A NEW METHOD OF CALCULATING VERY SMALL CROSS SECTIONS*

O. Mazonka, J. Błocki

Institute for Nuclear Studies, Świerk, Poland

and C. Jarzyński

Theoretical Division, Los Alamos National Laboratory, USA

(Received July 4, 1998)

Langevin equations are used to model many processes of physical interest, including low-energy nuclear collisions. We develop a general method for computing probabilities of very rare events (*e.g.* small fusion cross-sections) for processes described by Langevin dynamics. As we demonstrate with numerical examples, our method can converge to the desired answer at a rate which is orders of magnitude faster than that achieved by using direct simulations of the process in question.

PACS numbers: 25.70.-z, 25.70.Jj, 24.60.Ky

Introduction

Langevin methods offer a powerful tool for the numerical study of low-energy nuclear processes, such as fission and heavy-ion fusion. The evolution of nuclei during such events is typically described using a few collective degrees of freedom, evolving under both conservative and non-conservative forces. The latter, arising from the coupling of the collective variables to the intrinsic nucleonic degrees of freedom, can be modeled by a noisy and a dissipative term in a Langevin description of the collective motion. Once such a (stochastic) equation of motion has been written down, it is straightforward to numerically simulate the process in question, using a random number generator to supply the noise. By repeating the simulation — with different sequences of random numbers — one obtains independent "realizations" of the process in question, reflecting the statistical distribution of events occurring during an experiment.

* Presented at the International Conference "Nuclear Physics Close to the Barrier", Warszawa, Poland, June 30–July 4, 1998.

The "direct simulation" method outlined above becomes impractical when studying rare outcomes. For instance, if we are interested in computing the very small cross-section for the fusion of two heavy nuclei, then the vast majority of realizations will end with the nuclei flying apart, and the number of simulations required to obtain a statistically significant number of fusion events may well be prohibitively large.

The basic idea which we shall present is essentially a dynamical variant of *importance sampling*, which amounts to gaining information about one probability distribution (a "target" distribution, T), by choosing randomly from another (a "sampling" distribution, S) defined on the same space, and then *biasing* — assigning weights to — the points sampled. The weights are assigned in such a way that the *biased average* of a quantity, over N points drawn independently from S, and the *unbiased average* of that quantity over N points drawn from T, converge to the same value in the limit of infinitely many samples, $N \to \infty$. If the biased average converges faster with N than the unbiased one, then importance sampling becomes a practical tool for increasing the efficiency of the numerical estimation of the desired average.

In our case, we are interested in Langevin trajectories describing (for instance) the collision of two heavy nuclei, with a very small probability for fusion. Our *target* ensemble, T, is then the statistical distribution of all such trajectories with, say, a given initial center-of-mass energy and impact parameter. The probability of fusion which we wish to compute is defined with respect to this ensemble of trajectories. Our *sampling* ensemble, S, is the distribution of trajectories evolving — from the same initial conditions — under a *modified* Langevin equation, which is far more likely to result in fusion. The scheme which we propose then involves running a number of simulations with the modified equation of motion (thus obtaining fusion events with good statistics), and then biasing each trajectory, so as to compute the desired probability for fusion.

1. Theory

1.1. Importance sampling

Importance sampling is based on a very simple idea, embodied by Eq. (3) below. Suppose we have some space (ζ-space) on which are defined two normalized probability distributions, $p_S(\zeta)$ and $p_T(\zeta)$, corresponding to "sampling" and "target" ensembles, S and T. Supposing furthermore that $p_S(\zeta) > 0$ whenever $p_T(\zeta) > 0$, let us introduce a *biasing function*

$$w(\zeta) = \frac{p_T(\zeta)}{p_S(\zeta)}, \tag{1}$$

defined at all points ζ for which $p_S(\zeta) > 0$. Now let $\langle \mathcal{O} \rangle_S$ and $\langle \mathcal{O} \rangle_T$ denote the averages of some observable $\mathcal{O}(\zeta)$ over the two distributions:

$$\langle \mathcal{O} \rangle_i \equiv \int d\zeta \, p_i(\zeta) \, \mathcal{O}(\zeta) = \lim_{N \to \infty} (1/N) \sum_{n=1}^{N} \mathcal{O}(\zeta_n^T) \quad , \qquad i = S, T \, . \tag{2}$$

From Eqs (1) and (2) above we can equally well express the desired average as:

$$\langle \mathcal{O} \rangle_T = \langle w \mathcal{O} \rangle_S = \lim_{N \to \infty} (1/N) \sum_{n=1}^{N} w(\zeta_n^S) \, \mathcal{O}(\zeta^S) \, , \tag{3}$$

where $\zeta_1^S, \zeta_2^S, \cdots$ is a sequence of points sampled independently from S. Thus, provided we can compute $w(\zeta)$ and $\mathcal{O}(\zeta)$ for any ζ, Eq. (3) gives us a prescription for determining the average of $\mathcal{O}$ over the target ensemble T, *using points drawn from the sampling ensemble* S. This prescription becomes a practical tool if a sampling distribution can be chosen for which the rate of convergence with the number of samples (N) is faster when using Eq. (3), than when sampling directly from T.

1.2. Statistical distributions of Langevin trajectories

The original and modified Langevin equations can be represented by the generic equation

$$\frac{dx}{dt} = v(x) + \xi(t) \, , \tag{4}$$

where $v = v_0$ in one case, and $v = v_0 + \Delta v$ in the other. As before, given some initial conditions $x(0) = x^0$, let $x(t)$ denote the trajectory evolving from those initial conditions, for a particular realization of the noise term. We are interested in the probability density $p[x(t)]$ for obtaining a particular trajectory $x(t)$. The ratio between these two probability densities is given by:

$$w[x(t)] \equiv \frac{p_T[x(t)]}{p_S[x(t)]} = \exp -\Delta A[x(t)] \, , \qquad \Delta A \equiv A_T - A_S \, , \tag{5}$$

where A_T and A_S are Langevin actions[1] for $v = v_0$ and $v = v_0 + \Delta v$, respectively. And

$$\Delta A[x(t)] = \frac{1}{D} \int_0^\tau dt \left(\frac{dx}{dt} - v_0 - \frac{1}{2} \Delta v \right) \Delta v \, . \tag{6}$$

Here, dx/dt, v_0 and Δv are evaluated along the trajectory $x(t)$, D is diffusion coefficient.

1.3. Computing probabilities of rare events

We now have an expression which allows us, in principle, to compute the probability for fusion — defined with respect to the *original* equation of motion — by running independent simulations with the *modified* equation of motion

$$P_{\rm fus} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \Theta[x_n^S(t)] \exp -\Delta A[x_n^S(t)] \,. \tag{7}$$

Here, $x_n^S(t)$ is the trajectory generated during the n'th simulation, using the modified Langevin equation; ΔA is computed for each trajectory; and Θ is equal to one or zero, depending on whether or not fusion occurred.

1.4. Efficiency analysis

Having derived an estimator for $P_{\rm fus}$ based on the idea of importance sampling, we now consider the question of efficiency. In particular, we establish a specific measure of "how much we gain" by using importance sampling, with a given choice of $\Delta v(x)$.

The validity of Eq. (7) does not depend on the form of $\Delta v(x)$. Therefore, for any additional drift term Δv, there will be some threshold value $N^*_{\Delta v}$ such that $P_{\rm fus}^{(N)}$ provides a "good" estimate of $P_{\rm fus}$ for $N \geq N^*_{\Delta v}$. That is, $N^*_{\Delta v}$ is the number of trajectories which we need to simulate (using the modified Langevin equation), in order to determine $P_{\rm fus}$ to some desired accuracy, using the method outlined above. Of course, $N^*_{\Delta v}$ can depend strongly on the form of $\Delta v(x)$. We can thus compare the efficiency of estimating $P_{\rm fus}$, for different drift terms $\Delta v(x)$. In particular — since the special case $\Delta v = 0$ is equivalent to computing $P_{\rm fus}$ using the original Langevin equation — let us define the *efficiency gain*, $E^G_{\Delta v}$, associated with a given $\Delta v(x)$, as follows:

$$E^G_{\Delta v} \equiv \frac{N^*_0}{N^*_{\Delta v}} \,. \tag{8}$$

The numerator is just the number of trajectories needed to accurately estimate $P_{\rm fus}$ by running simulations with the original Langevin equation ($\Delta v = 0$); the denominator is the number needed using modified equation, for a given $\Delta v(x)$. Thus, $E^G_{\Delta v}$ is the factor by which we reduce the computational effort, by making use of importance sampling — again, for a given $\Delta v(x)$.

For a given additional drift term Δv, let us define

$$f[x(t)] \equiv w[x(t)]\Theta[x(t)] = \Theta \exp -\Delta A \,. \tag{9}$$

We get the following result for the efficiency gain of our importance sampling method, for a particular choice of $\Delta v(x)$:

$$E^G_{\Delta v} = \frac{N^*_0}{N^*_{\Delta v}} = \frac{P_{\rm fus}(1-P_{\rm fus})}{\sigma^2_{f,\Delta v}} \cong \frac{P_{\rm fus}}{\sigma^2_{f,\Delta v}} = \frac{\langle f\rangle_S}{\langle f^2\rangle_S - \langle f\rangle^2_S}\,, \tag{10}$$

where we have written $\sigma^2_{f,\Delta v}$ to explicitly specify that this is the variance of f for trajectories simulated with the additional drift term Δv (*i.e.* sampled from S).

We will use this result in Section 3 below, to compute the efficiency gain of the importance sampling method for particular examples.

2. Practical matters

In this section, we discuss a number of practical issues related to the actual implementation of the importance sampling method derived above.

We can rewrite Eq. (6) as:

$$\Delta A = \frac{1}{2D}\int_0^\tau dt\,(2\hat{\xi} + \Delta v)\Delta v\,. \tag{11}$$

This expression for ΔA lends itself to a convenient implementation of the method, as follows. When simulating a given trajectory $x(t)$ evolving under modified equation, we simultaneously integrate the following equation of motion for a new variable $y(t)$, satisfying the initial condition $y(0) = 0$:

$$\frac{dy}{dt} = \frac{\Delta v}{2D}(2\hat{\xi} + \Delta v)\,, \tag{12}$$

for the same realization of the noise term $\hat{\xi}(t)$. (Note that this equation is coupled to the equation of motion for x, since Δv in general depends on x.) Eq. (11) then implies that $\Delta A = y(\tau)$. Thus, at the end of the simulation, we use $x(\tau)$ to determine whether or not fusion has occurred, and if so, then we take $\Delta A = y(\tau)$ when assigning the bias $\mathrm{e}^{-\Delta A}$ to this event.

Often (see for instance Section 3 below), the evolution of our system is such that, once a trajectory $x(t)$ enters the region R which defines fusion, its chance for subsequently escaping that region is negligible: R effectively possesses an absorbing boundary. If this is true for both the original and modified evolution, it becomes convenient to define Δv to be zero everywhere within R. Then, if a trajectory $x(t)$ (evolving under the modified Langevin equation) crosses into R at some time $\tau' < \tau$, we can stop the simulation at that point in time, and take $\Theta = 1$, $\Delta A = y(\tau')$. This saves time, by eliminating the need to continue with the simulation.

We have, to this point, assumed that the stochastic noise $\hat{\xi}(t)$ is independent of x. More generally, we might have a diffusion coefficient which depends on the instantaneous configuration of the system: $D = D(x)$. When implementing the method using the additional variable $y(t)$, the only difference is that in Eq. (12) D is evaluated along $x(t)$ rather than being a constant.

Let us now drop the assumption that the system evolves in one dimension. The variables x, v_0, Δv, and $\hat{\xi}$ now become vectors, and D becomes a symmetric matrix whose elements reflect the correlations between the different components of the stochastic force:

$$\langle \hat{\xi}_i(t) \hat{\xi}_j(t+s) \rangle = D_{ij} \delta(s) \,. \tag{13}$$

Eq. (12) generalizes to the following evolution equation for y:

$$\frac{dy}{dt} = \frac{1}{2} (2\hat{\boldsymbol{\xi}} + \boldsymbol{\Delta v})^T D^{-1} \boldsymbol{\Delta v} \,. \tag{14}$$

Eq. (14) implicitly assumes that D is invertible, *i.e.* $\det(D) \neq 0$. If this is not the case, then — first — we must make sure that the projection of $\boldsymbol{\Delta} v$ onto the subspace spanned by the null eigenvectors of D is zero. Assuming this condition is satisfied, we can view Eq. (14) as pertaining only to the subspace spanned by the non-zero eigenvectors of D.

3. Numerical results

In this section we describe numerical experiments which we have carried out to test our method, using a simplified model of heavy ion collisions [2]. This model was previously studied by Aguiar *et al* in 1990[3], using Langevin simulations. For our example, we considered the collision of two ^{100}Zr nuclei. In this mass-symmetric case — for this simple model — the shape of the system is defined by two equal spheres connected by a cylinder. There are two macroscopic ("collective") variables parametrizing the shape: (1) the *relative distance* ρ between the sphere centers, which is the distance s divided by the sum of radii of the two spheres: $\rho = s/2R$; and (2) the *window opening* α, which is the square of the ratio of the cylinder radius to the radius of the sphere: $\alpha = (r_{\mathrm{cyl}}/R)^2$. The collective coordinates ρ and α are represented by the variables $\nu = \sqrt{\alpha}$ and $\sigma = \rho^2 - 1$. The evolution of the colliding nuclei is then represented by a Langevin trajectory in (σ, ν)-space. Fig. 1 depicts 30 such trajectories, all starting from a configuration of two touching spheres ($\sigma = 0, \nu = 0$), with a center-of-mass energy equal to 0.8 MeV above the interaction barrier. This energy is about 2.5 MeV below the "extra push" energy, so most of the trajectories (28 of them) lead

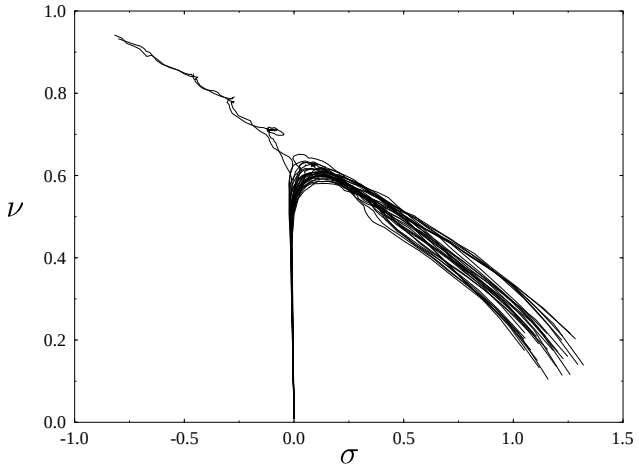

Fig. 1. Trajectories representing evolution of the system in (σ, ν)-parameter space. Two trajectories of 30 lead to fusion when the energy is about 2.5 MeV below the "extra push". See details in text.

to reseparation of the system (fission), and only two trajectories lead to a compound nucleus (fusion).

From Fig. 1 we have the following picture of the physical process occuring, in the context of this simplified model: first the window opening between the two nuclei grows rapidly; then around a saddle point, at $(\sigma, \nu) \sim (0.0, 0.6)$, the combination of deterministic and stochastic forces determines the ultimate fate of the nuclei, either fusion or reseparation; and finally the system evolves toward its destiny, with σ decreasing in the case of fusion, or increasing with reseparation. This suggests that, if we are to add an additional force to increase the likelihood of fusion, then it would be best to localize such a force in the vicinity of the saddle point. We have chosen an additional force along the negative σ direction, whose strength is a Gaussian function of (σ, ν), with a peak at $(0.0, 0.6)$.

In Figs 2 and 3 we show excitation functions — fusion probability plotted against center-of-mass energy above the barrier — as computed by both direct simulation and importance sampling. Each point was obtained using 1000 trajectories, and the result is displayed with error bars, as estimated from the numerical data. The solid line represents an analytical formula which closely approximates the fusion probability over the region shown. We see that, for approximately the same computational effort, our importance sampling method gives significantly better results than direct simulation. For the point corresponding to 0.5 MeV above the barrier, the error bar in Fig. 2 is about 5.5 times bigger than that in Fig. 3. The efficiency gain of the

importance sampling approach is therefore about 30 ($\sim 5.5^2$, see Eq. (10) in this case: we would need to launch about 30×10^3 trajectories evolving under the original Langevin equation to get the same degree of accuracy obtained in Fig. 3 with 10^3 trajectories.

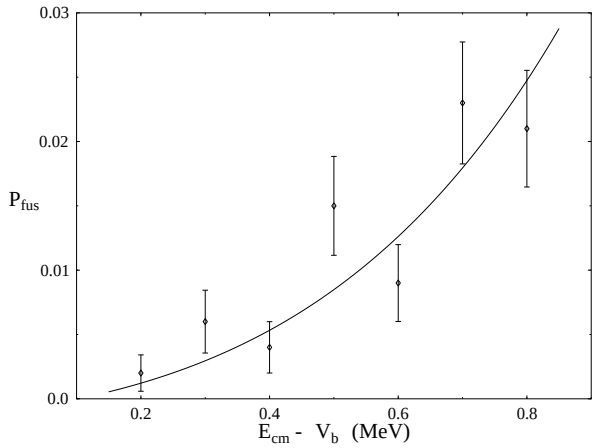

Fig. 2. Excitation function computed using direct simulation, with 1000 trajectories for each point. (The solid line is an analytical estimate extracted from a much larger number of simulations.)

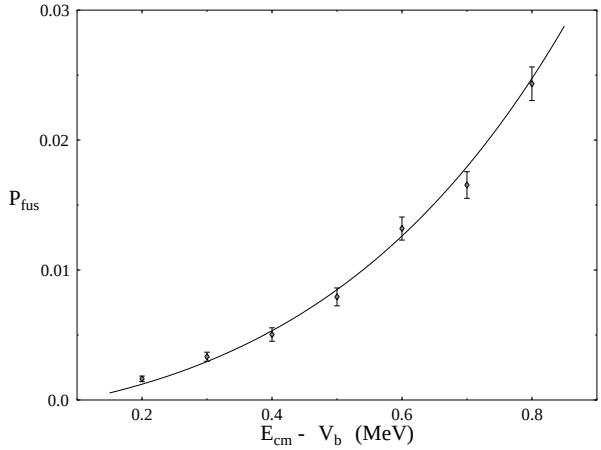

Fig. 3. Same as Fig. 2 but computed using importance sampling instead of direct simulation.

The gain in efficiency becomes more dramatic when we go to very small probabilities. To show this we considered the reaction ^{110}Pd+^{110}Pd, for which the extra push energy is 25.5 MeV. Launching 250000 trajectories with an initial center-of-mass energy of 1 MeV above the barrier, we obtained a probability of fusion $P_{\rm fus} = (6.970 \pm 0.268) \times 10^{-13}$. This was computed using our importance sampling method; about 88% of the trajectories evolving under the modified Langevin equation went to fusion. Using Eq. (10), our result gives an efficiency gain of $E^G_{\Delta v} = 3.5 \times 10^9$! We cannot compare our estimate of $P_{\rm fus}$ directly to an estimate obtained from simulating with the original Langevin equation, since we would need to run $\sim 10^{12}$ trajectories to have a decent chance of observing even a single fusion event. Importance sampling is indispensible in this case: calculating $P_{\rm fus}$ using direct simulations is not practical.

4. Conclusions

In this paper we have developed a method for computing the probabilities of rare events — exemplified by the fusion of two nuclei — for processes described by Langevin dynamics. The method, based on the idea of importance sampling, is straightforward to implement, quite general, and can lead to a very large increase in computational efficiency. For these reasons we believe it represents a very practical tool for using numerical simulations to compute small probabilities. Indeed, with our method, we were easily able to estimate a fusion probability, within a schematic model of nuclear collisions (see the end of Section 3), that would have been essentially impossible to estimate from direct simulations of the process in question. We see every reason to expect similar results when combining the method with more realistic semiclassical models of nuclear dynamics.

This work was performed as a part of the research program supported by the Polish State Committee for Scientific Research KBN Grant No. 2-P03B-143-10.

REFERENCES

[1] 1 F.W. Wiegel, *Introduction to Path-Integral Methods in Physics and Polymer Science*, World Scientific, Philadelphia 1986. Original references are: S. Chandrasekhar, *Rev. Mod. Phys.* **15**, 1 (1943); L. Onsager, S. Machlup, *Phys. Rev.* **91**, 1505 (1953) and *Phys. Rev.* **91**, 1512 (1953).

[2] 2 W. J.Świątecki, *Phys. Scr.* **24**, 113 (1981).

[3] 3 C.E. Aguiar, V.C. Barbosa, R. Donangelo, *Nucl. Phys.* **A514**, 205 (1990).