Energy fluctuations induced by the Nosé thermostat

D. M. Bylander and Leonard Kleinman

Department of Physics, University of Texas, Austin, Texas 78712
(Received 20 April 1992; revised manuscript received 29 June 1992)

The Nosé thermostat is a method for simulating a canonical ensemble in molecular-dynamics calculations. It is generally agreed that for large values of Q, the parameter determining the response of the thermostat to the system, the approach to equilibrium is slow and for intermediate values it is rapid. However, for small Q Nosé found the thermostat oscillates rapidly and does not equilibrate with the system; we found it equilibrates but yields incorrect values for energy fluctuations, whereas Cho and Joannopoulos found it equilibrates very slowly, eventually yielding correct energy fluctuations. Here we demonstrate how these differences arise.

The Nosé[1,2] thermostat is a method for simulating a canonical ensemble at a predetermined temperature in molecular-dynamics calculations. Although all physical averages can be calculated from any ensemble, fluctuations are well known to be ensemble dependent.[3]. Furthermore, in some cases thermostats can become essential, e.g., to keep electrons on the Born-Oppenheimer surface in Car-Parinello calculations.[4] We[5] recently performed quantum-mechanical molecular-dynamics calculations of the density dependence of the properties of liquid Na using the Nosé thermostat to set the temperature. However, we found that although all ordinary average values that we calculated were reasonable, the total energy fluctuations needed to obtain the specific heat,

$$C_v = (\langle E^2 \rangle - \langle E \rangle^2)/NkT^2 , \qquad (1)$$

were not. We noted that our thermostat "mass" Q was relatively small, resulting in a tight coupling between the thermostat and the system, which caused the kinetic energy fluctuations to be very small and fairly rapid. We speculated that if the coupling was substantially reduced and the time of integration extended to include several oscillations of the thermostat, that the specific heat would be correctly obtained. In this paper we cast the thermostat equation of motion in a new form and from detailed computations show that for small Q the specific heat is not correctly obtained from Eq. (1) although equilibrium is quickly achieved. For intermediate Q a numerically reasonable value for C_v is obtained and for a sufficiently large Q that the kinetic energy fluctuations, which conserve total energy, are not directly reflected in the total energy, the integration time to achieve ergodicity is prohibitively long. After this work was completed we became aware of recent work by Nosé[6] and Cho and Joannopoulos[7] (CJ) which agrees with our conclusions for large and intermediate Q but for small Q Nosé finds that the thermostat oscillates 12 times faster than we found and does not reach equilibrium with the system whereas CJ find that equilibrium is reached after an extremely long time and that not only average values but also fluctuations are correctly obtained. We end this paper with an explanation for these seemingly contradictory results.

Nosé[1,2] postulates a Hamiltonian of virtual variables

$$H^* = \Sigma_i p_i^2/2ms^2 + \varphi(\mathbf{q}) + p_s^2/2Q + gkT \ln s , \qquad (2)$$

where φ is the potential energy of the system. The equations of motion are obtained from the Hamiltonian formalism and then may be used as obtained, or transformed back into real (primed) variables via $\mathbf{q} = \mathbf{q}'$, $\mathbf{p}_i = s\mathbf{p}_i'$, $s = s'$, $p_s = sp_s'$, and $dt = s\,dt'$. Nosé proved that if a microcanonical ensemble is taken for the extended system (i.e., $H^* = E$), the physical system will satisfy a canonical ensemble providing the extended system is ergodic and g is chosen to be[8] $3N + 1$ or $3N$ depending on whether virtual or real variables are used. More recently it was noted[9,10] that the equations in real variables do not depend on s and $p_s = Q\,ds/dt$ independently, but only on

$$\zeta = s^{-1}\frac{ds}{dt} . \qquad (3)$$

The equations of motion in the real variables may then be written[10]

$$\frac{d^2\mathbf{q}_i}{dt^2} = -\frac{1}{m}\frac{\partial\varphi}{\partial\mathbf{q}_i} - \zeta\frac{d\mathbf{q}_i}{dt} , \qquad (4)$$

$$\frac{d\zeta}{dt} = \frac{1}{Q}\left[\sum_i m\left|\frac{d\mathbf{q}_i}{dt}\right|^2 - 3(N-1)kT\right] . \qquad (5)$$

Nosé[6], using arguments we find hard to justify,[11] obtained for the thermostat's period

$$t_1 = 2\pi[Q/6(N-1)kT]^{1/2} \qquad (6)$$

in the small-Q limit and

$$t_2 = [2C_v/3(N-1)k]^{1/2}t_1 \qquad (7)$$

in the large-Q limit. Let us differentiate (5) and substitute (4) to obtain

$$\frac{d^2\zeta}{dt^2} = \frac{2}{Q}\sum_i m\frac{d^2\mathbf{q}_i}{dt^2}\cdot\frac{d\mathbf{q}_i}{dt}$$

$$= \frac{-2}{Q}\sum_i\left[\frac{d\mathbf{q}_i}{dt}\cdot\frac{\partial\varphi}{\partial\mathbf{q}_i} + \zeta m\left|\frac{d\mathbf{q}_i}{dt}\right|^2\right] . \qquad (8)$$

The first term in the square brackets represents the interchange of kinetic and potential energy among the N particles; its average value at equilibrium will be zero. The second term averages at equilibrium to 2ζ times the total kinetic energy, i.e., $3(N-1)\zeta kT$. If we ignore the fluctuations in both terms, (8) represents a simple harmonic oscillator with period t_1. We may substitute for $\sum_i m(d\mathbf{q}_i/dt)^2$ in (8) from (5) to obtain

$$\frac{d^2\zeta}{dt^2}+2\zeta\frac{d\zeta}{dt}+\frac{6}{Q}(N-1)kT\zeta=-\frac{2}{Q}\sum_i\frac{d\mathbf{q}_i}{dt}\cdot\frac{\partial\varphi}{\partial\mathbf{q}_i}\ .$$

(9)

Thus we see that ζ obeys a nonlinear differential equation with an inhomogeneous term equal to $2/Q$ times the sum of the rates at which the interatomic forces are doing work on the atoms. We normally start a calculation with $\zeta=0$ and the atoms at equilibrium positions with random velocities. Thus the forces on the atoms tend initially to become opposite to the velocities, causing ζ to grow negatively. This negative ζ can be seen from Eq. (4) to result in an acceleration of the particles, opposing the deceleration caused by the atomic forces. After equilibrium is reached the inhomogeneous term becomes stochastic, representing the random conservative interchange of kinetic and potential energy of the atoms. This stochastic driving force causes ζ to fluctuate and it is this fluctuating ζ in Eq. (4) which causes the total energy of the system (excluding the thermostat) to fluctuate.

The real and virtual formulation of the equations appear to be equivalent in most cases. However, in the virtual formulation the velocity autocorrelation function is very difficult to obtain without approximation.[7] The real formulation has been extended to include more than one ζ; this enabled the method to be applied to systems such as classical spins where canonical variables do not exist.[12,13]

We have examined how the choice of Q effects averages as well as fluctuations. To speed the process, rather than performing quantum-mechanical calculations, we have applied the interatomic potential of Paskin and Rahman[14] to liquid Na consisting of 18 360.6 Å^3 cubic supercells containing 432 atoms. The molecular dynamics were performed using Eqs. (4) and (5) with time steps $\tau=50$ a.u. $=1.2094\times10^{-15}$ s as in Ref. 1 where we chose $Q=2\times10^4 m_e a_0^2$. Here we use that Q and Q's 10^3 and 10^5 times larger. For the two smaller Q's, calculations were performed at 450, 550, and 650 K; but for the largest Q, where ergodicity was not obtained, even with over 115 000 time steps, only a 550 K calculation was performed. The kinetic energy and total energy at 550 K for the smaller two Q's are plotted in Figs. 1 and 2. Note the different scales of the kinetic energy fluctuations in the two cases. The 500 and 1000 time steps during which the systems were reaching equilibrium are not shown. The $Q=2\times10^9 m_e a_0^2$ calculation was started very close to equilibrium by taking the last atomic positions and velocities of the $Q=2\times10^4 m_e a_0^2$ calculation as the starting point. Nevertheless, in Fig. 3 one sees large damped sinusoidal oscillations in the total energy and the same oscillations with the rapid kinetic-potential energy interchanges superposed in the kinetic energy. (We also started this system off in the usual manner. After about 1000 time steps the system reached quasiequilibrium and began damped oscillations which were 40% larger than those in Fig. 3.) Finally after about 68 000 time steps, the amplitude of the oscillations stopped decreasing and became fairly random as shown in Fig. 4.

In Table I the pressure, average values of kinetic, potential and total energy, average values of their squares, and their fluctuations, $\langle x^2\rangle-\langle x\rangle^2$, are listed for different values of Q and temperature. In addition, t_1 or t_2 from Eq. (6) or (7) is given in units of 50 a.u. time steps for comparison with Figs. 1–4 and the specific heat calculated from Eq. (1) is compared with that calculated by fitting the total energies at $T=450$, 550, and 650 K with
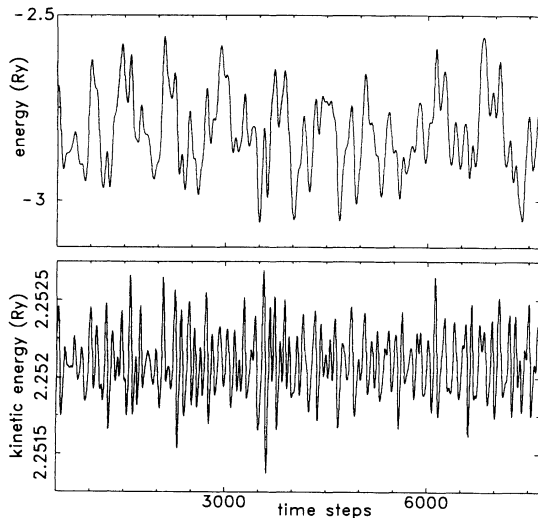

FIG. 1. Plot of kinetic energy and total energy of liquid Na 432-atom supercell at 550 K with $Q=2\times10^4 m_e a_0^2$.
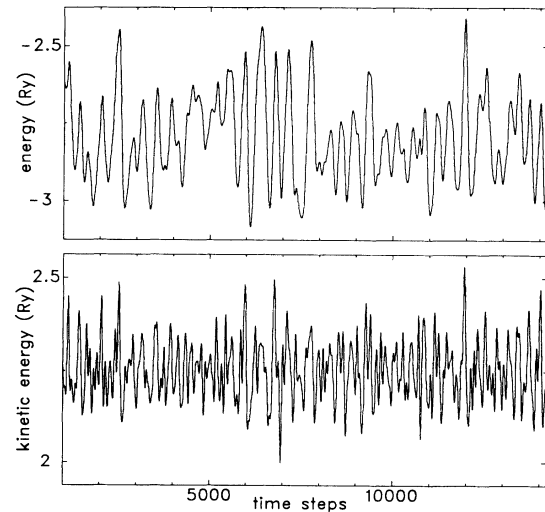

FIG. 2. Plot of kinetic energy and total energy of liquid Na 432-atom supercell at 550 K with $Q=2\times10^7 m_e a_0^2$.
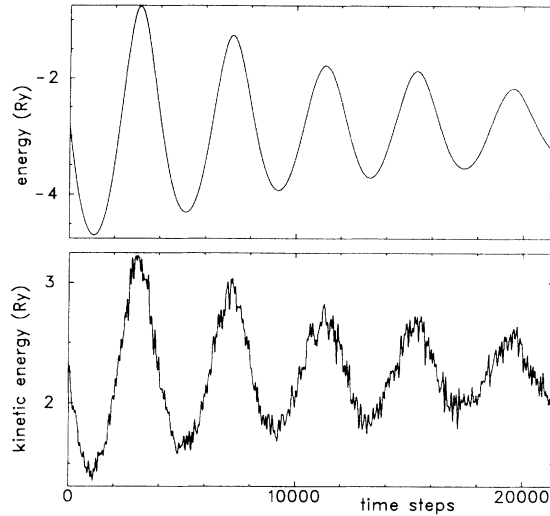
FIG. 3. Initial points in plots of kinetic energy and total energy of liquid Na 432-atom supercell at 550 K with $Q = 2 \times 10^9 m_e a_0^2$.

a quadratic and differentiating with respect to T. To reduce the dependence of the results on the starting and finishing times over which the averages are taken, for the smaller Q's, we average between the time step at which equilibrium has been reached, n, and $N - 1000$, between $n + 1$ and $N - 999$, $n + 2$ and $N - 998$, etc., and then average these 1000 averages. The same procedure but with 5000 averages is used for the largest Q. The first thing to note is that average values are not very sensitive to the choice of Q but fluctuations are. Since the atomic velocities are forced back to their correct average value by ζ whenever they fluctuate away from it, and the response of ζ to these fluctuations is controlled by Q, one would expect the kinetic energy fluctuations to be most Q dependent. Indeed, one finds $\langle (\Delta E_K)^2 \rangle$ increases by a
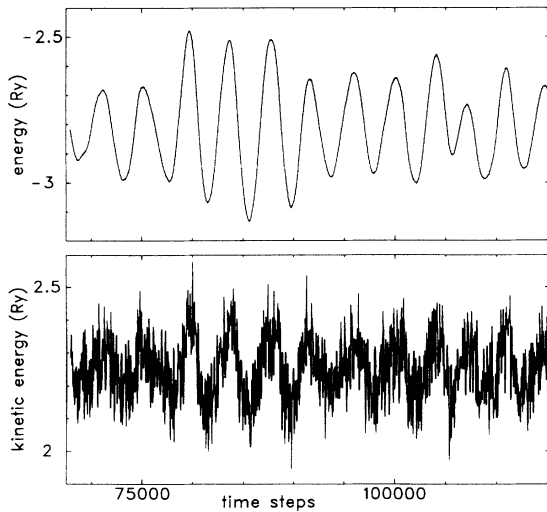

FIG. 4. Continuation of Fig. 3 plot after equilibrium has been reached.

TABLE I. Pressure and various energy averages at various temperatures and thermostat masses Q. The energies are for the 432-atom atom supercell. The periods t_1 and t_2 (for $Q = 2 \times 10^9$) from Eqs. (6) and (7) are given in units of $\tau = 50$ a.u. time steps. The specific heats are calculated from the energy fluctuations and from fitting $\langle E \rangle$ at the three temperatures.

T (K)	450	450	550	550	550	650	650
$Q (m_e a_0^2)$	2×10^4	2×10^7	2×10^4	2×10^7	2×10^9	2×10^4	2×10^7
t_1/τ or t_2/τ	8.3737	264.80	8.3737	264.80	4100	8.3737	264.80
$\langle P \rangle$ (bars)	676.5	689.3	2024.9	2095.3	2028.5	3648.8	3453.1
$\langle E_K^2 \rangle$ (Ry^2)	3.395247	3.400486	5.071911	5.078317	5.079947	7.083909	7.093634
$\langle E_K \rangle$ (Ry)	1.842619	1.842591	2.252090	2.252116	2.251924	2.661561	2.661535
$\langle (\Delta E_K)^2 \rangle$ (Ry^2)	4.2×10^{-8}	0.005343	4.4×10^{-8}	0.006289	0.008787	4.1×10^{-8}	0.009865
$\langle E_P^2 \rangle$ (Ry^2)	31.885227	32.039779	25.731255	25.545726	25.696364	20.242968	20.366273
$\langle E_P \rangle$ (Ry)	-5.646051	-5.659780	-5.071488	-5.053239	-5.067969	-4.497638	-4.511150
$\langle (\Delta E_P)^2 \rangle$ (Ry^2)	0.007335	0.006666	0.011259	0.010506	0.012050	0.014223	0.015794
$\langle E^2 \rangle$ (Ry^2)	14.473445	14.581885	7.960287	7.865137	7.953646	3.385423	3.447707
$\langle E \rangle$ (Ry)	-3.803432	-3.817189	-2.819398	-2.801122	-2.816046	-1.836076	-1.849615
$\langle (\Delta E)^2 \rangle$ (Ry^2)	0.007353	0.010954	0.011282	0.018851	0.023533	0.014247	0.026630
$C_v [\langle (\Delta E)^2 \rangle]$ (cal/mol K)	4.1692	6.2110	4.2823	7.1554	8.9323	3.8718	7.2371
$C_v (\partial E / \partial T)$ (cal/mol K)			7.1561	7.1569			

TABLE II. Energy fluctuations in Ry^2 using different time steps, integration algorithms, and variables: (I) 50 a.u., Verlet, real; (ii) 10 a.u., sixth-order Gear, real; (III) 2 a.u., sixth-order Gear, real; (IV) 10 a.u., sixth-order Gear, virtual.

	I	II	III	IV
$\langle(\Delta E_K)^2\rangle$	4.4×10^{-8}	3.2×10^{-8}	6.6×10^{-8}	4.7×10^{-8}
$\langle(\Delta E_P)^2\rangle$	11.26×10^{-3}	9.14×10^{-3}	9.04×10^{-3}	10.64×10^{-3}
$\langle(\Delta E)^2\rangle$	11.28×10^{-3}	9.16×10^{-3}	9.06×10^{-3}	10.66×10^{-3}

factor of 2×10^5 on increasing Q by a factor of 10^5 at 550 K or by a factor of 10^3 at 650 K. On the other hand, these changes in Q increase $\langle(\Delta E)^2\rangle$ by only a factor of 2.

The specific heat obtained by differentiating the three-point $E(T)$ curve is only accurate at the central point, 550 K, where it shows essentially no Q dependence. Since all average values in Table I are fairly Q independent, we think we are safe in assuming that $C_v=7.16$ cal/mol K is close to being the correct specific heat at 550 K for the force constant model we are using. The specific heat calculated from Eq. (1) is much too small for the smallest Q and too large for the largest Q. However, we note in Fig. 4 that there are three successive oscillations in the energy much larger than the rest. We believe these oscillations are real but fairly unusual. If one begins averaging from the 90 640th time step, $C_v=5.967$ cal/mol K is obtained. Since the result depends on the point from which one starts counting the data after equilibrium has been achieved, one must conclude that we have not followed the system for a sufficiently long time to reach ergodicity. The periods of ζ with increasing Q at 550 K (estimated by counting oscillations) are 104τ, 155τ, and 3915τ and those of the kinetic energy are 88τ, 130τ, and 100τ. The 100τ oscillations are seen on top of the 3915τ oscillations of the entire system in Fig. 4. In a microcanonical calculation we found a kinetic energy period of 105τ. Thus for the two smaller Q's it appears that the interaction of the thermostat with the kinetic energy causes large ζ and smaller kinetic-energy E_K deviations away from their natural frequencies of oscillation.[15] Interestingly for $Q=2\times10^7m_e a_0^2$, where ζ does not oscillate at the natural frequency (265τ) predicted by Eq. (6), E does. Also from a comparison of the two C_v's in Table I we conclude that $\langle(\Delta E)^2\rangle$ is correctly obtained at intermediate Q whereas a comparison of $\langle(\Delta E_K)^2\rangle$ with the theoretical value of 0.00784 Ry indicates it is not. This value is obtained from the canonical ensemble average at 500 K,

$$\langle(\Delta E_K)^2\rangle=\tfrac{3}{2}(N-1)(kT)^2 . \tag{10}$$

The Nosé thermostat is known not to generate a canonical ensemble for hard spheres[7] or for the one-dimensional harmonic oscillator[9] (although the extension of Refs. 12 and 13 does). Whether it does or not for large systems has been open to question.

Because a system with no lower bound to its energy is not ergodic and the $gKT\ln s$ in the Nosé Hamiltonian has no lower bound,[16] we believe the correct question should

be as follows: Does the Nosé thermostat result in large systems which accurately approximate a canonical ensemble and, if so, what are the necessary conditions for this to occur? Since the small Q case absolutely failed to be canonical for us and for Nosé[6] but eventually became so for CJ,[7] we examine it further. Because the quantum-mechanical calculations of Ref. 5 were so time consuming, we used the Verlet algorithm and the time step τ which is not small with respect to t_1, although it is compared to the smallest oscillation we could observe in ζ, E_K or E_P. We therefore repeated the calculation with time steps $\tau/5$ and $\tau/25$ with real variables and $\tau/5$ with virtual variables using the same sixth-order Gear predictor-corrector method used in Ref. 7. The results are shown in Table II. To the accuracy of the calculations these are identical; however, the τ time-step calculations showed no sign of $t_1=8.37\tau$ oscillations whereas the other three calculations had oscillations in $\dot{s}/s$ and kinetic energy of period $8.38\tau\pm0.07\tau$ after the virtual variable period was divided by $\langle s\rangle=0.9831$ to yield real time. These oscillations are of about the same amplitude in the kinetic energy as longer period $89\tau\pm6\tau$ oscillations and of much smaller amplitude in $\dot{s}/s$ than its $100\tau\pm\tau$ oscillations.

This explains why Nosé reported very rapid t_1 oscillations for small Q which we did not originally see. The question of whether equilibrium is reached appears to be semantics. It appears to be reached in the sense that no average values will change with increased integration time but it is not thermal equilibrium. Next, we gave the system a small drift velocity by adding a constant to the velocity of each atom at the end of the $\tau=10$ a.u., real variable run of Table II. This gave the system, which does not distinguish between drift and random velocities, an initial temperature 20 K above the thermostat's 550 K. The kinetic energy with real variables is plotted in Fig. 5; the virtual variable plot looks almost identical. Although the drift velocity has similar 8.37τ oscillations, together with longer period fluctuations, when it is subtracted from each atomic velocity, $\langle E_K\rangle$ decreases by about 0.08 Ry but the fluctuations in Fig. 5 are essentially unchanged. We see that $(\Delta E_K)^2$ in this figure is about 4×10^4 times larger than that in Fig. 1 and is fairly constant. The oscillations in Fig. 5 are between instantaneous temperatures of 570 and 530 K. Had we changed the velocities instantaneously to 570 K while maintaining zero-drift velocity, the system would have quickly damped to 550 K without large oscillations. When we used the Verlet algorithm with the same $\tau/5$ time step and $d\mathbf{q}_i/dt$ obtained from the nth and $n+1$st time steps
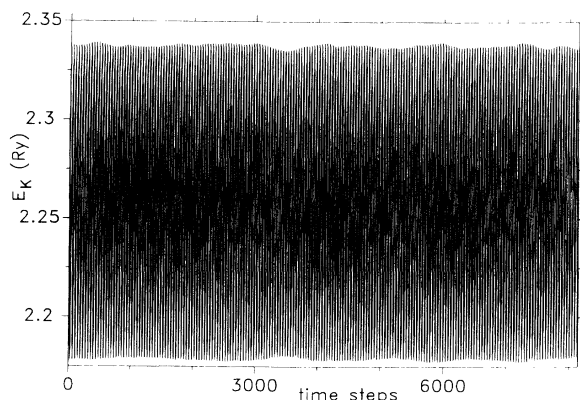
FIG. 5. A plot of kinetic energy when the 432 atoms are given a drift velocity vs standard time steps τ. The time steps used in integrating the equations of motion are $\tau/5$ and $Q = 2 \times 10^4 m_e a_0^2$.

FIG. 6. A plot of kinetic energy vs standard time steps τ when there are only 32 atoms per supercell and $Q = 2 \times 10^4 \sqrt{431/31} \, m_e a_0^2$. The time steps used in integrating the equations of motion are $\tau/10$.

rather than the $n - 1$st and $n + 1$st, $(\Delta E_K)^2$ dropped quickly below 10^{-7} Ry^2. With the fourth-order Gear predictor-corrector the E_K oscillations slowly increased and with the correct Verlet algorithm a result similar to Fig. 5 was produced. Summing Eq. (4) over all the atoms, we note that if there is no initial-drift velocity, none will ever develop but if there is an initial-drift velocity, the center-of-mass momentum is no longer a constant of the motion. Thus the Nosé thermostat is not Galilean invariant. In the center-of-mass frame the results are fairly insensitive to the accuracy of the integration scheme; in any other frame the results are extremely sensitive to the accuracy of the integration scheme, and probably are meaningless. That the center-of-mass frame is the only correct one can be seen in the following way. The Nosé Hamiltonian, Eq. (2), is easily seen to conserve the virtual momentum $\Sigma \mathbf{p}_i = s \Sigma \mathbf{p}_i'$. If this momentum conservation of the system is included with the energy conservation of the system plus thermostat in deriving the partition function, the result is extremely peculiar and definitely not canonical unless $\Sigma \mathbf{p}_i = 0$ in which case one obtains Nosé's canonical partition function times $\delta(\Sigma \mathbf{p}_i')$. This $\delta(\Sigma \mathbf{p}_i')$ is unphysical in that a physical system exchanges momentum with the thermal reservoir but otherwise causes no problems.[8] This has previously been asserted by Wallace[17] and just confirmed by Cho and Joannopoulos.[18]

Finally we attempted to check the assertion of Cho and Joannopoulos[7] that even for very small Q the system would eventually equilibrate, yielding all the correct moments of the kinetic energy, and that this equilibration could be hastened by having a small number of atoms in the supercell. We took $N = 32$ as did CJ, cut off the interatomic potential more rapidly to fit the smaller supercell, took a time step of 5 a.u., $T = 550$ K, and scaled Q to $2 \times 10^4 \sqrt{31/431} m_e a_0^2$ because the right-hand side of Eq.
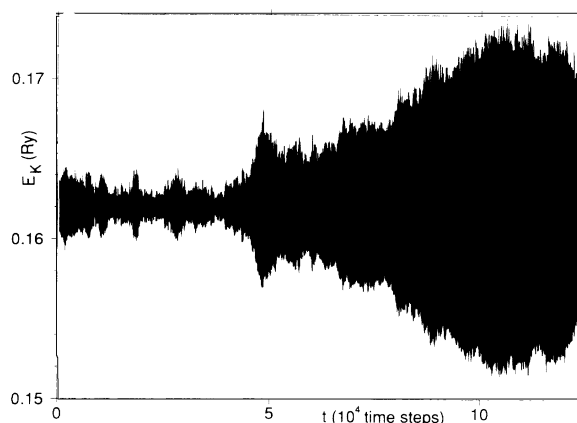
(5) is the deviation of the kinetic energy from its average value (which is proportional to $N - 1$) divided by Q, although Eq. (6) would suggest the scaling factor should be $31/431$. The kinetic energy which is plotted in Fig. 6 looks entirely different than that in Fig. 1 or in the corresponding cases in Table II where we used shorter time steps than in Fig. 1. The oscillations have a period in perfect agreement with the theoretical value $t_1 = 16.17\tau$. Between time steps 900 001 and 1 246 090 where the fluctuations are largest we find $\langle E_K \rangle = 0.161\,983\,7$ Ry in perfect agreement with the theoretical value of $0.161\,976\,4$ Ry but $\langle (\delta E_K)^2 \rangle = 4.035 \times 10^{-5}$ Ry^2 compared to the theoretical 5.642×10^{-4} Ry^2. Whether or not $\langle (\delta E_K)^2 \rangle$ and higher moments will eventually become correct after many millions of time steps we cannot say. What seems peculiar is not that it is taking the system a long time to obtain a good sample of phase space but that the magnitude of the fluctuations although ebbing and flowing has a general tendency to increase with time. We know that if the quasiergodic hypothesis holds, the Nosé thermostat yields a canonical ensemble in the center-of-mass frame. We have mentioned two systems (hard spheres and the one-dimensional harmonic oscillator) where the hypothesis fails. The question remaining to be answered is, in an infinite system of density ρ which is approximated by periodic supercells containing N atoms, does the quasiergodic hypothesis either fail or hold or are there regions of (ρ, N, Q) space in which it holds and other regions in which it fails?

[1]Shuichi Nosé, Mol. Phys. **52**, 256 (1984).

[2]Shuichi Nosé, J. Chem. Phys. **81**, 511 (1984).

[3]Duane C. Wallace, in *Electronic Structure, Dynamics, and Quantum Structural Properties of Condensed Matter*, edited by Jozef T. Devreese and Piet Van Camp (Plenum, New York, 1985).

[4]Peter E. Blochl and M. Parrinello, Phys. Rev. B **45**, 9413 (1992).

[5]D. M. Bylander and Leonard Kleinman, Phys. Rev. B **45**, 9663 (1992).

[6]Shuichi Nosé, Prog. Theor. Phys. Suppl. **103**, 1 (1991).

[7]K. Cho and J. D. Joannopoulos, Phys. Rev. A **45**, 7089 (1992).

[8]With momentum conservation $3N$ becomes $3N-3$.

[9]W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).

[10]Shuichi Nosé, Mol. Phys. **57**, 187 (1986).

[11]For small Q, Nosé combined real time and virtual momentum variables and then assumed the momentum fluctuations could be ignored. For large Q he replaced $\delta T/\delta s$, which can be infinite, by $[\langle(\delta T)^2\rangle/\langle(\delta s)^2\rangle]^{1/2}$.

[12]Aurel Bulgac and Dimitri Kusnezov, Phys. Rev. B **42**, 5045 (1990).

[13]D. Kusnezov, A. Bulgac, and W. Bauer, Ann. Phys. (N.Y.) **204**, 155 (1990).

[14]A. Paskin and A. Rahman, Phys. Rev. Lett. **16**, 300 (1966). We use their potential *LRO*-1 but rather than truncating it abruptly as they did, we smoothly reduce it to zero between $r=6.6605$ and 90051 Å.

[15]For $Q=2\times10^9$ both E and ζ oscillate with a period of 3915τ which is 4.5% smaller than the predicted value of Eq. (7). This may be due to the approximate nature of Eq. (7).

[16]We thank Professor N. G. van Kampen for calling this to our attention.

[17]Duane C. Wallace (private communication).

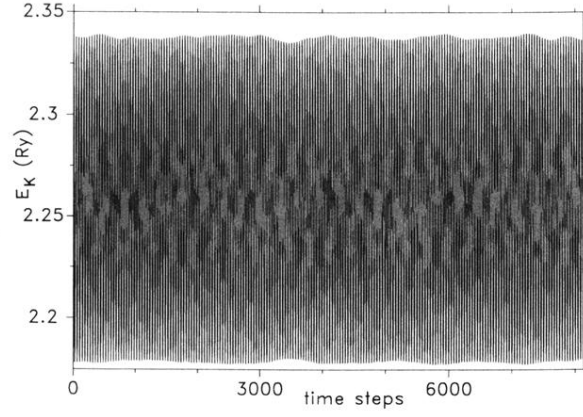[18]K. Cho and J. D. Joannopoulos (private communication).

FIG. 5. A plot of kinetic energy when the 432 atoms are given a drift velocity vs standard time steps τ. The time steps used in integrating the equations of motion are $\tau/5$ and $Q = 2 \times 10^4 m_e a_0^2$.