Application of the Nosé-Hoover method to optimization problems

I. Fukuda
Fujitsu Limited, 1-9-3, Nakase, Mihama-ku, Chiba-shi, Chiba 261-8588, Japan
(Received 8 August 2000; revised manuscript received 2 November 2000; published 13 June 2001)

A solution for continuous optimization problems is proposed using the Nosé-Hoover method. The proposed method aims for compatibleness, which has been a problem in many past solutions, between two requirements: searching with a high probability for finding candidates for the optimal points, and searching quickly in a feasible region. The Nosé-Hoover equation is used, where coordinates of a physical system are treated as the decision variables in a given optimization problem and a potential function is replaced by $-k_BT$ times the logarithm of an arbitrary density function for coordinate variables. The density can be set such that the visiting weight of the orbits to the equation has high values at areas where the objective function of the problem has low (high) values. Furthermore, a high value for the speed of the orbits can be set independently. Under an assumption of ergodicity, these values for the visiting weight and speed of the orbits are realized by long-time limits. Consequently, the two requirements can be satisfied. In numerical simulations assuming an objective function, the finite-time validity of the properties formulated with the long-time limits and the applicability of the proposed method to actual optimization problems were confirmed.

DOI: 10.1103/PhysRevE.64.016203 PACS number(s): 05.45.−a, 45.10.Db, 02.30.Hq, 02.60.Pn

I. INTRODUCTION

Many investigations into optimization problems have been done, and corresponding solutions have been applied, in a variety of areas, including energy minimization [1] in physics and chemistry, and design problems in engineering [2]. The methods for solving these problems must be able to find an optimal point of a given function by searching across a space while avoiding traps, such as local minima, and to find the point efficiently. The space to be searched is called the feasible region and the function whose optimal (minimal or maximal) point is sought is called the objective function.

Two kinds of optimization methods have been used: global and local. Global methods, including genetic algorithm [3], the simulated annealing method [4], the trajectory method [5], the tunneling method [6], and the filled function method [7], have been developed [8] and have widely been used to search for global optimal points. Many local methods [9], including those developed for the unconstrained optimization problem and constrained nonlinear programming problem, have been studied and improved primarily for the purpose of finding local optimal points efficiently. However, since the requirements for both global search and efficient search are difficult to satisfy with a single general method, respective efforts have usually been made to solve each problem individually [10].

As a solution, one proposal is to use a method that is designed both to conduct a concentrated search for candidates for the optimal points and to search quickly in a feasible region. To construct such a method, the Nosé-Hoover method, which has been used in the area of molecular dynamics (MD) [11], is studied in this work. Nosé [12] conceived a method for temperature control of physical systems defined by individual Hamiltonians. In the method, the Nosé equations are defined (in terms of ''virtual'' variables and ''real'' variables) from the ''extended system'' constructed by adding a single degree of freedom to the physical system. It has been shown that using the equations the canonical distribution at the intended temperature for the physical system is obtained under suitable conditions. By transforming the equation, Hoover [13] derived the Nosé-Hoover equation, which has a simpler form and leads to intuitive understanding, and developed the non-Hamiltonian formalism. These extended system methods have been actively studied, and the methods have been applied not only to MD but also to various areas, including lattice gauge theory, reactive dynamics, vibrational relaxation, and a system of classical spins (see Ref. [14]). In this paper, the Nosé-Hoover method is applied to optimization problems.

An optimization method using the Nosé-Hoover equation with a potential energy given by a deformed objective function is proposed. In this method, the optimal point can be found by evaluating the values of the objective function one by one; these values are obtained during the traveling-through of the orbit to the ordinary differential equation (ODE). This method is used to try to solve the continuous optimization problems defined in Euclidean spaces. Section II shows that a value for the visiting weight of the orbits and a value for the speed of the orbits, where these two values are set arbitrarily and independently from each other, can be realized by long-time limits. These mechanisms enable both a concentrated search for candidates for the optimal points and a quick search (Sec. III). Accordingly, a trade-off relation between searching weight and searching speed, which has been a problem in many past solutions, does not exist (Sec. IV). In Sec. V, the finite-time applicability of this method is examined in numerical simulations using a simple objective function. Section VI summarizes the results of this study.

II. THE NOSÉ-HOOVER METHOD

For a classical physical system expressed with coordinates $q \equiv (q_1, \ldots, q_n)$, momenta $p \equiv (p_1, \ldots, p_n)$, identical masses ($\equiv$ unity), and potential energy $V(q)$, the Nosé-Hoover equation [13] is represented by

$$\dot{q}_i = p_i, \quad i = 1, \ldots, n,$$

$$\dot{p}_i = -D_i V(q) - \zeta p_i, \quad i = 1, \ldots, n, \tag{1}$$

$$\dot{\zeta} = (\|p\|^2 - \beta)/Q,$$

where β and Q are positive parameters. (A dot over a letter denotes time differentiation.) The introduction of variable ζ, which corresponds to a friction coefficient (but can be negative as well as positive), realizes a mechanism for decreasing the speed, $\|\dot{q}\| = \|p\|$, (when $\|p\|^2 \gtrsim \beta$) and for increasing the speed (when $\|p\|^2 \lesssim \beta$). As a result, the physical system can conform to the canonical distribution [13,14] at temperature $\beta/(nk_B)$ (k_B: Boltzmann's constant). Parameter Q can be considered to control the speed of the response between a heat bath at that temperature and the physical system.

Suppose that coordinates $q = (q_1, \ldots, q_n)$ are decision variables in a continuous optimization problem defined in an n-dimensional Euclidean space $\mathbf{R}^n$. Let D be a feasible region (a subset consisting of all $q \in \mathbf{R}^n$ satisfying a given constraint condition in the problem) and U an objective function in the problem. The objective function U is linked to a potential function V in Eq. (1) in this and the following sections. In this method, during an integration process for ODE (1), we evaluate the values for $U(q(t))$ individually in order to be able to adopt the best point as an ''optimal solution.'' In the following paragraphs, the speed and visiting weight are defined for *orbit* $t \mapsto q(t)$ to Eq. (1), and analyzed to show that the system enables the setting of these two values arbitrarily and independently and that these values can be realized by long-time limits.

It is shown that the canonical ensemble averages of functions of (q,p) are obtained under fulfillment of the canonical distribution to the physical system [12–14]. According to this, an equation regarding the speed,

$$\lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau \|dq/dt\|^2 dt = \int_{\mathbf{R}^n} \|p\|^2 \exp\left[-\frac{n}{2\beta} \|p\|^2 \right] dp \Big/ \int_{\mathbf{R}^n} \exp\left[-\frac{n}{2\beta} \|p\|^2 \right] dp = \beta, \tag{2}$$

is derived. The left-hand side of this equation is a long-time average value of the square of the *speed of the orbit*, while the right-hand side can be interpreted as a *setting value* (or predetermined value) of (the square of, this term is sometimes omitted,) the speed of the orbit. Namely, it is possible to set an arbitrary value for a speed of the orbits and realize the value by a long-time average. A desirable property on the speed can thereby be obtained. Furthermore, the physical system also has thermal fluctuations derived from the canonical distribution. The effect of such fluctuations is that effective escape from traps (e.g., local minima of U) may be expected. This effect would give a different feature from that given by an orbit with a constant speed (e.g., yielded by the Gaussian constraint method [14] or its analog).

Regarding the visiting weight, the following is similarly obtained for an arbitrary area $A_1 \subset \mathbf{R}^n$:

$$\lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau \chi_{A_1}(q(t)) dt = \int_{A_1} \exp\left[-\frac{n}{\beta} V(q) \right] dq \Big/ \int_{\mathbf{R}^n} \exp\left[-\frac{n}{\beta} V(q) \right] dq, \tag{3}$$

where

$$\chi_{A_1}: \mathbf{R}^n \to \mathbf{R}, \quad q \mapsto \begin{cases} 1 & \text{for } q \in A_1 \\ 0 & \text{otherwise.} \end{cases}$$

The left-hand side of this equation represents a long-time limit value of the visiting weight (the rate of sojourn time) of the orbit into A_1, and the right-hand side can be interpreted as the setting value of the visiting weight. Now, a simple link, $V = U$, means that the setting value of the visiting weight is automatically determined by β, which defines the setting value of the speed [Eq. (2)]. Thus, such a link is not suitable for the purpose of arbitrary and independent setting for these values. However, using an arbitrary ''density function'' ρ_Q, the replacement of

$$V = -(\beta/n) \ln \rho_Q, \tag{4}$$

enables arbitrary setting and its realization for the visiting weight.

Equation (4) brings an arbitrary invariant density with respect to the q variable (assuming that ρ_Q is a smooth, positive, integrable function on an open set of $\mathbf{R}^n$). That is, the flow generated by Eq. (1) has an invariant density (the density of an invariant measure for the flow with respect to Lebesgue measure dx on $\mathbf{R}^{2n+1}$) $\rho(q,p,\zeta) = \rho_Q(q) \exp[-(n/2\beta)(\|p\|^2 + Q\zeta^2)]$. On the basis of this fact and an assumption of ergodicity in the following sense, the properties of the orbits are stated simply. That is, the above long-time limits exist for ρdx—almost everywhere by Birkhoff's individual ergodic theorem [15]; and if the flow is ergodic with respect to the measure ρdx then Eq. (2) and the following hold for almost every initial value:

$$\lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau \chi_{A_1}(q(t)) dt = \text{const} \times \int_{A_1} \rho_Q(q) dq. \tag{5}$$

Therefore, an arbitrary setting value for the visiting weight of the orbits into each area, where that value is independent of the setting value for the speed, can be realized by a long-

time limit. Accordingly, the trade-off relation between searching weight and searching speed can be eliminated (Sec. IV). This paper focuses on a basic study on the effect of Eq. (4), without further extension of the method. Regarding Eqs. (2) and (5), they have been formulated using infinite time. Every solution to ODE (1) is assumed to be defined for $-\infty<t<\infty$, and ergodicity is also assumed.

III. APPLICATION TO OPTIMIZATION PROBLEM

In order to apply the Nosé-Hoover method to optimization problems, a method for setting the visiting weight and speed of the orbits is examined. In the rest of this paper, minimization problems are discussed ($-U$ is used for maximization). Suppose that objective function U is a smooth, bounded below map from $\mathbf{R}^n$ into $\mathbf{R}$.

If we use the mechanism for arbitrarily setting the visiting weight of the orbits in order to set high weight values to areas where U has low values, then we can perform concentrated searching for candidates for the optimal points; in other words, we can search for them at high probabilities of attainment. Here, candidates for the optimal points (*candidates*) simply means a set of points that have sufficiently low values for U (points in a neighborhood of a minimal point may also be suitable: an appropriate definition is made in each case). To realize such a setting, let density ρ_Q be linked to U such that ρ_Q has a higher value, as a point shifts to where the value of U is lower [i.e., $U(q_1)\geqslant U(q_2) \Rightarrow \rho_Q(q_1)\leqslant\rho_Q(q_2)$]. To link them, for example, let us introduce function $\Theta_U:\mathbf{R}\rightarrow\mathbf{R}$. In addition, regarding the constraint condition, let us introduce function $U_D:\mathbf{R}^n\rightarrow\mathbf{R}$ so as to constrain the orbits into the feasible region D. Let $\rho_Q(q) = \exp[-\Theta_U(U(q)+U_D(q))]$ (assuming that the conditions for ρ_Q are met). Then, from Eq. (5), the setting value of the visiting weight of the orbits into area A_1 is expressed by

$$P_Q(A_1)\equiv \int_{A_1}\rho_Q(q)dq \Big/ \int_{\mathbf{R}^n}\rho_Q(q)dq = \text{const}\times\int_{A_1}\exp[-\Theta_U(U(q)+U_D(q))]dq. \tag{6}$$

Suppose that Θ_U is an increasing function. To satisfy the constraint, let us set U_D such that it is a constant (e.g., zero) inside D and has large values outside D. As a result, for A_1 outside D, if $P_Q(A_1)\sim 0$ (by $\rho_Q\sim 0$) holds then the visiting weight (viz., rate of sojourn time) of the orbits falls to almost zero. Namely, the constraint condition is effectively satisfied. Inside D, in contrast, U_D has been set as a constant and U has to be considered. The link stated above between ρ_Q and U is valid by the assumption that Θ_U is an increasing function. Thus, Eq. (6) shows that P_Q, which can be understood as a set function for representing *searching weight* for each area in $\mathbf{R}^n$ or as a probability for that area, also has a high value at A_1 where U has low values. These relationships are emphasized as $\Theta_U(u)$ increases rapidly with increasing u. Adjustment of Θ_U enables deformation of proportions for undulation of U (see Sec. V). In this way, the setting value for the visiting weight of the orbits inside D can be determined by U and Θ_U.

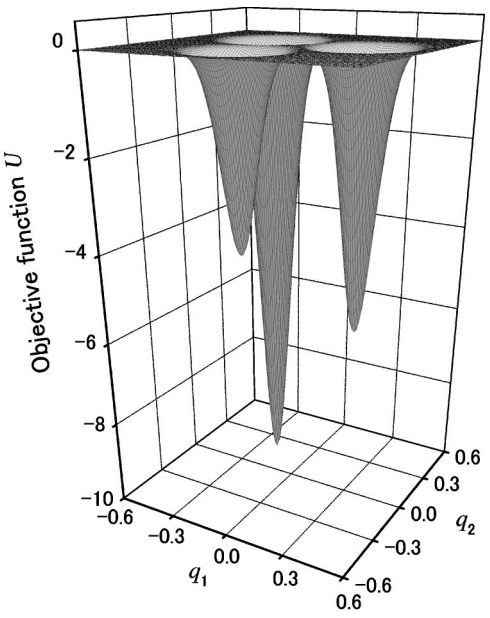

FIG. 1. Objective function U. Feasible region is a disk $\|q\| \leqslant 0.5$.

Let us consider the speed of the orbits as *searching speed* defined in the feasible region. This is estimated by $\|dq/dt\|^2$ ($\propto$ kinetic energy), and its value can be set by $\beta>0$ [Eq. (2)]. A high searching speed is desirable.

IV. TRADE-OFF BETWEEN SEARCHING WEIGHT AND SEARCHING SPEED

We require Eq. (4) in order to set searching weight arbitrarily, but we do not require it if we only set a high searching weight for candidates (Sec. III). A simple link $V=U$ using a low value for β serves the latter purpose [Eq. (3)]. If this is done, however, searching speed decreases [Eq. (2)]. Conversely, a high-probability search for candidates is not done if the link with a high value for β is used.

In approaches based on stochastic methods, the Metropolis algorithm (MA) [16] is a representative technique that is capable of adjusting searching weight. [In the algorithm, the acceptance probability of q' subsequent to point q is defined by $\min\{1,\exp[-\Delta U(q)/T]\}$, where $\Delta U(q)\equiv U(q')-U(q)$ and $T>0$ is a control parameter called temperature. Consider the searching weight to be determined by the equilibrium distribution $\exp(-U/T)\times\text{const.}$] Such a *trade-off relation* between the searching weight and the searching speed becomes a problem when using MA to solve optimization problems. If the value of parameter T is lowered, then an almost ideal distribution is obtained as $T\rightarrow 0$, which enables high-probability searches for candidates. When T is low, however, under adoption of *local generation* [a procedure where q' is selected from a certain defined neighborhood $N(q)$ of q], escape from *local trap* becomes difficult [i.e., as point q approaches near q_0 such that $U(q_0)<U(q_0')$ for all

TABLE I. Parameters for setting values.

Case No.	Visiting weight α_1	Visiting weight α_2	Square of speed β
1	1/600	2.0	40.0
2	1/900	−14.0	40.0
3	1/600	2.0	10.0

$q_0' \in N(q_0) - \{q_0\}$, most of the selected points are q''s with $\Delta U(q) > 0$; the acceptance probabilities of these points thus become $\exp[-\Delta U(q)/T]$ and low as $T \to 0$]. This thereby decreases the searching speed. Conversely, if T is high, the situation is analogous to that mentioned in the preceding paragraph; the acceptance procedure in the limit of $T \to \infty$ corresponds to a random search.

The simulated annealing method [4] employs temperature scheduling to solve this problem. The goal is efficient searching during a gradual change of the distribution toward the ideal one by lowering the value of control parameter T. Variations of the algorithm and successful results on their application have been reported (e.g., Ref [17]). From a theoretical aspect, convergence to the ideal distribution, as is discussed in Refs. [18,19], ensures effectiveness for long-time searching. However, as long as the local generation is assumed and an acceptance rule is defined such that the rule makes a displacement for $\Delta U(q) > 0$ smaller (viz., accep-

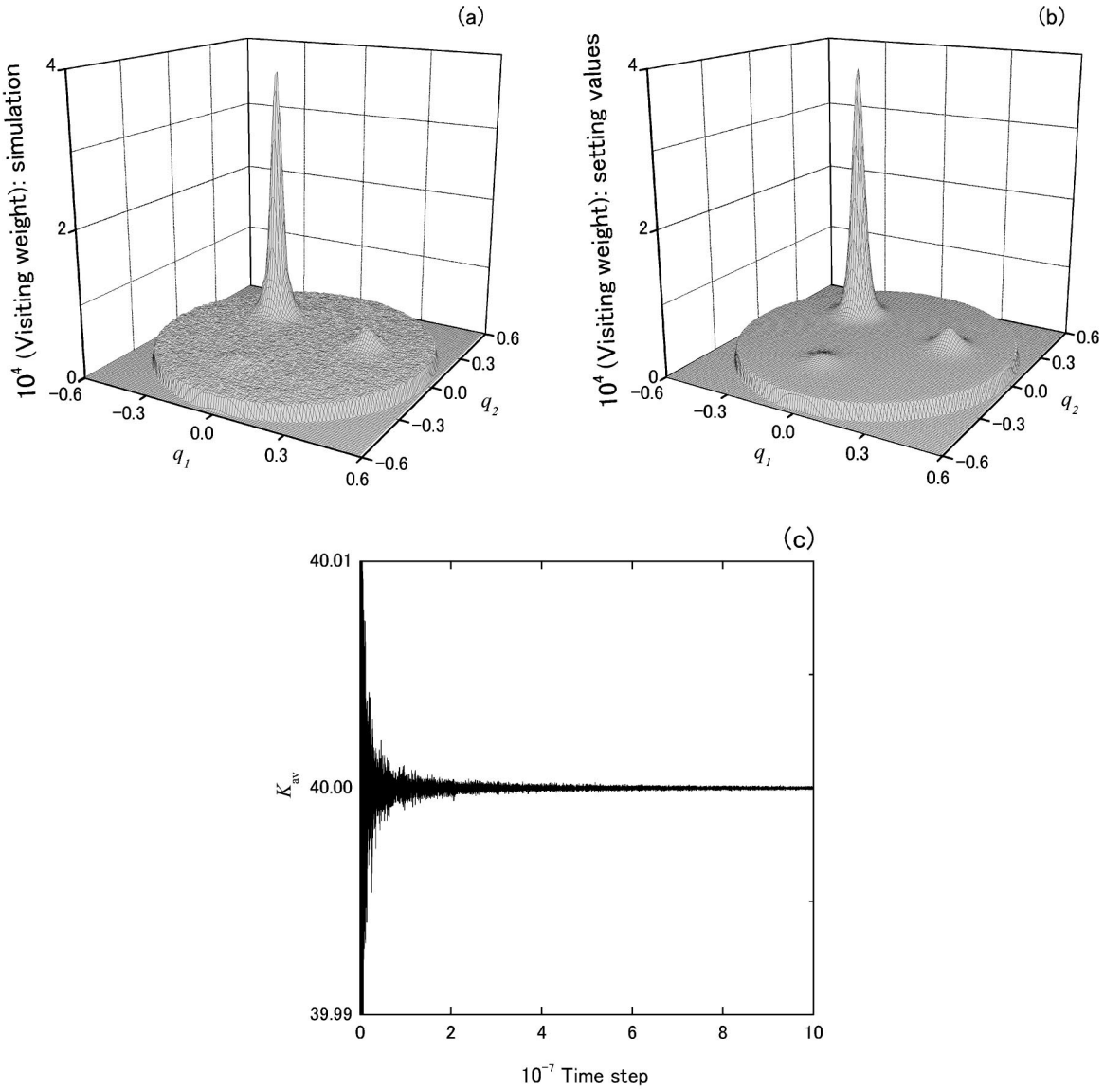

FIG. 2. (a) Simulation results for visiting weight of orbit in case 1. The frequency of visits by the orbit, n_ν/N, for each mesh ν is shown. n_ν: the number of times that the orbit passes mesh ν. $N \equiv 10^9$: the total number of time steps. (b) Setting values for the visiting weight of orbit in case 1. The value is the integral of normalized ρ_Q on each mesh. The points for the highest, second, and third peaks correspond to the global minimal, second lowest minimal, and third minimal points of the objective function, respectively. (c) Simulation results for the square of speed of orbit in case 1. The time average for this square of speed, K_{av}, is shown as a function of time steps. The setting value is 40.0.

tance probability of q' with $\Delta U(q)>0$ becomes smaller) as a certain defined control parameter is varied, the problem of the local trap cannot be avoided and becomes increasingly noticeable in the process, in principle. Owing to such a reason, more devices are usually employed for a more effective resolution [20].

The Nosé-Hoover method stated in the previous sections does not have such a trade-off relation, because the searching weight and the searching speed can be set independently (by the visiting weight and speed, respectively, of the orbits). This follows from the fact that the invariant density derived from the simple replacement by Eq. (4) has the form separated into the q component and the (p,ζ) component without common term k_BT. At the beginning of the process, we can set a high searching speed and set an advantageous searching weight for resolving a problem. Moreover, the method does not result in a smaller displacement as time progresses. The mechanism for increasing the speed (Sec. II) works effectively when the small displacement continues. A more formal statement is that $\|\Delta q(t)\| \equiv \|dq/dt\|\Delta t$, which is the principal quantity in $\|q'-q\| \equiv \|q(t+\Delta t)-q(t)\|$ (q' is the next point computed by a numerical integration), does not end in zero (for almost every initial value), because the long-time average of $\|\Delta q(t)\|^2$ becomes a setting value multiplied by $\Delta t^2>0$. The infinite time used in the theoretical formulation is required to represent a probabilistic assurance, such as what is called the law of large numbers.

V. NUMERICAL SIMULATION

As described in the previous sections, the effectual properties of the orbits were formulated using infinite time. In comparison, only finite time is available for applying the method to an actual optimization problem; we define an initial condition to the ODE and continue a numerical integration of the equation until a suitable predetermined stopping condition is satisfied. This section describes an examination of the application of the method to a simple optimization problem. In numerical simulations on the ODE, the finite-time applicability of the method was examined by verifying the validity within a finite time for Eqs. (2) and (5). Furthermore, the effect expected from differences in setting values for the visiting weight and speed of the orbits was examined.

A. Simulation conditions

Objective function U was set as the summation of three Gaussian form functions defined on a plane (viz., $n=2$), and feasible region D was a disk with its center at the origin and with a radius of r_0,

$$U(q_1,q_2) = -\sum_{m=1}^{3} a_m \exp(-b_m\|q-c_m\|^2), \tag{7}$$

$$D = \{q \equiv (q_1,q_2) \in \mathbf{R}^2 | \|q\| \leqslant r_0\}. \tag{8}$$

The parameters were $a_1=10.0$, $a_2=6.0$, $a_3=4.0$, $b_1=b_2=b_3=100.0$, $c_1=(-0.212,0.212)$, $c_2=(0.290,0.0776)$, $c_3=(-0.0776,-0.290)$, and $r_0=0.5$. (See Fig. 1.) Quantities q and $U(q)$, and time t were treated as dimensionless, for the sake of simplicity.

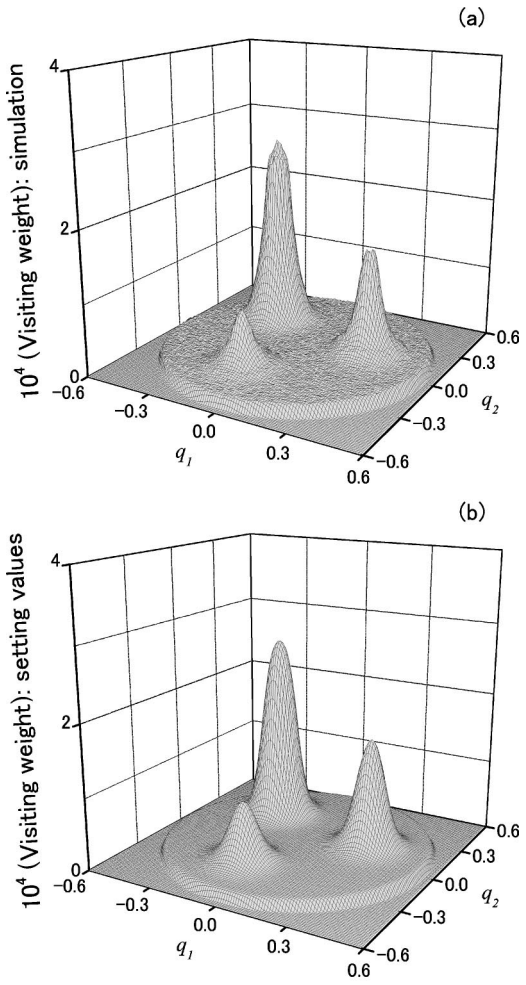

FIG. 3. Visiting weight of orbit in case 2: (a) simulation results, and (b) setting values (the points for the highest, second, and third peaks correspond to the global minimal, second lowest minimal, and third minimal points of the objective function, respectively). Calculations are similar to those in case 1 [Figs. 2(a) and 2(b)].

In an application of ODE (1), the functions for constraining the orbits into D and for setting the values of the visiting weight of the orbits, as described in Sec. III, were defined as follows:

$$U_D(q) = \begin{cases} 0 & \text{for } q \in D \\ d_1(\|q\|-r_0)^{d_2} & \text{otherwise,} \end{cases} \tag{9}$$

$$\Theta_U(u) = \alpha_1(u-\alpha_2)^3. \tag{10}$$

The parameters were $d_1=10^{10}$, $d_2=8$ for U_D, and $Q=10^{-3}$ in Eq. (1). Using the parameters for determining setting values, α_1 and α_2 for the visiting weight of the orbits and β for the square of the speed of the orbits, three cases were simulated (Table I). For case 1 and case 2 the setting values of the speed were identical but those of the visiting

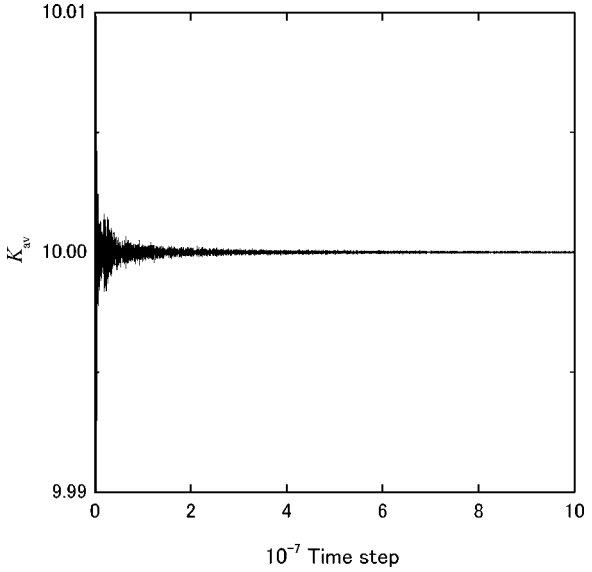

FIG. 4. Simulation results for the square of speed of orbit in case 3. The time average for this square of speed, K_{av}, is shown as a function of time steps. The setting value is 10.0.

weight were different. Conversely, for case 1 and case 3, the setting values of the visiting weight were identical, but those of the speed were different.

In numerical integrations of the ODE, the fourth-order Runge-Kutta method with a unit time step of $\Delta t=5\times10^{-5}$ was used along with an initial value of $q_1(0)=0.0$, $q_2(0)=0.1$, $p_1(0)=\sqrt{\beta/2}$, $p_2(0)=0.0$, and $\zeta(0)=0.0$.

Both sides of Eq. (5) were evaluated on each small square mesh (the sizes were 0.5 for the radius of D and 0.005 for the side of the mesh). The left-hand side of the equation for mesh ν was computed using the *frequency of visits* (viz., visiting weight on discrete time) by an orbit, n_ν/N, where n_ν is the number of n's such that $q(n\Delta t)\in\nu$ and N is the total number of time steps. The right-hand side, the setting value of the visiting weight, was evaluated by the integral of the density ρ_Q [viz., $P_Q(\nu)$ in Eq. (6)]. Regarding Eq. (2), a time average, $K_{\mathrm{av}}(t)\equiv(1/t)\int_0^t\|dq/ds\|^2ds$, was calculated for estimating the left-hand side of the equation. The right-hand side, the setting value of the speed, is equal to input parameter β.

B. Simulation results

In case 1, for the visiting weight of the orbit, Fig. 2(a) shows simulation results (by the frequency of visits with $N=10^9$) and Fig. 2(b) shows the setting values. The results of the simulation are visibly in good agreement with the setting values. Notice that the frequency of visits outside D is about zero. Namely, the constraint was effectively satisfied. For the speed, Fig. 2(c) shows simulation results for K_{av} as a function of time steps. It converged to a setting value of 40.0. In addition, simulations were performed for several initial values chosen randomly within $q\in D$, $\|p\|<10\sqrt{\beta}$, and $|\zeta|<10$. In these cases, results similar to case 1 were obtained. These results confirm the finite-time validity for Eqs. (2) and (5) along with the independence from the initial values.

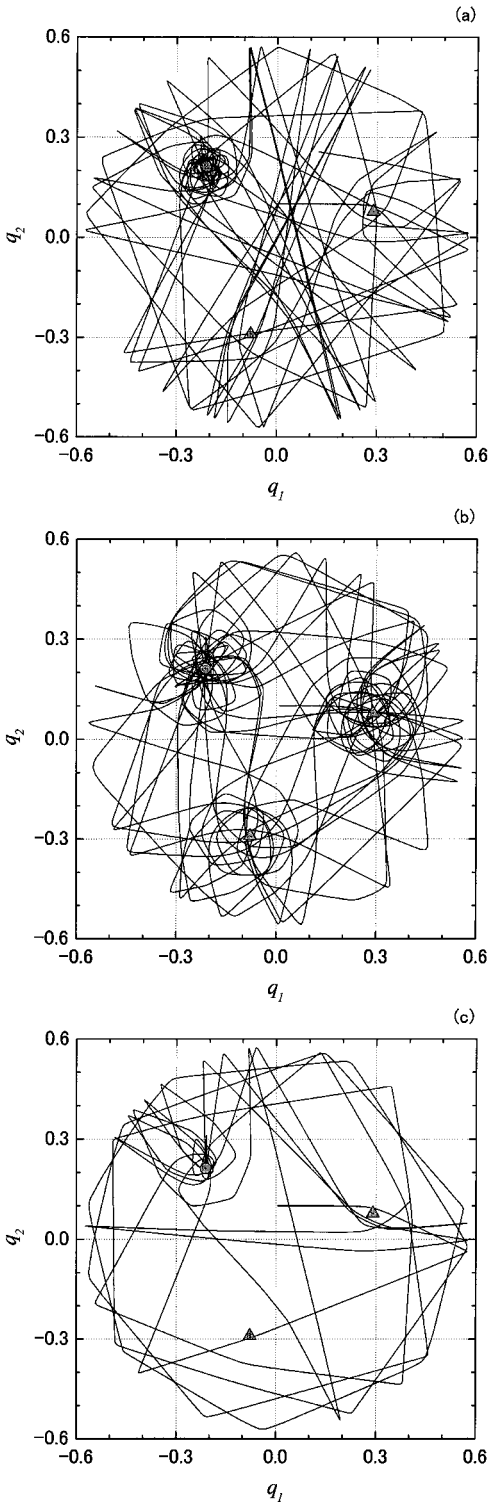

FIG. 5. Trajectories for 2×10^5 time steps. The symbol ○ indicates the global minimal point, and △ indicates local minimal point: (a) case 1, (b) case 2, and (c) case 3.

In case 2, Fig. 3 shows simulation results and the setting values for the visiting weight of the orbit. An agreement between them can be seen. Furthermore, the simulation results for the speed showed convergence to the setting value, similar to case 1. In case 3, the simulation results for the

speed converged to a setting value of 10.0 (Fig. 4) and those for the visiting weight showed agreement with the setting values, as in case 1. These results confirm that the visiting weight and speed can be set independently and be realized as individual setting values.

The difference in frequency of visits between case 1 [Fig. 2(a)] and case 2 [Fig. 3(a)] was caused by a difference in the setting values of the visiting weight of the orbits between the two cases. In a comparison of Figs. 2(b) and 3(b), the setting values for case 1 can be seen to reflect the value of the objective function more sensitively than those for case 2. [That is, much larger values for the visiting weight are assigned to areas where U has low values. The points for the highest peak, second peak, and third peak in Figs. 2(b) and 3(b) correspond to the points for the global minimum, second lowest minimum, and third minimum of U, respectively.] The simulation results until $N=10^9$ time steps show that the orbit in case 1 certainly visits near the optimal (global minimal) point at a high frequency of visits, while the orbit in case 2 visits all three minimal points more equally. This means that a difference in the setting value of the visiting weight of the orbits indeed affects the frequency of visits. Such an effect appeared within a shorter length of time. Figure 5 shows trajectories for 2×10^5 time steps for the orbits and the minimal points otherwise determined. As indicated in Figs. 2(a) and 3(a), behavior for searching minima in case 1 [Fig. 5(a)] is different from that in case 2 [Fig. 5(b)].

Case 1 and case 3 are different concerning the setting value of the speed of the orbits. The trajectories show that searching in case 1 [Fig. 5(a)] was performed over a larger area than that done in case 3 [Fig. 5(c)] for the same period of time. This means that the searching speed in case 1, which employs a larger setting value for the speed, is indeed larger than that in case 3.

The theoretical formulation of the properties of the orbits has been done on the assumption that the system is ergodic. A mathematical proof of ergodicity for a given system is widely known to be difficult. The agreement between simulation results and setting values shown above for both the visiting weight and the speed of the orbits supports ergodicity in this system. In a naive sense, ergodicity can be expected as the system grows complicated (e.g., the system is not an ideal one such as an integrable system, it consists of many degrees of freedom, or it has strong nonlinearity). This is also supported in numerical studies on the Nosé-Hoover equation and the Nosé equation [21–23]. Accordingly, ergodicity is expected for more complicated systems than that used in the present study. For given n and ρ_Q, under adoption of Eq. (4), the effect of parameters Q and β on the ergodicity is an important subject.

VI. CONCLUSION

A solution for continuous optimization problems has been proposed using the Nosé-Hoover method. This method enables compatibleness, which has been a problem in many past solutions, of the following two requirements: searching with a high probability for finding candidates for the optimal points, and searching quickly in a feasible region. Namely, this method does not have trade-off relation between searching weight and searching speed.

In the Nosé-Hoover equation used in this work, coordinates of a physical system are regarded as the decision variables in a given optimization problem and a potential function is linked to the objective function. To set the visiting weight of the orbits to the equation independently from the speed of the orbits, the potential function has been replaced by $-k_BT$ times the logarithm of an arbitrary density function for coordinate variables. It has been demonstrated that the density function gives an invariant density with respect to the coordinate variables and that the function can determine the visiting weight of the orbits. To search with a high probability for finding candidates for the minimal points, a method for setting the density function has been introduced so that the visiting weight has high values at areas where the objective function has low values. The speed of the orbits can be set by an arbitrary parameter. With these devices and an assumption of ergodicity, it has been shown that the values of the visiting weight and speed of the orbits can be set arbitrarily and independently, and that these values are realized by a long-time limit and a long-time average, respectively.

For an objective function given by the sum of three Gaussian form functions on a plane, numerical simulations were carried out. The finite-time validity of the orbits' properties formulated using the long-time limits has been verified. Furthermore, in simulations for several setting values of the visiting weight and speed of the orbits, the expected effect of differences in those values on the frequency of visits and on searching speed has been confirmed. These results show the possibilities for efficient resolution using the proposed method on actual optimization problems.

ACKNOWLEDGMENTS

The author thanks Dr. Umebu for valuable discussions, and his encouragement and critical reading of the first version of this manuscript. Author also thanks Kazuyoshi Minami, Shigeru Kameda, Shoichi Masuda, and Ichiro Suzuki for their useful discussions.

[1] *Global Minimization of Nonconvex Energy Functions: Molecular Conformation and Protein Folding*, edited by P.M. Pardalos, D. Shalloway, and G. Xue (American Mathematical Society, Providence, 1996).

[2] *Optimization in Computer-Aided Design*, edited by J.S. Gero (North-Holland, Amsterdam, 1985).

[3] D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning* (Addison-Wesley, Reading, MA, 1989).

[4] S. Kirkpatrick, C.D. Gelatt, Jr., and M.P. Vecchi, Science **220**,

671 (1983).
[5] F.H. Branin, Jr., IBM J. Res. Dev. 504 (1972).
[6] A.V. Levy and A. Montalvo, SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput. **6**, 15 (1985).
[7] R.P. Ge, J. Comput. Math. **5**, 1 (1987).
[8] A. Törn and A. Žilinskas, *Global Optimization*, Lecture Notes in Computer Science (Springer-Verlag, Heidelberg, 1989).
[9] *Optimization*, edited by G.L. Nemhauser, A.H.G. Rinnooy Kan, and M.J. Todd, Handbooks in Operations Research and Management Science, Vol. 1 (Elsevier, Amsterdam, 1989), Chaps. 1, 3.
[10] H.A. Eschenauer *et al.*, in *Proceedings of an International Seminar Organized by Deutsche Forschungsanstalt für Luft- und Raumfahrt, Bonn, 1989*, edited by H.W. Bergmann (Springer-Verlag, Heidelberg, 1989).
[11] M.P. Allen and D.J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).
[12] S. Nosé, J. Chem. Phys. **81**, 511 (1984).
[13] W.G. Hoover, Phys. Rev. A **31**, 1695 (1985).
[14] S. Nosé, Prog. Theor. Phys. Suppl. **103**, 1 (1991), and references therein.
[15] K. Inoue, Seminar on Probability **57**, 36 (1989).
[16] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[17] N.E. Collins, R.W. Eglese, and B.L. Golden, Am. J. Math. Manage. Sci. **8**, 209 (1988), and references therein.
[18] S. Geman and D. Geman, IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-6**, 721 (1984).
[19] H. Nishimori and J. Inoue, J. Phys. A **31**, 5661 (1998).
[20] E. Aarts and J. Korst, *Simulated Annealing and Boltzmann Machines* (Wiley, New York, 1989), Chap. 6.
[21] H.A. Posch, W.G. Hoover, and F.J. Vesely, Phys. Rev. A **33**, 4253 (1986).
[22] K. Cho and J.D. Joannopoulos, Phys. Rev. A **45**, 7089 (1992).
[23] F.D. Di Tolla and M. Ronchetti, Phys. Rev. E **48**, 1726 (1993).