ns.

Method 4.1 (Euler-Maruyama). An approximate of the solutions of $dX = a(X,t)\,dt + b(X,t)\,dW_t$ is given by

$$X_1 = X_0 + a(X_0,t)h + b(X,t)\sqrt{h}\Delta_h.$$
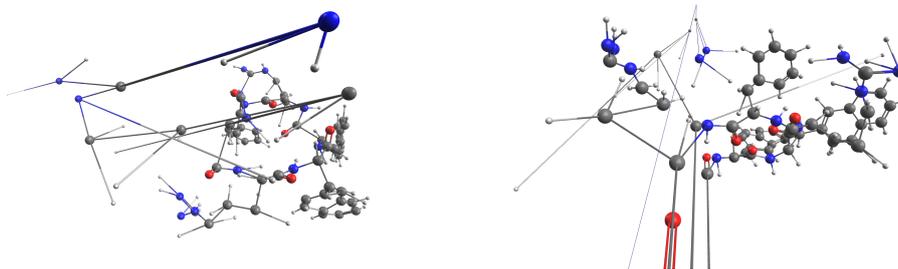
FIGURE 16. Outtakes: That is how it looks like, when the energy is exploding, i.e. the bonds

The additional term is a first order approximation of the increments of the Brownian motion

$$W_{t+h} - W_t \sim \mathcal{N}(0, h)$$

where Δ_h are normal distributed random numbers.

Definition 4.2 (Mean square consistency). Let $X(t, \omega)$ be the exact solution of a given SDE and let $X_h(t, \omega)$ be the numerical approximation. We call a method mean square consistent of order $p \in \mathbb{R}$ if

$$\mathrm{E}\left(|X_h - X|^2\right) < K \cdot h^p.$$

There are other concepts of consistency as well. In particular, mean square consistency does insure a path wise approximation.

4.1. **First order quasi-symplectic methods.** We aim at solving the Langevin equation. In order to reuse ideas from the last section on symplectic methods we rewrite the Langevin equation in a Hamilton like notation. Let M be the mass matrix with diagonal entries $M_{3i,3i} = M_{3i+1,3i+1} = M_{3i+2,3i+2} = m_i$. In the new variables, the equation reads

$$d\mathbf{Q} = \nabla_P \mathcal{H}(\mathbf{Q}, \mathbf{P})\, dt$$
$$d\mathbf{P} = -\nabla_Q \mathcal{H}(\mathbf{Q}, \mathbf{P})\, dt - \gamma M \mathbf{P}\, dt + \sigma\, d\mathbf{W}_t.$$

In the deterministic case, symplectic methods where necessary to preserve the fundamental symplectic structure. Since the flow of the Langevin equation is not symplectic, we introduce the idea of quasi-symplectic methods from Milstein [MT03]. Before explaining the idea we present the definition.

Definition 4.3 (Quasi symplectic methods, [MT03]). We call a numerical method $\Phi_h : \mathbb{R}^{2dN} \times \Omega$ *quasi symplectic* if

   (i) Φ_h is symplectic in the frictionless case $\gamma = 0$,

   (ii) $\mathrm{D}_{(Q,P)}\, \Phi_h(Q, P, \omega)$ does not depend on Q, P.

The second requirement in Definition 4.3 is motivated by the contraction of the space under the Langevin flow. Let $\varphi_t(\cdot, \cdot, \omega)$ be the flow of the Langevin equation and let D_0 be some closed bonded domain in phase space then the volume contracts due to the transformation formula

$$\int_{\varphi_t(D_0, \omega)} \mathrm{d}\mathbf{Q} \, \mathrm{d}\mathbf{P} = \int_{D_0} \left| \mathrm{D}_{(Q,P)} \, \varphi_t(\mathbf{Q}, \mathbf{P}, \omega) \right| \mathrm{d}\mathbf{Q} \, \mathrm{d}\mathbf{P}.$$

The Jacobian in the case of the Langevin equation is given by

$$\mathrm{D}_{(X,V)} \, \varphi_t = \mathrm{e}^{-\gamma \, t \, r(M) \cdot (t - t_0)} \, .$$

The calculation is similar to the proof for the symplecticity of Hamilton flows in Theorem 3.7. The existence of the Jacobian is granted by the diffeomorphism Theorem A.1. For the natural assumptions $\gamma > 0$ and $\mathrm{trace}(M) > 0$ we can observe a contraction in phase space. To approximate geometric quantities, the second requirement should at least ensure that the whole phase space is contracting equally under the numerical flow.

The idea to get such schemes is to apply an operator splitting like

$$(I) \quad \begin{cases} \mathrm{d}\mathbf{Q}_I = \nabla_P \mathcal{H} \, \mathrm{d}t \\ \mathrm{d}\mathbf{P}_I = -\nabla_Q \mathcal{H} \, \mathrm{d}t + \sigma \, \mathrm{d}W_t \end{cases} \qquad (II) \quad \begin{cases} \mathrm{d}\mathbf{Q}_{II} = \mathbf{0} \, \mathrm{d}t \\ \mathrm{d}\mathbf{P}_{II} = -\gamma \mathbf{P}_{II} \, \mathrm{d}t \end{cases} \, .$$

The first system is a Hamiltonian system with additive noise. For systems of this type symplectic first order mean-square methods exist [MRT02, Section 4]. The second system can be solved exactly.

Lemma 4.4 (Quasi-symplectic first order schemes, [MT03, Lemma 2.1]). *Let Φ^I be a symplectic method of first mean-square order for (I) and Φ^{II} be a first order method for (II). Then*

$$\Phi^I_{h/2} \circ \Phi^{II}_h \circ \Phi^I_{h/2}$$

and

$$\Phi^I_h \circ \Phi^{II}_h$$

are a quasi-symplectic first mean-square order method.

A natural choice for Φ^I is given by the stochastic version of the Störmer-Verlet scheme.

Method 4.5 (Stochastic Störmer-Verlet, [MRT02, Method 4.1]). For $0 \leq \alpha \leq 1$ the scheme defined by

$$\mathbf{Q}_\alpha = \mathbf{Q}_0 + \alpha h \nabla_P \mathcal{H}(\mathbf{Q}_0, \mathbf{P}_0),$$
$$\mathbf{P}_{\det} = \mathbf{P}_0 - h \nabla_Q \mathcal{H}(\mathbf{Q}_\alpha, \mathbf{P}_0),$$
$$\mathbf{Q}_1 = \mathbf{Q}_\alpha + (1 - \alpha) h \nabla_P \mathcal{H}(\mathbf{Q}_\alpha, \mathbf{P}_1),$$
$$\mathbf{P}_1 = \mathbf{P}_0 + \sigma \Delta_h,$$

is of first mean-square order and symplectic. For $\alpha = \,^1\!/_2$ we call the method the *stochastic Störmer-Verlet* scheme. Taking $\alpha = 1$ gives the *stochastic symplectic Euler* method.

Method 4.6 (Quasi-symplectic Euler)**.** Let Φ_h^I denote the stochastic symplectic Euler method 4.5 and let Φ^{II} be denote the explicit Euler method. We then call

$$\Phi_h^I \circ \Phi_h^{II}$$

the *quasi-symplectic Euler* scheme.

Method 4.7 (Quasi-symplectic Störmer-Verlet)**.** Let Φ_h^I denote the stochastic Störmer-Verlet Method 4.5 and let Φ^{II} be the exact flow. We then call

$$\Phi_{h/2}^I \circ \Phi_h^{II} \circ \Phi_{h/2}^I$$

the *quasi-symplectic Störmer-Verlet* scheme.

4.2. **Numerical examples.**

Example 4.8 (Stochastic pendulum)**.** To demonstrate the phase space contraction under different schemes we consider the following system

$$\begin{cases} \mathrm{d}q = -p\,\mathrm{d}t \\ \mathrm{d}p = -sin(q)\,\mathrm{d}t - \gamma p\,\mathrm{d}t + \sigma\,\mathrm{d}W_t. \end{cases}$$

We take the parameters $\gamma = 0.1, \sigma = 0.1$ and step size $h = 0.1$.

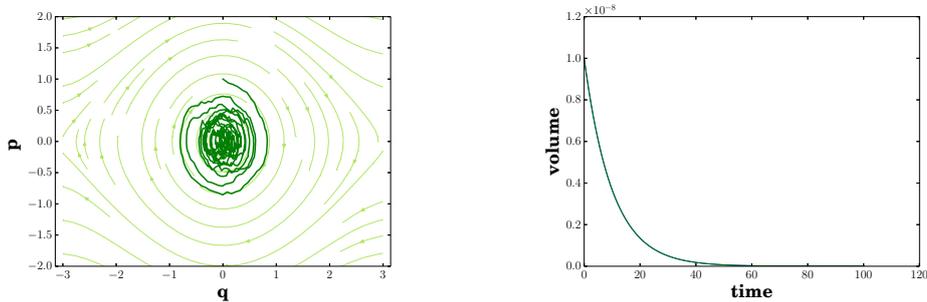*Euler-Maruyama*
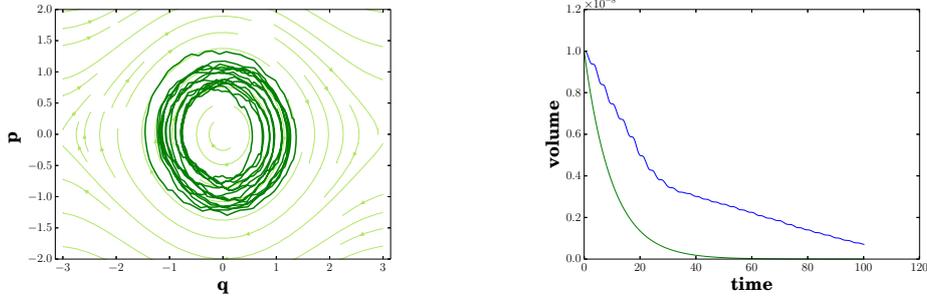
Quasi-symplectic Euler
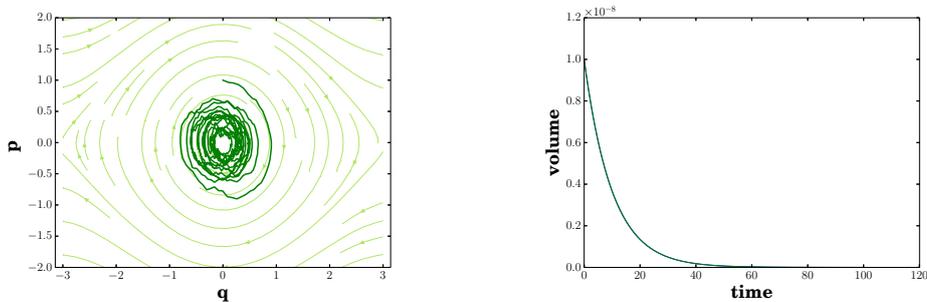
Quasi-symplectic Verlet

FIGURE 17.   The contraction under the quasi-sympectic schemes does perfectly fulfil the analytical prediction, whereas the Euler-Maruyama schemes fails at this point.
.

5. CONCLUSION

The main focus of the thesis relies on symplectic integration. In both, simple test cases as in large complex systems, these methods have proven their effectiveness. We presented their mathematical foundation in the deterministic case and introduced the generalisation of quasi-symplectic methods.

ACKNOWLEDGEMENT

APPENDIX A. DIFFEOMORPHISMTHEOREM

Theorem A.1 (Diffeomorphismtheorem, [KS12, Page 397])**.** *Suppose that the coefficients $a(\mathbf{x}, t), b(\mathbf{x}, t)$ satisfy the global Lipschitz and linear*

growth conditions

$$\left\|a(\mathbf{x},t) - a(\mathbf{y},t)\right\| + \left\|b(t,\mathbf{x}) - b(t,\mathbf{y})\right\| \leq K\left\|\mathbf{x} - \mathbf{y}\right\|,$$

$$\left\|a(t,\mathbf{x})\right\|^2 + \left\|b(t,\mathbf{x})\right\|^2 \leq K^2(1 + \left\|\mathbf{x}\right\|^2),$$

for every $0 \leq t < \infty, \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^n$, where K is a positive constant. Then there exist a strong solution X of

$$\mathrm{d}X = a(X_t,t)\,\mathrm{d}t + b(X_t,t)\,\mathrm{d}W_t.$$

If additionally the coefficients a,b have bounded and continuous derivatices of all order up to $k \geq 1$, then there exist a version $\widetilde{X}$ of X such that for every $t \geq 0$ the map

$$\mathbf{x} \mapsto \widetilde{X}_t(\mathbf{x},\omega)$$

is almost surely a $\mathcal{C}^{k-1}$ diffeomorphism.

References

[All04]   Michael P Allen, *Introduction to molecular dynamics simulation*, Computational soft matter: from synthetic polymers to proteins **23** (2004), 1–28.

[AWS+05]   Christian Appelt, Axel Wessolowski, J Arvid Söderhäll, Margitta Dathe, and Peter Schmieder, *Structure of the antimicrobial, cationic hexapeptide cyclo (rrwwrf) and its analogues in solution and bound to detergent micelles*, Chembiochem **6** (2005), no. 9, 1654–1662.

[Ein05]   A Einstein, *On the movement of small particles suspended in stationary liquids required by the molecular-kinetic theory of heat*, Annalen der Physik **17** (1905), 549–560.

[Eva12]   L.C. Evans, *An introduction to stochastic differential equations*, American Mathematical Society, 2012.

[FS01]   D. Frenkel and B. Smit, *Understanding molecular simulation: From algorithms to applications*, Computational science series, Elsevier Science, 2001.

[GKZ10]   M. Griebel, S. Knapek, and G. Zumbusch, *Numerical simulation in molecular dynamics: Numerics, algorithms, parallelization, applications*, Texts in Computational Science and Engineering, Springer Berlin Heidelberg, 2010.

[HLW06]   E. Hairer, C. Lubich, and G. Wanner, *Geometric numerical integration: Structure-preserving algorithms for ordinary differential equations*, Springer Series in Computational Mathematics, Springer Berlin Heidelberg, 2006.

[KK13]   R. Korn and E. Korn, *Optionsbewertung und portfolio-optimierung: Moderne methoden der finanzmathematik*, Vieweg+Teubner Verlag, 2013.

[KP92]   P. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, Springer–Verlag, 1992.

[KS12]   Ioannis Karatzas and Steven Shreve, *Brownian motion and stochastic calculus*, vol. 113, Springer Science & Business Media, 2012.

[LL87]   L.D. Landau and E.M. Lifshitz, *Fluid mechanics*, no. Bd. 6, Elsevier Science, 1987.

[MRT02]   GN Milstein, Yu M Repin, and MV Tretyakov, *Numerical methods for stochastic systems preserving symplectic structure*, SIAM Journal on Numerical Analysis **40** (2002), no. 4, 1583–1604.

[MT03]   GN Milstein and Michael V Tretyakov, *Quasi-symplectic methods for langevin-type equations*, IMA journal of numerical analysis **23** (2003), no. 4, 593–626.

[OBJ+11]   Noel M O'Boyle, Michael Banck, Craig A James, Chris Morley, Tim Vandermeersch, and Geoffrey R Hutchison, *Open babel: An open chemical toolbox*, Journal of cheminformatics **3** (2011), no. 1, 1.

[Øks03]   B. Øksendal, *Stochastic differential equations: An introduction with applications*, Hochschultext / Universitext, Springer, 2003.

[RCC+92]   Anthony K Rappé, Carla J Casewit, KS Colwell, WA Goddard Iii, and WM Skiff, *Uff, a full periodic table force field for molecular mechanics and molecular dynamics simulations*, Journal of the American chemical society **114** (1992), no. 25, 10024–10035.

[RLG+14]   Julien Roche, John M Louis, Alexander Grishaev, Jinfa Ying, and Adriaan Bax, *Dissociation of the trimeric gp41 ectodomain at the lipid–water interface suggests an active role in hiv-1 env-mediated membrane fusion*, Proceedings of the National Academy of Sciences **111** (2014), no. 9, 3425–3430.

Eigenständigkeitserklärung

Hiermit erkläre ich dass ich die Bachelorarbeit selbst geschrieben habe und keine anderen als die angegebenen Quellen verwendet habe.

<div align="right">

Kaiserslautern, 1. August 2016
Steffen Plunder

</div>