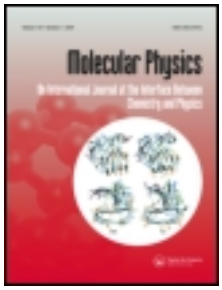

Molecular Physics: An International Journal at the Interface Between Chemistry and Physics

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tmph20

A simple and effective Verlet-type algorithm for simulating Langevin dynamics

Niels Grønbech-Jensen [a b c] & Oded Farago [a d e]

[a] Department of Mechanical and Aerospace Engineering , University of California , Davis , California , 95616 , USA

[b] Department of Chemical Engineering and Materials Science , University of California , Davis , California , 95616 , USA

[c] Computational Research Division , Lawrence Berkeley National Laboratory , Berkeley , California , 94720 , USA

[d] Department of Biomedical Engineering , Ben Gurion University of the Negev , Be'er Sheva , 84105 , Israel

[e] Ilse Katz Institute for Nanoscale Science and Technology , Ben Gurion University of the Negev , Be'er Sheva , 84105 , Israel

Accepted author version posted online: 09 Jan 2013.Published online: 14 Feb 2013.

To cite this article: Niels Grønbech-Jensen & Oded Farago (2013): A simple and effective Verlet-type algorithm for simulating Langevin dynamics, Molecular Physics: An International Journal at the Interface Between Chemistry and Physics, 111:8, 983-991

To link to this article: http://dx.doi.org/10.1080/00268976.2012.760055

Molecular Physics, 2013
Vol. 111, No. 8, 983–991, http://dx.doi.org/10.1080/00268976.2012.760055

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

A simple and effective Verlet-type algorithm for simulating Langevin dynamics

Niels Grønbech-Jensen[a,b,c,*] and Oded Farago[a,d,e]

[a]*Department of Mechanical and Aerospace Engineering, University of California, Davis, California 95616, USA;* [b]*Department of Chemical Engineering and Materials Science, University of California, Davis, California 95616, USA;* [c]*Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA;* [d]*Department of Biomedical Engineering, Ben Gurion University of the Negev, Be'er Sheva 84105, Israel;* [e]*Ilse Katz Institute for Nanoscale Science and Technology, Ben Gurion University of the Negev, Be'er Sheva 84105, Israel*

(*Received 17 November 2012; final version received 6 December 2012*)

We present a revision to the well known Störmer–Verlet algorithm for simulating second order differential equations. The revision addresses the inclusion of linear friction with associated stochastic noise, and we analytically demonstrate that the new algorithm correctly reproduces diffusive behaviour of a particle in a flat potential. For a harmonic oscillator, our algorithm provides the exact Boltzmann distribution for any value of damping, frequency and time step for both underdamped and overdamped behaviour within the usual stability limit of the Verlet algorithm. Given the structure and simplicity of the method, we conclude that this approach can trivially be adapted for contemporary applications, including molecular dynamics with extensions such as molecular constraints.

Keywords: molecular dynamics; simulating Langevin equations; Störmer–Verlet algorithms; computational statistical mechanics

1. Introduction

In molecular dynamics (MD) simulations, Newton's equations of motion are solved numerically to produce trajectories of the system in the microcanonical (N, V, E) ensemble. The most commonly used scheme for this purpose is the Verlet method, which is based on truncated Taylor expansions for the evolution of a particle with mass m, coordinate $r(t)$, velocity $v(t)$ and force $f(r, t)$ [1, 2]. Introducing the discrete-time variables $r^n = r(t_n)$, $v^n = v(t_n)$ and $f^n = f(r^n, t_n)$, the so-called velocity explicit Verlet (or velocity Verlet) scheme reads

$$r^{n+1} = r^n + v^n dt + \frac{dt^2}{2m} f^n, \tag{1}$$

$$v^{n+1} = v^n + \frac{dt}{2m}(f^n + f^{n+1}). \tag{2}$$

By considering two successive time steps and eliminating the velocity variables from the equations, the more original form of the Störmer–Verlet method [3, 4] (here called the position-Verlet method) is found

$$r^{n+1} = 2r^n - r^{n-1} + \frac{dt^2}{m} f^n, \tag{3}$$

with the associated velocity calculated by the central difference

$$v^n = \frac{r^{n+1} - r^{n-1}}{2dt}. \tag{4}$$

The Verlet scheme is accurate to second order in dt, which means that the deviation, per time step, of the computed trajectory from the true (analytic) one, scales with the third power of dt. The merits of the Verlet scheme that make it so widely popular for MD simulations include its convenience, efficiency and time reversibility, which ensures that the error of the total energy of long-time integrations is bounded and does not drift or diffuse, as illustrated by the exact solution to the Verlet method applied to a harmonic oscillator (see, e.g. below).

The most frequently used ensemble in statistical-mechanics is the canonical (N, V, T) ensemble where the temperature of the system, rather than its energy, is constant. A variety of methods for conducting MD simulations in the canonical ensemble have been proposed over the years. A very appealing class of such methods include integrators for Langevin dynamics (LD) simulations. In LD, two forces are added to the conservative force field - a friction force proportional to the velocity with friction coefficient $\alpha \geq 0$, and thermal white noise $\beta(t)$. Explicitly, the Langevin equation of motion is given by [5]

*Corresponding author. Email: ngjensen@ucdavis.edu

$$\dot{r} = v, \tag{5}$$

$$m\dot{v} = f(r,t) - \alpha v + \beta(t). \tag{6}$$

In order to satisfy the dissipation-fluctuation theorem, it is often assumed that the stochastic force is Gaussian distributed, and has the statistical properties [5]

$$\langle \beta(t) \rangle = 0, \tag{7}$$

$$\langle \beta(t)\beta(t') \rangle = 2\alpha k_B T \delta(t-t'), \tag{8}$$

where k_B is the Boltzmann's constant and T is the thermodynamic temperature of the heat bath.

The difficulty to develop accurate integrators for LD stems from the non-analytic nature of $\beta(t)$, which invalidates the Taylor expansion commonly used for the derivation of the Verlet scheme. The most naive way to overcome this difficulty is to replace the delta-function correlated noise with a set of rectangular pulses of mean-squared size $\sqrt{2k_BT\alpha/dt}$, each of which acting over the centred time interval $(t_n - dt/2, t_n + dt/2)$. Employing this approximation for $\beta(t)$ yields the famous Brooks-Brünger-Karplus (BBK) scheme [6], which unfortunately turns out to exhibit a simulated temperature that differs by $\mathcal{O}(dt)$ from the correct one [7]. This disappointing result can be attributed to the effect of a small time step dt, which results in a stochastic velocity change of order $\sqrt{dt}$ that overshadows the 'regular' (deterministic) linear velocity change of order dt. Statistical analysis of the small time interval shows, in fact, that the opposite is true, since the average over all noise realisations vanishes. In order to develop a reliable integrator for LD one therefore needs to carefully treat the coupling between the stochastic and analytic contributions. This is, unfortunately, not easily accomplished with a discretised approximation for $\beta(t)$. A different approach to the problem has been introduced by Ermak and Buckholz (EB) [8]. The EB method is based on a numerical integration of the formal solution to the Langevin equation. This gives

$$v^{n+1} = v^n \exp(-\alpha dt/m) + \frac{1}{m}\int_{t_n}^{t_{n+1}} \exp[-\alpha(t_{n+1}-t')/m] \times [f(t') + \beta(t')]\, dt', \tag{9}$$

$$r^{n+1} = r^n + \frac{mv^n}{\alpha}[1 - \exp(-\alpha dt/m)] + \frac{1}{\alpha}\int_{t_n}^{t_{n+1}} \{1 - \exp[-\alpha(t_{n+1}-t')/m]\} \times [f(t') + \beta(t')]\, dt'. \tag{10}$$

Several features distinguish the EB method from Verlet-type schemes such as BBK. First, the friction force is accounted for separately from the other two forces, via exponentially decaying functions. Second, this scheme considers an ensemble of trajectories instead of a single one, thereby requiring two random Gaussian numbers per time step. The one appearing in the velocity equation, $\int_0^{dt} \exp[-\alpha(dt-t')/m]\beta(t')dt'$, represents the stochastic change in velocity during the time interval. The other, $\int_0^{dt}\{1 - \exp[-\alpha(dt-t')/m]\}\beta(t')dt'$, appears in the position equation and characterises the random displacement. These two numbers are different, but *correlated*. Third, there is formally no need to compute the new position before the new velocity as in the velocity-Verlet scheme and, in principle, one can interchange the order by which the equations are evolved. The last point is particularly important since it is directly related to the most critical issue that remains unsolved with the EB approach, which is how to perform the numerical integration over the deterministic force $f(t)$. The appearance of exponential 'weight functions' in the integrals in Equations (9) and (10) opens the possibility for using a variety of linear combinations involving f^{n-1}, f^n and f^{n+1}. The most popular integration schemes for the deterministic force constitute the van Gunsteren-Berendsen [9] and the Langevin impulse [10] methods.

More recently, Ricci and Ciccotti (RC) introduced a formalism for a systematic derivation of numerical integrators for LD [11]. The essential steps of the formalism are to (1) express the integrator as an exponential operator, (2) split the time interval dt into a series of even smaller time steps and (3) use the Suzuki–Trotter expansion for time-ordered exponential operators to find an approximation for the evolution operator corresponding to dt. By using mid-point splitting, RC arrived at a scheme that require only one random number per dt. The same formalism was later used to develop a new family of integrators that more carefully treat the coupling between the stochastic and deterministic components of the dynamics [12, 13]. A characteristic feature of these new schemes is that they require two *independent* random numbers per time step, which is different from the EB family of schemes, where two correlated random numbers are involved.

Returning for a moment to MD simulations of Newtonian dynamics, it is important to recall that, due to its computational efficiency and time reversibility, the Verlet method is regarded as superior to other integration methods that may produce more accurate trajectories. In the case of LD, the trajectories include a random component and, therefore, their accuracy (which can only be defined in statistical terms) becomes an issue of even less importance. The efficiency of different methods for canonical ensemble simulations of many-particles systems must be tested according to their ability to reproduce measurable statistical quantities, such as the Boltzmann distribution corresponding to the temperature of the heat bath. To this end, none of the existing methods has managed to

demonstrate exact thermodynamic response [7]. We here present a novel Verlet-type scheme for LD simulations that is very simple to implement and which yields *correct* statistical-mechanical behaviour of a particle diffusing in both a flat and a harmonic potential.

2. Derivation of the algorithm

In the spirit of the simplicity of the Verlet algorithm in its traditional forms, we here arrive at a new useful method through a straight-forward derivation. Starting with the continuous-time Langevin Equations (5)–(8), we integrate Equation (6) over a (small) time interval dt between two times, t_n and $t_{n+1} = t_n + dt$:

$$\int_{t_n}^{t_{n+1}} m\dot{v}\, dt' = \int_{t_n}^{t_{n+1}} f\, dt' - \int_{t_n}^{t_{n+1}} \alpha\dot{r}\, dt' + \int_{t_n}^{t_{n+1}} \beta(t')\, dt'. \tag{11}$$

With no approximation, this can be written

$$m(v^{n+1} - v^n) = \int_{t_n}^{t_{n+1}} f\, dt' - \alpha(r^{n+1} - r^n) + \beta^{n+1}, \tag{12}$$

where

$$\beta^{n+1} \equiv \int_{t_n}^{t_{n+1}} \beta(t')\, dt' \tag{13}$$

is a Gaussian random number with $\langle\beta^n\rangle = 0$ and $\langle\beta^n\beta^l\rangle = 2\alpha k_B T\, dt\, \delta_{n,l}$.

Integrating over Equation (5) yields

$$\int_{t_n}^{t_{n+1}} \dot{r}\, dt' = r^{n+1} - r^n = \int_{t_n}^{t_{n+1}} v\, dt', \tag{14}$$

which can be approximated with

$$r^{n+1} - r^n \approx \frac{dt}{2}(v^{n+1} + v^n). \tag{15}$$

An equation similar to Equation (15) has been used by Ricci and Ciccotti in one of the forms of their scheme (see Equation (18) in Ref. [11]). It introduces errors that scale with dt^3 in the deterministic trajectory and in the variance of the stochastic component (see discussion in Ref. [14]).

Inserting v^{n+1} from Equation (12) into Equation (15) provides a convenient pair of equations

$$r^{n+1} - r^n = bdt\, v^n + \frac{bdt}{2m}\int_{t_n}^{t_{n+1}} f\, dt' + \frac{bdt}{2m}\beta^{n+1}, \tag{16}$$

$$v^{n+1} - v^n = \frac{1}{m}\int_{t_n}^{t_{n+1}} f\, dt' - \frac{\alpha}{m}(r^{n+1} - r^n) + \frac{1}{m}\beta^{n+1}, \tag{17}$$

where

$$b \equiv \frac{1}{1 + \frac{\alpha dt}{2m}}. \tag{18}$$

For any given realisation of β^n, Equation (16) is correct to second order in dt, while Equation (17) is exact as written. We now approximate the integral of the deterministic force f such that both equations are correct to second order in dt:

$$r^{n+1} = r^n + bdt\, v^n + \frac{bdt^2}{2m} f^n + \frac{bdt}{2m}\beta^{n+1}, \tag{19}$$

$$v^{n+1} = v^n + \frac{dt}{2m}(f^n + f^{n+1}) - \frac{\alpha}{m}(r^{n+1} - r^n) + \frac{1}{m}\beta^{n+1}. \tag{20}$$

We first notice that when $\alpha = 0$, the above Equations (19) and (20) reduce to the standard velocity-Verlet scheme given in Equations (1) and (2). Second, we notice the very reasonable feature of the representation of the damping, which is calculated as the integral of the actual path that the object has travelled during the time step dt. Third, the noise is represented as a *single* stochastic variable for each time step, realised by a single aggregated impulse during dt that influences the dynamics over the time step. In this regard, Equations (19) and (20) constitute a simple functional Verlet-type scheme for solving stochastic Langevin equations. Unlike the aforementioned EB-type family of schemes [8–10], our method does not employ two stochastic variables. This is a consequence of our main objective; namely, to produce the correct statistical-mechanics for large ensembles (spatial or temporal), while focusing less on the detailed dynamics *within* each time step.

Before proceeding to analysing the behaviour of the developed scheme, we rewrite the method in a couple of equivalent different forms that can be useful and that can illustrate close connections to previously published work. First, we observe that by inserting Equation (19) into Equation (20), the equations can be rewritten

$$r^{n+1} = r^n + bdt v^n + \frac{bdt^2}{2m} f^n + \frac{bdt}{2m}\beta^{n+1}, \tag{21}$$

$$v^{n+1} = av^n + \frac{dt}{2m}(af^n + f^{n+1}) + \frac{b}{m}\beta^{n+1}, \tag{22}$$

where

$$a \equiv \frac{1-\frac{\alpha dt}{2m}}{1+\frac{\alpha dt}{2m}}. \tag{23}$$

This form reveals that the derived method presented here parallels one mentioned in passing by Melchionna (Equation (41) in Ref. [12]), where the contribution of the friction force has been integrated similarly, but with a different noise term that includes two independent random numbers per time step. Second, we rewrite the scheme in the form given only by displacement coordinates. This is accomplished by subsequently inserting the expressions for v^n (22) and r^n (21) into that of r^{n+1} given in Equation (21). The result is

$$r^{n+1} = 2br^n - ar^{n-1} + \frac{bdt^2}{m} f^n + \frac{bdt}{2m}(\beta^{n+1} + \beta^n) \tag{24}$$

with the associated velocity given by

$$v^n = \frac{a}{b^2}\frac{r^{n+1} - r^{n-1}}{2dt} - \frac{\alpha}{2m}\frac{bdt^3}{m} f^n + \frac{bdt}{2m}(a\beta^{n+1} - \beta^n). \tag{25}$$

Equation (24) shows that our formulation is also in close proximity to the BBK scheme [6], which is of a position-Verlet-type. (Notice that it is essential to have the proper form for the accompanying velocity in order to determine if a position formulation is the same as a velocity explicit form.) However, unlike our Equation (24), the BBK scheme employs only a single stochastic number for the two time steps covered in the displacement equation. In fact, the methods of Melchionna and BBK, while not identical, are closely related, since the use of a single stochastic number in the position representation (BBK) translates into applying two random numbers in the velocity equation (Melchionna). Conversely, in our scheme, we have a single random number in the velocity formulation, and two in the position representation covering two time steps.

3. Linear analysis

In order to evaluate the general applicability of the above method, we calculate key statistical measures of the numerical scheme for some characteristic linear cases [15], and compare them to known statistical-mechanical behaviour of the true Langevin system.

3.1. Thermal diffusion: $f = 0$

The first basic property to investigate is the diffusive behaviour of a particle moving in a flat potential ($f = 0$) at temperature T. In this case, Equation (22) reads

$$v^{n+1} = av^n + \frac{b}{m}\beta^{n+1}. \tag{26}$$

By using the same equation to express v^n in terms of v^{n-1} and β^n, and by repeating this procedure until reaching the initial velocity v^0, one arrives at the relation

$$v^n = a^n v^0 + \frac{b}{m}\sum_{k=0}^{n-1} a^k \beta^{n-k}, \tag{27}$$

$$= a^n v^0 + \frac{b\sqrt{2\alpha k_B T dt}}{m}\sum_{k=0}^{n-1} a^k \sigma^{n-k}, \tag{28}$$

where σ is a standard Gaussian random number with $\langle\sigma\rangle = 0$ and $\langle\sigma^2\rangle = 1$, the superscript denoting different realisations of this variable such that $\langle\sigma^n\sigma^l\rangle = \delta_{nl}$. Summing over the random numbers in Equation (28) yields another Gaussian random number, and in combination with Equation (18), one arrives at

$$v^n = a^n v^0 + \sqrt{1-a^{2n}}\sqrt{\frac{k_B T}{m}}\,\sigma. \tag{29}$$

For large n, beyond the transients from the initial condition ($a^n \ll 1$), we find that the velocity is characterised by a Gaussian (Maxwell–Boltzmann) distribution with zero mean and

$$\langle (v^n)^2\rangle = \frac{k_B T}{m}, \tag{30}$$

which results in reproducing the exact expectation for the average kinetic energy (thermal energy)

$$E_k = \frac{1}{2}m\langle (v^n)^2\rangle = \frac{1}{2}k_B T. \tag{31}$$

Complementing this result, we turn to the displacement coordinate Equation (21) for $f = 0$:

$$r^n = r^{n-1} + bdtv^{n-1} + \frac{bdt}{2m}\beta^n. \tag{32}$$

Inserting Equation (28) for v^{n-1}, and repeating the procedure until reaching the initial position r^0, we obtain the result

$$r^n = r^0 + \frac{m}{\alpha}(1-a^n)v^0 + \frac{b}{\alpha}\left[\frac{\alpha dt}{2m}\beta^n + \sum_{k=1}^{n-1}\left(\frac{1}{b} - a^k\right)\beta^{n-k}\right]. \tag{33}$$

By using the definitions of a (23) and b (18), we find that for large n (such that $a^n \ll 1$), Equation (33) can be written

as

$$r^n = r^0 + \frac{m}{\alpha}v^0 + \sqrt{ndt\frac{2k_BT}{\alpha}}\,\sigma. \tag{34}$$

The second term in Equation (34), which is the transient ballistic displacement, matches exactly the corresponding value predicted by the Langevin solution [second term on the right-hand side of Equation (10) for $dt \to \infty$]. More importantly, we find that our scheme results in a simulated diffusion coefficient

$$D = \lim_{ndt\to\infty} \frac{\langle r^n - r^0\rangle^2}{2ndt} = \frac{k_BT}{\alpha}, \tag{35}$$

which agrees with Einstein's fluctuation-dissipation relationship for LD [5]. Notice that the exact results for fluctuations and diffusion obtained here are independent of the magnitudes of the time step dt, damping $\alpha > 0$, and temperature T.

3.2. Thermal harmonic oscillator: $f = -\kappa r$

Encouraged by the performance of the method for $f = 0$, we now turn to analysing the method applied to a damped thermal harmonic oscillator, where $f = -\kappa r$ represents a linear Hooke's spring with spring constant $\kappa > 0$. Our Equations (21) and (22) now read

$$\begin{pmatrix} r^{n+1} \\ v^{n+1}\end{pmatrix} = \mathcal{V}\begin{pmatrix} r^n \\ v^n\end{pmatrix} + \mathcal{N}\beta^{n+1}, \tag{36}$$

where

$$\mathcal{V} = \begin{pmatrix} 1 - b\dfrac{\Omega_0^2dt^2}{2} & bdt \\ -b\Omega_0^2dt\left(1 - \dfrac{\Omega_0^2dt^2}{4}\right) & a - b\dfrac{\Omega_0^2dt^2}{2}\end{pmatrix}, \tag{37}$$

$$\mathcal{N} = \frac{b}{m}\begin{pmatrix} \dfrac{dt}{2} \\ 1 - \dfrac{\Omega_0^2dt^2}{4}\end{pmatrix}, \tag{38}$$

and where $\Omega_0 = \sqrt{\kappa/m}$ is the resonance frequency of the undamped continuous-time oscillator. The matrix $\mathcal{V}$ has the eigenvalues $\Lambda_\pm$:

$$\Lambda_\pm = b\left(1 - \frac{\Omega_0^2dt^2}{2}\right) \pm \sqrt{b^2\left(1 - \frac{\Omega_0^2dt^2}{2}\right)^2 - a}, \tag{39}$$

from where we can evaluate the formal stability limit $\Omega_0 dt < 2$, consistent with the requirement $|\Lambda_\pm| < 1$. Notice that $\Lambda_+\Lambda_- = a$ for all parameter values.

We analyse the basic thermodynamic properties of the thermal harmonic oscillator by perpetuating the map Equation (36) from initial conditions r^0 and v^0,

$$\begin{pmatrix} r^n \\ v^n\end{pmatrix} = \mathcal{V}\begin{pmatrix} r^{n-1} \\ v^{n-1}\end{pmatrix} + \mathcal{N}\beta^n, \tag{40}$$

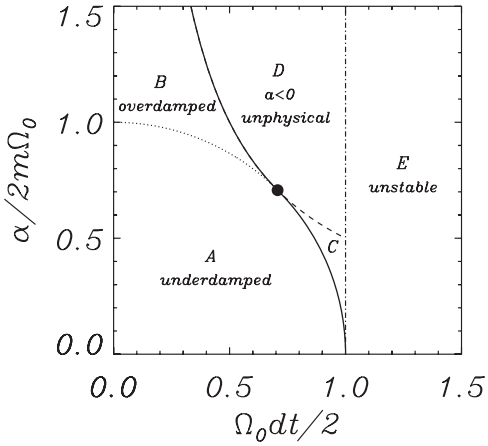

Figure 1. Sketch of the characteristic regimes for the Verlet integrator applied to the damped harmonic oscillator. Regimes are defined by the eigenvalues given in Equation (39). Solid curve is the boundary between physical (left) and unphysical (right) behaviour. Vertical dash-dotted line is the formal stability boundary of the Verlet method, $\Lambda_- = -1$. Marker • indicates where $\Lambda_\pm = 0$. Regimes are: (a) Underdamped, $\Lambda_\pm$ complex; (b) Overdamped, $0 < a < \Lambda_\pm < 1$ real; (c) $-1 < \Lambda_\pm < -a < 0$ real; (d) $\Lambda_\pm$ real, Λ_- negative, Λ_+ positive ($a < 0$) and (e) Unstable, $\Lambda_- < -1$.

$$= \mathcal{V}^n\begin{pmatrix} r^0 \\ v^0\end{pmatrix} + \sum_{k=0}^{n-1}\mathcal{V}^k\mathcal{N}\beta^{n-k}, \tag{41}$$

where the kth iteration is conveniently given by the unitary transformation

$$\mathcal{V}^k = \mathcal{U}\begin{pmatrix} \Lambda_+^k & 0 \\ 0 & \Lambda_-^k\end{pmatrix}\mathcal{U}^{-1}. \tag{42}$$

The two characteristic parameters of the Verlet method applied to a damped harmonic oscillator are $\alpha/2m\Omega_0$ and $\Omega_0 dt$. Figure 1 displays the five different regimes of simulated oscillator behaviour as a function of those two parameters. The two physically relevant dynamical regimes, underdamped (A) and overdamped (B), are considered first.

3.2.1. Regime A: underdamped dynamics

The oscillator is always stable in this regime, where $(\alpha/2m)^2 < \Omega_0^2(1 - \Omega_0^2dt^2/4)$, and the eigenvalues can be conveniently written

$$\Lambda_\pm = \sqrt{a}\,e^{\pm i\Omega_V dt}. \tag{43}$$

Here, Ω_V is the underdamped Verlet-oscillator resonance frequency given by

$$\sqrt{a}\cos\Omega_V dt = b\left(1 - \frac{\Omega_0^2dt^2}{2}\right), \tag{44}$$

$$\sqrt{a}\sin\Omega_V dt = bdt\sqrt{\Omega_0^2\left(1 - \frac{\Omega_0^2dt^2}{4}\right) - \left(\frac{\alpha}{2m}\right)^2}, \tag{45}$$

such that

$$\mathcal{V} = \begin{pmatrix} \sqrt{a}\cos\Omega_V dt + 1 - b & bdt \\ -bdt\left[\frac{a}{b^2dt^2}\sin^2\Omega_V dt + \left(\frac{\alpha}{2m}\right)^2\right] & \sqrt{a}\cos\Omega_V dt + a - b \end{pmatrix}. \tag{46}$$

The diagonalising transformation is then given by

$$\mathcal{U} = \frac{1}{|u|}\begin{pmatrix} 1 & 1 \\ -\frac{\alpha}{2m} + i\frac{\sqrt{a}}{bdt}\sin\Omega_V dt & -\frac{\alpha}{2m} - i\frac{\sqrt{a}}{bdt}\sin\Omega_V dt \end{pmatrix} \tag{47}$$

with $|u|$ being the normalisation of the eigenvectors of $\mathcal{V}$ that appear as the column vectors in $\mathcal{U}$. The kth time step can then be expressed from Equation (42) in the explicit form:

$$\mathcal{V}^k = \frac{bdt}{\sqrt{a}\sin\Omega_V dt}\sqrt{a}^k\begin{pmatrix} \frac{\sqrt{a}}{bdt}\sin\Omega_V dt\cos\Omega_V kdt + \frac{\alpha}{2m}\sin\Omega_V kdt & \sin\Omega_V kdt \\ -\left(\frac{a}{b^2dt^2}\sin^2\Omega_V dt + \left(\frac{\alpha}{2m}\right)^2\right)\sin\Omega_V kdt & \frac{\sqrt{a}}{bdt}\sin\Omega_V dt\cos\Omega_V kdt - \frac{\alpha}{2m}\sin\Omega_V kdt \end{pmatrix}, \tag{48}$$

from where we obtain the explicit form of the expression of interest in Equation (41)

$$\mathcal{V}^k\mathcal{N}\beta^{n-k} = \frac{bdt\sqrt{2\alpha k_B T}}{2m\sqrt{a}\sin\Omega_V dt}\sqrt{a}^k\begin{pmatrix} \sqrt{a}\sin\Omega_V dt\cos\Omega_V kdt + (1+\sqrt{a}\cos\Omega_V dt)\sin\Omega_V kdt \\ \left(1 - \frac{\Omega_0^2 dt^2}{4}\right)\frac{2}{dt}[\sqrt{a}\sin\Omega_V dt\cos\Omega_V kdt - (1-\sqrt{a}\cos\Omega_V dt)\sin\Omega_V kdt] \end{pmatrix}\sigma^{n-k}. \tag{49}$$

The thermodynamic limit of $\langle(r^n)^2\rangle$ and $\langle(v^n)^2\rangle$ can now be found by inserting Equation (49) into Equation (41), whereafter the square sum can be evaluated since each term in the summation is an independent Gaussian random number. Taking the limit $n \to \infty$, this yields for the variance of the displacement:

$$\begin{aligned}\langle(r^n)^2\rangle &= \frac{\alpha k_B T b^2 dt^3}{2m^2 a\sin^2\Omega_V dt}\left[\frac{1}{2}(1+a+2\sqrt{a}\cos\Omega_V dt)\sum_{k=0}^{\infty}a^k - \frac{1}{2}(1+a\cos 2\Omega_V dt + 2\sqrt{a}\cos\Omega_V dt)\sum_{k=0}^{\infty}a^k\cos 2\Omega_V kdt\right. \\ &\quad \left. + \sqrt{a}\sin\Omega_V dt(1+2\sqrt{a}\cos\Omega_V dt)\sum_{k=0}^{\infty}a^k\sin 2\Omega_V kdt\right] \\ &= \frac{\alpha k_B T b^2 dt^3}{2m^2 a\sin^2\Omega_V dt}\left[\frac{1}{2}\frac{4b(1-\frac{\Omega_0^2 dt^2}{4})}{1-a} - \frac{(1-a)\sqrt{a}\cos\Omega_V dt + \frac{1}{2}(1-a^2)}{1+a^2-2a\cos 2\Omega_V dt}\right] \\ &= \frac{\alpha k_B T b^2 dt^3}{2m^2 a\sin^2\Omega_V dt}\,\frac{\Omega_0^2(1-\frac{\Omega_0^2 dt^2}{4}) - (\frac{\alpha}{2m})^2}{\frac{\alpha dt}{2m}\Omega_0^2} \\ &= \frac{k_B T}{m\Omega_0^2}.\end{aligned} \tag{50}$$

We similarly obtain the result for the statistical variance of the velocity:

$$\begin{aligned}\langle(v^n)^2\rangle &= \frac{2\alpha k_B T b^2 dt}{m^2 a\sin^2\Omega_V dt}\left(1-\frac{\Omega_0^2 dt^2}{4}\right)^2\left[\frac{1}{2}(1+a-2\sqrt{a}\cos\Omega_V dt)\sum_{k=0}^{\infty}a^k - \frac{1}{2}(1+a\cos 2\Omega_V dt - 2\sqrt{a}\cos\Omega_V dt)\right. \\ &\quad \left.\times\sum_{k=0}^{\infty}a^k\cos 2\Omega_V kdt + \sqrt{a}\sin\Omega_V dt(1-\sqrt{a}\cos\Omega_V dt)\sum_{k=0}^{\infty}a^k\sin 2\Omega_V kdt\right] \\ &= \frac{2\alpha k_B T b^2 dt}{m^2 a\sin^2\Omega_V dt}\left(1-\frac{\Omega_0^2 dt^2}{4}\right)^2\left[\frac{b\frac{\Omega_0^2 dt^2}{2}}{1-a} + \frac{(1-a)\sqrt{a}\cos\Omega_V dt - \frac{1}{2}(1-a^2)}{1+a^2-2a\cos 2\Omega_V dt}\right] \\ &= \frac{2\alpha k_B T b^2 dt^2}{m^2 a\sin^2\Omega_V dt}\left(1-\frac{\Omega_0^2 dt^2}{4}\right)\frac{\Omega_0^2(1-\frac{\Omega_0^2 dt^2}{4}) - (\frac{\alpha}{2m})^2}{4\frac{\alpha}{2m}} = \frac{k_B T}{m}\left(1-\frac{\Omega_0^2 dt^2}{4}\right).\end{aligned} \tag{51}$$

The two variances are noteworthy, since evaluation of the average potential and kinetic energies gives

$$E_p = \frac{1}{2}\kappa\langle(r^n)^2\rangle \;=\; \frac{1}{2}k_BT, \tag{52}$$

$$E_k = \frac{1}{2}m\langle(v^n)^2\rangle \;=\; \frac{1}{2}k_BT\left(1-\frac{\Omega_0^2dt^2}{4}\right). \tag{53}$$

These hold true for any parameter choice in the underdamped regime. Notably, Equation (52) implies that our method produces the *exact* statistical distribution regardless of time step, frequency (potential curvature), damping parameter or temperature. Therefore, it is reasonable to expect that this method will provide correct Boltzmann distribution in thermodynamic equilibrium when simulating complex many-particle systems, that may have a multiple of participating frequencies. To our knowledge, this important feature has not previously been reported for a numerical integrator of the Langevin equation [7].

Complementing the Boltzmann distribution of the displacement is the variance of the velocity, Equation (53). Despite the obvious discrepancy between the true kinetic energy of the Langevin equation and the one shown for our algorithm, we submit that the presented result is the best possible for a method that builds on the discretised Verlet formalism. The reduction in the variance of the velocity by a factor of $(1-\Omega_0^2dt^2/4)$ does not arise from the treatment of friction and noise, as implied by the fact that this factor depends on neither damping nor temperature. Instead, the observed deviation arises from the approximation of the potential curvature introduced by the discrete integrator. This is consistent with the well-known inherent discrepancy between displacement and associated velocity that causes the periodic deviations with magnitude dt^2 from strict energy conservation in a simulated harmonic oscillator. This can be explicitly demonstrated by using the Verlet Equations (1) and (2) for an undamped harmonic oscillator with initial conditions r^0 and v^0. The result is

$$\begin{pmatrix} r^n \\ v^n \end{pmatrix} = \begin{pmatrix} r^0 \\ v^0 \end{pmatrix}\cos\Omega_V ndt + \frac{1}{\sqrt{\Omega_0^2\left(1-\frac{\Omega_0^2dt^2}{4}\right)}} \times \begin{pmatrix} v_0 \\ -\Omega_0^2(1-\frac{\Omega_0^2dt^2}{4})r_0 \end{pmatrix}\sin\Omega_V ndt, \tag{54}$$

which illustrates that the Verlet velocity (momentum) is depressed and is not exactly the conjugated coordinate to the displacement, and that the discrepancy is related to the proportionality seen between the two thermodynamic expressions (52) and (53).

3.2.2. Regime B: overdamped dynamics

This regime is defined by the requirements $(\alpha/2m)^2 > \Omega_0^2(1-\Omega_0^2dt^2/4)$ and $\alpha dt/2m < 1$ $(a>0)$. The latter condition is imposed to ensure $\Lambda_\pm > 0$, which is necessary for physically meaningful dynamics, where only monotonic decay is possible in the overdamped regime for $T=0$.

The matrix $\mathcal{V}$ appearing in Equation (36), which is now expressed by

$$\mathcal{V} = \begin{pmatrix} \sqrt{a}\cosh\lambda_V dt + 1 - b & bdt \\ -bdt[(\frac{\alpha}{2m})^2 - \frac{a}{b^2dt^2}\sinh^2\lambda_V dt] & \sqrt{a}\cosh\lambda_V dt + a - b \end{pmatrix}, \tag{55}$$

has the eigenvalues

$$\Lambda_\pm = \sqrt{a}\,e^{\pm\lambda_V dt}, \tag{56}$$

where λ_V is given by

$$\sqrt{a}\cosh\lambda_V dt = b\left(1-\frac{\Omega_0^2dt^2}{2}\right), \tag{57}$$

$$\sqrt{a}\sinh\lambda_V dt = bdt\sqrt{\left(\frac{\alpha}{2m}\right)^2 - \Omega_0^2\left(1-\frac{\Omega_0^2dt^2}{4}\right)}. \tag{58}$$

The matrix $\mathcal{V}$ can be diagonalised [see Equation (42)] using the transformation matrix

$$\mathcal{U} = \begin{pmatrix} \frac{1}{|u_+|} & \frac{1}{|u_-|} \\ \frac{1}{|u_+|}\left[-\frac{\alpha}{2m}+\frac{\sqrt{a}}{bdt}\sinh\lambda_V dt\right] & \frac{1}{|u_-|}\left[-\frac{\alpha}{2m}-\frac{\sqrt{a}}{bdt}\sinh\lambda_V dt\right] \end{pmatrix} \tag{59}$$

with $|u_+|$ and $|u_-|$ being the normalisations of the two real eigenvectors of $\mathcal{V}$. This leads to the explicit results:

$$\mathcal{V}^k = \frac{bdt}{\sqrt{a}\sinh\lambda_V dt}\sqrt{a}^k \begin{pmatrix} \frac{\sqrt{a}}{bdt}\sinh\lambda_V dt\cosh\lambda_V kdt + \frac{\alpha}{2m}\sinh\lambda_V kdt & \sinh\lambda_V kdt \\ -\left(\left(\frac{a}{2m}\right)^2 - \frac{a}{b^2dt^2}\sinh^2\lambda_V dt\right)\sinh\lambda_V kdt & \frac{\sqrt{a}}{bdt}\sinh\lambda_V dt\cosh\lambda_V kdt - \frac{\alpha}{2m}\sinh\lambda_V kdt \end{pmatrix} \tag{60}$$

and

$$\mathcal{V}^k\mathcal{N}\beta^{n-k} = \frac{bdt\sqrt{2\alpha k_BT}}{2m\sqrt{a}\sinh\lambda_V dt}\sqrt{a}^k \begin{pmatrix} \sqrt{a}\sinh\lambda_V dt\cosh\lambda_V kdt + (1+\sqrt{a}\cosh\lambda_V dt)\sinh\lambda_V kdt \\ \left(1-\frac{\Omega_0^2dt^2}{4}\right)\frac{2}{dt}[\sqrt{a}\sinh\lambda_V dt\cosh\lambda_V kdt - (1-\sqrt{a}\cosh\lambda_V dt)\sinh\lambda_V kdt] \end{pmatrix}\sigma^{n-k}. \tag{61}$$

Following the same procedure used in the underdamped case above, we insert Equation (61) into Equation (41), and evaluate the variances of the displacement and velocity in the thermodynamic limit for $n \to \infty$ by calculating the square sum of the amplitudes of the independent stochastic numbers of each term. After some calculations, we arrive at

$$\begin{aligned}
\langle (r^n)^2 \rangle &= \frac{\alpha k_B T b^2 dt^3}{2m^2 a \sinh^2 \lambda_V dt} \left[-\frac{1}{2}(1 + a + 2\sqrt{a}\cosh\lambda_V dt) \sum_{k=0}^{\infty} a^k + \frac{1}{2}(1 + a\cosh 2\lambda_V dt + 2\sqrt{a}\cosh\lambda_V dt) \right. \\
&\quad \left. \times \sum_{k=0}^{\infty} a^k \cosh 2\lambda_V k dt + \sqrt{a}\sinh\lambda_V dt (1 + 2\sqrt{a}\cosh\lambda_V dt) \sum_{k=0}^{\infty} a^k \sinh 2\lambda_V k dt \right] \\
&= \frac{\alpha k_B T b^2 dt^3}{2m^2 a \sinh^2 \lambda_V dt} \left[-\frac{1}{2} \frac{4b(1 - \frac{\Omega_0^2 dt^2}{4})}{1 - a} + \frac{(1 - a)\sqrt{a}\cosh\lambda_V dt + \frac{1}{2}(1 - a^2)}{1 + a^2 - 2a\cosh 2\lambda_V dt} \right] \\
&= \frac{\alpha k_B T b^2 dt^2}{2m^2 a \sinh^2 \lambda_V dt} \frac{(\frac{\alpha}{2m})^2 - \Omega_0^2(1 - \frac{\Omega_0^2 dt^2}{4})}{\frac{\alpha}{2m}\Omega_0^2} \\
&= \frac{k_B T}{m\Omega_0^2},
\end{aligned} \tag{62}$$

which is exactly the same result as for the underdamped case Equation (50). We similarly obtain the result for the variance of the velocity:

$$\begin{aligned}
\langle (v^n)^2 \rangle &= \frac{2\alpha k_B T b^2 dt}{m^2 a \sinh^2 \lambda_V dt} \left(1 - \frac{\Omega_0^2 dt^2}{4}\right)^2 \left[-\frac{1}{2}(1 + a - 2\sqrt{a}\cosh\lambda_V dt) \sum_{k=0}^{\infty} a^k + \frac{1}{2}(1 + a\cosh 2\lambda_V dt - 2\sqrt{a}\cosh\lambda_V dt) \right. \\
&\quad \left. \times \sum_{k=0}^{\infty} a^k \cosh 2\lambda_V k dt - \sqrt{a}\sinh\lambda_V dt (1 - \sqrt{a}\cosh\lambda_V dt) \sum_{k=0}^{\infty} a^k \sinh 2\lambda_V k dt \right] \\
&= \frac{2\alpha k_B T b^2 dt}{m^2 a \sinh^2 \lambda_V dt} \left(1 - \frac{\Omega_0^2 dt^2}{4}\right)^2 \left[\frac{b\frac{\Omega_0^2 dt^2}{2}}{1 - a} + \frac{(1 - a)\sqrt{a}\cosh\lambda_V dt - \frac{1}{2}(1 - a^2)}{1 + a^2 - 2a\cosh 2\lambda_V dt} \right] \\
&= \frac{2\alpha k_B T b^2 dt^2}{m^2 a \sinh^2 \lambda_V dt} \left(1 - \frac{\Omega_0^2 dt^2}{4}\right) \frac{\Omega_0^2(1 - \frac{\Omega_0^2 dt^2}{4}) - (\frac{\alpha}{2m})^2}{4\frac{\alpha}{2m}} \\
&= \frac{k_B T}{m} \left(1 - \frac{\Omega_0^2 dt^2}{4}\right),
\end{aligned} \tag{63}$$

which is also identical to the comparable underdamped result from Equation (51). The averages of potential and kinetic energies are therefore given by Equations (52) and (53) also for the overdamped regime. Thus, we can now conclude that our method produces the *exact* Boltzmann distribution for any physical dynamics, underdamped or overdamped, regardless of time step, frequency (potential curvature), damping parameter or temperature.

3.2.3. *Regime C: $\Omega_V = \pi/dt$*

This somewhat unphysical regime, is characterised by conditions similar to the overdamped regime B; i.e. $\alpha dt/2m < 1$ ($a > 0$) and $(\frac{\alpha}{2m})^2 > \Omega_0^2(1 - \frac{\Omega_0^2 dt^2}{4})$. However, while regime B corresponds to $\Omega_0 dt < \sqrt{2}$, regime C is defined for $\Omega_0 dt > \sqrt{2}$, resulting in $\Lambda_\pm < 0$. Thus, this regime is typically reached for large time steps when simulating lightly damped dynamics, and it is the precursor for violating formal stability, given by $\Omega_0 dt < 2$ (see Figure 1). At $T = 0$ and with initial conditions $r^0 \neq 0$, the dynamics is characterised by r^n and v^n alternating their signs every time step (i.e. $\Omega_V = \pi/dt$) with an exponentially decaying envelope. For $T > 0$, the thermodynamic properties of this regime can, therefore, be evaluated similarly to that of the overdamped regime, since we here can write

$$\mathcal{V}^k = (-1)^k \mathcal{U} \begin{pmatrix} \Lambda_+^k & 0 \\ 0 & \Lambda_-^k \end{pmatrix} \mathcal{U}^{-1}, \tag{64}$$

where $\mathcal{U}$ and $\Lambda_\pm$ are given by Equations (56)–(59). Apart from the alternating sign change in Equation (64), this is identical to the corresponding mapping in regime B. This implies that Equations (62) and (63) from regime B for the variances of r^n and v^n also apply here. Thus, we conclude that the thermodynamic results, Equation (52) and Equation (53), are found also in this regime, which is physically unreasonable, yet numerically accessible.

3.2.4. Regime D: unphysical

This unphysical regime, $\alpha dt/2m > 1$ (yet within the formal stability criterion $\Omega_0 dt < 2$), is typically reached for large time steps when simulating moderate to strongly damped dynamics (see Figure 1). Since this regime involves one positive and one negative eigenvalue, we can map this onto the analysis from regime B with $\sqrt{a} \to \sqrt{|a|}$ and $\lambda_V \to -\lambda_V$ to complete the analysis and obtain the correct distributions.

4. Discussion and conclusion

We have presented a new Verlet-type algorithm for simulating Langevin dynamics. The method, written in the three different forms: Equations (19) and (20), Equations (21) and (22) or Equations (24) and (25), is aligned with other published methods that have demonstrated first and second order accuracy for both simulated trajectories and derived thermodynamic quantities [6–15]. Linear analysis demonstrates that the method of this paper is robust and capable of providing *exact* representation of both diffusive behaviour in a flat potential and the Boltzmann distribution in a harmonic potential regardless of damping and frequency (curvature of the potential). The exact distribution is obtained for any time step subject to the usual Verlet stability limit and the condition $\alpha dt < 2m$, which is a necessary requirement for a meaningful attenuated trajectory. The method is very simple and in the usual Verlet structure, which means that it can be readily implemented for any Langevin application, including molecular dynamics of many-particle systems with and without molecular constraints and other commonly used modelling features.

As a final note, we underline that a Verlet-simulated oscillator may not measure the exact temperature from the variance of the velocity [see Equation (53)]. This deficiency, which is common to all Verlet-type methods, arises from the known discrepancy between displacement and momentum as conjugated variables. This means that using the variance of the velocity for precisely assessing the temperature of a simulated system may be counterproductive. The appropriate criterion for obtaining the desired temperature is the achievement of correct statistical sampling, which is indeed given by our method.

Acknowledgements

NGJ acknowledges a very useful discussion with Zhaojun Bai. This project was supported in part by the US Department of Energy Project No. DE-NE0000536 000.

References

[1] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithm to Application* (Academic Press, San Diego, 1996).
[2] W.C. Swope, H.C. Andersen, P.H. Berens, and K.R. Wilson, J. Chem. Phys. **76**, 637 (1982).
[3] E. Hairer, C. Lubich, and G. Wanner, Acta Numerica **12**, 399 (2003).
[4] L. Verlet, Phys. Rev. **159**, 98 (1967).
[5] G. Parisi, *Statistical Field Theory* (Addison Wesley, Menlo Park, 1988).
[6] A. Brünger, C.L. Brooks, and M. Karplus, Chem. Phys. Lett. **105**, 495 (1984).
[7] W. Wang and R.D. Skeel, Mol. Phys. **101**, 2149 (2003).
[8] D.L. Ermak and H. Buckholz, J. Comput. Phys. **35**, 169 (1980).
[9] W.F. van Gunsteren and H.J.C. Berendsen, Mol. Phys. **45**, 637 (1982).
[10] R.D. Skeel and J.A. Izaguirre, Mol. Phys. **100**, 3885 (2002).
[11] A. Ricci and G. Ciccotti, Mol. Phys. **101**, 1927 (2003).
[12] S. Melchionna, J. Chem. Phys. **127**, 044108 (2007).
[13] G. Bussi and M. Parinello, Phys. Rev. E **75**, 056707 (2007).
[14] F. Thalmann and J. Farago, J. Chem. Phys. **127**, 124109 (2007).
[15] B. Mishra and T. Schlick, J. Chem. Phys. **105**, 299 (1996).