

# Stochastic Effects; Application in Nuclear Physics

Oleg Mazonka

Świerk, February 11, 2000



# Stochastic Effects; Application in Nuclear Physics

(February 11, 2000)

Stochastic effects in nuclear physics refer to the study of the dynamics of nuclear systems evolving under stochastic equations of motion. In this dissertation we restrict our attention to classical scattering models. We begin with introduction of the model of nuclear dynamics and deterministic equations of evolution. We apply a Langevin approach — an additional property of the model, which reflects the statistical nature of low energy nuclear behaviour. We then concentrate our attention on the problem of calculating tails of distribution functions, which actually is the problem of calculating probabilities of rare outcomes. Two general strategies are proposed. Results and discussions follow. Finally in the appendix we consider stochastic effects in nonequilibrium systems. A few exactly solvable models are presented. For one model we show explicitly that stochastic behaviour in a microscopic description can lead to ordered collective effects on the macroscopic scale. Two others are solved to confirm the predictions of the fluctuation theorem.

## Contents

I. Introduction .....	<b>1</b>
1. Introductory concepts .....	1
Fluctuations in classical description of Heavy Ion dynamics .....	1
Fluctuation-dissipation theorem .....	3
Probabilities of rare events .....	3
2. Outline of chapters .....	3
II. Model of nuclear dynamics .....	<b>5</b>
1. Introduction .....	5
2. Shape parameterization .....	5
3. Kinetic and potential energies; shell corrections .....	7
Kinetic energy .....	7
Potential energy .....	8
Shell effects .....	9
4. Equations of motion .....	10
Generic deterministic equation .....	10
Dissipation .....	11
5. Fluctuations .....	12
Fokker-Planck equation .....	12
Derivation of fluctuating force .....	13
Diffusion term for Fermi particles .....	15
Langevin approach .....	16
Implementation of the Langevin equation .....	18
6. Additional topics .....	19
Exchange of particles .....	19
Dissipative fluctuations .....	20
Parameterization of the potential .....	20
III. Computation of small probabilities .....	<b>24</b>
1. Introduction .....	24
The probability .....	24
2. Polymer method .....	25
Polymer analogy .....	25
Action .....	25
Metropolis sweeps .....	25
Computation of work .....	26
Advantages and disadvantages of the polymer method .....	27
3. Examples .....	28
Analytically solvable model .....	28

Simplified model of HIC .....	30
4. Importance Sampling Method .....	32
Langevin dynamics .....	33
Importance sampling .....	34
Biasing Langevin trajectories .....	35
Langevin action .....	35
Compensation for modifications .....	37
Efficiency gain .....	37
5. Application of ISM: practical matters .....	39
Calculating weights .....	39
Cut-off .....	40
$D$ depending on $x$ .....	40
Multidimensionality .....	40
Inertial motion .....	41
Dissipative motion .....	42
Elimination of degenerate directions .....	42
Extension to the non-Markovian case .....	43
6. Examples .....	44
Schematic model .....	44
Exactly solvable model .....	46
IV. Calculations .....	<b>49</b>
Introduction .....	49
Units and constants .....	49
1. Fusion probabilities .....	49
2. Cross sections .....	51
3. Calculation of small probabilities .....	54
V. Summary and conclusions .....	<b>56</b>
VI. Appendix: Solvable models .....	<b>58</b>
1. Introduction .....	58
2. Feynman's ratchet and pawl .....	58
Introduction .....	58
Zero external load .....	59
Nonzero external load .....	65
The system as heat engine and refrigerator .....	68
Linear Response .....	71
Summary .....	74
Explicit expressions .....	75
3. The piston model .....	76
Introduction .....	76
The model .....	76
The main result .....	77
Fluctuation theorem relation .....	79
Derivation .....	80
4. Dragged harmonic oscillator .....	81
Introduction .....	81

Introduce and solve the model .....	82
Energy conservation and entropy generation .....	85
Transient and steady-state fluctuation theorems .....	85
Detailed fluctuation theorem .....	87
Generalization .....	89
Nonequilibrium work relation for free energy differences .....	89
Explicit expressions .....	91

# I. Introduction

In conventional nonequilibrium physics one studies how a system, which is in thermal equilibrium with a large heat bath, responds to a small change in the external parameters like magnetic field, pressure etc. The situation in heavy-ion collisions(HIC) is quite different and in a way unique. A closed microscopic system without contact to an external bath evolves irreversibly in time from a pure state, at the beginning, to a mixture of states as seen by a highly incomplete measurements. From a generalized point of view one may regard all degrees of freedom other than the collective ones as the “heat bath” which absorbs information and thus produces the entropy. However, one should keep in mind that this “heat bath” depends on the choice of macroscopic variables and on their dynamical evolution. Furthermore, different stages of equilibration may be achieved because the colliding nuclei are pulled apart by electric and centrifugal forces before the macroscopic variables can equilibrate. The great variety of dissipative phenomena seen in heavy-ion collisions together with the challenge of statistical physics far from equilibrium has aroused the interest of a big community of physicists.

The subject of this thesis are stochastic models and stochastic models in nuclear physics — low energy nuclear collision systems which evolve under stochastic equations. Most of the focus will be on methods of calculating small probabilities, although at the end we touch on the more or less related topic of random walk dynamics as a fundamental study of stochastic effects.

This introductory chapter is divided into two sections. The first is a thumbnail sketch of concepts related to stochastic effects in general as well as to dissipative dynamics. The purpose here is to place into context the topics covered in this thesis, and to introduce notation and terminology. The second part of this introduction provides brief outlines of chapters 2,3,4 and 5, the main part of the thesis.

## 1. Introductory concepts

**Fluctuations in classical description of Heavy Ion dynamics.** One of the goals of this thesis is the incorporation of statistical fluctuations into a description of Heavy Ion(HI) dynamics. As we emphasize, and as has been realized independently by others [1], such fluctuations play an important role in processes — such as the creation of super-heavy elements by fusion — with very small cross sections.

At the classical level, a quantitative description of such fluctuations may be achieved with a Fokker-Planck equation. However, this equation does not translate immediately

into a semi-classically valid description of the corresponding quantal problem: quantal (Fermi-Dirac) statistics leads to a suppression of fluctuations, which does not vanish in the semi-classical limit. The problem thus is one of providing a semi-classical description of fluctuations which is consistent with the fermionic nature of nucleons.

It has been realized that a semi-classical Fokker-Planck equation which obeys Fermi-Dirac statistics can be derived, with a simple modification of the relationship between the dissipative and the diffusive terms in the equation. The “trick” involves replacing the classical one-body density of states with the quantal many-body density, which contains the relevant quantal statistics.

From our quantal Fokker-Planck equation, a Langevin description of HI dynamics (with fluctuations included), suitable for implementation in numerical simulations, is easily derived.

We have tested this approach by starting with a schematic model of HI collisions [2], then adding the stochastic forces, and finally simulating the resulting dynamics numerically. From these simulations we were able to obtain an excitation function for fusion, in symmetric collisions. Our next step was to apply the same procedure to a more realistic model [3], and to include asymmetric collisions. This allowed us to produce excitation functions and distributions of experimentally observed quantities, which could then be compared directly with experimental data.

The derivation of a quantally valid, semi-classical Fokker-Planck (or Langevin) description of nuclear dissipation, is not the end of the story. As mentioned, statistical fluctuations are expected to play an important role precisely in processes for which the interesting outcome – e.g. fusion of superheavy elements – occurs with very low probability. In such situations, direct numerical simulation of the process does not provide a realistic method of attacking the problem: if the probability of fusion is of the order, say,  $10^{-10}$ , then prohibitively many simulations would be needed to accurately assess the cross-section.

We have developed new methods of getting around this problem, and have obtained very encouraging results. Our approach makes use of a one-to-one correspondence between the statistical distributions of Langevin trajectories, and the thermal distribution of so-called “directed polymers”. This analogy was pointed out more than fifty years ago by Chandrasekhar [4], and provides a powerful formalism for analyzing relative probabilities of Langevin trajectories.

One of the methods involves adding an extra, non-physical force acting on the nuclear collective degrees of freedom, in our numerical Langevin simulations. This force is chosen so as to increase the probability for fusion by orders of magnitude, enabling us to obtain good statistics from a reasonable number of simulations; then we analytically correct for the inclusion of this unphysical force, thus finally obtaining the desired cross-section. We have tested this method in different models and have obtained fusion probabilities in situations where they are extremely small. We have found that we could get good statistics with far fewer simulations than is needed to accurately extract the fusion probability by direct Langevin simulation.

In a related method, we constrain the Langevin trajectory to end in a desired region of a configuration space (e.g. describing fusion), then we apply methods developed in the context of physical chemistry [5] to sample the distribution of these constrained trajec-



ries. From this distribution — making use of the analogy between Langevin trajectories and directed polymers — we are finally able to extract the desired information (fusion probability).

Such methods have not been applied before to study the collective nuclear dynamics, and become an indispensable tool in the development of practical numerical codes for studying low-energy nuclear processes with very small cross-sections.

**Fluctuation-dissipation theorem .** The Fluctuation-dissipation theorem (FDT) is one of the most fundamental results of nonequilibrium statistical mechanics. The very first example of it, the well-known Einstein relation, was obtained long before FDT was formulated. The significance of FDT is the fact that it relates two different macroscopic phenomena, fluctuation and dissipation, to the same microscopic origin. When going to a macroscopic description, microscopic fluctuations turn into macroscopic fluctuations on the one hand, and friction in macroscopic coordinates on the other. Friction redistributes the energy of the collective motion to the microscopic degrees of freedom. The connection between dissipation and macroscopic fluctuations is then derived from a microscopic description, under the condition that the physical system is continually evolving toward a state of equilibrium. This condition is essentially a statement of the Second Law. In other words, having defined the dissipation mechanism we strictly determine the character of fluctuations by requiring consistency with the Second Law. To simulate the stochastic noise we use *white Wiener noise*. The choice is not arbitrary because the friction in linear form leads to white noise fluctuation, i.e. having no memory. Such a choice of relation between dissipation and fluctuations is a classical example of relation for Brownian particle, however the concept of Brownian motion is far beyond the concept of Brownian particle.

**Probabilities of rare events.** The evolution of stochastic systems can be presented as bunch (family) of trajectories. With rare events we understand a number of final states of the system corresponding to a definite physical outcome, which happens rarely either among the ensemble of systems or with a big number of repetitions of evolution of the system. For example, motion of the Brownian particle confined in two-well potential can be considered like a stochastic trajectory for finite period of time. The probability to pass over the barrier from one well to the other can be small enough to be estimated directly.

## 2. Outline of chapters

In chapter II we first describe the model ([3,6]), which we will use in numerical calculations for comparing with realistic experimental data. We start from the parameterization of nuclear shapes and the definitions of kinetic and potential energies within this parameterization. The description is constrained by 3 geometrical collective coordinates, and one corresponding to charge asymmetry. In calculating the potential energy we take into account the shell effects in a phenomenological form proposed by Myers and Swiatecki [61], [8]. Next we introduce equations of motion, which are in fact the Lagrange-Rayleigh equations. The motion in two collective coordinates is assumed to be over damped due to the fact that the inertia terms in equations are much smaller than the dissipative ones. The dissipation mechanism is taken as one-body dissipation described by two basic formulae: the window and wall formula.

Then we present the details of incorporating stochastic effects into the model. Namely, what kind of fluctuations have to be included in the deterministic equations of motion. FDT says that the fluctuation force has to be taken in a form of white Wiener noise, since the Rayleigh's dissipation function has a quadratic form. The only quantity to be determined is the covariance matrix (diffusion tensor). We find the relation between this quantity and the dissipation tensor using the concept of equilibrium state.

The next chapter is devoted to the problem of calculation small probabilities. First we describe the polymer method. This method allows us to deal with a stochastic trajectory like a polymer on potential surface, extracting useful information from such an analogy. Two examples follow: an exactly solvable problem and a schematic model of HIC. Another method based on importance sampling of Langevin trajectories (ISM), can be applied as well for calculation small probabilities. We describe the theory of this method in detail, including efficiency analysis. In the next section we discuss the practical matters of application of the method. Again the same two examples are presented to show how ISM works.

Chapter IV consists mainly of results obtained from the study of the model presented in chapter II using the ISM and compared with real experimental data. Discussion and conclusions are made in there.

In the appendix we present three highly idealized, but exactly solvable, models of thermodynamical systems far from equilibrium. The first one is a microscopic heat engine (also shown to be working as a refrigerator). Actually it is the discrete analog of Feynman's ratchet and pawl in which the macroscopic directed motion is explicitly driven by a temperature difference between two reservoirs. Two other models were considered to confirm the predictions of the *Fluctuation Theorem*, — a number of structurally related theoretical results derived in the field of nonequilibrium statistical mechanics in recent years [89–91, 82, 92–95].

## II. Model of nuclear dynamics

### 1. Introduction

The experimental results on deep-inelastic collisions have stimulated tremendous theoretical activity. A variety of methods and concepts, appropriate to classical and statistical mechanics, fluid dynamics, and thermodynamics, have been applied in addition to the traditional methods of quantal nuclear theory. A common feature of these models is that at some stage the detailed, quantum-mechanical description of the many-body system is abandoned in favor of performing statistical averages over certain microscopic degrees of freedom, while treating certain macroscopic degrees of freedom classically. [12–31]

The various classical scattering models are characterized by three basic ingredients, namely:

- a) the collective degrees of freedom ( $q_i$ ) which are treated explicitly, and their associated inertial parameters (momenta);
- b) the potential energy  $V(q_i)$ , including the nucleus-nucleus potential;
- c) the assumed dissipative forces (friction) which remove the energy from the collective degrees of freedom.

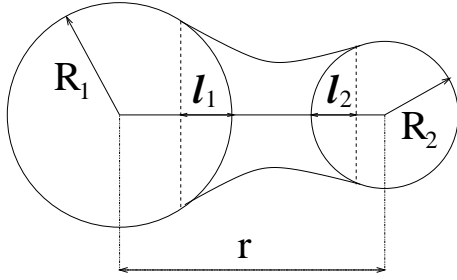
Once these points are settled, it is straightforward in principle to compute the time evolution of the collective variables by solving the generalized Lagrange equations:

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = - \frac{\partial \mathcal{F}}{\partial \dot{q}_i}. \quad (1)$$

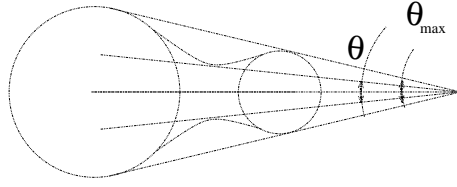
Here  $\mathcal{L}(q_i, \dot{q}_i) = T(q_i, \dot{q}_i) - V(q_i)$  is the Lagrangian of the system, and

$$\mathcal{F}(q_i, \dot{q}_i) = -\frac{1}{2} \frac{d}{dt} \{T(q_i, \dot{q}_i) + V(q_i)\}$$

is the Rayleigh's dissipation function. In practice solving the coupled differential Eqs.1 is usually a complicated numerical problem, and therefore one retains only the minimum number of degrees of freedom  $q_i$  which is considered absolutely essential. The question as to which degrees of freedom should be included in a description of nuclear fission and nucleus-nucleus collision has been discussed some time ago [9,10]. Attention was focused on three coordinates, namely a separation coordinate, a mass asymmetry coordinate, and a neck coordinate.



**Fig 2.1** Parameterization of the nuclear shape (precession shape).  $r$  is the distance between centers of spheres,  $R_1$  and  $R_2$  are their radii, and  $l_1$  and  $l_2$  are the widths of their segments being inside the neck.



**Fig 2.2** Window opening, —  $\alpha = \frac{\sin \theta}{\sin \theta_{\max}}$ .

## 2. Shape parameterization

Here we briefly outline the shape parameterization of two colliding nuclei, which already has been introduced in [7] and used in [3] and [6].

The axially symmetric shapes of fixed volume consists of two generally unequal spheres modified by a smoothly fitted portion of a third quadratic surface of revolution. A set of dimensionless degrees of freedom specifying the configuration are the following (Fig 2.1 and 2.2):

- |    |                    |                                       |     |
|----|--------------------|---------------------------------------|-----|
| 1) | distance variable  | $\rho = r/(R_1 + R_2)$ ,              |     |
| 2) | neck variable      | $\lambda = (l_1 + l_2)/(R_1 + R_2)$ , | (2) |
| 3) | asymmetry variable | $\Delta = (R_1 - R_2)/(R_1 + R_2)$ .  |     |

In the above,  $r$  is the distance between the centers of the spheres, whose radii are  $R_1$  and  $R_2$ . The quantities  $l_1$ ,  $l_2$  are the distances from the inner tips of the two spheres to the respective junction points with the middle quadratic surface of revolution. The natural boundaries of the configurational space  $(\rho, \lambda, \Delta)$  turn out to be given by [7]:

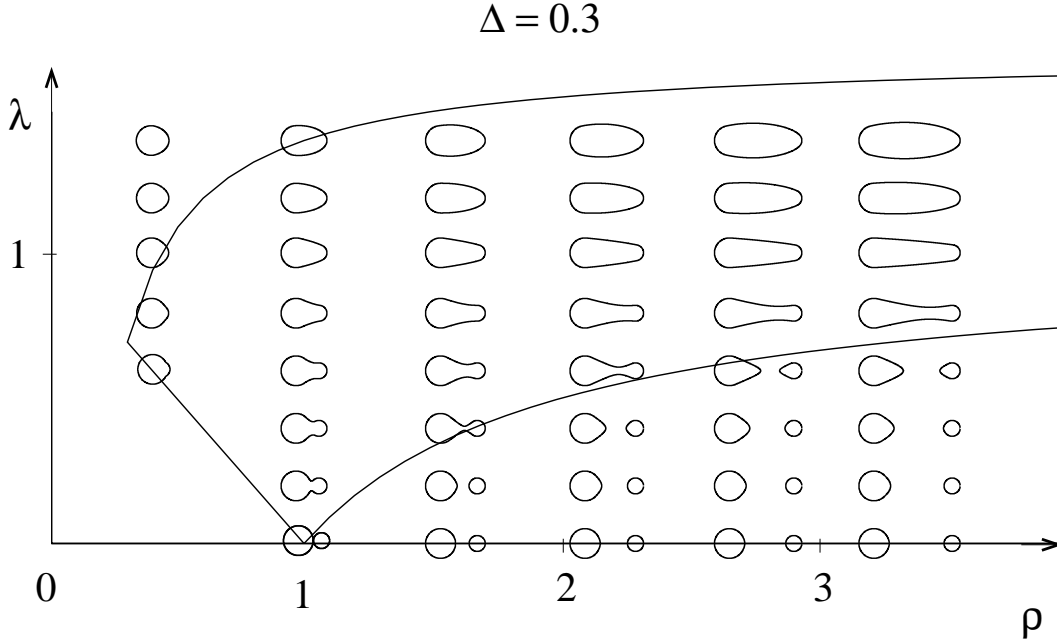
$$\rho \geq |\Delta|, \quad -1 \leq \Delta \leq 1, \quad 2 - (1 + \rho^{-1})|\Delta| \geq \lambda \geq \max\{0, 1 - \rho\}. \quad (3)$$

As may be readily verified, the upper boundary for  $\lambda$  corresponds to egg-like shapes for which the middle quadratic has just covered up completely the smaller sphere. At the lower boundary we have separated spheres for  $\lambda = 0$  and portions of intersecting spheres (without any middle quadratic surface) for  $\lambda = 1 - \rho$ . At scission the center quadratic degenerates into a cone, which implies  $\lambda_{sc} = 1 - \rho_{sc}^{-1}$ .

In addition to the macroscopic variables  $(\rho, \lambda, \Delta)$  there are three rotational degrees of freedom  $(\theta_1, \theta_2, \theta_{rel})$  and three angular velocities  $(\omega_1, \omega_2, \omega_{rel})$  connected with the rotation of sphere 1 and sphere 2 and the rotation of the shape as a whole.

As a parameter, which controls the transition from separate nuclei to compound nucleus, we introduce the window opening parameter  $\alpha$  (Fig 2.2):

$$\alpha = \frac{\sin \theta}{\sin \theta_{\max}}. \quad (4)$$



**Fig 2.3** Shapes in  $(\rho, \lambda)$ -space are shown for the fixed parameter  $\Delta = 0.3$ . The straight line  $\rho + \lambda = 1$  corresponds to two intersecting spheres. There are no shapes under that line. Another line  $\lambda = 1 - \rho^{-1}$  is the scission line. Configurations below this line are two separated fragments. The third line is upper boundary  $\lambda = 2 - |\Delta|(1 + \rho^{-1})$ , which corresponds to the situation when smaller of the spheres is enclosed by the neck. Configurations above this line consist of the part of one sphere and the part of ellipsoid, which is the neck.

### 3. Kinetic and potential energies; shell corrections

**Kinetic energy.** The collective kinetic energy of gas or liquid is given by

$$T = \frac{1}{2} \int \rho_d(\mathbf{r}) [u(\mathbf{r})]^2 d^3\mathbf{r} = \frac{1}{2} \rho_d \int_{\text{shape}} [u(\mathbf{r})]^2 d^3\mathbf{r}, \quad (5)$$

where  $\rho_d(\mathbf{r})$  is the matter density and  $u(\mathbf{r})$  the collective velocity at position  $\mathbf{r}$ . In our case the matter density is uniformly distributed within the shape and zero outside. The velocity field  $u(\mathbf{r})$  has to be chosen in accordance with the collective degrees of freedom.

The time-dependent changes of the shape result in a collective flow which rearrange the matter to maintain a uniform density distribution. The boundary condition is that the velocity component normal to the surface is equal to the normal component of the surface velocity. The assumption of incompressible irrotational flow together with the boundary condition define uniquely the whole velocity field and thereby the kinetic energy for the shape degrees of freedom. However, the solution of the resulting Laplace equation is numerically too costly for practical applications. So it is convenient to use the Werner-Wheeler approximation [59]. It has been found that in the dynamical trajectory calculations the motion in the  $\Delta$  direction was invariably overdamped to such an extent that the component of the inertia tensor  $M_{ij}$  associated with  $\Delta$  could be safely neglected. The kinetic energy is then reduced to the form

$$T = \frac{1}{2} M_{\rho\rho} \dot{\rho}^2 + M_{\rho\lambda} \dot{\rho} \dot{\lambda} + \frac{1}{2} M_{\lambda\lambda} \dot{\lambda}^2. \quad (6)$$

The matrix element  $M_{\rho\rho}$  of the mass tensor for two separated spheres is exactly the reduced mass.

In the rotational degrees of freedom the kinetic energy is equal to:

$$T_r = \frac{1}{2}(I_{\text{rel}}\omega_{\text{rel}}^2 + I_1\omega_1^2 + I_2\omega_2^2), \quad (7)$$

where  $I_1$  and  $I_2$  are the rigid moments of inertia of sphere 1 and 2, respectively and  $I_{\text{rel}} = I_{\text{tot}} - I_1 - I_2$  where  $I_{\text{tot}}$  is the rigid moment of inertia of the whole system. Two spheres can rotate independently from the relative rotation, but due to the tangential friction, after some time all the frequencies  $\omega_{\text{rel}}$ ,  $\omega_1$  and  $\omega_2$  are more or less equal and the body rotates as a rigid one.

The mass tensor is calculated from Eq.5 and equal to (see, for example, [65])

$$M_{\varrho\varsigma} = \rho_d \pi \int \left( \frac{1}{2} B_{\varrho} B_{\varsigma} + P^2 A_{\varrho} A_{\varsigma} \right) dz \quad \varrho, \varsigma = \rho, \lambda, \Delta, \quad (8)$$

where  $P = \bar{\rho}$  is the equation of the nuclear surface in cylindrical coordinates  $(\bar{\rho}, z, \phi)$ , and

$$A_{\varrho} = -\frac{1}{P^2} \int^z \frac{\partial P^2}{\partial \varrho} dz' - \frac{\pi}{V_0} \int \frac{\partial P^2}{\partial \varrho} z'' dz'', \quad (9)$$

$$B_{\varrho} = -\frac{1}{2} \frac{1}{P^2} \frac{\partial P^2}{\partial z} \int^z \frac{\partial P^2}{\partial \varrho} dz' + \frac{\partial P^2}{\partial \varrho}. \quad (10)$$

$V_0$  is the total volume of the system

$$V_0 = \pi \int P^2 dz \quad (11)$$

Moments of inertia are given by

$$I_{\text{tot}} = I_{\text{rel}} + I_1 + I_2, \quad (12)$$

$$I_{\text{tot}} = \rho_d \pi \int \left( \frac{1}{2} P^4 + P^2 z^2 \right) dz - \frac{\rho_d \pi^2}{V_0} \left( \int P^2 z dz \right)^2, \quad (13)$$

$$I_{1,2} = \rho_d \frac{4\pi}{15} R_{1,2}^5. \quad (14)$$

**Potential energy.** The potential energy of the shape under question is calculated as a sum of the nuclear part and the Coulomb part. The nuclear part using a double folding procedure developed by Krappe, Nix and Sierk [60] can be written as:

$$V_n = -\frac{C_s}{8\pi^2 r_0^2 a^3} \iint \left( \frac{1}{a} - \frac{2}{\sigma} \right) \exp(-\sigma/a) d^3\mathbf{r} d^3\mathbf{r}', \quad (15)$$

where  $\sigma = |\mathbf{r} - \mathbf{r}'|$ ,  $C_s = a_s(1 - k_s I^2)$  and  $I = (N - Z)/A$ . The parameters  $r_0$ ,  $a$ ,  $a_s$  and  $k_s$  are taken from the fit done in [60]. For axially symmetric shapes formula 15 reduces to the three dimensional integral of the following type:

$$V_n = \frac{C_s}{4\pi r_0^2} \iiint \left( 2 - \left[ \left( \frac{\sigma}{a} \right)^2 + 2 \frac{\sigma}{a} + 2 \right] \exp(-\sigma/a) \right) \frac{P_2(z, z') P_2(z', z)}{\sigma^4} dz \, dz' \, d\phi, \quad (16)$$

where

$$P_2(z, z') = P(z) \left( P(z') - P(z') \cos \phi - \frac{dP}{dz}(z - z') \right),$$

$$\sigma^2 = P(z)^2 + P(z')^2 - 2P(z)P(z') \cos \phi + (z - z')^2.$$

A similar procedure has to be done in calculating the Coulomb part of the potential which for an axially symmetric shape can be written as:

$$V_c = \frac{\pi}{3} \rho_z \rho_{z'} \iiint \frac{P_2(z, z') P_2(z', z)}{\sigma} dz \, dz' \, d\phi, \quad (17)$$

where  $\rho_z$  and  $\rho_{z'}$  are the charge densities.

**Shell effects.** The effect of shell structure on the nuclear deformation energy has been understood in broad outline since the sixties and, in more quantitative detail, with the advent of the Strutinsky shell correction method. This macroscopic-microscopic method has provided useful insights into the role of shell effects. It has also served as the basis for relatively accurate extrapolations into new domains of the periodic table and nuclear deformations, complementing self-consistent calculations of the Hartree-Fock type.

The importance of shell effects in the fusion of nuclei was demonstrated experimentally as an enhanced fusion probability when using Pb or Pb-like targets from the transuranic region. Shell effects are included in calculations in a phenomenological form proposed by Myers and Swiatecki [61], [8]. In that approach the shell correction  $S_0(N, Z)$  to the potential energy is written in the form

$$S_0(N, Z) = (5.8 \text{ MeV}) \left( \frac{F_N + F_Z}{(A/2)^{2/3}} - 0.325 A^{1/3} \right), \quad (18)$$

where

$$F_N = q_N(N - N_{i-1}) - \frac{3}{5}(N^{5/3} - N_{i-1}^{5/3})$$

with

$$q_N = \frac{3}{5} \frac{N_i^{5/3} - N_{i-1}^{5/3}}{N_i - N_{i-1}}.$$

Here  $N_{i-1}$  and  $N_i$  correspond to the numbers of closed shell neutrons and  $N$  is the actual number of neutrons between  $N_{i-1}$  and  $N_i$  for a system of mass  $A$ . The same formulas for  $F_Z$  and  $q_Z$  describe the shell effect for protons. For neutron and proton numbers  $N$  and  $Z$  corresponding to closed shells  $N_{i-1}$  and  $Z_{i-1}$  the shell effect is strongest and is equal to

$$S_0(N = N_{i-1}, Z = Z_{i-1}) = -5.8 \cdot 0.325 A^{1/3}. \quad (19)$$

For example, in the case of  $^{208}\text{Pb}$  this gives about -11.2 MeV.

The above shell correction  $S_0(N, Z)$  refers to spherical shapes. For deformed shapes the shell correction is attenuated according to [61], [8]:

$$S(N, Z) = S_0(N, Z) \left( 1 - 2 \frac{\text{dist}^2}{a^2} \right) \exp \left( - \frac{\text{dist}^2}{a^2} \right), \quad (20)$$

where

$$\text{dist}^2 = \int \frac{d\Omega}{4\pi} (r(\theta, \phi) - R_0)^2$$

and  $r(\theta, \phi)$  is the radius vector describing the given shape and  $R_0$  is the radius of the equivalent sphere.

One remaining problem is how to interpolate in a smooth way between the sum of the shell corrections  $S_1$  and  $S_2$  of the colliding nuclei in the entrance channel and the shell correction  $S_c$  of the compound shape. We adopted an interpolation in terms of the degree of communication between the two nuclei as specified by the “window opening” parameter  $\alpha$  of [7], defined by Eq.4

$$S = (1 - \alpha)(S_1 + S_2) + \alpha S_c. \quad (21)$$

For separated shapes (below the scission line)  $\alpha$  is equal to zero and we have only  $S_1 + S_2$ . When the neck loses its concavity at  $\alpha = 1$  the shell correction is equal to the compound nucleus shell correction  $S_c$ , and this is used for convex shapes with  $\alpha > 1$ .

Here we do not consider that shell corrections attenuate with growing temperature as was done in [62].

#### 4. Equations of motion

In this section we introduce the dynamical deterministic equations of motion.

**Generic deterministic equation.** The system is described by the classical Rayleigh-Lagrange equations 1 with the following ingredients:

$$\mathcal{L} = \frac{1}{2} \dot{q} M \dot{q} - V(q) \quad (22a)$$

$$\mathcal{F} = \frac{1}{2} \dot{q} \Gamma \dot{q}, \quad (22b)$$

where  $M$  is the mass tensor and  $\Gamma$  the dissipation tensor. If one defines

$$x \equiv \begin{pmatrix} q \\ \dot{q} \end{pmatrix}, \quad v \equiv \begin{pmatrix} \dot{q} \\ -M^{-1}(\dot{M} + \Gamma) - \partial_q V(q) \end{pmatrix}, \quad (23)$$

where  $\partial_q$  is the gradient operator, then Eq.1 can be rewritten as

$$\boxed{\dot{x} = v.} \quad (24)$$

We call Eq.24 a generic deterministic equation to distinguish it from stochastic equation of motion where the stochastic noise term is added.



In our case the quantities  $q$  form the set of the macroscopic variables  $q = (\rho, \lambda, \Delta, \Delta_Z, \theta, \theta_1, \theta_2)$ . The first three are the shape parameters, the last three are rotational degrees of freedom and  $\Delta_Z$  is an additional degree of freedom, charge asymmetry, equal to  $\frac{Z_1^{1/3} - Z_2^{1/3}}{Z_1^{1/3} + Z_2^{1/3}}$ . We consider the motion in  $\Delta$  and  $\Delta_Z$  directions to be overdamped. So altogether there are five second order and two first order differential equations.

**Dissipation.** The tensor  $\Gamma$  in Eq.22b is a dissipation tensor, which can be found after defining the mechanism of the energy dissipation. We restrict ourselves to one-body dissipation [11], arising from collisions of independent particles with the moving boundary of the nucleus. In the one-body dissipation model [11], the energy flow from collective to intrinsic motion is attributed to the interaction of individual nucleons with the mean field produced by all nucleons in the system. In a simplified picture this interaction may be viewed as mediated by collisions between nucleons and a moving container wall, or by the passage of nucleons from one fragment to the other through a window.

There are two limiting cases in which two different simple formulas for the rate of the dissipated energy can be derived. The first one is so called the mononuclear regime when the system of colliding ions can be considered as a monosystem with a thick neck. In that case the gas of nucleons can be considered as a relaxed Fermi gas and the rate of the energy dissipation is given by the following wall formula [11]:

$$\dot{E}_{\text{wall}} = \rho_m \bar{v} \oint dS (\dot{n} - D)^2, \quad (25)$$

where  $\rho_m$  is the mass density of nucleons,  $\bar{v}$  is their average speed (equal to three quarters of the Fermi velocity in the Fermi gas model),  $dS$  is an element of nuclear surface,  $\dot{n}$  is the normal velocity of walls and  $D$  is the overall drift velocity of the gas of nucleons ensuring the invariance of Eq.25 against translations and rotations.

In the second limiting case, the dinuclear regime, when two ions are either separated or connected by a thin neck, Eq.25 can not be applied as we are dealing with two Fermi gases separated by the collective velocity. In that case the so-called “wall plus window” formula can be applied and it reads as follows:

$$\dot{E}_{\text{w+w}} = \rho_m \bar{v} \int_1 dS (\dot{n} - D_1)^2 + \rho_m \bar{v} \int_2 dS (\dot{n} - D_2)^2 + \frac{1}{4} \rho_m \bar{v} S_w (u_t^2 + 2u_r^2) + \frac{16}{9} \frac{\rho_m \bar{v}}{S_w} \dot{V}_1^2. \quad (26)$$

The first two terms correspond to the wall formula (Eq.25) but are calculated for each fragment separately with drift velocities  $D_1$  and  $D_2$  for each of the gases. The third term is associated with the dissipation due to the exchange of particles through the window of the area  $S_w$ . The components of the relative velocity of two fragments  $u_t$  and  $u_r$  correspond to the velocities parallel and perpendicular to the window. The last term in Eq.26 corresponds to the dissipative resistance against the asymmetry changes [63], [64] with  $\dot{V}_1$  being the rate of the change of the volume of fragment 1. Eqs.25 and 26 express the rate of the dissipated energy in the two limiting cases of the mononuclear and dinuclear regimes. In general when the situation is in between these two limiting cases a smooth transition between formulas 25 and 26 is used [3]:

$$\dot{E} = f \dot{E}_{\text{wall}} + (1 - f) \dot{E}_{\text{w}+\text{w}} \quad (27)$$

with a form factor  $f$  going to 1 for sphere or spheroid like shapes and to 0 at scission.

Eq.27 is the additional dynamical equation which can be added to the generic deterministic equation (24) describing the evolution of the collective coordinates.

There is another equation, for the time derivative of the difference in the excitation energies of both parts of the nuclear system:

$$\frac{d}{dt}(E_1^* - E_2^*) = \dot{E}_1 - \dot{E}_2 + 2T_0 \dot{S}_{21}, \quad (28)$$

where first two terms are defined in previous subsection and the last one is due to the temperature feedback with  $T_0 = (T_1 + T_2)/2$  being the average temperature of the first and second fragment.  $\dot{S}_{21}$  is the entropy flux taken from Feldmeier [65].

## 5. Fluctuations

In this section we present how to introduce fluctuations into deterministic equations of motion as well as how to connect their magnitudes to dissipation in specific situations.

**Fokker-Planck equation.** Let  $Q$  and  $q$  denote collective and intrinsic degrees of freedom, and  $P$  and  $p$  their associated momenta correspondingly. Assume an inertia  $M$  for the collective coordinate, and  $m$  for the internal <sup>1</sup>. Hamiltonian dynamics for an ensemble of trajectories  $F(Q, P, q, p, t)$  can be described by the Liouville equation:

$$\frac{\partial F}{\partial t} = \hat{L}F, \quad (29)$$

$$\text{where } \hat{L} = -\frac{P}{M} \frac{\partial}{\partial Q} + \frac{\partial H}{\partial Q} \frac{\partial}{\partial P} - \frac{p}{m} \frac{\partial}{\partial q} + \frac{\partial H}{\partial q} \frac{\partial}{\partial p},$$

and  $H$  is the Hamiltonian of the system.

If we want to describe the system by only collective variables, we follow the distribution  $f(Q, P, t) = \int dq dp F(Q, P, q, p, t)$ . The equation for  $f(Q, P, t)$  can easily be found:

$$\frac{\partial f}{\partial t} = -\frac{P}{M} \frac{\partial f}{\partial Q} + \frac{\partial}{\partial P} \left( \left\langle \frac{\partial H}{\partial Q} \right\rangle f \right). \quad (30)$$

We used here that  $\frac{\partial^2}{\partial p \partial q} H = 0$ ,  $F \rightarrow 0$  when  $Q$  or  $P \rightarrow \pm\infty$  and definition

$$\langle \cdot \rangle \equiv \frac{\int dq dp (\cdot) F(Q, P, q, p, t)}{\int dq dp F(Q, P, q, p, t)}$$

Now we would like Eq.30 not to depend on  $F$ . Let us consider  $Q$  to be *slow* parameters and assume that the evolution of the fast variables  $q$  under  $H$  is chaotic and ergodic when

---

<sup>1</sup>We will deal with this variables as if they were one-value quantities. Generally speaking, they are not, but the generalization can be achieved easily.

$Q$  is held fixed. We also assume that the time required for the Hamiltonian  $H$  to change significantly is much bigger than the Lyapunov time associated with the fast chaos. So the last term in Eq.30 contains information about the average drift and diffusion in the  $P$  direction. The average drift could be in the form of a conservative force,  $\frac{\partial V(Q)}{\partial Q}$ , plus a term caused by coupling to the internal degrees of freedom,  $G$ , which (we will show later) is a dissipative term. Then Eq.30 takes the Fokker-Planck form:

$$\frac{\partial f}{\partial t} = \hat{l}f + \frac{\partial}{\partial P}(Gf) + \frac{1}{2} \frac{\partial^2}{\partial P^2}(Df), \quad (31)$$

$$\text{where } \hat{l} = -\frac{P}{M} \frac{\partial}{\partial Q} + \frac{\partial V}{\partial Q} \frac{\partial}{\partial P}$$

is the Liouville operator.

Now let us take an initial distribution which is uniform on a single energy shell in the full phase space (slow + fast degrees). We will call this distribution  $F_U(Q, P, q, p)$ , where the subscript  $U$  specifies the total energy:

$$F_U(Q, P, q, p) = \delta[U - H(Q, P, q, p)], \quad (32)$$

Under time evolution, the distribution  $F_U$  will not change (Liouville's theorem):  $\hat{L}F_U = 0$ . When we project out the fast degrees of freedom, we get:

$$f_U(Q, P) = \int dq dp \delta[U - H(Q, P, q, p)]. \quad (33)$$

But if we view the slow degrees of freedom as "parameters" which determine the Hamiltonian for the fast degrees, then the right side of Eq.33 just represents the density of states of the fast degrees, when the slow ones are "frozen" at  $(Q, P)$ . So:  $f_U(Q, P) = \Sigma$ , where  $\Sigma$  is the density of energy states in Eq.31 above. Now we demand that the Fokker-Planck equation (31) has the property that  $f_U$  is invariant with time, because  $F_U$  is invariant with time [46]. Due to the fact that  $\hat{l}f_U = 0$  we find:

$$\frac{\partial}{\partial P}(G\Sigma) + \frac{1}{2} \frac{\partial^2}{\partial P^2}(D\Sigma) = 0,$$

and

$$\boxed{G = -\frac{1}{2\Sigma} \frac{\partial}{\partial P}(D\Sigma)}. \quad (34)$$

Hence one can rewrite the Eq.31:

$$\frac{\partial f}{\partial t} = \hat{l}f + \frac{1}{2} \frac{\partial}{\partial P} \left[ \Sigma D \frac{\partial}{\partial P} \left( \frac{f}{\Sigma} \right) \right]. \quad (35)$$

This Fokker-Planck equation is analogous to the energy diffusion equation derived by C.Jarzynski [48–51].

**Derivation of fluctuating force.** In the absence of coupling of the collective motion to the internal degrees of freedom of the nucleus, the evolution in  $(Q, P)$ -space is governed by the following deterministic, conservative equations of motion:

$$\frac{dQ}{dt} = \frac{P}{M}, \quad \frac{dP}{dt} = -\frac{dV}{dQ}. \quad (36)$$

An ensemble of trajectories evolving under these equations may then be described by a phase space density  $f(Q, P, t)$  which evolves according to the Liouville equation:

$$\frac{\partial f}{\partial t} = -\frac{P}{M} \frac{\partial f}{\partial Q} + \frac{dV}{dQ} \frac{\partial f}{\partial P}. \quad (37)$$

When we allow for coupling to the internal nucleonic degrees of freedom, then the latter act as a kind of a heat bath, exerting both a dissipative and a fluctuating force on the collective degree of freedom. The effects of these forces may be described by adding a term to the evolution equation for  $f$ , which now becomes a Fokker-Planck equation:

$$\frac{\partial f}{\partial t} = -\frac{P}{M} \frac{\partial f}{\partial Q} + \frac{dV}{dQ} \frac{\partial f}{\partial P} + \frac{1}{2} \frac{\partial}{\partial P} \left[ \Sigma D \frac{\partial}{\partial P} \left( \frac{f}{\Sigma} \right) \right]. \quad (38)$$

Here,  $\Sigma$  denotes the density of states for the nucleonic degrees of freedom, and  $D$  is a momentum diffusion coefficient. The form of the last term in Eq.38 is determined by detailed balance. This last term may be easily rewritten as follows:

$$\text{last term} = -\frac{\partial}{\partial P} \left[ \frac{1}{2\Sigma} \frac{\partial}{\partial P} (\Sigma D) f \right] + \frac{1}{2} \frac{\partial^2}{\partial P^2} (Df) = -\frac{\partial}{\partial P} (F_{fric} f) + \frac{1}{2} \frac{\partial^2}{\partial P^2} (Df), \quad (39)$$

where

$$F_{fric} \equiv -G = \frac{1}{2\Sigma} \frac{\partial}{\partial P} (\Sigma D). \quad (40)$$

The quantity  $F_{fric}$  plays the role of a momentum drift term, that is, an average force resulting from the coupling to the nucleonic "heat bath". We now show explicitly that  $F_{fric}$  has the form of a frictional force.

The total energy of the collective + internal degrees of freedom may be written as:

$$\mathcal{E} = \frac{P^2}{2M} + V(Q) + E^*, \quad (41)$$

where  $E^*$  is the excitation energy of the internal degrees of freedom. Since the total energy  $\mathcal{E}$  is conserved, we can treat the excitation energy  $E^*$  as a function of  $Q$  and  $P$  -  $E^* = E^*(Q, P)$  - which implies that, for fixed  $Q$ ,

$$\frac{\partial}{\partial P} = -\frac{P}{M} \frac{\partial}{\partial E^*}. \quad (42)$$

We can then write the average force,  $F_{fric}$ , exerted by the internal degrees of freedom on the collective degree, as:

$$F_{fric} = -\frac{1}{2\Sigma} \frac{\partial}{\partial E^*}(\Sigma D) \frac{P}{M} = -\gamma \dot{Q}, \quad (43)$$

where  $\dot{Q} \equiv dQ/dt = P/M$ , and

$$\boxed{\gamma = \frac{1}{2\Sigma} \frac{\partial}{\partial E^*}(\Sigma D)}. \quad (44)$$

Eq.43 reveals us that  $F_{fric}$  is indeed a friction-like force, whereas Eq.44 is a fluctuation-dissipation relation between the friction coefficient  $\gamma$ , and the momentum diffusion coefficient  $D$ .

Since the nucleonic degrees of freedom form a many-particle system, the density of states  $\Sigma(E^*)$  grows very rapidly, as a function of excitation energy  $E^*$ . (Roughly,  $\Sigma \sim e^{\sqrt{E^*}}$ .) Thus, for sufficiently large excitation energy, the derivative appearing in Eq.44 may be dominated by the derivative of  $\Sigma$ , i. e. we may treat  $D$  as a constant and with this approximation, we get

$$\gamma \approx \frac{D}{2} \frac{\partial}{\partial E^*} \ln \Sigma = \frac{D}{2T}, \quad (45)$$

where  $T$  is the temperature of the nucleus. Equation 45 is a version of the Einstein relation, originally derived for Brownian motion.

The Fokker-Planck approach outlined above is appropriate for the description of an ensemble of trajectories in the collective coordinate phase space. In a numerical simulation of a single trajectory, we take a Langevin approach instead. Then the equations of motion for the trajectory look like the following:

$$\frac{dQ}{dt} = \frac{P}{M}, \quad \frac{dP}{dt} = -\frac{dV}{dQ} + F_{fric} + \tilde{F}_{fluc}. \quad (46)$$

Here,  $F_{fric} = -\gamma \dot{Q}$  is computed, for instance, with the wall formula [11], whereas  $\tilde{F}_{fluc}$  is a rapidly fluctuating stochastic force. This stochastic force is responsible for the momentum diffusion described by the coefficient  $D(Q, P)$ . The force  $\tilde{F}_{fluc}$  may be simulated numerically, by repeatedly giving the collective momentum a random kick,  $\delta P$ , i. e. a discontinuous change in the collective momentum. The value of  $\delta P$  is chosen randomly from a Gaussian distribution, with a mean value and variance given by:

$$\overline{\delta P} = 0 \quad (47)$$

$$\overline{(\delta P)^2} = D \delta t \quad (48)$$

where  $\delta t$  is the (small) time step between kicks.

The same result (Eq.45) one can obtain using only Lagrange and Langevin equations and not incorporating ensemble distribution equations as shown in [57].

**Diffusion term for Fermi particles.** In the more general case, where we do not treat  $D$  as a constant, equation 44 may be inverted to yield the following expression for the momentum diffusion term, in terms, of the friction coefficient:

$$D = \frac{2}{\Sigma} \int_0^{E^*} dE \Sigma \gamma. \quad (49)$$

The wall formula for one-body dissipation gives the rate of energy transfer from the collective to the nucleonic degrees of freedom as:

$$\frac{dE^*}{dt} = \rho \bar{v} \oint da \dot{n}^2, \quad (50)$$

where  $\rho$  is the mass density of the nucleus,  $\bar{v}$  is the average speed of nucleons, and the final factor on the right side is an integral, over the surface of the nucleus, of the square of the normal outward velocity of a surface element. For a given surface element,  $\dot{n}$  can be written as  $\dot{Q} \partial n / \partial Q$ , where  $(\partial n / \partial Q) dQ$  gives the normal outward displacement of the surface element accompanying an infinitesimal change  $dQ$  in the value of the collective coordinate. Since the rate at which energy is dissipated is equal to  $\gamma \dot{Q}^2 (= -F_{fric} \dot{Q})$ , we get, after dividing both sides of 50 by  $\dot{Q}^2$ , the following expression for the friction coefficient:

$$\gamma = \rho \bar{v} \oint da \left( \frac{\partial n}{\partial Q} \right)^2 \equiv \bar{v} \cdot I(Q). \quad (51)$$

(Note that  $\bar{v}$ , the average nucleonic speed, is a function of excitation energy.) We then have the following expression for the momentum diffusion coefficient:

$$D = \frac{2I}{\Sigma} \int_0^{E^*} dE \Sigma \bar{v}. \quad (52)$$

We treat  $D$  as a function of both  $Q$  (since  $I$  depends on  $Q$ ), and  $P$  (since the excitation energy is determined from  $P$ , for a given  $Q$ ).

If the excitation energy satisfies two conditions: 1)  $E^* \ll \varepsilon_F A^{1/3}$  and 2)  $E^* \gg \varepsilon_F A^{-1}$  then it is appropriate to use the formula for level density of Fermi gas [54]:

$$\Sigma = \frac{6^{1/4}}{12} \frac{g_0}{(g_0 E^*)^{5/4}} \exp \left\{ 2 \left( \frac{\pi^2}{6} g_0 E^* \right)^{1/2} \right\}, \quad (53)$$

where  $g_0 \equiv g(\varepsilon_F) \approx \frac{3}{2} \frac{A}{\varepsilon_F}$  is a sum (protons and neutrons) of the density of one-particle levels on the Fermi surface and  $\varepsilon_F$  is the Fermi energy. The temperature is given by:

$$T^{-1} = \frac{1}{\Sigma} \frac{\partial \Sigma}{\partial E^*} = -\frac{5}{4} E^{*-1} + \left( \frac{\pi^2}{6} \frac{g_0}{E^*} \right)^{1/2} \quad (54)$$

Supposing that  $\bar{v}$  does not change too much with growing  $E^*$  one obtains:

$$D = 8I\bar{v}E^* \left( -1 + \alpha^{1/4} \exp \left( -\alpha^{1/2} \right) \text{Erfi} \left( -\alpha^{1/4} \right) \right), \quad (55)$$

where  $\alpha = \frac{2\pi^2}{3} g_0 E^*$  and function Erfi is the imaginary error function defined as:  $\text{Erfi}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{t^2} dt$ .

**Langevin approach.** Let  $x(t)$  be an evolution trajectory of a system governed by deterministic rule  $\dot{x} = v(x, t)$ . If stochasticity of the system is taken into account then let  $\xi(x, t)$  be a *white Wiener noise*<sup>2</sup>. with a mean value being zero:  $\langle \xi(x, t) \rangle = 0$  and covariance

$$\langle \xi(x, t) \otimes \xi(x, t + s) \rangle = D(x, t) \delta(s),$$

where  $D$  is the covariance matrix and the sign  $\otimes$  corresponds to matrix multiplication. An ensemble of systems is defined by distribution function  $f(x, t)$ . We consider quantities  $x, v, \xi, \partial_x \equiv \frac{\partial}{\partial x}$  to be vectors and  $D$  a matrix. Then the system can be described by 4 possible equations ( $\partial_t \equiv \frac{\partial}{\partial t}$ ):

Deterministic dynamics	Stochastic Wiener noise process
1) generic deterministic equation $\dot{x} = v$	3) Langevin equation $\dot{x} = v + \xi$
2) Liouville-like <sup>*)</sup> equation $\partial_t f = -\partial_x(vf)$	4) Fokker-Planck equation $\partial_t f = -\partial_x(vf) + \frac{1}{2} \partial_{xx}^2(Df)$

<sup>\*)</sup> This equation is exactly Liouville's only in case of nondissipative Newtonian dynamics.

In this subsection we would like to argue why we choose Langevin approach for study the dynamical system.

The first two approaches are completely deterministic. They allow us to certain estimate properties of the system. For example, given a model of colliding nuclei, one can calculate extra-push, extra-extra-push energies, averages of excitation energy, nucleons transfer and so on. However, they do not give a probabilistic description of what we observe in experiments. Therefore we use the probabilistic approach, which allows us to deal with distributions and ensembles. This is accomplished with the random force in the Langevin equation or diffusive term in the Fokker-Plank equation (FPE)<sup>3</sup>.

---

<sup>2</sup>We use the term “white Wiener noise” to point out that the stochastic noise has no time memory and is described by Wiener differential [68]; which means that in equation  $dx = v dt + \xi dW$  the following is satisfied  $\langle \xi_t \xi_{t+s} \rangle \sim \delta(s)$  and  $dW^2 = dt$  and  $\langle \xi^n \rangle = 0$ , when  $n > 2$ .

<sup>3</sup>It is worthwhile to note that there are two interpretations of the Langevin equation and, hence, two types of FPE. One is Ito's representation, where it is assumed that  $\frac{d}{dt}\xi = \frac{\partial}{\partial t}\xi$ . The other one is Stratonovich's, where  $\frac{d}{dt}\xi = \frac{\partial}{\partial t}\xi + \dot{x} \frac{\partial}{\partial x}\xi$  [56]. Klimontovich [58] distinguishes yet the “kinetic form” of FPE.

In previous subsections we used the FPE to establish the fluctuation-dissipation relation from the condition of detailed balance [55]. But in practical implementation (numerical calculations) the use of the FPE is problematic because the dynamical description demands a parameterization of the distribution. And the number of parameters grows exponentially with the dimensionality of the phase space ( $x$ -space). For instance, in the same model of heavy-ion collisions, the configurational space is of dimensionality 4; consider 2 directions to be overdamped; plus 1 dimension of excitation energy. Altogether 7. It is clear that numerical evolution of the FPE is almost impossible to the appropriate order of precision. Thus we are left with the Langevin equation only.

**Implementation of the Langevin equation.** Let us summarize the derivation of the appropriate Langevin equation.

The starting point is the deterministic macroscopic equation

$$\langle \dot{x} \rangle = v(\langle x \rangle) \quad (56)$$

describing the average behavior of the collective coordinates. Adding the noise we obtain the “microscopic” stochastic equation:

$$\boxed{\dot{x} = v(x) + \xi(x, t)}. \quad (57)$$

However

$$\langle \dot{x} \rangle = \langle v(x) \rangle \neq v(\langle x \rangle) \quad (58)$$

in the general case. Fortunately, the choice of the Lagrangian  $\mathcal{L}$  and Rayleigh dissipation function  $\mathcal{F}$  in the form 22a cause the equation of motion to be linear. So the equality  $\langle v(x) \rangle = v(\langle x \rangle)$  holds.

The next step is the determination of the stochastic force  $\xi(x, t)$ . The linearity of the friction term sets the condition of white noise:  $\xi(t)\xi(t+s) \sim \delta(s)$ . The assumption that microscopic fluctuations are small over infinitesimal intervals of time says that this noise is of Wiener type.

One can treat Eq.57 as a simulation of the corresponding FPE. We have derived the relation between the diffusion coefficient  $D$  in Eq.38 and the friction force (Eq.40). Hence using the Ito representation one gets:

$$x(t + \Delta t) = x(t) + v(x(t))\Delta t + \int_t^{t+\Delta t} \xi(x(t), t') dt' \quad (59)$$

with

$$\langle \xi(x, t)\xi(x, t+s) \rangle = D(x)\delta(s) \quad \text{or} \quad \xi dt = D dW, \quad (60)$$

where  $dW$  is the Wiener differential ( $dw^2 = dt$ ). The diffusion term in the corresponding FPE is

$$\frac{1}{2} \left( \frac{\partial}{\partial x} \right)^2 Df. \quad (61)$$



Alternatively, we can use the Stratonovich representation:

$$x(t + \Delta t) = x(t) + v(x(t)) \Delta t + \int_t^{t+\Delta t} \xi(x(t'), t') dt' \quad (62)$$

with

$$\langle \xi(x, t) \xi(x, t + s) \rangle = C^2(x) \delta(s). \quad (63)$$

The integral in the last term in Eq.62 is presented as:

$$\frac{1}{2} [\xi(x(t), t) + \xi(x(t + \Delta t), t + \Delta t)] \Delta t,$$

which is not very convenient since one has to know  $\xi$  at the next point of integration. The diffusion term in the corresponding FPE

$$\frac{1}{2} \left( \frac{\partial}{\partial x} C \right)^2 f = \frac{1}{2} \left( \frac{\partial}{\partial x} \right)^2 C^2 f - \frac{1}{2} \frac{\partial}{\partial x} \left( C f \frac{\partial}{\partial x} C \right). \quad (64)$$

consists of an additional systematic friction-like part.

Ito representation is simpler to use than the Stratonovich one. However, both approaches are equally valid.

## 6. Additional topics

**Exchange of particles.** When the atomic nucleus is idealized as a system of weakly interacting nucleons in a time-dependent mean-field potential well, the nature of the macroscopic dynamics describing the time evolution of the nuclear shape is believed to be intimately related to the nature of the nucleonic motions inside the nucleus. Here we concentrate our attention on the motion of nucleons through a small window area  $\sigma$  (dinuclear regime). One of the effects of such a motion is the dissipation in macroscopic degrees of freedom described by the window formula. Another effect is fluctuations in number of exchanged particles. These fluctuations are not connected to the thermal fluctuations considered above and touch only one degree of freedom — asymmetry. We assume that the motion of nucleons is chaotic and the correlation time is infinitesimally small. Then for sufficient small period of time the process of exchanging nucleons can be treated as a Poisson process. And because of finite number of particles the distribution undergo the binomial law. Let us consider the number of particles flown from nucleus 1 to nucleus 2 during a time interval  $\Delta t$  to be  $N_1$  and from 2 to 1 to be  $N_2$ . Then the change of the atomic mass number of nucleus 1 is equal to:

$$\Delta A_1 = g = N_2 - N_1 \quad (65)$$

This quantity can be rewritten as a sum of average term and fluctuating term with zero mean value:

$$\Delta A_1 = \tilde{g} = \bar{g} + \Delta L, \quad \langle \Delta L \rangle = 0 \quad (66)$$

The diffusion coefficient  $D_A$  is defined by relation:

$$\langle \Delta L^2 \rangle = D_A \Delta t = \langle \tilde{g}^2 \rangle - \bar{g}^2 \quad (67)$$

Quantities  $N_1$  and  $N_2$  are independent and are taken from binomial distributions. So, the following equivalencies:

$$\langle N_1 N_2 \rangle = \langle N_1 \rangle \langle N_2 \rangle \quad \text{and} \quad \langle N_i^2 \rangle = \langle N_i \rangle^2 + \langle N_i \rangle \left( 1 - \frac{\langle N_i \rangle}{A_i} \right) \quad (68)$$

lead us to the explicit expression of  $\langle \Delta L^2 \rangle$ :

$$\langle \Delta L^2 \rangle = \langle N_1 \rangle + \langle N_2 \rangle - \left( \frac{\langle N_1 \rangle^2}{A_1} + \frac{\langle N_2 \rangle^2}{A_2} \right) \quad (69)$$

We denote  $A_1$  and  $A_2$  atomic mass numbers and  $A = A_1 + A_2$ . It has been shown [11,63,66] that the average number of exchanged particles is given by:

$$\langle N_1 \rangle \approx \langle N_2 \rangle \approx \frac{1}{4} \sigma \bar{v} n \Delta t, \quad (70)$$

where  $\bar{v}$  is the average particle speed and  $n$  is the concentration of particles – number per unit volume. This gives us immediately the expression for  $D_A$ :

$$\boxed{D_A = \frac{1}{2} \sigma \bar{v} n - \frac{1}{16a} \sigma^2 \bar{v}^2 n^2 \Delta t} \quad \text{here} \quad a = \frac{A_1 A_2}{A_1 + A_2}. \quad (71)$$

The limit when two containers can be treated as infinite reservoirs of particles is achieved when the time interval  $\Delta t$  is small enough:

$$\Delta t \ll \frac{8a}{\sigma \bar{v} n} \quad (72)$$

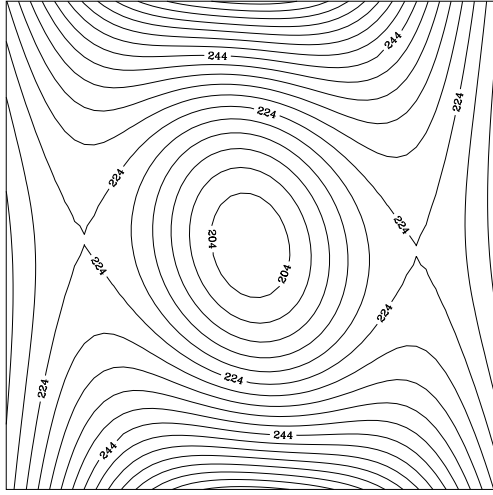
Going to our asymmetry variable  $\Delta$  we use the following relations:

$$\dot{\Delta} = f \dot{A}_1, \quad D_\Delta = f^2 D_A, \quad (73)$$

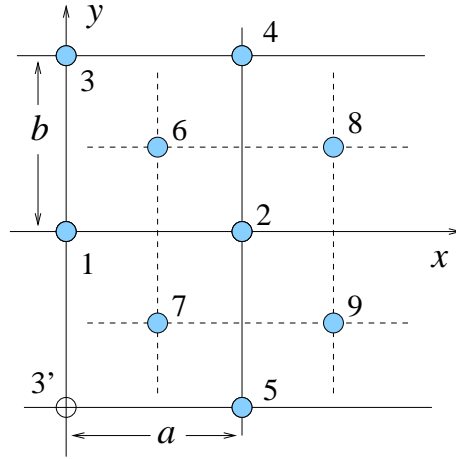
$$\text{where} \quad f = \left. \frac{\partial \Delta}{\partial A_1} \right|_A = \frac{2}{3} A \frac{A_1^{-2/3} A_2^{-2/3}}{(A_1^{1/3} + A_2^{1/3})^2} \quad (74)$$

In Chapter IV we will see what contribution the fluctuations of this kind give to the dynamics; and how they are important in comparison to thermal fluctuations.

**Dissipative fluctuations.** The same scheme can be applied to the fluctuations connected to collision of nucleons with the wall of the container. In derivation of wall and window formulae there have been used only the average quantities. We take into account that the number of particles colliding with the wall and the number of particles crossing the window can fluctuate.



**Fig 2.4** The potential shown on  $(\Delta, \Delta'_Z)$ -subspace, [X-axis  $\Delta=(-0.5, 0.5)$  and Y-axis  $\Delta'_Z=(-0.05, 0.05)$  ], for the system  $^{86}\text{Kr} + ^{136}\text{Xe}$  in the point  $\rho = 1.26$ ,  $\lambda = 0.05$ , which corresponds to the shape at the beginning of the trajectory when nuclei just start to feel nuclear interaction. There are no shell included. Contour lines correspond to levels of the potential energy in MeV with respect to the potential energy of separated nuclei.



**Fig 2.5** We choose points to determine the coefficients in polynomial expansion of the potential around  $(0,0)$ -point in  $(\Delta, \Delta_Z)$ -subspace. Point 1 is the center of coordinate system.

**Parameterization of the potential.** In the above the dynamical coalescence and reparation model was described. This model is more realistic then the schematic one we use in Chapter III to test methods of calculating small probabilities. Simulations made with a more realistic model are time-consuming due to the fact that the potential does not have simple analytic form and can be presented only as a three-dimensional integral, Eqs.16 and 17. Each trajectory in numerical simulations consists of hundreds points; and at each point one has to calculate conservative forces in four directions –  $(\rho, \lambda, \Delta, \Delta_Z)$ . It is clear that calculation of millions of trajectories, for different reactions and sufficient statistics, becomes in practice impossible.

Here we propose the way which can speed up the calculations. Let us parameterize the potential to avoid calculation of three-dimensional integral<sup>4</sup>. The simplest way is to make the four-dimensional lattice in the phase space. Then having values of the potential on the lattice one can interpolate the value in any intermediate point. Simple estimation show that a lattice of size  $100^4$  takes about 200 Mb of the computer memory and a lot of time to calculate the potential. We can use the property that the potential energy is symmetric under transformation  $(\Delta, \Delta_Z) \rightarrow (-\Delta, -\Delta_Z)$ . First, let us note that the parameter  $\Delta_Z$  is strongly related to  $\Delta$  i.e. the potential grows fast when going from minimum in direction  $(\Delta = 1, \Delta_Z = -1)$  and grows slowly in direction  $(\Delta = 1, \Delta_Z = 1)$ . So, let us choose new parameters  $(\Delta, \Delta'_Z)$ , where  $\Delta'_Z = \Delta_Z - \Delta$ . This pair of parameters

<sup>4</sup>An analogous parameterization of the conservative forces for the same model is in [69].

keeps the same property of symmetry. If we do not take into account shell effects then the potential around the point  $(\Delta = 0, \Delta'_Z = 0)$  looks like a smooth potential well (See Fig. 2.4) in the directions  $x = \Delta$  and  $y = \Delta'_Z$ . Therefore we have chosen parameterization in that subspace as a 4'th degree polynomial:

$$z = f_1 + f_2x^2 + f_3xy + f_4y^2 + f_5x^4 + f_6x^3y + f_7x^2y^2 + f_8xy^3 + f_9y^4. \quad (75)$$

This is the expansion of the potential  $z$  around  $(0,0)$ -point. The symmetry property eliminates the odd degree terms. The potential in the subspace  $(\rho, \lambda)$  is parameterized by a bilinear interpolation of the lattice points. Eq.75 can be rewritten:

$$z = \vec{f} \cdot \vec{p}(x, y), \quad (76)$$

where

$$\vec{p}(x, y) = (1, x^2, xy, y^2, x^4, x^3y, x^2y^2, xy^3, y^4)$$

and  $\vec{f} = (f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8, f_9)$ . Now let us choose the 9 points as shown on Fig.2.5 giving the values  $\vec{z} = (z_1, z_2, z_3, z_4, z_5, z_6, z_7, z_8, z_9)$ . Then the coefficients  $\vec{f}$  can be found from the equation:

$$\hat{M} \cdot \vec{f} = \vec{z}, \quad (77)$$

where

$$\hat{M} = \begin{pmatrix} \vec{p}(0,0) \\ \vec{p}(a,0) \\ \vec{p}(0,b) \\ \vec{p}(a,b) \\ \vec{p}(a,-b) \\ \vec{p}(\frac{a}{2}, \frac{b}{2}) \\ \vec{p}(\frac{a}{2}, -\frac{b}{2}) \\ \vec{p}(\frac{3a}{2}, \frac{b}{2}) \\ \vec{p}(\frac{3a}{2}, -\frac{b}{2}) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & a^2 & 0 & 0 & a^4 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & b^2 & 0 & 0 & 0 & 0 & b^4 \\ 1 & a^2 & ab & b^2 & a^4 & a^3b & a^2b^2 & ab^3 & b^4 \\ 1 & a^2 & -ab & b^2 & a^4 & -a^3b & a^2b^2 & -ab^3 & b^4 \\ 1 & \frac{a^2}{4} & \frac{ab}{4} & \frac{b^2}{4} & \frac{a^4}{16} & \frac{a^3b}{16} & \frac{a^2b^2}{16} & \frac{ab^3}{16} & \frac{b^4}{16} \\ 1 & \frac{a^2}{4} & -\frac{ab}{4} & \frac{b^2}{4} & \frac{a^4}{16} & -\frac{a^3b}{16} & \frac{a^2b^2}{16} & -\frac{ab^3}{16} & \frac{b^4}{16} \\ 1 & \frac{9a^2}{4} & \frac{3ab}{4} & \frac{b^2}{4} & \frac{81a^4}{16} & \frac{27a^3b}{16} & \frac{9a^2b^2}{16} & \frac{3ab^3}{16} & \frac{b^4}{16} \\ 1 & \frac{9a^2}{4} & -\frac{3ab}{4} & \frac{b^2}{4} & \frac{81a^4}{16} & -\frac{27a^3b}{16} & \frac{9a^2b^2}{16} & -\frac{3ab^3}{16} & \frac{b^4}{16} \end{pmatrix} \quad (78)$$

The exact solution for  $\vec{f}$  is the following.

$$\begin{aligned} f_1 &= z_1; \\ f_2 &= \frac{-18z_1 + 18z_2 - 2z_3 + z_4 + z_5 + 2z_6 + 2z_7 - 2z_8 - 2z_9}{12a^2}; \\ f_3 &= \frac{-z_4 + z_5 + 16z_6 - 16z_7}{6ab}; \\ f_4 &= \frac{-42z_1 - 18z_2 + 2z_3 - 3z_4 - 3z_5 + 30z_6 + 30z_7 + 2z_8 + 2z_9}{12b^2}; \\ f_5 &= \frac{6z_1 - 6z_2 + 2z_3 - z_4 - z_5 - 2z_6 - 2z_7 + 2z_8 + 2z_9}{12a^4}; \\ f_6 &= \frac{-3z_6 + 3z_7 + z_8 - z_9}{3a^3b}; \end{aligned} \quad (79)$$

$$\begin{aligned}
f_7 &= \frac{2z_1 - 2z_2 - 2z_3 + z_4 + z_5}{2a^2b^2}; \\
f_8 &= \frac{2z_4 - 2z_5 - 5z_6 + 5z_7 - z_8 + z_9}{3ab^3}; \\
f_9 &= \frac{30z_1 + 18z_2 + 10z_3 + 3z_4 + 3z_5 - 30z_6 - 30z_7 - 2z_8 - 2z_9}{12b^4}.
\end{aligned}$$

This simple trick greatly speeds up the calculation especially when we need to get sufficient statistics.

# III. Computation of small probabilities

## 1. Introduction

In the world of microphysics experimental results are often presented in form of the distributions of the measured quantities and not as precise values of these quantities. This reflects the statistical nature of physical phenomena. In nuclear physics, due to the microscopic and quantal nature of nuclei, probabilities and statistical distributions become very relevant. If one takes radioactive decay, for example, it is well known that the nuclear ensemble does not decay at a definite time but is spread over time according to the exponential decay law.

One way to attack the problem theoretically is to use a Monte Carlo method which is very often applied in multidimensional problems. Using Monte Carlo simulations one can efficiently estimate distributions of values we are interested in, in regions of the most probable outcomes. However, sometimes it is important also to calculate distributions and probabilities of rare outcomes. The direct simulation method in this case becomes impractical. This chapter is devoted to that problem.

Experimentalists are able to measure probabilities down to  $10^{-35}$ . It is obvious that no one can calculate such values using direct methods of Monte Carlo. So it is important to have methods capable of giving results comparable with the experimental data. The problem of calculating small probabilities was also studied in different fields. It plays a significant role in chemistry [32–41].

**The probability.** Each realization of a stochastic process (trajectory) is defined by a sequence of random numbers. So one is able to estimate (analytically or numerically) how probable one trajectory is, in comparison to others. That means that each trajectory has its own statistical weight  $p$  which we suppose to be known. The probability of some outcome  $B$  is the ratio of the integral over realizations leading to  $B$ , to the integral over all possible realizations:

$$P_B = \frac{\int_B p(\chi) d\chi}{\int p(\chi) d\chi}, \quad (80)$$

where  $\chi$  is an index defining a trajectory uniquely.

## 2. Polymer method

**Polymer analogy.** A particularly simple and convenient way to describe a stochastic physical process is the Markov description. If one uses a Markov chain of states as a stochastic description of the transition path, then the path can be treated as a polymer in which subsequent beads represent states of the system at subsequent time slices. The analogy between such a polymer and randomly accelerated particle is well known. The transition probability between neighboring states (beads) defines an ensemble of representative paths or polymer configurations. This analogy suggests two novel computational approaches to stochastic process. The first is the possibility to sample the ensemble of realizations by moving the whole path through the phase space according to the defined distribution. The second is the possibility of rewriting the problem of calculating probabilities, as the problem of calculating Helmholtz free energies for a polymer thermalized by a heat reservoir. In this approach, different conditions of the process correspond to different free energies. For instance, the equation 80 can be written down as:

$$P_B = e^{-(F_B - F)}, \quad (81)$$

where  $F_B$  is the free energy of the polymer constrained by the condition  $B$  (in our case this condition is that the last bead correspond to the stage  $B$ ), and  $F$  is the free energy of the unconstrained polymer. As will be shown below the free energy concept arises from the analogy between the statistical distribution of trajectories and the canonical distribution of polymer configurations.

**Action.** The Markov chain representing the path can be described by an action, analogous to a discretized quantum path integral action [52]. If  $\chi$  is a given realization of a process (a given configuration of the polymer) then one can write down the statistical weight  $p(\chi)$  as:

$$p(\chi) = e^{-S(\chi)}. \quad (82)$$

The value  $S(\chi)$  depending on the configuration variable  $\chi$  is called an action. One can treat this value as the energy of the polymer in the configuration  $\chi$ . The distribution  $e^{-S(\chi)}$  corresponds to a canonical Boltzmann distribution with temperature equal to 1. For a Markovian chain  $p(\chi)$  is a product of statistical weights for subsequent time slices and in  $S(\chi)$  this is a sum of subsequent actions (additivity).

The free energy is given by:

$$e^{-F_C} = \int_C e^{-S(\chi)} d\chi, \quad (83)$$

where  $C$  is a fixed external condition constraining the polymer. The definition 83 immediately leads to equation 81.

**Metropolis sweeps.** By changing simultaneously the entire configuration of the polymer in a proper way one can sample the canonical distribution  $e^{-S(\chi)}$ . This algorithm

is known as Metropolis Monte Carlo algorithm. Let us define the probability of changing configuration  $\chi$  into configuration  $\zeta$  in this way:

$$P(\chi \rightarrow \zeta) = \tilde{P}(\chi \rightarrow \zeta) + \delta(\chi - \zeta)R(\chi). \quad (84)$$

The first term

$$\tilde{P}(\chi \rightarrow \zeta) = \eta(\chi, \zeta) \min[1, e^{-(S(\zeta)-S(\chi))}] \quad (85)$$

is the Metropolis Monte Carlo probability rule [42] with a generating probability  $\eta(\chi, \zeta)$  and acceptance probability  $\min[1, e^{-(S(\zeta)-S(\chi))}]$ . The second term in 84 is the probability of rejecting a trial move:

$$R(\chi) = 1 - \int \tilde{P}(\chi \rightarrow \chi') d\chi'. \quad (86)$$

The generating function  $\eta(\chi, \zeta)$  has to satisfy two conditions. The first is that the function  $\eta$  should be symmetric and should not change the measure of space, i.e. if we put  $\tilde{P}(\chi \rightarrow \zeta) = \eta(\chi, \zeta)$  then the exploration of phase space should be uniform. The second is that the function  $\eta$  should generate ergodic, i.e. exhaustive, exploration of the space.

The Metropolis rule allows us to avoid the explicit calculation of the rejection probability,  $R(\chi)$ . The only consideration is the efficiency, which strongly depends on the generating distribution  $\eta(\chi, \zeta)$ . A new chosen configuration  $y$  should not be too close to  $x$ , otherwise exploration of the phase space will be too slow. On the other hand if chances that  $S(y) \gg S(x)$  are high then the number of rejections predominates over acceptations and changes of configuration occur rarely.

The probability  $P(\chi \rightarrow \zeta)$  is normalized by definitions 84-86 and conserves the canonical ensemble:

$$\int d\chi e^{-S(\chi)} P(\chi \rightarrow \zeta) = e^{-S(\zeta)} \quad (87)$$

**Computation of work.** Equation 81 immediately leads to the following method of calculating  $P_B$ . The difference in free energy is equal to the work which needs to be done to constrain the polymer to the condition  $B$  in an adiabatic way [43].

Let us first introduce a new parameter  $\lambda$  which changes from 0 to 1 and corresponds to an unconstrained polymer when  $\lambda = 0$ , and constrained with a condition  $B$  when  $\lambda = 1$ . Now let us choose a sequence of conditions which continuously switches the parameter  $\lambda$  from 0 to 1. When  $\lambda$  is changing infinitely slowly then the total work performed on the system is equal to the Helmholtz free energy difference. In this case the system is in quasistatic equilibrium with the reservoir (having temperature equal to 1) throughout *the switching process*. Hence:

$$\Delta F = \int_0^1 d\lambda \left\langle \frac{\partial S_\lambda}{\partial \lambda} \right\rangle_\lambda. \quad (88)$$

This result is a well-established identity [44].

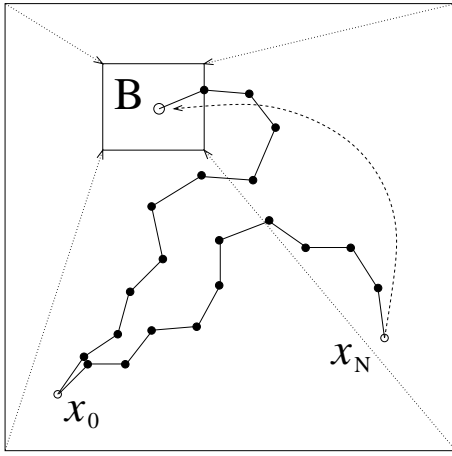


For example, we are interested in the case when the last bead of the polymer is inside a region  $B$ . We can compose the sequence of states where the last bead is confined within a box with hard walls. Initially the box is so big that the polymer is considered to be unconstrained. Then the box shrinks slowly toward the region  $B$  (Fig 3.1).

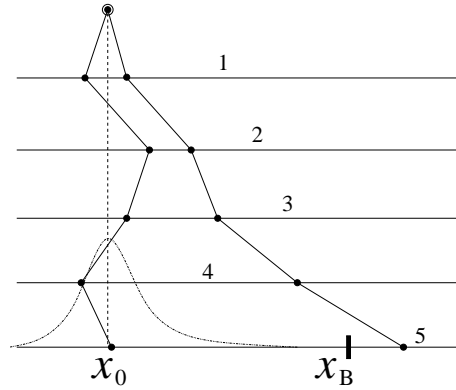
Of course, practically we can not reach the adiabatic limit. And each simulation would give us different result of performed work  $W$ . The relation of free energy difference and performed work in nonequilibrium regime was recently found out by C.Jarzynski [45], [47]:

$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta F}, \quad (89)$$

where  $\beta$  is the reversed temperature (equal to 1 in our case) and the average is defined over an ensemble of independent simulations.



**Fig 3.1** Schematic presentation of the process of constraining the polymer in the region  $B$ . Initially the box is so big that the polymer is considered to be unconstrained. Then the box shrinks slowly toward the region  $B$ .



**Fig 3.2** Simple analytic model of a free falling particle being kicked in horizontal direction. The question is: what is the probability that the particle ends beyond the border  $x_B$ .

**Advantages and disadvantages of the polymer method .** Once the algorithm of the calculations is set then the computing effort does not depend on how small the probability of interest is. One can calculate equally well either some value of probability or a value of several orders of magnitude smaller as the process of calculating work  $W$  does not depend on the value of  $\Delta F$ . This is the main advantage of the polymer method.

Changing the external parameter  $\lambda$  in an adiabatic way means that the polymer is in equilibrium in each moment of time. In other words, the characteristic time of polymer relaxation is much smaller then the characteristic time of parameter changing. The polymer explores the phase space at each intermediate value of  $\lambda$ . This implies a large number of evaluations of  $S(\chi)$  for different configurations of the polymer. So if the calculation of  $S(\chi)$  takes, say, several minutes then the method becomes almost impractical.

Another difficulty appears when the potential energy of the polymer has a non-trivial landscape. In this case the relaxation time of the polymer is the time needed for it to

explore all local minima. If barriers separating local minima are high enough then the probability of passing through a barrier is small and the characteristic relaxation time becomes extremely large.

One solution to this problem is to use “smart” function  $\eta(\chi, \zeta)$  which allows long-ranged jumps to penetrate barriers. This function must have information about potential relief and this makes the task depend on the specific situation<sup>5</sup>.

Another way to deal with the problem is to divide the potential surface into local minima regions (if it is possible, of course, to divide it into regions between which the probability of passing is negligible) and calculate the work separately for moving the polymer from one potential well to another. Often the algorithm of calculating the work  $W$  is such that  $W$  can only grow up to the point after which the polymer quickly relaxes to the new state — falls into the next potential well. While falling the polymer does negative work which is not taken into account. To account for this, one can calculate the work in reverse direction. So the true quantity  $W$  will be the difference:  $W = W_{forth} - W_{back}$ .

The third way is to lift the main (absolute) minimum in potential energy so that it would be comparable with the other minima. Then making the usual direct simulations of the stochastic process we calculate probabilities of different outcomes. And finally we compensate the real probability by the factor  $e^{-W}$ , where  $W$  is the work done when changing the potential. For example, if there are two possible outcomes of the stochastic process  $A$  and  $B$ , computation of the work gives the quantity  $W$  and simulations with the modified potential give probabilities  $P'_A$  and  $P'_B$  of events  $A$  and  $B$ . The real probabilities  $P_A$  and  $P_B$  can then be expressed as:

$$\frac{P_B}{P_A} = e^{-(F_B - F_A)} = e^{-(F_B - F'_A) - W} = \frac{P'_B}{P'_A} e^{-W}, \quad (90)$$

where  $F_A$ ,  $F_B$  are free energies for conditions  $A$  and  $B$ , and  $F'_A$  is free energy with the lifted potential energy in the region  $A$ .

### 3. Examples

**Analytically solvable model .** It is instructive to show how the method works in a simple exactly solvable model. Let us choose the process of discrete steps with Gaussian deviation:

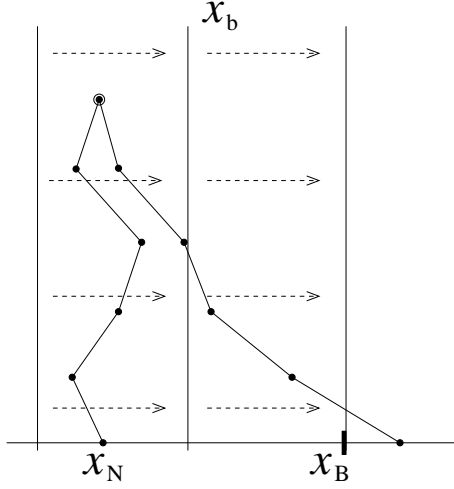
$$P(x_{i+1}) = \mathcal{N} \exp - \frac{(x_{i+1} - x_i)^2}{2\sigma^2}, \quad x_0 = 0. \quad (91)$$

The trajectory is presented as a set of  $N$  numbers  $\{x_i | i = 1, \dots, N\}$ . This process is actually the simulation of a pure Wiener process which also will be considered later in this chapter. We can look at this process as a free particle falling down with a constant speed with stochastic kicks in the horizontal direction (Fig.3.2). As a small probability

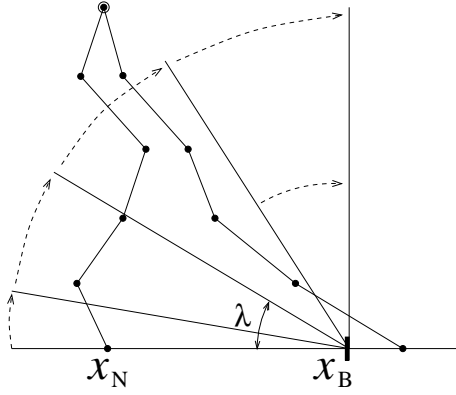
---

<sup>5</sup>Remember that the function  $\eta$  has to satisfy the two conditions mentioned above (just after Eq.86).

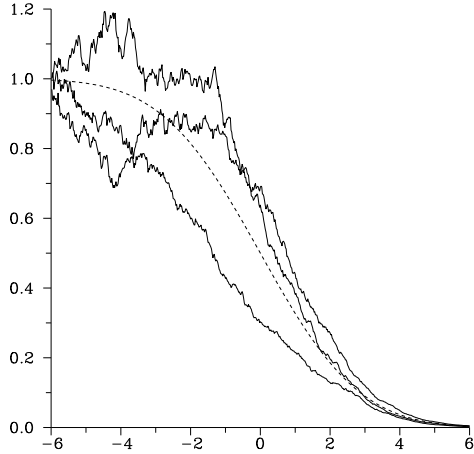
event let us define the event  $B$  when the particle falls far away from initial position  $x_0 = 0$ , say,  $x_N > x_B$ .  $x_N$  is the coordinate of the last  $N$ 'th bead.



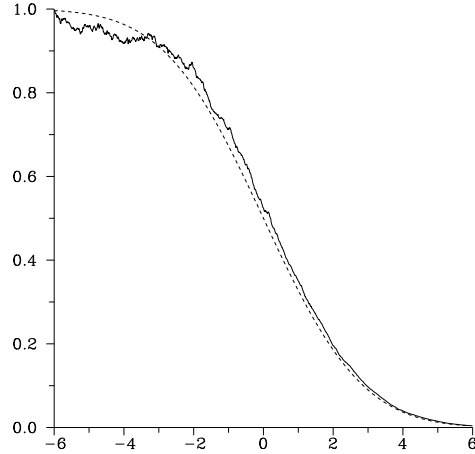
**Fig 3.3** Producing the work  $W$  consists of moving the last bead slowly to  $x_B$  not allowing it to jump behind the moving border  $x_b$ . One can draw the dependence of work from the parameter  $x_b$  (Fig 3.5, 3.6 below).



**Fig 3.4** The “board” rises from horizontal to vertical state. So the probability for  $x_N$  to be far away to the left of  $x_B$  becomes smaller and smaller, until it is exactly zero when potential energy is  $+\infty$  if  $x_N < x_B$ .



**Fig 3.5** Three samples of computing work process. X-axis shows the position of the border crawling from -6 to 6 ( $h = 0.01$ ). Y-axis corresponds to  $e^{-W}$ , i.e. the probability, where  $W$  is the work done up to a moment. Dashed line presents the exact (adiabatic limit) behaviour.



**Fig 3.6** The same dependence as on previous picture, but the border is moving with a speed which is ten times slower ( $h = 0.001$ ). We see that the numerical curve approaches the analytic solution as adiabatic criterion begins to be satisfied.

The action is deduced from the product of step probabilities  $P(x_i)$ :

$$S = \frac{1}{2\sigma^2} \sum_{i=1}^N (x_i - x_{i-1})^2 \quad (92)$$

The exact solution of the given model is:

$$P(x_N > x_B) \equiv F(x_B) = \frac{1}{2}(1 - \text{Erf}(x_B/\sqrt{2\sigma^2 N})) \quad (93)$$

For numerical simulations we have taken  $N = 5$ ,  $\sigma = 1$  and  $x_B = 6$ . The function  $\eta$  has been chosen as follows:

$$x_i \rightarrow x_i + \sum_{j=1}^i \delta x_j, \quad \delta x_j = \sigma_\eta \mathcal{R}, \quad (94)$$

where  $\mathcal{R}$  is a random number taken from a uniform distribution on  $[-\frac{1}{2}; \frac{1}{2}]$ , and  $\sigma_\eta = 1$  — the range of the trial function. The quantity  $\sigma_\eta$  is selected in a way that the number of acceptances and rejections were compatible. Otherwise the polymer would make steps either too small or too rare, resulting in a long relaxation time.

We have taken two examples of *the switching process*. The first (Fig 3.3) is the process of shrinking the space for  $x_5$  from  $[-\infty; \infty]$  to  $[6; \infty]$  by moving the border  $x_b$  from a large negative value to 6, not allowing  $x_5 < x_b$ . The process is the following. We give time to the polymer to relax — a number of trial moves ( $N_{sweep} = 100$ ). Then we shift the border by a step  $h$  ( $h = 0.001$ ) at the same time shifting the last bead  $x_5$  by  $h$ ; and then we calculate the work equal to the difference of the action before and after shifting. The work  $W$  is calculated along with the border  $x_b$ , and can be plotted and compared with the exact solution  $F(x_b)$ , which is presented in Figures 3.5 and 3.6. The border starts from the point -6 (far enough from  $x_0 = 0$  to consider it to be  $-\infty$ ) and goes to 6. Out of 10 samples of the switching process the deduced result is  $P_B = 3.86 \cdot 10^{-3} \pm 4.7 \cdot 10^{-4}$ . The exact result is  $P_B^0 = 3.645 \cdot 10^{-3}$ .

Another example of choosing a parameter  $\lambda$  is schematically shown in Figure 3.4. We put the last bead on the “board” in the gravitational field. One end of the board is fixed at the point  $x_B$  and the other is lifted up as the board goes from horizontal position to vertical. We associate the parameter  $\lambda$  with the angle of tilt of the board, for instance. During the switching process the probability of  $x_5$  being  $< x_B$  becomes smaller and finally, when the board gets to the vertical state,  $x_5$  is definitely bigger than  $x_B$ . The result of simulations (out of 10 samples) is  $P_B = 3.640 \cdot 10^{-3} \pm 3.4 \cdot 10^{-4}$ . The parameter  $\lambda$  was changing from 0 to 1 with the step  $h = 10^{-4}$ . Gravitational field has been chosen:

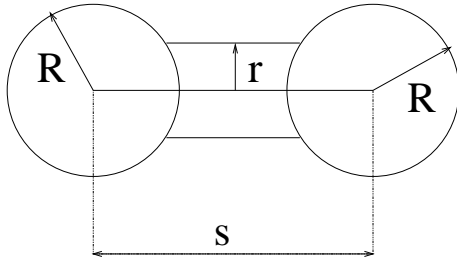
$$(x_B - x_N) \arctan\left(\lambda \frac{\pi}{2}\right), \quad \text{if } x_N < x_B.$$

**Simplified model of HIC .** Here we describe numerical experiments which we have carried out to test the method, using a simplified model of heavy ion collisions [2]. This model was previously studied by Aguiar *et al* in 1990 [67], using Langevin simulations. In this mass-symmetric case - for this simple model - the shape of the system is defined by two equal spheres connected by a cylinder (Fig.3.7). There are two macroscopic (“collective”) variables parameterizing the shape: (1) the *relative distance*  $\rho$  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*  $\alpha$ , which is the square of the ratio of the cylinder radius to the radius of the sphere:  $\alpha = (r_{cyl}/R)^2$ .

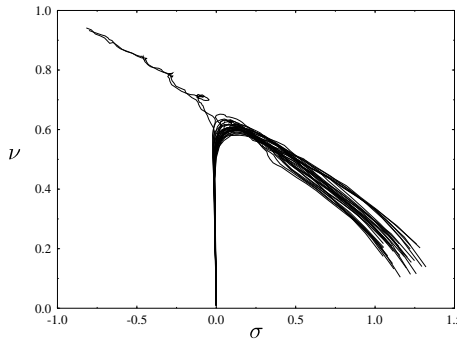
After some approximations for the potential, dissipation and kinetic energy terms, one obtains the following coupled differential equations for the time evolution of the system (see Ref. [2] for details):

$$\begin{aligned} \mu \frac{d^2 \sigma}{d\tau^2} + \nu^2 \frac{d\sigma}{d\tau} + \nu - X &= \hat{\xi}_1 \\ \frac{d\nu}{d\tau} - \frac{2\nu + 3\nu^2 - \sigma}{4\nu(\sigma + \nu^2)} &= \hat{\xi}_2. \end{aligned} \quad (95)$$

Here, the collective coordinates  $\rho$  and  $\alpha$  are represented by the variables  $\nu = \sqrt{\alpha}$  and  $\sigma = \rho^2 - 1$ ; the constant  $\mu$  is a reduced mass,  $\tau$  is a reduced time,  $X$  is conservative force, and  $\hat{\xi}_1$  and  $\hat{\xi}_2$  are stochastic forces (Gaussian white noise), related to the dissipative terms by a fluctuation-dissipation relation. The evolution of the colliding nuclei is then represented by a Langevin trajectory in  $(\sigma, \nu)$ -space. Fig.3.8 depicts 30 such trajectories of the collision of two  $^{100}\text{Mo}$  nuclei. All trajectories start from a configuration of two touching spheres ( $\sigma = 0, \nu = 0$ ), with a center-of-mass energy equal to 4 MeV above the interaction barrier. This energy is about 5 MeV below the “extra push” energy, so most of the trajectories (28 of them) lead to reseparation of the system (fission), and only two trajectories lead to a compound nucleus (fusion).



**Fig 3.7** Shape parameterization in a schematic model of W.Swiatecki (symmetric case) Two spheres are connected by a cylinder.

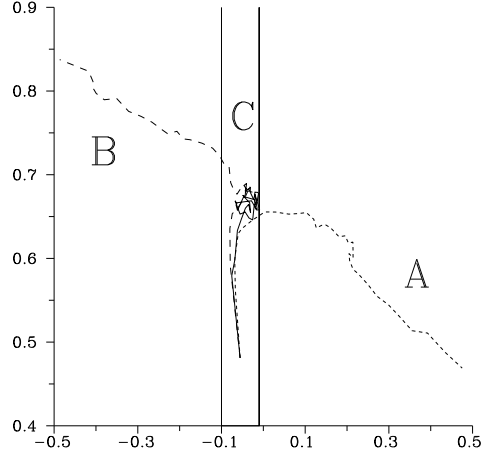


**Fig 3.8** Simulations. Thirty trajectories simulated using the schematic model of nuclear collisions, Eq.95. The system is  $^{100}\text{Zr} + ^{100}\text{Zr}$ , at 0.8 MeV above the interaction barrier. Two trajectories lead to fusion; the rest to reseparation.

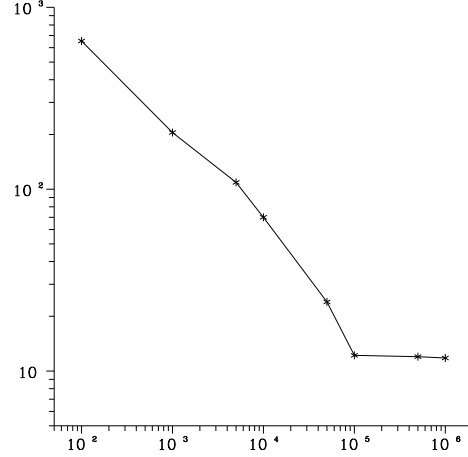
From Fig.3.8 we have the following picture of the physical process occurring, in the context of this simplified model: first the window opening between the two nuclei grows rapidly; then around a saddle point the combination of deterministic and stochastic forces determines the ultimate fate of the nuclei, either fusion or reseparation; and finally the system evolves toward its destiny, with  $\sigma$  decreasing in the case of fusion, or increasing with reseparation.

In polymer representation two outcomes correspond to two stable states (Fig.3.9). Let us call them  $A$ (reseparation) and  $B$ (fusion). During the switching process the polymer eventually will pass through saddle point in the phase space of polymer configuration. It is clear that this saddle point being unstable state corresponds roughly to the situation

when most of the beads lay near the saddle point of the parameter phase space. So let  $C$  denote the position of the polymer constrained between two walls  $x_A$  and  $x_B$  near the saddle point of the parameter phase space. The absence of one of the wall, while being in  $C$ , immediately leads to relaxation of the polymer either to  $A$  or  $B$ .



**Fig 3.9** Three states of the polymer corresponding to 3 stochastic trajectories:  $A$ -reseparation,  $B$ -fusion, and  $C$ -remaining on the saddle point. The state  $C$  is unstable; relaxation goes whether  $C \rightarrow A$  or  $C \rightarrow B$ . The quantities of interest are free energy differences between  $F_A$ ,  $F_B$  and  $F_C$ .



**Fig 3.10** Dependence of  $W_{AC}$  on number of sweeps. Work  $W$  depends strongly on number of simulations  $N$ . Adiabatic limit is achieved when  $N \rightarrow \infty$ . For estimation of true value of  $W$  one has to take sufficient number of sweeps of the polymer.

The ratio between the probability of fusion  $P_B$  and reseparation  $P_A$  can be calculated from:

$$\frac{P_B}{P_A} = e^{-(W_{AC} - W_{BC})}, \quad (96)$$

where  $W_{AC}$  and  $W_{BC}$  are works needed to transfer the polymer from  $A$  to  $C$  and from  $B$  to  $C$ .

For numerical simulations it has been chosen the “board” algorithm. Changing the potential for separate beads we force the polymer to slide down toward  $C$ . The additional potential has been taken in a form:

$$S_B(\lambda) = \sum_{\sigma_i > x_B} i(\sigma_i - x_B) \arctan\left(\lambda \frac{\pi}{2}\right), \quad (97)$$

and the analogous formula for  $A$ . Parameter  $\lambda$  switches from 0 to 1. The algorithm of sweeps differs from the one used in the analytically solvable model above. Here we shift randomly one bead for one step rather than the whole polymer.

This procedure has given us the probability for energy 0.1 MeV equal to  $P_B = 2.57 \cdot 10^{-5}$ . The result obtained from direct sampling of stochastic trajectories is  $2 \cdot 10^{-5}$  calculated from 100000 simulations. The different speed of changing parameter  $\lambda$  gives the different values of work. But adiabaticity can be achieved after some critical number of simulations. This can be seen from Fig.3.10.

#### 4. Importance Sampling Method

Another look at the problem of calculating small probabilities is biasing Langevin trajectories or importance sampling method (ISM). It will be used later on for obtaining results in modeling low-energy nuclear collisions. This method is simpler in use than the outlined above, and is based on Langevin description. The scheme which we propose then involves running a number of simulations with the modified equation of motion and then biasing each trajectory, so as to compute the probability of the desired process.

**Langevin dynamics.** Langevin equations are used to model many processes of physical interest. Disregarding many intrinsic degrees of freedom and remaining with only a few collective variables of motion one describes a system with conservative and non-conservative forces. The other part of Langevin equation – fluctuation term – introduces noise in the dynamics. Actually with Langevin equations we understand the dynamical equation (or system of equations) where together with deterministic term there are stochastic terms usually in a form of white noise.

Schematically Langevin equation can be written down in the following way:

$$\frac{dx}{dt} = v_0(x) + \hat{\xi}(t). \quad (98)$$

Here,  $x$  is the collective variable<sup>6</sup>,  $v_0(x)$  is a “drift” term which embodies the deterministic forces – both conservative and dissipative – acting on the collective degrees of freedom, and  $\hat{\xi}(t)$  is a stochastic, white noise term:

$$\langle \hat{\xi}(t)\hat{\xi}(t+s) \rangle = D\delta(s), \quad (99)$$

where  $\langle \dots \rangle$  denotes an average over realizations of  $\hat{\xi}(t)$ , and  $D > 0$  is a diffusion constant.

A single realization of this process is described by a trajectory  $x(t)$ . Let us define a whole trajectory  $x(t)$  as  $\chi$ . (We separate  $x$ -space from  $\chi$ -space because  $x$  defines only a state of the system in a given time and  $\chi$  defines the whole evolution corresponding to one realization of the process.) The statistical ensemble of realizations of  $\hat{\xi}(t)$  – together with Eq.98 and the initial condition  $x(0)$  – defines a statistical ensemble of trajectories  $\chi$ . Given the above Langevin equation (Eq.98), along with the initial condition, we are interested in computing the probability that the final point of the trajectory,  $x(\tau)$ , falls within a given region  $B$  in  $x$ -space. Now let us introduce a functional  $\Theta[\chi]$  which is equal to 1 if the trajectory yields a  $B$  event<sup>7</sup>, and 0 otherwise.<sup>8</sup>:

---

<sup>6</sup>More generally, the number of variables required to specify the instantaneous state of the system will be greater than one:  $x \rightarrow \vec{x} = (x_1, \dots, x_d)$ . In the case of overdamped motion, this vector specifies the instantaneous configuration of the colliding nuclei. For evolution in which inertial effects are important, the vector  $\vec{x}$  will include both configurational variables, and their associated momenta.

<sup>7</sup> $B$  event is an element of  $x$ -space being inside the region  $B$

<sup>8</sup>Of course, in our case, the functional  $\Theta[\chi]$  is really just a function of the final point,  $x(\tau)$ . More

$$\Theta[\chi] = \begin{cases} 1, & x(\tau) \in B; \\ 0, & x(\tau) \notin B. \end{cases} \quad (100)$$

Then the probability of a  $B$  event,  $P_B$ , is just the average of this functional over the ensemble of trajectories  $\chi$ :

$$P_B = \langle \Theta \rangle. \quad (101)$$

We can compute  $P_B$  by sampling randomly from the ensemble  $\chi$  (i.e. repeatedly simulating trajectories evolving under Eq.98), and counting how many lead to the region  $B$ :

$$P_B \cong P_B^{(N)} = \frac{1}{N} \sum_{n=1}^N \Theta[\chi_n], \quad (102)$$

where  $\chi_n$  is the  $n$ 'th of  $N$  independent simulations of the process. However, since the number of simulations needed to compute  $P_B$  to a desired accuracy grows as the inverse of the probability itself (see Eq.125 later on), this “direct sampling” method becomes impractical for very small values of  $P_B$ .

**Importance sampling.** Suppose we have some space ( $\zeta$ -space) in which are defined some observable  $\mathcal{O}(\zeta)$  and two normalized probability distributions,  $p_1(\zeta)$  and  $p_2(\zeta)$ , supposing furthermore that  $p_2(\zeta) > 0$  whenever  $p_1(\zeta) > 0$ . Let  $\langle \mathcal{O} \rangle_1$  and  $\langle \mathcal{O} \rangle_2$  denote the averages over these two distributions. Importance sampling is based on a simple idea, embodied by the next formula:

$$\langle \mathcal{O}(\zeta) \rangle_1 \equiv \int \mathcal{O}(\zeta) p_1(\zeta) d\zeta = \int \mathcal{O}(\zeta) \frac{p_1(\zeta)}{p_2(\zeta)} p_2(\zeta) d\zeta \equiv \left\langle \mathcal{O}(\zeta) \frac{p_1(\zeta)}{p_2(\zeta)} \right\rangle_2 \quad (103)$$

Let us introduce a *biasing function*

$$w(\zeta) = \frac{p_1(\zeta)}{p_2(\zeta)}, \quad (104)$$

defined at all points  $\zeta$  for which  $p_2(\zeta) > 0$ . If we are interested in computing  $\langle \mathcal{O} \rangle_1$ , then we can do so by repeatedly sampling from the distribution  $p_1(\zeta)$ :

$$\langle \mathcal{O} \rangle_1 = \lim_{N \rightarrow \infty} (1/N) \sum_{n=1}^N \mathcal{O}(\zeta_n^1), \quad (105)$$

where  $\zeta_1^1, \zeta_2^1, \dots$  is a sequence of points sampled independently from  $p_1(\zeta)$ . However, it follows from Eqs.103 and 104 above that we can equally well express the desired average as:

---

generally, however, we could have introduced a criterion which depends on the entire trajectory, in which case  $\Theta[\chi]$  would be a genuine functional.



$$\langle \mathcal{O} \rangle_1 = \langle w \mathcal{O} \rangle_2 = \lim_{N \rightarrow \infty} (1/N) \sum_{n=1}^N w(\zeta_n^2) \mathcal{O}(\zeta_n^2), \quad (106)$$

where  $\zeta_1^2, \zeta_2^2, \dots$  is a sequence of points sampled independently from  $p_2(\zeta)$ . Thus, provided we can compute  $w(\zeta)$  and  $\mathcal{O}(\zeta)$  for any  $\zeta$ , Eq.106 gives us a prescription for determining the average of  $\mathcal{O}$  over the ensemble 1, using points drawn from the sampling ensemble 2. This prescription becomes a useful 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.106, than when sampling directly from the ensemble 1 (Eq.105).

**Biassing Langevin trajectories.** Let us now consider how we might apply importance sampling to the problem which interests us, namely, computing the probability of rare events in Langevin dynamics. Let us therefore consider a modified version of Eq.98:

$$\frac{dx}{dt} = v_0(x) + \Delta v(x) + \hat{\xi}(t), \quad (107)$$

where the modification consists of adding an extra drift term,  $\Delta v$ , chosen to increase the probability of simulating a rare B event. Now, for a given trajectory  $x(t)$  let  $p_1[\chi]$  denote the probability density that we will obtain this particular trajectory, under the original Langevin equation (Eq.98), and let  $p_2[\chi]$  denote the probability density for obtaining this trajectory, under the modified equation (Eq.107)<sup>9</sup>. Finally, let  $w$  denote the ratio of these two densities<sup>10</sup>:

$$w[\chi] \equiv \frac{p_1[\chi]}{p_2[\chi]}. \quad (108)$$

Then, by Eq.106, we can express the probability  $P_B$  (defined with respect to the *original* Langevin equation) as:

$$P_B = \langle w \Theta \rangle_2 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N w[\chi'_n] \cdot \Theta[\chi'_n], \quad (109)$$

where  $\chi'_n$  is the  $n$ 'th of  $N$  independent realizations obeying the modified Langevin equation. This means that, if we know how to compute  $w$  for any trajectory  $x(t)$ , then we can estimate  $P_B$  by running  $N$  simulations with the modified Langevin equation, adding together the weights  $w$  for all those trajectories which lead to  $B$ , and dividing by the total number of simulations,  $N$ . In the limit  $N \rightarrow \infty$ , this estimate converges to the correct value of  $P_B$ .

**Langevin action.** The original and modified Langevin equations, Eqs.98 and 107, can be represented by the generic equation

---

<sup>9</sup>In this case  $\chi$  stands for a solution of equation 107

<sup>10</sup>Note that while the values of  $p_1[\chi]$  and  $p_2[\chi]$  depend on the *measure* chosen on path space, the ratio  $w[\chi]$  is independent of that measure.

$$\boxed{\frac{dx}{dt} = v(x) + \hat{\xi}(t)}, \quad (110)$$

where  $v = v_0$  in one case, and  $v = v_0 + \Delta v$  in the other. Given some initial conditions  $x(0) = x_0$ , let  $\chi$  denote the trajectory evolving from those initial conditions, for a particular realization of the noise term. We are interested in the probability density  $p[\chi]$  for obtaining a particular trajectory  $\chi$ . To introduce a measure on the space of all possible trajectories, let us discretize the trajectory. Thus, let  $(x_0, x_1, \dots, x_M)$  and  $\chi_{discr}$  denote a set of points specifying the state of the system, after time intervals  $\delta t = \tau/M$ :

$$x_m = x(m \delta t), \quad m = 0, \dots, M. \quad (111)$$

We can then ask for the probability density  $p(x_1, \dots, x_M | x_0)$  that the trajectory passes through the sequence of points  $(x_1, x_2, \dots, x_M)$  at times  $\delta t, 2\delta t, \dots, \tau$ , given the initial point  $x_0$ . [Note that this is a probability distribution in  $M$ -dimensional  $(x_1, \dots, x_M)$ -space, with  $x_0$  acting as a parameter of the distribution.]

On the small enough scale of time deviation of a stochastic trajectory from the deterministic one satisfies Gauss distribution:

$$p(x_m, x_{m+1}) = (2\pi D\delta t)^{-1/2} \exp -\frac{(x_{m+1} - x_{m+1}^{det}(x_m))^2}{2D\delta t}, \quad (112)$$

where  $x_{m+1}^{det}(x_m)$  is the solution of equation 110 without a noise term:  $\frac{dx}{dt} = v(x)$  after time interval  $\delta t$ , starting from point  $x_m$ . Then an explicit expression for  $p$  is given by [53]:

$$p[\chi_{discr}] = (2\pi D\delta t)^{-M/2} \exp -A[\chi_{discr}], \quad (113a)$$

where

$$A[\chi_{discr}] \equiv \frac{1}{2D\delta t} \sum_{m=0}^{M-1} [x_{m+1} - x_{m+1}^{det}]^2. \quad (113b)$$

[So far we treat the diffusion coefficient  $D$  as a single value (for one dimensional case) and not depending on  $x$ .] Eq.113a is strictly speaking valid (in a sense that it does not depend on the way of discretization) only in the limit  $M \rightarrow \infty$  (with  $\tau$  fixed), although it constitutes a good approximation if  $\sqrt{D\delta t}$  is small in comparison with the length scale set by variations in  $v(x)$ .

Considering the limit  $M \rightarrow \infty$  one obtains:

$$p[\chi] = \mathcal{N} \exp \int_0^\tau dt \left( \frac{dx}{dt} - v(x) \right)^2. \quad (114)$$

Here,  $dx/dt$  and  $v$  are evaluated along the trajectory  $x(t)$ . The normalizing factor  $\mathcal{N}$  is divergent. But it does not matter as will be seen below.

Eq.114 is a general expression for the probability density of obtaining a particular trajectory  $\chi$ , launched from  $x_0$ . Now let  $p_1[\chi]$  and  $p_2[\chi]$  denote this probability density,

for the specific Langevin processes described by Eqs.98 and 107, respectively. Then the ratio between these two probability densities is given by:

$$w[\chi] \equiv \frac{p_1[\chi]}{p_2[\chi]} = \exp -\Delta A[\chi] \quad (115a)$$

$$\Delta A \equiv A_1 - A_2, \quad (115b)$$

where  $A_1$  and  $A_2$  are computed, using  $v = v_0$  and  $v = v_0 + \Delta v$ , respectively. Normalizing factors  $\mathcal{N}$  in both cases 1 and 2 are the same due to the fact that probabilities are calculated from different distributions but for the same trajectory. For Gauss function in Eq.113 that means only the shift of the distribution in  $\chi$ -space. The final result is:

$$w[\chi] = \exp -\Delta A \quad (116a)$$

$$\Delta A[\chi] = \frac{1}{D} \int_0^\tau dt \left( \frac{dx}{dt} - v_0 - \frac{1}{2} \Delta v \right) \Delta v. \quad (116b)$$

**Compensation for modifications.** Combining Eqs.109 and 116, we now have an expression which allows us, in principle, to compute the probability – defined with respect to the original equation of motion (Eq.98) – by running independent simulations with the modified equation of motion (Eq.107):

$$P_B = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \Theta[\chi_n^2] \cdot \exp -\Delta A[\chi_n^2]. \quad (117)$$

Here,  $\chi_n^2$  is the trajectory generated during the  $n$ 'th simulation, using the modified Langevin equation;  $\Delta A$  is computed for each trajectory; and  $\Theta$  is, as before, equal to one or zero, depending on whether or not  $B$  event occurred.

For a given set of original and modified Langevin equations, and for a large number  $N$  of trajectories simulated under the modified equations, we thus have the following estimator for  $P_B$ :

$$P_B \cong P_B^{(N)} = \frac{1}{N} \sum_{n=1}^N \Theta_n \exp -\Delta A_n, \quad (118)$$

where  $\Theta_n \equiv \Theta[\chi_n^2]$ , and  $\Delta A_n \equiv \Delta A[\chi_n^2]$ . The estimator  $P_B^{(N)}$  converges to the true value  $P_B$  in the limit of infinitely many trajectories:  $\lim_{N \rightarrow \infty} P_B^{(N)} = P_B$ .

**Efficiency gain.** Having derived an estimator for  $P_B$  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.117 does not depend on the form of  $\Delta v(x)$ . Therefore, for any additional drift term  $\Delta v$ , there will be some threshold value  $N_{\Delta v}^*$  such that  $P_B^{(N)}$  provides a “good” estimate of  $P_B$  for  $N \geq N_{\Delta v}^*$ . That is,  $N_{\Delta v}^*$  is the number of trajectories which

we need to simulate (with the modified Langevin equation), in order to determine  $P_B$  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  $P_B$ , for different drift terms. In particular – since the special case  $\Delta v = 0$  is equivalent to computing  $P_B$  using the original Langevin equation – let us define the *efficiency gain*,  $E_{\Delta v}^G$ , associated with a given  $\Delta v(x)$ , as follows:

$$E_{\Delta v}^G \equiv \frac{N_0^*}{N_{\Delta v}^*}. \quad (119)$$

The numerator is just the number of trajectories needed to accurately estimate  $P_B$  by running simulations with the original Langevin equation ( $\Delta v = 0$ ); the denominator is the number needed using Eq.118, for a given  $\Delta v(x)$ . Thus<sup>11</sup>,  $E_{\Delta v}^G$  is the factor by which we reduce the computational effort, by making use of importance sampling – again, for a given  $\Delta v(x)$ .

Let us derive an expression for  $E_{\Delta v}^G$  in terms of quantities extracted directly from numerical simulations. For a given additional drift term  $\Delta v$ , let us define

$$f[\chi] \equiv w[\chi] \cdot \Theta[\chi] = \Theta \exp -\Delta A. \quad (120)$$

Our method of computing  $P_B$  then amounts to computing the average of  $f$ , by sampling trajectories  $\chi$  from the sampling ensemble  $S$ <sup>12</sup>:  $\langle f \rangle_S = P_B \cong P_B^{(N)} = (1/N) \sum_n f_n$ , where  $f_n = f[\chi_n^S]$ . The statistical error in our result – the expected amount by which  $P_B^{(N)}$  differs from  $P_B$  – is then given by the usual formula for the standard deviation of the mean:

$$\sigma_{P_B}^{(S)} = \frac{\sigma_f}{\sqrt{N}}, \quad (121)$$

where  $\sigma_f^2$  is the variance of the quantity  $f$  over the sampling ensemble,

$$\sigma_f^2 = \langle f^2 \rangle_S - \langle f \rangle_S^2. \quad (122)$$

If we want to compute  $P_B$  to a desired relative accuracy  $r$  – in the sense that we want the ratio  $\sigma_{P_B}/P_B$  (expected error / desired average) to fall below  $r$  – then we get the following expression for the minimum number of trajectories needed:

$$N_{\Delta v}^* = \frac{\sigma_f^2}{r^2 P_B^2} = \frac{1}{r^2} \cdot \frac{\langle f^2 \rangle_S - \langle f \rangle_S^2}{\langle f \rangle_S^2}. \quad (123)$$

In other words, if we simulate more than this many trajectories, then we can expect the statistical error in our final estimate of  $P_B$  to be no greater than  $r$  times  $P_B$ .

---

<sup>11</sup>We are assuming that simulating a single trajectory with the modified Langevin equation takes as much time as simulating one with the original Langevin equation.

<sup>12</sup>Here we will use  $S$  to signify an ensemble of trajectories  $\chi$  obtained from modified Langevin equation 107 and  $T$  to signify an ensemble of trajectories  $\chi$  obtained from original Langevin equation 98

In the case of direct sampling from the target ensemble  $T^{13}$ , we estimate  $P_B$  by computing the average of  $\Theta$ . The expected statistical error of this average making use of the fact that  $\Theta^2 = \Theta$  is just

$$\sigma_{P_B}^{(T)} = \left[ \frac{\langle \Theta^2 \rangle_T - \langle \Theta \rangle_T^2}{N} \right]^{1/2} = \left[ \frac{P_B(1 - P_B)}{N} \right]^{1/2}. \quad (124)$$

By setting this expected statistical error equal to  $rP_B$ , we get

$$N_0^* = \frac{1}{r^2} \cdot \frac{1 - P_B}{P_B} \cong \frac{1}{r^2 P_B}, \quad (125)$$

for  $P_B \ll 1$ . (If we simulate this many trajectories, then the expected number of fusion events is  $1/r^2$ , and the expected statistical error in this number is  $1/r$ .)

Combining Eqs.123 and 125, we get the following result for the efficiency gain of our importance sampling method, for a particular choice of  $\Delta v(x)$ :

$$E_{\Delta v}^G = \frac{N_0^*}{N_{\Delta v}^*} = \frac{P_B(1 - P_B)}{\sigma_f} \cong \frac{P_B}{\sigma_f} = \frac{\langle f \rangle_S}{\langle f^2 \rangle_S - \langle f \rangle_S^2}. \quad (126)$$

Eq.126 gives the efficiency gain of importance sampling, with a particular choice of  $\Delta v$ , in terms of averages which can be estimated from simulations performed under the modified Langevin equation alone (i.e. sampling only from  $S$ , not  $T$ ). An expression for  $E_{\Delta v}^G$  in terms of averages estimated using both the original and modified equations of motion, is:

$$E_{\Delta v}^G = \left[ \frac{\sigma_{P_B}^{(T)}}{\sigma_{P_B}^{(S)}} \right]^2 = \frac{\langle \Theta \rangle_T - \langle \Theta \rangle_T^2}{\langle f^2 \rangle_S - \langle f \rangle_S^2}. \quad (127)$$

We will use these results below, to compute the efficiency gain of the importance sampling method for particular examples.

## 5. Application of ISM: practical matters

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

**Calculating weights.** We obtained an expression for  $\Delta A$  as a functional of the trajectory  $x(t)$ . Since  $x(t)$  satisfies Eq.107, we can rewrite Eq.116b as:

$$\Delta A = \frac{1}{2D} \int_0^\tau dt (2\hat{\xi} + \Delta v)\Delta v. \quad (128)$$

This expression for  $\Delta A$  leads itself to a convenient implementation of the method proposed here, as follows. When simulating a given trajectory  $x(t)$  evolving under Eq.107,

---

<sup>13</sup>See the previous footnote.

we simultaneously integrate the following equation of motion for a new variable  $y(t)$ , satisfying the initial condition  $y(0) = 0$ :

$$\boxed{\frac{dy}{dt} = \frac{\Delta v}{2D}(2\hat{\xi} + \Delta v)}, \quad (129)$$

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  $\Delta v$  in general depends on  $x$ .) Eq.128 then implies that

$$\Delta A = y(\tau). \quad (130)$$

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 weight  $e^{-\Delta A}$  to this event.

**Cut-off.** All the formalism was derived for a fixed evolution time  $\tau$ . But usually it is not true – the time of evolution can differ for different realizations. Often the evolution of our system is such that, once a trajectory  $x(t)$  enters the region  $B$  which defines fusion, its chance for subsequently escaping that region is negligible:  $B$  effectively possesses an absorbing boundary. It becomes convenient to define both  $v_0$  and  $\Delta v$  to be zero everywhere within  $B$ . Then, if a trajectory  $x(t)$  (evolving under the modified Langevin equation) comes into  $B$  at some time  $\tau' < \tau$ , we can stop the simulation at that point in time, and take  $\Theta = 1$ ,  $\Delta A = y(\tau')$ . This makes the prescription consistent with formulas derived for fixed  $\tau$ .

**$D$  depending on  $x$ .** We have, to this point, assumed that the stochastic noise  $\hat{\xi}(t)$  is independent of  $x$ . That is,  $D = \text{const}$ . More generally, we might have a diffusion coefficient which depends on the instantaneous configuration of the system:  $D = D(x)$ , i.e.

$$\langle \hat{\xi}(t)\hat{\xi}(t+s) \rangle_{x(t)=x} = D(x)\delta(s). \quad (131)$$

In this case Eq.114 becomes:

$$p(\chi) = \mathcal{N}' \exp \int_0^\tau dt \frac{1}{D(x)} \left( \frac{dx}{dt} - v(x) \right)^2. \quad (132)$$

Eq.115 remains unchanged since the factor  $\mathcal{N}'$  in Eq.132 is the same for  $p_1$  and  $p_2$ ; and Eq.116 changes only in that  $1/D$  is brought inside the integral in Eq.116b, with  $D$  evaluated along the trajectory  $x(t)$ . When implementing the method using the additional variable  $y(t)$ , the only difference is that  $D$  is evaluated along  $x(t)$  rather than being a constant, in Eq.129

**Multidimensionality.** Let us now drop the assumption that the system evolves in one dimension. The state of the system is now described by a vector  $\vec{x} = (x_1, \dots, x_d)$ , evolving under a set of coupled Langevin equations,

$$\frac{dx_i}{dt} = v_i(\vec{x}) + \hat{\xi}_i(t), \quad i = 1, \dots, d, \quad (133)$$

where  $\vec{v} = \vec{v}_0$  (original equations of motion) or  $\vec{v} = \vec{v}_0 + \Delta\vec{v}$  (modified equations of motion). The diffusion coefficient  $D$  becomes a symmetric matrix whose elements reflect the correlations between the different components of the stochastic force:

$$\langle \hat{\xi}_i(t) \hat{\xi}_j(t+s) \rangle = D_{ij} \delta(s). \quad (134)$$

[For simplicity, we assume that this matrix is a constant. The generalization to  $D_{ij} = D_{ij}(\vec{x})$  is as straightforward as in the one-dimensional case.]

The simplest case of multi-dimensional evolution occurs when the components  $\hat{\xi}_i$  are mutually uncorrelated. Then  $D$  is a diagonal matrix, and the generalization of Eq.129 is given by:

$$\frac{dy}{dt} = \sum_{i, D_{ii} \neq 0} \frac{\Delta v_i}{2D_{ii}} (2\hat{\xi}_i + \Delta v_i). \quad (135)$$

The sum is taken over values of  $i$  for which  $D_{ii} \neq 0$ , corresponding to those directions along which there is a non-zero stochastic force. Along those directions for which  $D_{ii} = 0$ , we must have  $\Delta v_i = 0$ .

To see this, note that if  $D_{ii} = 0$  for a particular value of  $i$ , then the equation of motion describing the evolution of  $x_i$  is deterministic ( $\hat{\xi}_i = 0$ ). If we now modify that equation by adding an additional non-zero term  $\Delta v_i$ , then any trajectory  $\mathbf{x}(t)$  obeying the modified equations of motion will not be a solution of the original equations – for any realization of the noise term  $\hat{\xi}(t)$  – and vice-versa (unless  $\Delta v_i$  happens to be zero exactly along the trajectory). This violates the condition state before Eq.104: in order for the importance sampling to be valid, our modified equations of motion have to be able to generate any trajectory which might be generated by the original equations of motion.

If  $D$  is not diagonal, then Eq.135 generalizes to the following evolution equation for  $y$ :

$$\frac{dy}{dt} = \frac{1}{2} (2\vec{\xi} + \Delta\vec{v})^T D^{-1} \Delta\vec{v}. \quad (136)$$

Eq.136 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  $\Delta\vec{v}$  onto the subspace spanned by the null eigenvectors of  $D$  is zero<sup>14</sup>. Assuming this condition is satisfied, we can view Eq.136 as pertaining only to the subspace spanned by the non-zero eigenvectors of  $D$ .

The results discussed in this generalization from one-dimensional to multi-dimensional evolution are based on the generalization of Gauss form in Eq.112, Eq.113. As with Eq.136 above, these equations pertain to the  $d'(\leq d)$ -dimensional subspace spanned by the non-zero eigenvectors of  $D$ .

**Inertial motion.** There is an example of multi-dimensional evolution worthy of particular mention. This is the case in which the evolution of the system is not over-damped, i.e. inertial effects are present. The evolution then occurs in the phase space of the system, and the equations of motion for  $\vec{x} = (\vec{q}, \vec{p})$  are typically of the form:

---

<sup>14</sup>The case when one or few of the eigenvalues of  $D$  are close to zero will be discussed later on.

$$\frac{d\vec{q}}{dt} = I^{-1}\vec{p} \quad (137a)$$

$$\frac{d\vec{p}}{dt} = \vec{F} - \Gamma I^{-1}\vec{p} + \vec{\xi}. \quad (137b)$$

Here,  $\vec{q}$  is a vector of variables specifying the configuration of the system;  $\vec{p}$  is the vector of associated momenta;  $I$  is an inertia tensor;  $\vec{F}$  is the vector of conservative forces acting on the system;  $\Gamma$  is a friction tensor; and  $\vec{\xi}$  is the vector of stochastic forces, whose associated diffusion tensor  $D$  is related to  $\Gamma$  by a fluctuation-dissipation relation. (Typically,  $I$ ,  $\vec{F}$ ,  $\Gamma$ , and  $D$  are all functions of  $\vec{q}$ .) In this case, the equations of motion for  $\vec{q}$  are deterministic (Eq.137a), therefore any additional drift terms must appear only in the momentum equations (Eq.137b), as an additional force,  $\Delta\vec{F}$ . The equation of motion for  $y(t)$ , Eq.136, then pertains only to momentum space (the  $\vec{p}$ -subspace of phase space).

**Dissipative motion.** It is often the situation that the dissipative and stochastic forces acting on the collective degrees of freedom depend on the “temperature” of the bi-nuclear system. This is another way of saying that these forces, at time  $t$ , depend on the total amount of the collective energy which has been dissipated up to that time. Since the energy dissipated, as a function of time, will differ from one realization to another, it seems we are faced with “memory-dependent” forces, i.e. forces which, at time  $t$ , depend not only on the instantaneous state of the system, but also on its *history*, up to time  $t$ . An easy way to deal with this situation is simply to expand our list of variables  $\vec{x}$  to include a new member,  $x_{d+1} = E_{\text{diss}}$ , denoting the collective energy dissipated. This new variable is initialized at zero, and evolves under an evolution equation which depends on the model used to describe the colliding nuclei. (For instance,

$$\frac{dE_{\text{diss}}}{dt} = (I^{-1}\vec{p})^T[\Gamma I^{-1}\vec{p} - \vec{\xi}] \quad (138)$$

in the case of collective evolution described by Eq.137.)

With the addition of this new variable, i.e.  $\vec{x} \rightarrow (x_1, \dots, x_d, x_{d+1})$ , we now again have a set of (coupled) Langevin equations, in which the drift and stochastic terms depend only on the instantaneous state of the system,  $\vec{x}$ . We can thus apply the method without further modification.

**Elimination of degenerate directions.** Let us now discuss the more complicated case when matrix  $D$  has zero or close to zero eigenvalues in given regions of the configuration space. Start from the equation of motion for  $\vec{x} = (\vec{r}, \vec{v})$ :

$$\frac{d\vec{r}}{dt} = \vec{v} \quad (139a)$$

$$\frac{d\vec{v}}{dt} = I^{-1}\vec{F} - I^{-1}\Gamma\vec{v} + \vec{\xi} + \vec{w}. \quad (139b)$$

Here,  $\vec{r}$  is a vector of variables specifying the configuration of the system;  $\vec{v}$  is the velocity vector;  $I$  is an inertia tensor;  $\vec{F}$  is the vector of conservative forces acting on the system;  $\Gamma$  is a dissipation tensor;  $\vec{\xi}$  is the vector of stochastic forces, as before, with  $\langle \vec{\xi}_i(t)\vec{\xi}_j(t+s) \rangle_x = D(x)_{ij}\delta(s)$ , and  $\vec{w}$  is an additional force which satisfies the condition following Eq. 136.



The main idea of introducing  $\vec{w}$  instead of  $\Delta\vec{v}$  is to scale the projections of  $\Delta\vec{v}$  onto eigenvectors of  $D$  proportionally to eigenvalues and then to compose a new vector  $\vec{w}$ . This automatically eliminates the projection of  $\Delta\vec{v}$  onto the subspace spanned by the null eigenvectors of  $D$  and fulfils a condition that when eigenvalue is close to zero then the associated projection is also small. Physically this corresponds to the situation that the system could not be pushed in the directions, in which the motion is deterministic or almost deterministic (the width of the stochastic force in these directions is small in comparison to the others).

Let  $\vec{V}^i$  and  $\lambda_i^2$  denote  $i$ 'th eigenvector and eigenvalue of matrix  $D$  respectively. The numbers  $\lambda_i^2$  are always not less than zero if  $\Gamma$  describes physical dissipation and,  $D$  and  $\Gamma$  are connected by fluctuation-dissipation theorem. One can treat the set [of  $\lambda_i^2$ ]  $\{\lambda_i | i = 1, \dots, d\}$  as a vector  $\vec{\lambda}$  and correspondingly  $\vec{V}^i$  as a matrix  $\hat{V}$ , in which columns are eigenvectors  $\vec{V}$ . Then the vector  $\vec{\lambda}\hat{\gamma}$ , where  $\hat{\gamma}$  are independent random numbers multiplying components of  $\vec{\lambda}$  and taken from the unit-width Gauss distribution, would mean stochastic force acting in coordinates of basis of eigenvectors  $\vec{V}$ . Multiplying from the left by  $\hat{V}$  and then  $\hat{I}^{-1}$  we cast this vector into coordinates of  $(\vec{r}, \vec{v})$ . Hence:

$$\delta\vec{\xi} = \hat{I}^{-1}\hat{V}\vec{\lambda}\hat{\gamma}\sqrt{\delta t} \quad (140)$$

is an increment of  $\vec{v}$  caused by stochastic force  $\vec{\xi}$  during the time  $\delta t$ . Eq.140 is just the general prescription of introducing a noise in multidimensional inertial dissipative dynamics.

The quantities  $\lambda$  are widths of the stochastic force  $\vec{\xi}$  in  $\vec{V}$  directions. Following the analogy let us define “the stochastic freedom” of the system as:

$$\vec{T}^i = \hat{I}^{-1}\vec{V}^i\lambda_i \quad (141)$$

The vectors  $\vec{T}$  and their absolute values determine directions and widths of the stochastic force now in  $(\vec{r}, \vec{v})$  coordinates. For this reason the additional force  $\vec{w}$  can be calculated by spanning  $\Delta\vec{v}$  onto vectors  $\vec{T}$ :

$$\vec{w} = \sum_i \frac{(\Delta\vec{v} \cdot \vec{T}^i)\vec{T}^i}{|\vec{T}^i|} \quad (142)$$

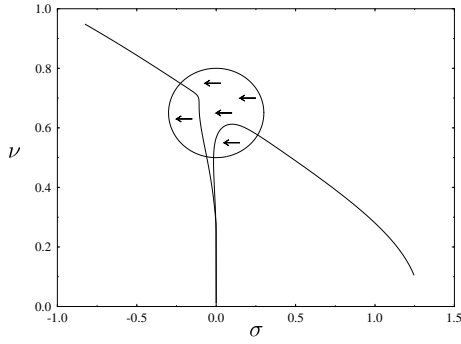
This definition of  $\vec{w}$  automatically will squeeze the components of  $\Delta\vec{v}$  which are forbidden or nearly forbidden by degenerate directions.

**Extension to the non-Markovian case.** The method of importance sampling is very general. It can easily be extended to the case of non-Markovian process. Any numerical study of a Langevin evolution or another continuous process has to be simulated by a discrete time steps. In practice a trajectory of any stochastic process  $\chi$  is defined by a set of random variables  $\{\zeta_1, \zeta_2, \dots, \zeta_M\}$ , in general, chosen from different (depending on time and history) distributions  $\{p_1(\zeta, t_1), p_2(\zeta, t_2, \zeta_1), p_3(\zeta, t_3, \zeta_2, \zeta_1), \dots, p_M(\zeta, t_M, \zeta_{M-1}, \dots, \zeta_1)\}$ . We do not care about the fact that the moments of time  $\{t_1, t_2, \dots, t_M\}$  can be given by a specific algorithm, which could give correlations on a shorter time intervals.

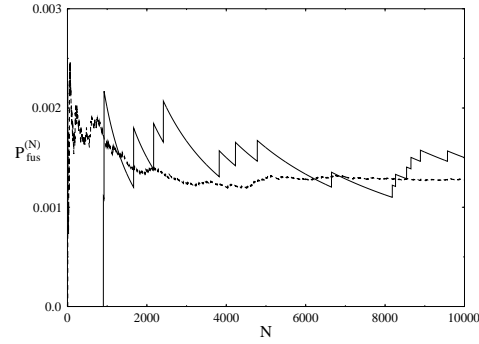
Then again introducing new sample distribution  $p^S(\zeta, t)$  in a way that sequence of most probable values  $\zeta$  leads to desired event let us define the biasing function

$$w(\chi) = \frac{\prod_{n=1}^M p_n(\zeta, t_n, \dots)}{\prod_{n=1}^M p_n^S(\zeta, t_n, \dots)}. \quad (143)$$

The distributions  $p_n^S$  and  $p_n$  are based on the same history  $(\zeta_{n-1}, \zeta_{n-2}, \dots, \zeta_1)$  and taken at the same time  $t_n$ . The random variable  $\zeta$  is sampled from  $p^S$ . However one has to know the probability  $p(\zeta)$  of the same value  $\zeta$  in the unbiased distribution  $p$  to evaluate the function  $w(\chi)$ .



**Fig 3.11** The circle indicates the region around the saddle point and arrows show the direction of the extra force, chosen to push the system toward fusion. Also shown are two (deterministic) trajectories, one ending in fusion and the other one in reseparation.



**Fig 3.12** Convergence of the  $P_{\text{fus}}^{(N)}$  with number of trajectories simulated ( $N$ ), for both direct simulation (solid line), and using importance sampling (dashed line).

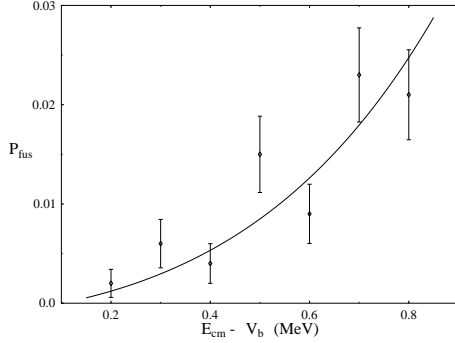
## 6. Examples

**Schematic model.** To test our method we again use the schematic model of heavy-ion collisions [2]. It was outlined briefly above (see subsection with Eqs.95). To increase efficiency we have to modify the stochastic equations (95). The physical behaviour around saddle point 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)  $\sigma$  direction, whose strength is a Gaussian function of  $(\sigma, \nu)$ , with a peak at  $(0.0, 0.6)$ . This leads to the following *modified* Langevin equations of motion:

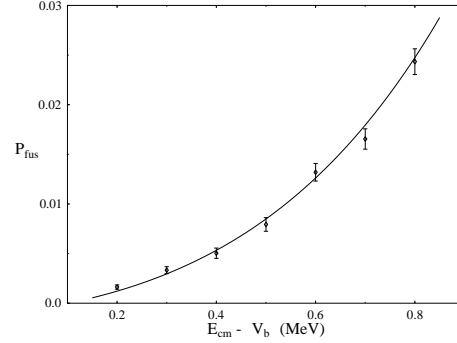
$$\begin{aligned} \mu \frac{d^2 \sigma}{d\tau^2} + \nu^2 \frac{d\sigma}{d\tau} + \nu - X &= \hat{\xi}_1 - A \exp \left\{ -\frac{\sigma^2 + (\nu - 0.6)^2}{0.02} \right\} \\ \frac{d\nu}{d\tau} - \frac{2\nu + 3\nu^2 - \sigma}{4\nu(\sigma + \nu^2)} &= \hat{\xi}_2, \end{aligned} \quad (144)$$

with  $A$  an externally adjustable parameter. Fig.3.11 schematically shows the region around the saddle point, where the additional force pushing the system toward fusion

is localized. Two deterministic trajectories (evolving under the original equations of motion, but without the stochastic terms) are also shown, to guide the eye. One of these was launched with an energy of 1 MeV above the barrier (leading to reseparation), the other one at 5 MeV above the barrier (leading to fusion) for the reaction  $^{100}\text{Zr} + ^{100}\text{Zr}$ .



**Fig 3.13** 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 3.14** Same as Fig.3.13, but computed using importance sampling instead of direct simulation.

To compare our importance sampling method to direct simulation of the original process, we first choose to compute the probability of fusion for trajectories starting with an energy 0.2 MeV above the barrier. We ran  $10^4$  independent trajectories under both the original and the modified Langevin equations, Eqs.95 and 144, respectively, and kept a running tally of the probability of fusion, as computed by the two methods (Eq.102 for the case of direct simulation, Eq.118 for importance sampling). We took the strength of the additional force to be  $A = 0.3\mu$  in these simulations, where  $\mu$  is a reduced mass. Fig.3.12 illustrates the difference between the rates of convergence of the two methods. The plots show the estimates  $P_{\text{fus}}^{(N)}$  as computed by each method, as a function of number of trajectories simulated,  $N$ . It is clear that the importance sampling (broken line) converges much faster than direct simulation (solid line): after about 5000 trajectories, the former has converged very close to its asymptotic value, whereas the latter is still “jumping around” after 10000 trajectories.

The sawtooth pattern exhibited by the direct simulation estimate is typical of the situation in which rare events contribute disproportionately to an ensemble average: in this particular case, only 15 out of the  $10^4$  “direct” trajectories (simulated under Eq.95) lead to fusion; each of these events causes a sudden jump in the  $P_{\text{fus}}^{(N)}$ , followed by a gradual decline as more non-fusion events accumulate. By contrast, under the modified Langevin equations, Eq.144, a total of 253 trajectories lead to fusion, i.e. the non-zero contributions to  $P_{\text{fus}}^{(N)}$  are distributed among many more events, with the consequence of smoother and faster convergence.

In Figs.3.13 and 3.14 we show excitation functions – fusion probability plotted against

center-of-mass energy above the barrier – as computed by both direct simulation and importance sampling, with the same additional force as used in Fig.3.12. 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.<sup>15</sup> Again 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.3.13 is about 5.5 times bigger than that in Fig.3.14. The efficiency gain of the importance sampling approach is therefore about 30 ( $\sim 5.5^2$ , see Eq.127) in this case: we would need to launch about  $30 \times 10^3$  trajectories evolving under the original Langevin equation to get the same degree of accuracy obtained in Fig.3.13 with  $10^3$  trajectories.

The gain in efficiency becomes more dramatic when we go to very small probabilities. To show this we considered the reaction  $^{110}\text{Pd} + ^{110}\text{Pd}$ , 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  $P_{\text{fus}} = (6.970 \pm 0.268) \times 10^{-13}$ . This was computed using our importance sampling method, with an additional force corresponding to  $A = 1.9\mu$ ; about 88% of the trajectories evolving under the modified Langevin equation went to fusion. Using Eq.126, our result gives an efficiency gain of  $E_{\Delta v}^G = 3.5 \times 10^9$ ! We cannot compare our estimate of  $P_{\text{fus}}$  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 indispensable in this case: calculating  $P_{\text{fus}}$  using direct simulations is not practical.

**Exactly solvable model.** To illustrate the efficiency gain which can be achieved by using the importance sampling method presented in this paper, it is instructive to consider an exactly solvable model. Consider a particle in 2 dimensions which falls at a uniform rate  $v_z$  from a height  $h$ , while experiencing random “kicks” in the horizontal direction:

$$\begin{cases} \dot{z} = -v_z \\ \dot{x} = \hat{\xi}(t) \end{cases}, \quad (145)$$

where  $\hat{\xi}(t)$  represents white noise corresponding to a diffusion constant of unit magnitude:  $\langle \hat{\xi}(t)\hat{\xi}(t+s) \rangle = \delta(s)$ . We assume that the initial horizontal location is zero, i.e.  $x(0) = 0$ , and are interested in the horizontal location of the particle when it hits the “ground” ( $z = 0$ ), at time  $\tau = h/v_z$ . The motion in the  $x$ -direction is a Wiener process, whose solution is a Gaussian distribution with a variance which grows linearly with time. Let us suppose we are interested in the probability that the particle will end to the right of the

---

<sup>15</sup>This was obtained by running a very large number of simulations for different values of energy above the barrier, and then fitting the results to an exponential multiplied by a second-order polynomial. The expected error associated with the curve itself is everywhere smaller than the smallest of the error bars shown in Fig.3.14.

fixed point  $x_0$ , i.e.  $x(\tau) > x_0$ . Since the ensemble distribution of  $x(\tau)$ -values is a Gaussian (with variance equal to  $\tau$ ), this probability is given in terms of an error function:

$$P[x(\tau) > x_0] \equiv F(x_0) = \frac{1}{2} [1 - \text{erf}(x_0/\sqrt{2\tau})].$$

For large values of  $x_0$ , this probability dies off very rapidly,  $P \sim e^{-x_0^2/2\tau}/x_0$ , therefore very many simulations would be needed to compute  $P$  to some desired relative accuracy  $r$ .

Let us now consider an additional force, in the form of a constant horizontal “wind”, pushing the particle in the direction of the point  $x_0$ . The modified equations of motion are then

$$\begin{cases} \dot{z} = -v_z \\ \dot{x} = \hat{\xi}(t) + w, \end{cases} \quad (146)$$

where  $w$  is the strength of the wind. From Eq.128, it follows that  $\Delta A = wx(\tau) - w^2\tau/2$ . From Eq.123, we can then compute the minimum number of simulations necessary to obtain  $P$  to a relative accuracy  $r$ , using importance sampling, for a particular value of wind strength:

$$N_w^* = \frac{1}{r^2} \left\{ e^{w^2\tau} \frac{F(x_0 + w\tau)}{F^2(x_0)} - 1 \right\}. \quad (147)$$

The efficiency gain is then:

$$E_w^G \equiv \frac{N_0^*}{N_w^*} = \frac{F(x_0) - F^2(x_0)}{e^{w^2\tau} F(x_0 + w\tau) - F^2(x_0)}. \quad (148)$$

For the case  $x_0 = 3.0$ ,  $\tau = 1.0$  ( $P = 1.35 \times 10^{-3}$ ), we have plotted efficiency gain as a function of wind strength, in Fig.3.15. We see that the optimal wind strength (at which we obtain maximal efficiency) is  $w_{opt} = 3.157$ . For this value of  $w$ , the number of trajectories needed to estimate  $P$  using direct simulation of the original Wiener process, is about 220 times the number needed to estimate  $P$  (to the same degree of accuracy) with importance sampling.

It is interesting to note that even for  $x_0 = 0$  (for which half the trajectories fall to the right of  $x_0$ ), we gain efficiency by using importance sampling. In this case, for  $\tau = 1$ , the maximal efficiency gain (about 1.8) is achieved with a wind strength  $w = 0.6125$ , as shown in Fig.3.16.

The maximal efficiency gain grows as the probability  $P$  becomes small. Using an asymptotic expression for the error function, we have

$$F(x_0) \rightarrow \left( \frac{\tau}{2\pi} \right)^{1/2} \frac{1}{x_0} \exp(-x_0^2/2\tau) \quad (\tau \text{ fixed}, x_0 \rightarrow \infty), \quad (149)$$

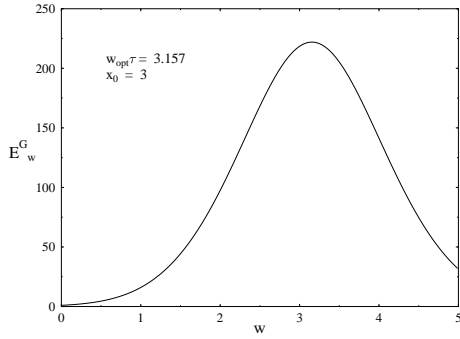
from which it follows that

$$N_w^* \rightarrow \frac{x_0}{r^2} \exp[(x_0 - w\tau)^2/2\tau]. \quad (150)$$

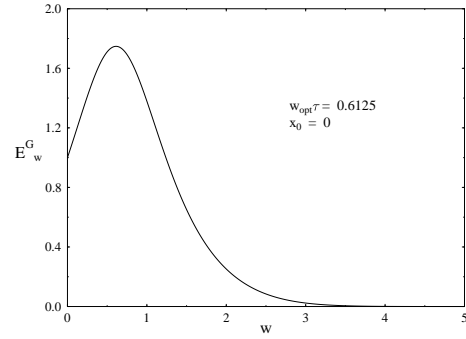
This formula nicely encapsulates the dramatic efficiency gain achieved for small values of  $P$  (i.e. large  $x_0$ ). Without importance sampling, i.e. setting  $w = 0$ , the number of trajectories needed grows exponentially in  $x_0^2$ :  $N_0^* \sim \exp(x_0^2/2\tau)$  (dominant contribution). However, using importance sampling, with the optimal wind value ( $w_{opt} = x_0/\tau$ ), the number needed grows linearly with  $x_0$ :  $N_w^* \sim x_0$ . Thus,

$$N_w^* \sim \sqrt{\ln N_0^*}. \quad (151)$$

Even when  $N_0^*$  is tremendously large,  $N_w^*$  may remain modest (for  $w = w_{opt}$ ). For instance, for  $x_0 = 6$  and  $\tau = 1$  we have  $P[x(\tau) > x_0] = .99 \times 10^{-9}$ . One would need  $\sim 10^{12}$  direct simulations to compute this probability to 10% accuracy. Using importance sampling, one can achieve the same accuracy with fewer than 700 trajectories. Numerical results (with  $N = 700$  and  $w = 6.0$ ) gave us  $P = (1.04 \pm 0.10) \times 10^{-9}$ .



**Fig 3.15** Efficiency gain as a function of wind strength, Eq.148, for  $x_0 = 3$  ( $P = 1.35 \times 10^{-3}$ ). At the optimal wind value, the gain is around 220.



**Fig 3.16** Same as Fig.3.15, but for  $x_0 = 0$ . Here,  $P = 0.5$ , so there would be no problem in estimating this probability from direct simulations. Nevertheless, there is an efficiency gain of nearly two, when using importance sampling.

## IV. Calculations

### Introduction

In this Chapter we present some calculated results obtained from the model described in Chapter II with including fluctuations and using the ISM for some cases. The using of ISM is inevitable for small enough probabilities.

#### Units and constants.

In our calculations we used the following units of measure:

- \* time  $\bar{s} = 10^{-22} s$ ,
- \* length fm =  $10^{-13} \text{cm}$ ,
- \* energy MeV,
- \* cross-section b =  $10^{-24} \text{cm}^2 = 100 \text{fm}^2$ .

And constants in these units:

- \*  $m = 1.036435 \text{ MeV } \bar{s}^2 / \text{fm}^2$ ,
- \*  $\hbar = 6.58222 \text{ MeV } \bar{s}$ ,
- \*  $e^2 = 1.439978 \text{ MeV fm}$ ,
- \*  $r_0 = 1.18 \text{ fm}$ ,

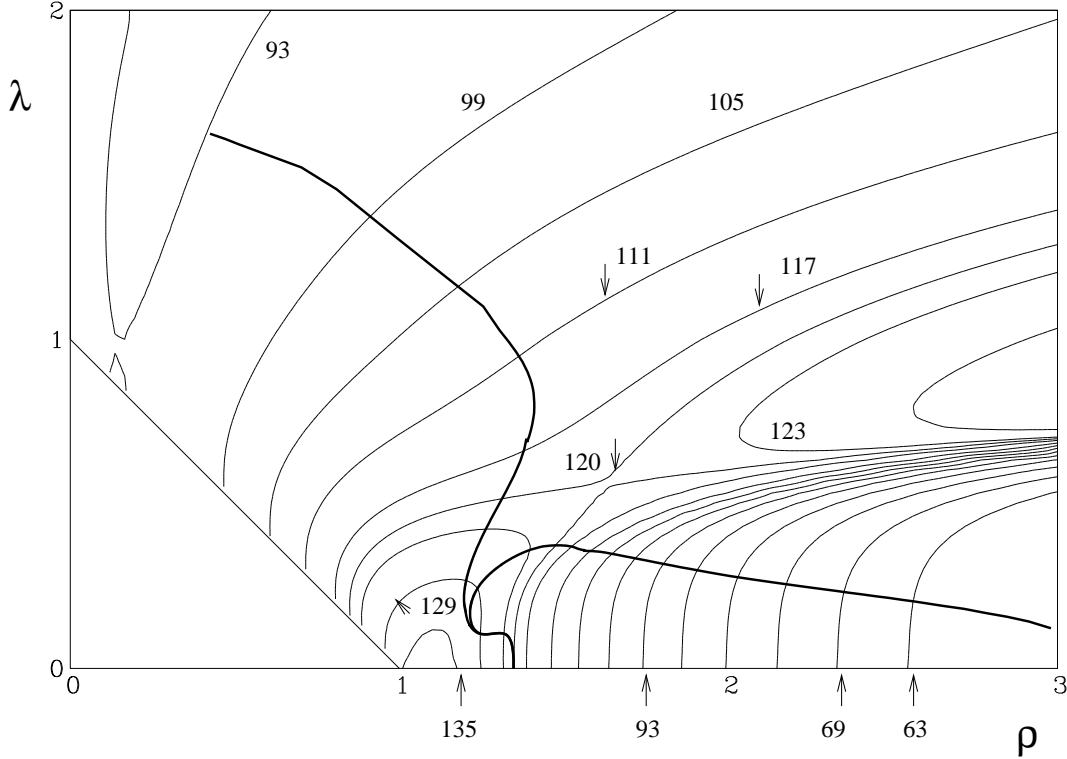
$m$  - average mass of nucleon,  $e$  - electron charge and  $r_0$  - radius of nucleon taken from Folding Model [60].

### 1. Fusion probabilities

Including Langevin dynamics into deterministic model allows us to calculate probabilities of fusion of heavy ions for given energies and impact parameters. Direct simulations correspond to numerical solving of Eq.57, i.e. launching a number of trajectories and counting those leading to fusion.

To present typical trajectories we have chosen the  $^{86}\text{Kr} + ^{70}\text{Ge}$  reaction. In Fig.4.1 a contour plot map of the potential energy with shell corrections included is presented. The energy is a function of  $\rho$  and  $\lambda$  at fixed values of the mass asymmetry  $\Delta$  and the charge asymmetry  $\Delta_Z$  equal to 0. Two deterministic trajectories projected on  $(\rho, \lambda)$  space for central collision ( $L = 0$ ) are shown. Both trajectories start from the same point in the configurational space but with different initial kinetic energies (in the center of mass system  $E_{\text{cm}} = 132 \text{ MeV}$  and  $134 \text{ MeV}$ ). As we see the trajectory with bigger energy goes to fusion while the other one reseparates. Fluctuations bring to our dynamics indeterminism

in a sense that for the same initial conditions for a given reaction trajectories can either go to fusion or to reseparation. This is shown in Fig.4.2. Four trajectories with included thermal fluctuations start with the same kinetic energy ( $E_{\text{cm}} = 132$  MeV) and one of them goes to fusion. One could say then that the probability of fusion would be in this case equal to  $1/4$ , whereas in the deterministic calculations (Fig.4.1) it is exactly equal to 0.



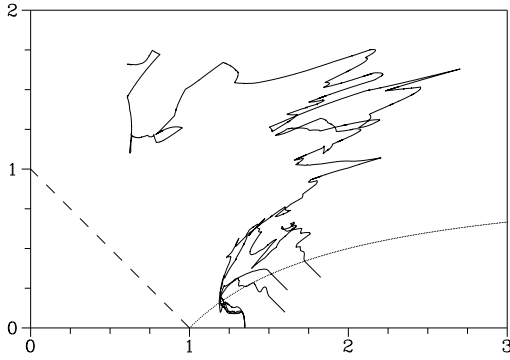
**Fig 4.1** Two deterministic trajectories on the contour plot potential map for the  $^{86}\text{Kr}+^{70}\text{Ge}$  reaction. The potential energy is shown in MeV with respect to separated nuclei in  $(\rho, \lambda)$ -space. The trajectories start with initial kinetic energy 132 and 134 MeV in the center of the mass system.

In Chapter II we considered three types of fluctuations: 1) thermal fluctuations; 2) asymmetry fluctuations; and 3) dissipative fluctuations. To compare the influence of these types of fluctuations we again take the same reaction  $^{86}\text{Kr}+^{70}\text{Ge}$  and investigate probabilities for fusion separately for each type in the energy region near “extra-push”. Three curves of excitation function are presented in Fig.4.3. Dotted curve corresponds to the asymmetry, dashed one to the dissipative and solid one to the thermal fluctuations. From this figure we learn that effects of asymmetry and dissipative fluctuations are rather small in comparison to thermal fluctuations — the width of excitation function corresponding to thermal fluctuations is much bigger than of the others.

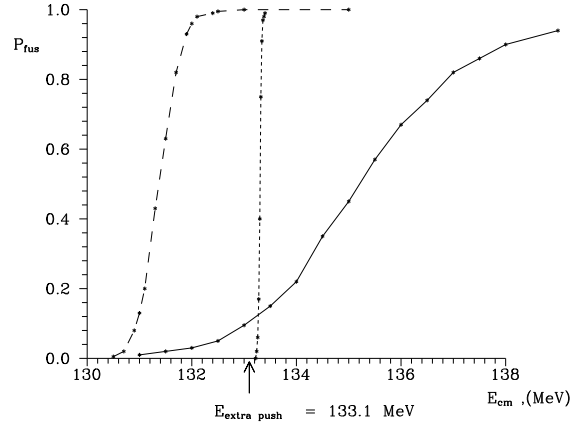
We have used the model for calculations of the near-threshold fusion probabilities for the  $^{86}\text{Kr}+^{136}\text{Xe}$  reaction studied recently at GSI Darmstadt by Christelle Stodel *et al* [104]. Cross sections for different channels were measured for this system at the near-



threshold energies. From the combined compound-residue cross sections, the dependence of the fusion probability on the excitation energy of the compound nucleus was deduced by the authors of Ref. [104], see Fig.4.4. The deduced fusion probability increases from  $P_{\text{fus}} \approx 2 \times 10^{-4}$  at  $E_{\text{cm}} = 195$  MeV to  $P_{\text{fus}} \approx 1$  at  $E_{\text{cm}} = 230$  MeV. This is an especially interesting reaction from the point of view of shell effects as in both colliding nuclei there are closed neutron shells. Therefore one can expect to see clearly the role of the shell effects in this case. And indeed in the deterministic calculations the value of the extra-push energy is equal to 224.2 MeV in calculations with shell effects included and 228.7 MeV in calculations without shells. This should be compared to experimental value of 224 MeV [106], the energy at which fusion probability drops down to  $1/2$ . In our calculations we attempted to check whether the rise of the fusion probability, observed already at quite low energies, can be explained in terms of the standard fusion dynamics with fluctuations determined by the dissipation-fluctuation theorem. In Fig.4.4 results of our calculations for the  $^{86}\text{Kr}+^{136}\text{Xe}$  reaction are compared with the experimental data. By using the importance sampling method, we were able to carry out conclusive calculations even at the lowest energies where the fusion probabilities drop down to few times  $10^{-5}$ , and thus to verify the near-threshold part of the excitation function. Results of the calculations show that the observed near-threshold fusion probabilities in the  $^{86}\text{Kr}+^{136}\text{Xe}$  reaction cannot be explained with the standard fusion dynamics model based on one-body dissipation. Clearly, fusion extends to lower energies than expected by the theoretical calculations.

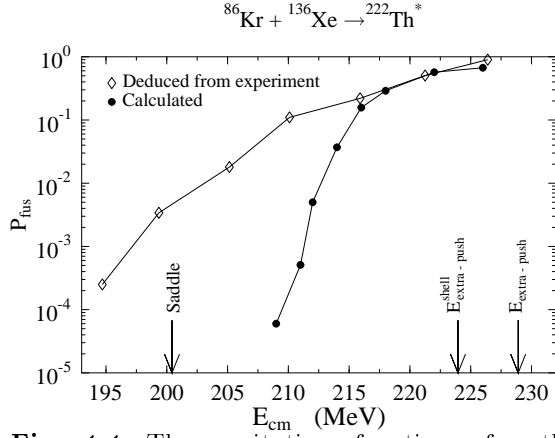


**Fig 4.2** Stochastic trajectories for the  $^{86}\text{Kr}+^{70}\text{Ge}$  reaction. Four trajectories with included thermal fluctuations start with the same kinetic energy 132 MeV and one of them goes to fusion.

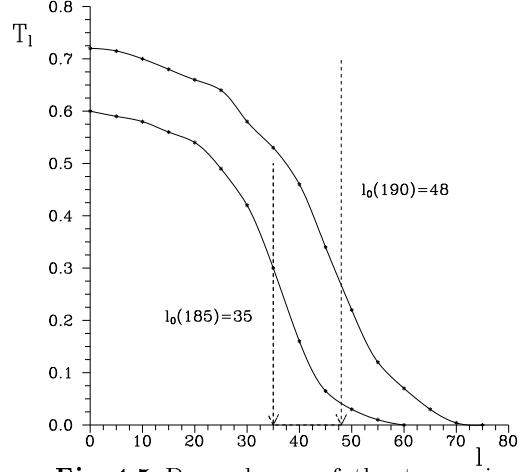


**Fig 4.3** The excitation functions for the  $^{86}\text{Kr}+^{70}\text{Ge}$  reaction for different kind of fluctuations. Dotted curve corresponds to the asymmetry, dashed one to the dissipative and solid one to the thermal fluctuations.

On the other hand independent on dissipation it is very hard to understand, on the basis of the present model, experimental excitation function extending to energies almost 6 MeV lower than the saddle point position (Fig.4.4). If one would forget dissipation and calculate the fusion probability as a probability of penetrating the saddle point barrier it would cut down this fusion probability by many orders of magnitude and the point at 195 MeV would be out of the picture.



**Fig 4.4** The excitation functions for the  $^{86}\text{Kr} + ^{136}\text{Xe}$  reaction.



**Fig 4.5** Dependence of the transmission coefficient  $T_l$  on angular momentum reaction. The upper curve is calculated for the colliding energy  $E_{cm} = 190$  MeV, and the lower one for 185 MeV. The dashed arrows show critical

## 2. Cross sections

The total fusion cross section at the kinetic energy  $E$  can be written as a sum over all partial waves:

$$\sigma_E = \pi \lambda^2 \sum (2l + 1) T_l = \sum \frac{d\sigma(l)}{dl} \quad (152)$$

where  $T_l$  is the transmission coefficient through the fusion barrier with angular momentum  $L = \hbar l$ , and the spin distribution (partial fusion cross section with respect to the orbital angular momentum) is defined as

$$d\sigma(l)/dl = \pi \lambda^2 (2l + 1) T_l \quad (153)$$

Here  $\lambda$  is the de Broglie wavelength:

$$\lambda^2 = \frac{\hbar^2}{2\mu E_{cm}}; \quad \mu = \frac{mA_1A_2}{A_1 + A_2} \quad (154)$$

In order to describe fusion of heavy ions well above the barrier, classical trajectory models including frictional forces but neglecting statistical fluctuations have been frequently used, e.g. Ref. [101–103] One applies a sharp cut-off model for the transmission coefficient

$$T_l = \begin{cases} 1, & \text{for } l < l_0, \\ 0, & \text{for } l > l_0, \end{cases} \quad (155)$$

i.e. fusion occurs for  $l < l_0$ , and the fusion excitation function reads

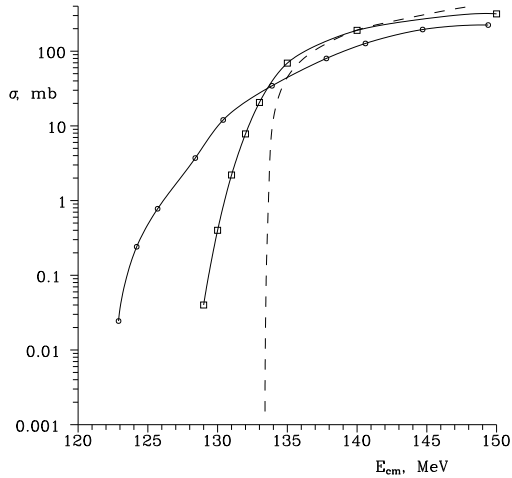
$$\sigma_E = \pi \lambda^2 \sum_{l=0}^{l_0} (2l + 1) = \frac{\pi \hbar}{2\mu E} (l_0(E) + 1)^2. \quad (156)$$

This formula has been applied in calculating fusion excitation functions throughout the periodic table with the surface friction model [102]. The critical  $l$ -value for fusion  $l_0$  was determined for each center-of-mass energy  $E$  by solving dynamical equations without the fluctuating forces; i.e. in this approximation the mean trajectory determines the fusion cross section. Technically,  $l_0$  is determined by running trajectories for each energy by diminishing the initial  $l$ -value, starting with a large one, until a trajectory with a particular  $l$ -value is captured behind the barrier.

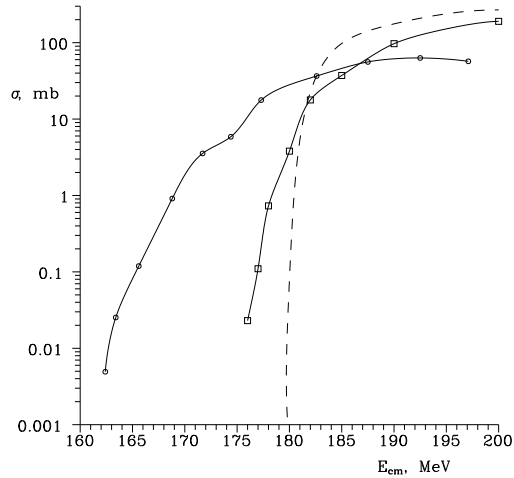
The corresponding spin distributions are of the triangular form:

$$\frac{d\sigma(l)}{dl} = \begin{cases} \pi\lambda^2(2l+1), & \text{for } l < l_0, \\ 0, & \text{for } l > l_0. \end{cases} \quad (157)$$

If statistical fluctuations are taken into account we obtain smooth dependence of the transmission coefficient (probability for fusion)  $T_l$  on angular momentum instead of the step function in deterministic case, Eq.155. This is shown in Fig.4.5 for the  $^{86}\text{Kr} + ^{104}\text{Ru}$  reaction. Two functions for energies 185 MeV and 190 MeV are presented; the dotted lines indicate values of  $l_0$  in the sharp cut-off calculations without fluctuations.



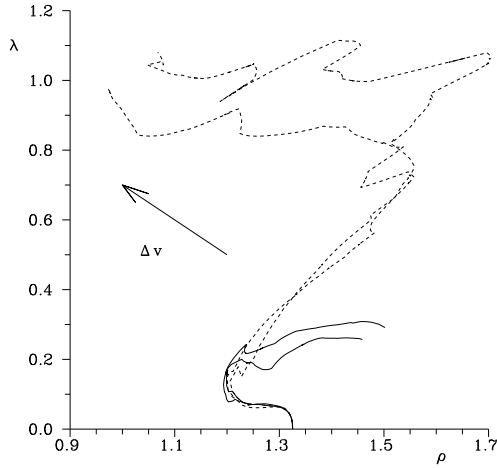
**Fig 4.6** The excitation functions for the the  $^{86}\text{Kr} + ^{70}\text{Ge}$  reaction. The curve lying on the circles represents the experimental data [105], and squares show our calculations. The dashed line correspond to the excitation function obtained by classical calculations without statistical fluctuations (It goes to 0 in approaching to the extra-push).



**Fig 4.7** The excitation functions for the the  $^{86}\text{Kr} + ^{104}\text{Ru}$  reaction. See the description for Fig 4.6

Two reactions  $^{86}\text{Kr} + ^{70}\text{Ge}$  and  $^{86}\text{Kr} + ^{104}\text{Ru}$  have been taken to calculate cross sections. Experimental data for these reactions were obtained at GSI, Darmstadt [105]. In Figs. 4.6 and 4.7 there are shown excitation functions. Dashed lines represent excitation function corresponding to deterministic calculations, (Eq.156). It is obvious that these

functions are limited by extra-push energy from below, while calculations with statistical fluctuations included extend the excitation function to lower energies.



**Fig 4.8** Stochastic trajectories for the the excitation functions for the  $^{86}\text{Kr} + ^{104}\text{Ru}$  reaction with initial kinetic energy 175 MeV. Two solid lines represent two trajectories leading to reseparation. Dashed lines correspond to trajectories with additional non-physical force  $\Delta v = 1$  included. This force is acting in  $(-1, 1)$  direction and squeezed to "stochastic freedom", see Eq.142.

$\Delta v$	$N$	$N_{\text{fus}}$	$P_{\text{fus}} \cdot 10^3$	Error $\cdot 10^3$
0 (Direct)	1000	0	-	-
0.2	1000	0	-	-
0.6	1000	37	0.4343	0.095
1.0	1000	280	0.4347	0.037
1.4	1000	696	0.4294	0.024
1.8	1000	910	0.4731	0.039
2.2	1000	973	0.5258	0.130

**Table 1** Calculation of probabilities using the importance sampling method.  $N$  is the number of simulated trajectories;  $N_{\text{fus}}$  - number of trajectories gone to fusion;  $P_{\text{fus}}$  - deduced probability.

### 3. Calculation of small probabilities

In calculating excitation function for lowest energies directly the number of trajectories which one should run in order to get statistically significant result for fusion probability grows exponentially. In this way the direct method becomes at some point completely impractical. Therefore in the excitation functions shown in Figs. 4.4, 4.6 and 4.7 points where cross sections are very small were calculated with an ISM method. As have been shown in the analytically solvable model (section III.6) number of trajectories needed to be run using ISM grows as a power function (see Eq.151). We have this type of the dependence when the additional force is chosen close to be an optimal one. Generally from one hand this additional force should be chosen in a way to increase probability of the trajectory to go toward fusion as much as possible. On the other hand this force should not disturb too much the distribution of trajectories, because this increases the width of the biasing function (Eq.116, 116b), and in this way it lowers a statistics.

In our calculation the direction of this additional force was always chosen in a way to lead to the compound system. In Fig.4.8 dynamical trajectories in case of  $^{86}\text{Kr} + ^{104}\text{Ru}$  reaction at  $E_{\text{cm}} = 175$  MeV are presented. Two solid trajectories correspond to calculations without any additional force, whereas the other two dashed trajectories correspond

to calculations with an additional force  $\Delta v = 1$ . In this way one can see how adding an additional force is changing the dynamical behaviour of the system going from reseparation to fusion.

In Table 1 results of the fusion probability for the same case as in Fig.4.8 as a function of the amplitude of the additional force is presented. One can see that in the region of  $\Delta v = (0.6, 1.4)$  we obtain stable results for the fusion probability with error bars being the lowest for  $\Delta v = 1.4$ . In this way one could say that the value  $\Delta v = 1.4$  is an optimal one in this case.

## V. Summary and conclusions

We have shown how one can introduce fluctuations into a deterministic description of Heavy Ion Dynamics. This is absolutely unavoidable in case one wants to reproduce experimentally observed distributions and not only their average values.

Chapter II is devoted to the description of the deterministic model of Heavy Ion Collisions with stochastic effects included. It consists of two parts: in the first one we have described deterministic model of HIC, and in the second one we elaborated statistical fluctuations. A Langevin description of HI dynamics, suitable for implementation in numerical simulations, is easily derived. This approach can be studied by starting with a deterministic model of HI collisions, then adding the stochastic forces, and finally simulating the resulting dynamics numerically. From these simulations one is able to obtain an excitation functions for fusion – experimentally observed quantities, which can then be compared directly with experimental data.

The derivation of Langevin description of nuclear dynamics is not the end of the story. Statistical fluctuations play an important role specifically in process for which the interesting outcome – e.g. fusion of heavy elements – occurs with very low probability. In such situations, direct numerical simulations of the process do not provide a realistic method of attacking the problem.

In connection to that in Chapter III there are presented new methods we have developed of getting around this problem obtaining very encouraging results. One approach makes use of a one-to-one correspondence between the statistical distributions of Langevin trajectories, and the thermal distributions of directed polymers. This analogy provides a powerful formalism for analyzing relative probabilities of Langevin trajectories. We have tested this method on both simple schematic but analytically solvable model, and the simplified model of HIC [2]. This method has turned out to be very powerful in a sense that the calculational effort does not depend on how small are the probabilities we want to calculate. However it demands a big number of sampling. And we did not use it for calculations with a realistic model as due to the complicated dynamics of this model the process of sampling takes a lot of the computer time.

The other method (called by us ISM) involves adding, in our numerical Langevin simulations, an extra, non-physical force acting on the nuclear collective degrees of freedom. This force is chosen so as to increase the probability for fusion by orders of magnitude, enabling us to obtain good statistics from a reasonable number of simulations; then we analytically correct for the inclusion of this unphysical force, thus finally obtaining the desired cross section. We have tested first this method in different models and have ob-

tained fusion probabilities in situations where they are extremely small. We have found that we could get good statistics with far fewer simulations that is needed to accurately extract the fusion probability by direct Langevin simulation. We were encouraged by the results we obtained from simplified model of nuclear collisions. So we used this method to calculate probabilities (cross sections) within more realistic model.

Results of the calculations for three different reactions:  $^{86}\text{Kr}+^{104}\text{Ru}$ ,  $^{86}\text{Kr}+^{70}\text{Ge}$  and  $^{86}\text{Kr}+^{136}\text{Xe}$  are presented in Chapter IV. Comparison with an experimental data shows that the model is working properly in the region of high probabilities of fusion but is falling down too fast for the subbarrier region.

In here we have presented ideas one can follow in many fields. As far as the process of fusion is concerned it is possible now to try to make estimates for the cross sections in the exotic superheavy region.

*We have shown how one can introduce fluctuations into a deterministic description of heavy ion dynamics. At the same time we have proposed methods for calculating small probability events within a reasonable statistics and not too much of the computational effort. Those methods are quite general and can be applied not only in nuclear physics but also in other fields of physics where one is dealing with a very small but still measurable cross sections.*

## VI. Appendix: Solvable models

At the time we were developing a model of nuclear dynamics with fluctuations some problem appeared which are not directly related to nuclear physics but strictly speaking to statistical physics. The description of those problems was added to this thesis.

### 1. Introduction

In this chapter we present three different models. The first is Feynman's ratchet and pawl discrete simplified model. It operates between two heat reservoirs. We solve exactly for the steady-state directed motion and heat flows produced, first in the absence and then in the presence of an external load. We show that the model can act both as a heat engine and as a refrigerator. We finally investigate the behavior of the system near equilibrium, and use our model to confirm general predictions based on linear response theory.

In recent years, a number of theoretical results pertaining to systems far from thermal equilibrium have been derived. While not identical to one another, these results bear a similar structure, and the term *fluctuation theorem* (FT) has come to refer to them collectively. The rest two highly idealized, but exactly solvable, models are shown mainly for applying the fluctuation theorem. One involves a piston inside an infinite cylinder, surrounded by two ideal gases at different temperatures. The other involves a particle dragged through a thermal environment by a uniformly translating harmonic force. For these models we are interested in the entropy production distribution function. We show that the solutions satisfy the relations of the fluctuation theorem.

### 2. Feynman's ratchet and pawl

**Introduction.** Feynman's ratchet and pawl system [70] is a well known (but not the earliest! [71]) example of a proposed “mechanical Maxwell's demon”, a device whose purpose is to convert into useful work the thermal motions present in a heat reservoir. The idea is beautifully simple: set up a ratchet and pawl so that a wheel is allowed to turn only in one direction, and then attach that wheel to a windmill whose vanes are surrounded by a gas at a finite temperature; see Fig.46-1 of Ref. [70]. Every so often, an accumulation of collisions of the gas molecules against the vanes will cause the wheel to rotate by one notch in the allowed direction, but presumably never in the forbidden direction. Such rectification of thermal noise could be harnessed to perform useful work



(such as lifting an flea against gravity), in direct violation of the Second Law. Of course, in order for statistical fluctuations to cause rotation at a perceptible rate, the ratchet and pawl must be microscopic, and this points to the resolution of the paradox. If thermal motions of the gas molecules are sufficient to cause the wheel to rotate a notch, then the thermal motion of the pawl itself will occasionally cause it to disengage from the ratchet, at which point the wheel could move in the “forbidden” direction. Feynman compared the rates of the two processes – rotation in the allowed and the forbidden directions – and found them to be equal when the system is maintained at a single temperature. Thus no net rotation arises, and the Second Law is saved.

Since the failure of the ratchet and pawl system to perform work arises from thermal fluctuations of the pawl, a natural solution to the problem is to reduce these fluctuations by externally cooling the pawl to a temperature below that of the gas. In this case the device does indeed operate as designed, but this no longer constitutes a violation of the Second Law: the ratchet and pawl is now effectively a microscopic heat engine, capitalizing on a temperature difference to extract useful work from thermal motions.

While the ratchet and pawl was introduced in Feynman's Lectures primarily for pedagogical purposes, recent years have seen a renewed interest in this system [72–86], largely due to the fact that analogous mechanisms have been proposed as simple models of motor proteins.

Our purpose in this section is to introduce an exactly solvable model of Feynman's microscopic heat engine. This model is discrete rather than continuous<sup>16</sup>, but it captures two essential features of the original example: (1) a periodic but asymmetric interaction potential between the ratchet and the pawl (corresponding to an asymmetric sawtooth shape for the ratchet's teeth), and (2) two “modes” of interaction (corresponding to the pawl being either engaged or disengaged from the ratchet). These features are sufficient for the model to reproduce the behavior discussed by Feynman. A different discrete model of noise-induced transport has been studied by Schimansky-Geier, Kschischo, and Fricke [84]; however, we believe ours to be the first in which (as in Feynman's example) the transport is explicitly driven by a temperature difference between two reservoirs.

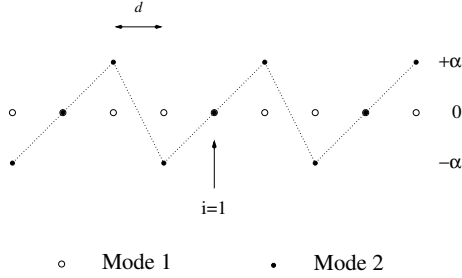
**Zero external load.** We begin this section by introducing our model in the absence of an external load, then pointing out the analogy with Feynman's ratchet and pawl. Following that, we solve the model.

Consider a particle which moves via discrete jumps between neighboring sites along a one-dimensional regular lattice, where  $d$  is the lattice spacing. We assume that the particle is coupled to a heat reservoir at temperature  $T_B$ , and that its jumps are thermal in nature. That is, the probability (per unit time) of making a jump to site  $i + 1$ , starting from site  $i$ , is related to the probability rate of the inverse process by the usual detailed balance relation:  $P_{i \rightarrow i+1}/P_{i+1 \rightarrow i} = \exp(-\Delta E/T_B)$ , where  $\Delta E = U_{i+1}^{(m)} - U_i^{(m)}$  is the instantaneous change in the particle's potential energy, associated with the jump from  $i$  to  $i + 1$ . We use

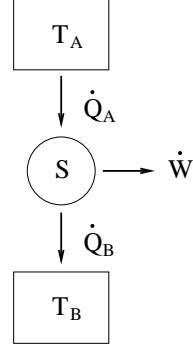
---

<sup>16</sup>Note that a particle (or, more generally, a reaction coordinate), evolving from one sufficiently deep potential (or free) energy minimum to another, behaves much as if hopping from one site to another on a discrete lattice. See, for instance, Fig.6 of Ref. [83].

the notation  $U_i^{(m)}$  to denote the potential energy of the particle at site  $i$ ; the superscript  $m$  denotes the “mode” of the potential, to be explained momentarily. The integer index  $i$  runs from  $-\infty$  to  $+\infty$ , and we are using units in which Boltzmann’s constant  $k_B = 1$ .



**Fig A.1** The potential energy  $U_i^{(m)}$  is shown for both modes,  $m = 1, 2$ , in the absence of an external load; the lattice spacing is  $d$ , and site 1 is labeled explicitly.



**Fig A.2** A schematic representation of our system ( $S$ ) in contact with two heat reservoirs (at temperatures  $T_A$  and  $T_B$ ). As implied by the arrows, we define  $\dot{Q}_A$  to be the net flow of heat from reservoir  $A$  to the system, and  $\dot{Q}_B$  to be the heat flow from the system to reservoir  $B$ . Therefore  $\dot{W} = \dot{Q}_A - \dot{Q}_B$  is the power delivered as work against an external load. (First we consider that there is no such load, hence  $\dot{Q}_A = \dot{Q}_B$ .)

Next, we assume that the potential energy function  $U_i^{(m)}$  has two possible *modes*,  $m = 1, 2$ , and that it changes stochastically between these two. In the first mode, the potential energy of each site is zero:  $U_i^{(1)} = 0$ . In the second mode, the potential energy is periodic, with period 3:  $U_i^{(2)} = \alpha \cdot [(i \bmod 3) - 1]$ , where  $\alpha$  is a positive constant with units of energy. As shown in Fig.A.1, the second mode is a discrete version of an asymmetric sawtooth potential. We assume that the stochastic process governing the changes between modes is also a thermal process, driven by a heat reservoir at temperature  $T_A$ . Thus, the probability rate of a change to mode 2, starting from mode 1, relative to the probability rate of the reverse mode change, is given by the detailed balance factor  $\exp(-\Delta E/T_A)$ , where  $\Delta E = U_i^{(2)} - U_i^{(1)}$  is now the (site-dependent) change in particle energy due to an instantaneous change from mode 1 to mode 2.

We now describe more precisely these two stochastic processes, the one governing the jumps of the particle, the other governing the change between modes. First, we assume that the processes are independent of one another, and that each is a Poisson process occurring at a rate  $\Gamma$ . That is, during every infinitesimal time interval  $\delta t$ , there is a probability  $\Gamma \delta t$  that the particle will attempt a jump to a neighboring site. An “attempt” looks as follows. First, the particle decides (randomly, with equal probability) whether to try jumping to the left ( $-d$ ) or to the right ( $+d$ ). Then the Metropolis algorithm [87] is used to satisfy detailed balance: if the value of  $\Delta E$  associated with the jump

is zero or negative, then the jump takes place; if  $\Delta E > 0$ , then the jump occurs with probability  $\exp(-\Delta E/T_B)$ . Similarly, during every infinitesimal time interval  $\delta t$ , there is a probability  $\Gamma \delta t$  that the mode will attempt to change. Again, the attempt is either accepted or rejected according to the Metropolis algorithm (at temperature  $T_A$ ).


We have introduced three parameters which we will view as being “internal” to the system:  $d$ ,  $\alpha$ , and  $\Gamma$ ; these essentially set the length ( $d$ ), energy ( $\alpha$ ), and time ( $\Gamma^{-1}$ ) scales relevant for the system. The two remaining parameters entering the model,  $T_A$  and  $T_B$ , we will view as “external”.

The analogy between our model and Feynman's ratchet and pawl runs as follows. First, the position of our particle corresponds to an angle variable,  $\theta$ , specifying the orientation of the ratchet wheel. When the particle accomplishes a net displacement of three lattice sites, either to the right or to the left, this is equivalent to the ratchet being displaced by one notch, or tooth<sup>17</sup>; see Fig.A.1. Since we want to keep track of the wheel over long intervals of time, possibly including many full rotations,  $\theta$  varies from  $-\infty$  to  $+\infty$ , rather than being a periodic variable from 0 to  $2\pi$ ; this is why the variable  $i$ , specifying the lattice site, takes on all integer values.

When a ratchet and pawl are “engaged” – that is, when the pawl actually presses against the teeth of the ratchet – then there exists an interaction energy, arising from the compression of the spring which holds the pawl against the ratchet, which has the form of a periodic sawtooth potential in the angle variable  $\theta$ . In our discrete model, the analogue of this interaction energy is the discrete sawtooth potential  $U_i^{(2)}$  described above; mode 2 thus corresponds to the situation in which the ratchet presses against the pawl. By contrast, mode 1 corresponds to the ratchet and pawl begin “disengaged”, as will occur every so often as a result of thermal fluctuations of the pawl. Of course, in Feynman's system, the potential energy of the disengaged mode is always greater than that of the engaged mode (due to the spring compression needed to actually place the pawl out of reach of the ratchet's teeth), whereas in our model this is not the case:  $U_i^{(1)} = 0$ . This, however, does not change the problem in any qualitatively significant way.

As mentioned, the motion of the particle from site to site is a thermal process occurring at temperature  $T_B$ , while the stochastic “flashing” between modes 1 and 2 is a thermal process occurring at temperature  $T_A$ . Thus, in the context of the analogy with the physical ratchet and pawl,  $T_B$  denotes the temperature of the gas surrounding the panes connected to the ratchet wheel, and  $T_A$  is the temperature at which the pawl is maintained.

Given this model, we would like to know whether a current arises: in the long run, will there be a net drift of the particle – corresponding to a non-zero angular velocity of the wheel in the original ratchet and pawl – and if so, what will be the rate and direction of the drift?

Our system of two coupled spins has six possible states, which we number with an index  $n$ , . The dynamics of the particle is described by stochastic jumps among these six states. This is a Markov process, and we can describe its statistics with a set

---

<sup>17</sup>Thus, the distance  $3d$  in our model translates to an angular interval  $\Delta\theta = 2\pi/N_{teeth}$ , where  $N_{teeth}$  is the number of teeth along the perimeter of the ratchet wheel.

of rate equations. Let  $p_n(t)$  denote the probability that the system is found in state  $n$  at time  $t$ . Then the evolution of these probabilities is governed by the following set of coupled equations:

$$\frac{dp_n}{dt} = \Gamma \sum_{n' \neq n} \left[ p_{n'} P(n' \rightarrow n) - p_n P(n \rightarrow n') \right] \quad , \quad n = 1, \dots, 6. \quad (158)$$

Here,  $\Gamma \cdot P(n \rightarrow n')$  is the probability rate at which the system, when its state is  $n$ , makes a transition to state  $n'$ .

Our rate equations can be expressed as:

$$\boxed{\frac{d\mathbf{p}}{dt} = \Gamma \mathcal{R} \mathbf{p}}, \quad (159)$$

where  $\mathbf{p} = (p_1, \dots, p_6)^T$ ,  $\mathcal{R}$  is the  $6 \times 6$  matrix

$$\mathcal{R} = \begin{pmatrix} -2 & \frac{1}{2} & \frac{1}{2} & \mu & 0 & 0 \\ \frac{1}{2} & -2 & \frac{1}{2} & 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & -1 - \mu & 0 & 0 & 1 \\ 1 & 0 & 0 & -\mu - \frac{\nu + \nu^2}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{\nu}{2} & -\frac{3 + \nu}{2} & \frac{1}{2} \\ 0 & 0 & \mu & \frac{\nu^2}{2} & \frac{\nu}{2} & -2 \end{pmatrix}, \quad (160)$$

and we have introduced the constants

$$\mu = e^{-\alpha/T_A} \quad \text{and} \quad \nu = e^{-\alpha/T_B}. \quad (161)$$

Note that  $\mu$  and  $\nu$  are monotonically increasing functions of  $T_A$  and  $T_B$ , and the temperature range  $0 < T_A$  ( $T_B$ )  $< \infty$  translates to  $0 < \mu$  ( $\nu$ )  $< 1$ . We thus think of  $\mu$  and  $\nu$  as “rescaled” temperatures.

As a consistency check, note that the elements in each column of  $\mathcal{R}$  add to zero:  $\sum_m \mathcal{R}_{mn} = 0$ . Thus,  $\mathcal{R}$  annihilates the row vector  $\mathbf{1}^T = (1, 1, 1, 1, 1, 1)$  by left multiplication:  $\mathbf{1}^T \mathcal{R} = \mathbf{0}^T$ , which means that total probability is conserved:

$$\frac{d}{dt} \sum_n p_n = \frac{d}{dt} (\mathbf{1}^T \cdot \mathbf{p}) = \mathbf{1}^T \cdot (\Gamma \mathcal{R} \mathbf{p}) = 0. \quad (162)$$

The long-time behavior of our system is governed by the steady-state distribution of probabilities,  $\bar{\mathbf{p}}$ , which is the null eigenvector of  $\mathcal{R}$ :

$$\mathcal{R} \bar{\mathbf{p}} = 0. \quad (163)$$

Determining  $\bar{\mathbf{p}}$  is a straightforward exercise in Jordan elimination, and leads to the following result:  $\bar{\mathbf{p}} = \mathbf{x}/N$ , where

$$x_1 = 52\mu + 28\mu^2 + 12\nu + 19\nu^2 + 5\nu^3 + 21\mu\nu + 2\mu\nu^2 + 8\mu^2\nu \quad (164)$$

$$x_2 = 36\mu + 16\mu^2 + 28\nu + 27\nu^2 + 5\nu^3 + 25\mu\nu + 8\mu\nu^2 + 2\mu^2\nu \quad (165)$$

$$x_3 = 44\mu + 19\mu\nu + 20\nu + 49\nu^2 + 15\nu^3 \quad (166)$$

$$x_4 = 64 + 20\nu + 48\mu + 15\mu\nu \quad (167)$$

$$x_5 = 24\mu + 40\nu + 20\nu^2 + 18\mu^2 + 30\mu\nu + 15\mu\nu^2 \quad (168)$$

$$x_6 = 22\mu^2 + 16\mu\nu + 44\mu\nu^2 + 14\mu^2\nu + 15\mu\nu^3 + 26\nu^2 + 10\nu^3, \quad (169)$$

and  $N(\mu, \nu) = \sum_{i=1}^6 x_i$  is a normalization factor.

When both temperatures go to zero,  $\mu, \nu \rightarrow 0$ , we get  $\bar{\mathbf{p}}^T = (0, 0, 0, 1, 0, 0)$ . This makes sense: in that limit, the system freezes to the state of lowest energy, which is state 4.

We now address the question of net drift. We first define a net current,  $J \equiv J_+ - J_-$ , where  $J_+$  is the rate at which particle drift from states 2 and 5 to 3 and 6, and  $J_-$  is the rate of the reverse transitions. By “rate”, we mean number of passes per unit time, averaged over an infinitely long interval of time. The current  $J$  represents the *net* average rate of particle drift. In the length units:

$$v = 3 d J, \quad (170)$$

where  $v$  denotes the (steady-state) average velocity of the particle.

Explicit expression for the quantities  $J_{\pm}$  are given by:

$$J_+ = \Gamma (\bar{p}_2 \mathcal{R}_{32} + \bar{p}_5 \mathcal{R}_{65}) \quad (171)$$

$$J_- = \Gamma (\bar{p}_3 \mathcal{R}_{23} + \bar{p}_6 \mathcal{R}_{56}). \quad (172)$$

(Recall that  $\Gamma \mathcal{R}_{mn}$  is the transition rate to state  $m$ , given that the system is found in state  $n$ ; therefore  $\bar{p}_n \Gamma \mathcal{R}_{mn}$  is the net rate at which transitions from  $n$  to  $m$  are observed to occur in the steady state.) Using our results for  $\bar{\mathbf{p}}$ , we compute  $J_{\pm}$ , take the difference to get  $J$ , and finally multiply by  $3 d$  to get, after some algebra:

$$\boxed{v(T_A, T_B) = -3 d \frac{\Gamma}{N} (\mu - \nu)(1 - \nu)(3\mu + 4).} \quad (173)$$

There are a number of things to note about this result. First, it implies that if  $T_A > T_B$  (i.e.  $\mu > \nu$ ), then there is a net flow of the particle to the left; if  $T_B > T_A$ , the particle drifts to the right. If the temperatures are equal, then there is no drift, in agreement with Feynman's analysis (as well as the Second Law!).

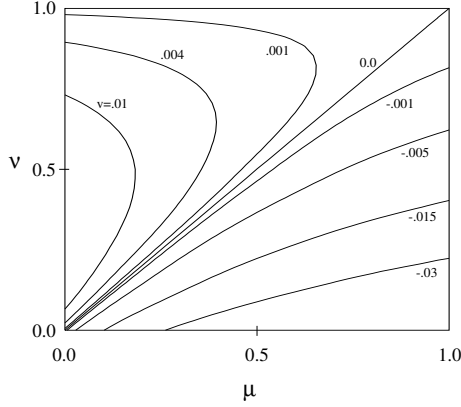
Next, notice that  $v \rightarrow 0$  as  $T_B \rightarrow \infty$  ( $\nu \rightarrow 1$ ). In that limit, the change in energy arising from a jump to the left or to the right becomes negligible in comparison to the temperature of the reservoir which drives those jumps; thus, from any lattice site, the particle is as likely to jump to the left as to the right, resulting in no net drift.

Finally, in the limit  $T_A \rightarrow \infty$  ( $\mu \rightarrow 1$ ), we get:

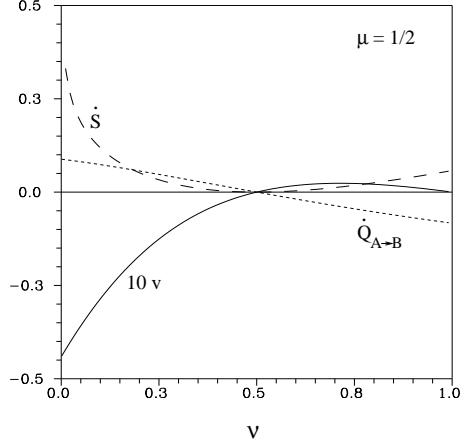
$$v = -21 d \frac{\Gamma}{N} (1 - \nu)^2, \quad T_A \rightarrow \infty. \quad (174)$$

This is the limit in which changes between the modes occur independently of location of the particle: every attempt to change the mode is accepted. This is analogous to the

situation studied by Astumian and Bier [73], where the “flashing” between the two modes of the potential is simply a Poisson process at a fixed rate, independent of the particle position.



**Fig A.3** A contour plot of the function  $v(\mu, \nu)$ , where  $\mu = \exp(-\alpha/T_A)$  and  $\nu = \exp(-\alpha/T_B)$  are the rescaled temperatures of the two reservoirs.



**Fig A.4** The drift  $v$  (multiplied by 10), heat flow  $\dot{Q}_{A \rightarrow B}$ , and rate of entropy production  $\dot{S}$ , plotted as functions of  $\nu$ , for fixed  $\mu = 1/2$ .

We can also use our expression for the steady-state probability distribution to compute the average rate at which heat is transferred between the two reservoirs and our system. Whenever the system makes a transition from state 1 to state 4, or from state 6 to state 3, its energy drops by  $\alpha$ ; this energy is released into the reservoir coupled to spin  $A$  (at temperature  $T_A$ ). Conversely, during the transitions  $4 \rightarrow 1$  and  $3 \rightarrow 6$ , the system absorbs energy  $\alpha$  from reservoir  $A$ . Since no other transitions involve a transfer of energy between that reservoir and the system, we can express the net rate at which the system absorbs energy from reservoir  $A$  as:

$$\dot{Q}_A = \alpha \Gamma (\bar{p}_4 \mathcal{R}_{14} + \bar{p}_3 \mathcal{R}_{63} - \bar{p}_1 \mathcal{R}_{41} - \bar{p}_6 \mathcal{R}_{36}). \quad (175)$$

We can similarly write down an expression for  $\dot{Q}_B$ , the rate at which heat flows from our system to reservoir  $B$ :

$$\dot{Q}_B = \alpha \Gamma (\bar{p}_5 \mathcal{R}_{45} + \bar{p}_6 \mathcal{R}_{56} + 2\bar{p}_6 \mathcal{R}_{46} - \bar{p}_4 \mathcal{R}_{54} - \bar{p}_5 \mathcal{R}_{65} - 2\bar{p}_4 \mathcal{R}_{64}). \quad (176)$$

Fig.A.2 illustrates the sign convention which we choose in defining  $\dot{Q}_A$  and  $\dot{Q}_B$ . After plugging in the values for the components of  $\bar{\mathbf{p}}$  and  $\mathcal{R}$ , we find that  $\dot{Q}_A = \dot{Q}_B$ , as we could have predicted, since in the steady state there is no net absorption of heat by the system, nor is any of the heat delivered as work against an external load. Thus, the particle drift is ultimately driven by a net passage of heat from  $A$  to  $B$ , by way of the system. The explicit expression for this heat flow is given by:

$$\dot{Q}_{A \rightarrow B} = \dot{Q}_A = \dot{Q}_B = 3 \frac{\alpha \Gamma}{N} (\mu - \nu) P(\mu, \nu), \quad (177)$$

where  $P(\mu, \nu) = 4 + 14\mu + 15\nu + 4\mu\nu + 5\nu^2 > 0$ .

The factor  $(\mu - \nu)$  in Eq.177 guarantees that the direction of the heat flow is from the hotter to the cooler reservoir (or not at all if  $T_A = T_B$ ). Since the same factor appears in Eq.173, and since all other factors in these equations are positive, these two equations show explicitly that the direction of the flow of heat determines the direction of the particle motion: when the heat flows from  $A$  to  $B$ , the particle drifts to the left; when it flows from  $B$  to  $A$ , the particle drifts rightward. The ratio  $v/\dot{Q}_{A \rightarrow B}$  then gives us the average displacement of the particle, per unit of heat passed (via the system) from reservoir  $A$  to reservoir  $B$ :

$$\lim_{\tau \rightarrow \infty} \frac{\Delta x}{\Delta Q} = \frac{v}{\dot{Q}_{A \rightarrow B}} = -\frac{d}{\alpha} \cdot \frac{(1 - \nu)(3\mu + 4)}{P(\mu, \nu)} < 0. \quad (178)$$

Here,  $\Delta x$  and  $\Delta Q$  are the net particle displacement, and the net heat transferred from  $A$  to  $B$ , respectively, over a time interval  $\tau$ .

We can also compute the rate at which entropy is produced during this process. The rate of entropy production associated with the flow of heat from reservoir  $A$  to the system is:  $\dot{S}_A = -\dot{Q}_A/T_A$ ; and for reservoir  $B$ :  $\dot{S}_B = \dot{Q}_B/T_B$ . The net entropy production rate is thus:

$$\dot{S} = \dot{S}_A + \dot{S}_B = \frac{T_A - T_B}{T_A T_B} \dot{Q}_{A \rightarrow B} \quad (179)$$

$$= 3 \frac{\Gamma}{N} \left( \ln \frac{\mu}{\nu} \right) (\mu - \nu) P(\mu, \nu) \geq 0. \quad (180)$$

Figs.A.3 and A.4 illustrate the results obtained in this subsection. Fig.A.3 is a contour plot of the drift  $v$ , as a function of the rescaled reservoir temperatures  $\mu$  and  $\nu$ . The contour  $v = 0$  runs along the diagonal,  $\mu = \nu$ , as well as along  $\nu = 1$  ( $T_B \rightarrow \infty$ ). The appearance of positive contours ( $v > 0$ ) to the left of the diagonal, and negative ones to the right, illustrates the point that the drift is rightward when  $T_B > T_A$  and leftward when  $T_A > T_B$ .

In Fig.A.4 we have fixed the value of  $T_A$  by setting  $\mu = 1/2$ , and have plotted  $v$ ,  $\dot{Q}_{A \rightarrow B}$ , and  $\dot{S}$  as functions of  $\nu$ . All three quantities hit zero at  $\nu = 1/2$ , where  $T_A = T_B$ : nothing interesting happens when the system is maintained at a single temperature. Note also that  $v$  and  $\dot{Q}_{A \rightarrow B}$  are opposite in sign (in agreement with Eq.178), while  $\dot{S}$  is always nonnegative (in agreement with the Second Law). Finally, note that  $v \rightarrow 0$  as  $\nu \rightarrow 1$  ( $T_B \rightarrow \infty$ ).

In plotting these two figures, we set all the internal parameters to unity:  $\alpha = d = \Gamma = 1$ .

**Nonzero external load.** In this subsection we add an external load to our simple model. In Feynman's example, this load is an flea, attached by a thread to the ratchet wheel: when the wheel rotates in the appropriate direction, the creature is lifted against gravity. In our model, we add a slope to the discrete potential:

$$U_i^{(m)} \rightarrow U_i^{(m)} + i f d, \quad (181)$$

where  $f$  is a real constant (and  $i$  is still the lattice site variable, *not*  $\sqrt{-1}$ ). Effectively,  $f$  is a constant external force which pulls the particle leftward if  $f > 0$ , rightward if  $f < 0$ .

The presence of an external load allows the system to perform *work*. If, in the steady state achieved for fixed  $\mathbf{x} = (f, T_A, T_B)$ , the particle experiences a drift  $v(\mathbf{x})$ , then the *power* (work per unit time) delivered against the external load is:

$$\dot{W}(\mathbf{x}) = f v(\mathbf{x}). \quad (182)$$

By conservation of energy, this must be balanced by heat lost by the reservoirs:

$$\dot{W} = \dot{Q}_A - \dot{Q}_B. \quad (183)$$

Our approach to solving for the steady-state behavior is the same as in the previous subsection. We are interested in the quantities  $v(\mathbf{x})$ ,  $\dot{Q}_A(\mathbf{x})$ , and  $\dot{Q}_B(\mathbf{x})$  which describe that behavior. As before, we map our model onto a two-spin, six-state system, whose statistical evolution is governed by the rate equations  $d\mathbf{p}/dt = \Gamma \mathcal{R} \mathbf{p}$  (Eq.159). There is then an associated steady state  $\bar{\mathbf{p}}$ , and the quantities in which we are interested are given, as in the previous subsection, by:

$$v = 3 d \Gamma (\bar{p}_2 \mathcal{R}_{32} + \bar{p}_5 \mathcal{R}_{65} - \bar{p}_3 \mathcal{R}_{23} - \bar{p}_6 \mathcal{R}_{56}) \quad (184a)$$

$$\dot{Q}_A = \alpha \Gamma (\bar{p}_4 \mathcal{R}_{14} + \bar{p}_3 \mathcal{R}_{63} - \bar{p}_1 \mathcal{R}_{41} - \bar{p}_6 \mathcal{R}_{36}) \quad (184b)$$

$$\dot{Q}_B = \alpha \Gamma (\bar{p}_5 \mathcal{R}_{45} + \bar{p}_6 \mathcal{R}_{56} + 2\bar{p}_6 \mathcal{R}_{46} - \bar{p}_4 \mathcal{R}_{54} - \bar{p}_5 \mathcal{R}_{65} - 2\bar{p}_4 \mathcal{R}_{64}). \quad (184c)$$

The only difference is that the presence of the term *ifd* in the potential changes the elements of  $\mathcal{R}$ , and therefore the vector of steady-state probabilities  $\bar{\mathbf{p}}$ .

Because the acceptance probability of an attempted move in the Metropolis scheme has the form  $\text{Prob.} = \min\{1, \exp -\Delta E/T\}$ , there is no general analytic expression for the matrix  $\mathcal{R}$  valid for all values of  $f$ . Instead, the elements of  $\mathcal{R}$  are piecewise analytic in  $f$ ; the analytic expression for  $\mathcal{R}_{ij}$  depends on the sign of the transition energy  $\Delta E_{j \rightarrow i}$ . A little thought reveals that the  $f$ -axis can be divided into four ranges of values, over each of which the elements of  $\mathcal{R}$  can be written as analytic functions of  $\alpha$ ,  $d$ ,  $f$ ,  $T_A$ , and  $T_B$ . These ranges are:

$$-\infty < f < -\alpha/d \quad (185a)$$

$$-\alpha/d < f < 0 \quad (185b)$$

$$0 < f < 2\alpha/d \quad (185c)$$

$$2\alpha/d < f < +\infty. \quad (185d)$$

In the two extreme ranges, (a) and (d), stretching to  $-\infty$  and  $+\infty$ , the slope is so steep that the potential energy function no longer has a sawtooth shape in mode 2. We will ignore these ranges and focus instead on the ones for which  $U_i^{(2)}$  *does* look like a discrete sawtooth.

For range (b), i.e.  $-\alpha/d < f < 0$ , the potential slopes downward with increasing  $i$ , and an explicit expression for  $\mathcal{R}$  is:



$$\mathcal{R}^b = \begin{pmatrix} -\frac{3}{2} - \frac{1}{2\sigma} & \frac{1}{2\sigma} & \frac{1}{2} & \mu & 0 & 0 \\ \frac{1}{2} & -\frac{3}{2} - \frac{1}{2\sigma} & \frac{1}{2\sigma} & 0 & 1 & 0 \\ \frac{1}{2\sigma} & \frac{1}{2} & -\frac{1}{2} - \mu - \frac{1}{2\sigma} & 0 & 0 & 1 \\ 1 & 0 & 0 & -\mu - \frac{\nu^2}{2\sigma} - \frac{\nu\sigma}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{\nu\sigma}{2} & -\frac{3}{2} - \frac{\nu\sigma}{2} & \frac{1}{2} \\ 0 & 0 & \mu & \frac{\nu^2}{2\sigma} & \frac{\nu\sigma}{2} & -2 \end{pmatrix}, \quad (186)$$

where

$$\sigma = e^{-fd/T_B}. \quad (187)$$

For range (c),  $0 \leq f < 2\alpha/d$ , the potential slopes upward with  $i$ , and we have

$$\mathcal{R}^c = \begin{pmatrix} -\frac{3}{2} - \frac{\sigma}{2} & \frac{1}{2} & \frac{\sigma}{2} & \mu & 0 & 0 \\ \frac{\sigma}{2} & -\frac{3}{2} - \frac{\sigma}{2} & \frac{1}{2} & 0 & 1 & 0 \\ \frac{1}{2} & \frac{\sigma}{2} & -\frac{1}{2} - \mu - \frac{\sigma}{2} & 0 & 0 & 1 \\ 1 & 0 & 0 & -\mu - \frac{\nu^2}{2\sigma} - \frac{\nu\sigma}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{\nu\sigma}{2} & -\frac{3}{2} - \frac{\nu\sigma}{2} & \frac{1}{2} \\ 0 & 0 & \mu & \frac{\nu^2}{2\sigma} & \frac{\nu\sigma}{2} & -2 \end{pmatrix}. \quad (188)$$

Note that both  $\mathcal{R}^b$  and  $\mathcal{R}^c$  reduce to the matrix  $\mathcal{R}$  of the previous subsection, when  $\sigma = 1$ , i.e.  $f = 0$ .

As in the previous subsection, the first order of business is to solve for the steady-state distribution of probabilities,  $\bar{\mathbf{p}}$ . This is again an exercise in Jordan elimination, only now the process is considerably more tedious: terms do not cancel as nicely as when  $f = 0$ . The final results for the steady state probability vectors  $\bar{\mathbf{p}}^b$  and  $\bar{\mathbf{p}}^c$  (corresponding to the two ranges of  $f$  values) are of the form

$$\bar{p}_n^j = \frac{P_n^j(\mu, \nu, \sigma)}{N^j(\mu, \nu, \sigma)} \quad , \quad j = b, c, \quad n = 1 \cdots 6, \quad (189)$$

where the  $P_n^j$ 's are finite polynomials in the variables  $\mu$ ,  $\nu$ , and  $\sigma$ , and  $N^j = \sum_{n=1}^6 P_n^j$  is a normalization factor. Explicit expression for the polynomials  $P_n^j$  are given in the last subsection of this section.

We can now use Eq.184 to obtain  $v(\mathbf{x})$ ,  $\dot{Q}_A(\mathbf{x})$ , and  $\dot{Q}_B(\mathbf{x})$ . The results for the two ranges of  $f$  values,  $j = b, c$ , are:

$$v^j(\mathbf{x}) = \frac{3}{2}d\Gamma \frac{X^j}{\sigma N^j} \quad (190a)$$

$$\dot{Q}_A^j(\mathbf{x}) = \alpha\Gamma \frac{Y^j}{N^j} \quad (190b)$$

$$\dot{Q}_B^j(\mathbf{x}) = \alpha\Gamma \frac{Y^j}{N^j} - \frac{3}{2}fd\Gamma \frac{X^j}{\sigma N^j} \quad , \quad (190c)$$

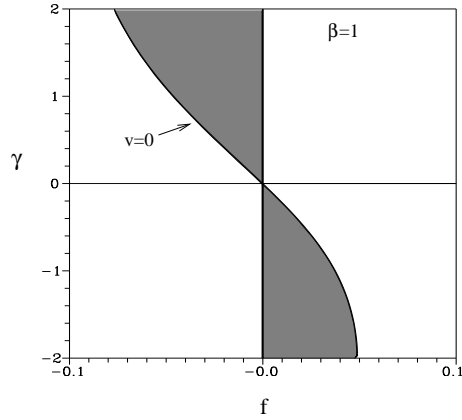
where  $X^j(\mu, \nu, \sigma)$  and  $Y^j(\mu, \nu, \sigma)$  are polynomials for which explicit expressions are presented in the last subsection of this section.

Note that the steady-state behavior described by Eq.190 clearly satisfies energy conservation (see Eqs.182 and 183):

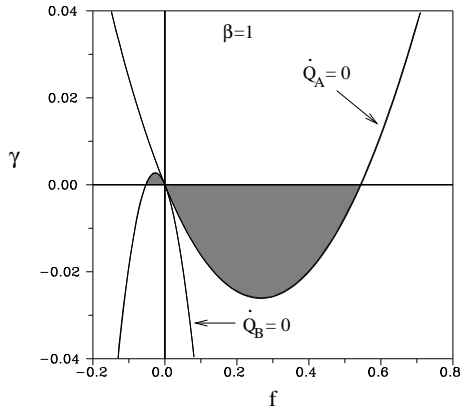
$$\dot{Q}_A - \dot{Q}_B = f v. \quad (191)$$

Since  $v$ ,  $\dot{Q}_A$ , and  $\dot{Q}_B$  are not independent (i.e. they satisfy the above relation), the “space of steady-state behaviors” is effectively *two-dimensional*.

Eq.190 is the central result of this section. For arbitrary temperatures  $T_A$  and  $T_B$ , and for any external load  $f$  such that the interaction potential retains its sawtooth shape ( $-\alpha/d < f < 2\alpha/d$ ), Eq.190 gives the directed motion and heat flows which describe the steady-state behavior of the system for those external parameter values. These results reduce to those of the previous subsection when we set  $f = 0$ . In the next subsection, we use the above results to show that our model can act both as a heat engine and as a refrigerator. In the following subsection, we consider the behavior of our system near equilibrium, and we use Eq.190 to confirm predictions based on a general, linear response analysis.



**Fig A.5** The contour  $v(f, \gamma) = 0$  is shown, for fixed average inverse temperature  $\beta = 1$ . The shaded regions are those for which  $f v > 0$ , i.e. for which the system behaves as a heat engine.



**Fig A.6** The contours  $\dot{Q}_A(f, \gamma) = 0$  and  $\dot{Q}_B(f, \gamma) = 0$  are shown, for fixed  $\beta = 1$ . In the shaded regions, there is a net flow of heat out of the colder reservoir; the system then acts as a refrigerator.

**The system as heat engine and refrigerator.** We can anticipate two different scenarios in which our system acts as a “useful” device:

(1)  $\dot{W} > 0$ . In this case, the system is a *heat engine*, causing the particle to drift up the potential energy slope, with efficiency  $\eta_{\text{eng}} = \dot{W}/\dot{Q}_>$ , where  $\dot{Q}_>$  is the rate of heat flow out of the hotter reservoir.

(2)  $\dot{W} < 0$  and  $\dot{Q}_< > 0$ , where  $\dot{Q}_<$  is the rate of heat flow out of the colder reservoir. Here the system is a *refrigerator*, with efficiency  $\eta_{\text{ref}} = \dot{Q}_</|\dot{W}|$ . The particle drifts down the potential slope, and the resulting energy liberated allows for a net transfer of heat from the colder to the hotter reservoir, without violating the Second Law.

We will now use the results derived in the previous subsection to show that our simple model indeed exhibits both these behaviors.

The system is a heat engine when  $v(f, T_A, T_B)$  and  $f$  are of the same sign; see Eq.182. To demonstrate that our model exhibits this behavior, let us introduce the variables

$$\beta = \frac{T_A + T_B}{2T_A T_B} \quad (192)$$

$$\gamma = \frac{T_A - T_B}{T_A T_B}, \quad (193)$$

and consider the behavior of our system in  $(f, \gamma)$ -space, for a fixed value of  $\beta$ . That is, we hold fixed the *average* inverse temperature of the reservoirs ( $\beta$ ), and consider the operation of the system in terms of the external load ( $f$ ) and the *difference* between inverse temperatures of the reservoirs ( $\gamma$ ). (We will use these variables again in the following subsection. Furthermore, because  $\dot{Q}_A$ ,  $\dot{Q}_B$ , and  $v$  are mutually dependent – see Eq.191 and the comment following it – we can generically explore a measurable fraction of the two-dimensional space of steady-state behaviors by varying only two independent external parameters,  $f$  and  $\gamma$ , while holding the third,  $\beta$ , fixed.)

In Fig.A.5 we plot the contour  $v(f, \gamma) = 0$ , having set  $\beta = 1$ , and with the internal parameter values  $\alpha = d = \Gamma = 1$ . (The range of  $\gamma$  values for this choice of  $\beta$  is  $-2 < \gamma < 2$ .) To the left of this contour, we have  $v > 0$ ; to the right,  $v < 0$ . The shaded region thus represents the values of  $(f, \gamma)$  for which  $v$  and  $f$  are of the same sign, i.e. where the system behaves as a heat engine.

We can understand the placement of the shaded region as follows. Consider a point  $P$  on the positive  $\gamma$ -axis:  $f = 0$ ,  $\gamma > 0$ . This corresponds to  $T_A > T_B$ , with no external load. We know that the particle then drifts leftward,  $v < 0$ , although no work is performed, since  $f = 0$ . Let us now imagine that we tilt the potential slightly “downward” ( $f < 0$ ). For a small enough tilt, we expect that the particle will continue to drift to the left, but now this drift is uphill, and therefore work is done against the external load:  $\dot{W} > 0$ . We conclude that points immediately to the left of the positive vertical axis will correspond to external parameters for which the system behaves as a heat engine. Fig.A.5 confirms this: the shaded region hugs the positive  $\gamma$ -axis from the left. (Similar reasoning applies for the negative  $\gamma$ -axis, where the region  $\dot{W} > 0$  appears to the right.) If we now continue to tilt the slope more and more downward ( $f$  increasingly negative), at fixed  $\gamma > 0$ , we expect the leftward drift to become progressively slower, until for some tilt we get  $v = 0$ . At this point the leftward “thermal force” exerted on the particle due to the temperature difference between the reservoirs, exactly balances the external load. This occurs at the boundary of the shaded region (the contour  $v = 0$ ): for more negative slopes, the particle slides down the slope, and the system no longer acts as a heat engine.

Our system is a refrigerator when  $\dot{Q}_< > 0$ , where  $\dot{Q}_<$  is the rate at which heat leaves the colder of the two reservoirs. In Fig.A.6 we plot the contours  $\dot{Q}_A(f, \gamma) = 0$  and  $\dot{Q}_B(f, \gamma) = 0$ , again for  $\beta = 1$ . These two contours are tangent at the origin, and divide the plane of  $(f, \gamma)$ -values as follows:  $\dot{Q}_A > 0$  for points lying above the contour  $\dot{Q}_A = 0$ , and  $\dot{Q}_A < 0$  for points lying below that contour. Similarly,  $\dot{Q}_B > 0$  ( $\dot{Q}_B < 0$ ) for points lying above (below) the contour  $\dot{Q}_B = 0$ . Now, above the horizontal axis ( $\gamma > 0$ ) we have  $T_A > T_B$ , hence  $\dot{Q}_< = -\dot{Q}_B$ . The small shaded region in the second quadrant therefore

represents values of  $(f, \gamma)$  for which reservoir  $B$  is the colder of the two reservoirs, *and* it is losing heat (since  $\dot{Q}_B < 0$ ). Hence in this region our system acts as a refrigerator. Below the horizontal axis,  $T_B > T_A$  and thus  $\dot{Q}_< = \dot{Q}_A$ . The larger shaded region in the fourth quadrant thus also represents refrigeration, only now reservoir  $A$  is the being drained of heat.

We can understand the general shape of the shaded regions by assuming that, when the temperatures are equal ( $\gamma = 0$ ) and the slope is very small, the quantities  $v$ ,  $\dot{Q}_A$ , and  $\dot{Q}_B$  are linear functions of the slope:

$$v(f, 0) = c_v f + O(f^2) \quad (194a)$$

$$\dot{Q}_A(f, 0) = c_A f + O(f^2) \quad (194b)$$

$$\dot{Q}_B(f, 0) = c_B f + O(f^2), \quad (194c)$$

with  $c_v, c_A, c_B \neq 0$ . (Then  $c_v < 0$ , since the particle cannot slide *up* the slope when the reservoir temperatures are equal.) Energy conservation implies that the difference between  $\dot{Q}_A$  and  $\dot{Q}_B$  must be quadratic in  $f$  (since  $\dot{Q}_A - \dot{Q}_B = \dot{W} = f v = c_v f^2$ ), hence

$$c_A = c_B. \quad (195)$$

Thus, for equal temperatures and a sufficiently small slope, one of the reservoirs will be losing heat and the other will be gaining it. If we now slightly lower the temperature of the reservoir which is losing heat, then we have a refrigerator: heat flows out of the colder reservoir. In our model, we have  $c \equiv c_A = c_B > 0$ , hence for points on the  $\gamma = 0$  axis immediately to the right of the origin, heat flows out of reservoir  $A$  and into reservoir  $B$ ; immediately to the left of the origin the reverse holds true. This explains why, just to the right (left) of the origin, the shaded region corresponding to refrigeration hugs the horizontal axis from below (above). If we continue to the right along the line  $\gamma = 0$ , increasing the value of  $f$ , then the particle will drift ever more rapidly to the left as the slope becomes ever more steeply inclined. The potential energy lost as the particle slides down the incline gets dissipated into the reservoirs; for sufficiently large  $f$ , the rate of dissipation is great enough that both reservoirs become heated:  $\dot{Q}_A < 0$ ,  $\dot{Q}_B > 0$ . This happens to the right of the point at which the contour  $\dot{Q}_A = 0$  crosses the horizontal axis with a positive slope; for values of  $f$  beyond this point the system can no longer operate as a refrigerator.

We can also understand why the two contours  $\dot{Q}_A = 0$  and  $\dot{Q}_B = 0$  “kiss” at the origin. The result  $c_A = c_B$  means that  $\dot{Q}_A$  and  $\dot{Q}_B$  are (to leading order) equal along the horizontal axis  $\gamma = 0$ , near the origin. However, they are also (exactly) equal along the vertical axis, since  $\dot{W} = 0$  when  $f = 0$ . Thus,  $\dot{Q}_A$  and  $\dot{Q}_B$  are equal, to leading order in  $f$  and  $\gamma$ , for a small region around the origin:  $\dot{Q}_A = \dot{Q}_B = c f + b \gamma$ . This implies that their contours are both tangent to the line  $\gamma = -c f / b$  at the origin.

Since we have explicit expressions for  $v(\mathbf{x})$ ,  $\dot{Q}_A(\mathbf{x})$ , and  $\dot{Q}_B(\mathbf{x})$ , we can analytically compute the thermal efficiency  $\eta$  associated with the steady-state operation of our model, when it acts as either a heat engine or a refrigerator. By the Second Law, these efficiencies must never exceed the Carnot efficiencies,  $\eta_{\text{eng}}^C$  and  $\eta_{\text{ref}}^C$  (which depend only on  $T_A$  and  $T_B$ ). Ideally, we could use our exact results to find the maximum *relative efficiency* ( $y = \eta / \eta^C$ ) which our model achieves, both when it acts as a heat engine and when it

acts as a refrigerator. Unfortunately, the expressions for  $v$ ,  $\dot{Q}_A$ , and  $\dot{Q}_B$  are sufficiently complicated (see the last subsection of this section) that we are not able to find these maxima analytically. However, at the end of the next subsection, we will present analytical results for the maximum relative efficiencies achieved when the system operates *close to equilibrium*.

**Linear Response.** When there is no external load and the reservoir temperatures are equal ( $T_A = T_B = \beta^{-1}$ ), our system is in thermal equilibrium, and there results no average particle drift or heat flow:  $v = \dot{Q}_A = \dot{Q}_B = 0$ . In the previous subsection we briefly considered the behavior of our system *near* equilibrium (see e.g. Eq.194). We now consider this case in more detail. We will present a general analysis essentially the same as that of Jülicher, Ajdari, and Prost [81], and then we will show that the exact results obtained for our model confirm the predictions of this analysis.

For a sufficiently small load  $f$  and inverse temperature difference  $\gamma$ , and a fixed value of  $\beta$  (characterizing the inverse temperature of the equilibrium state around which we expand), we expect to be in the *linear response* regime: the particle drift and heat flows depend linearly on  $f$  and  $\gamma$ . Let us introduce the quantity

$$\Phi = \frac{1}{2}(\dot{Q}_A + \dot{Q}_B) \quad (196)$$

– roughly, a “heat flux” from  $A$  to  $B$  – and let us write, to leading order in  $f$ ,  $\gamma$ :

$$\begin{pmatrix} v \\ \Phi \end{pmatrix} = \begin{pmatrix} \partial v / \partial f & \partial v / \partial \gamma \\ \partial \Phi / \partial f & \partial \Phi / \partial \gamma \end{pmatrix} \begin{pmatrix} f \\ \gamma \end{pmatrix}, \quad (197)$$

with the derivatives of  $v$  and  $\Phi$  – re-expressed as functions of  $(f, \gamma, \beta)$  rather than  $(f, T_A, T_B)$  – evaluated at equilibrium ( $f = \gamma = 0$ ). As per the arguments given at the end of the previous subsection,  $\dot{Q}_A$  and  $\dot{Q}_B$  are equal, to leading order in  $f$  and  $\gamma$ , near equilibrium.

The rate of entropy production is then:

$$\begin{aligned} \dot{S} &= -\frac{\dot{Q}_A}{T_A} + \frac{\dot{Q}_B}{T_B} \\ &= -\beta \dot{W} + \gamma \Phi \\ &= (f \quad \gamma) \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} f \\ \gamma \end{pmatrix}, \end{aligned} \quad (198)$$

where  $M_{11} = -\beta \partial v / \partial f$ ,  $M_{12} = -\beta \partial v / \partial \gamma$ ,  $M_{21} = \partial \Phi / \partial f$ , and  $M_{22} = \partial \Phi / \partial \gamma$ , evaluated at equilibrium. The Second Law implies that

$$\det \mathbf{M} \geq 0, \quad (199)$$

whereas Onsager's reciprocity relation [88] predicts that  $M_{12} = M_{21}$ , or

$$-\beta \frac{\partial v}{\partial \gamma} \Big|_{\text{eq}} = \frac{\partial \Phi}{\partial f} \Big|_{\text{eq}}. \quad (200)$$

Also, the diagonal elements of  $\mathbf{M}$  must be positive: the particle must slide *down* the potential slope when the temperatures are equal ( $M_{11} > 0$ ), and there must be a flow of heat from the hotter to the colder reservoir when the slope is zero ( $M_{22} > 0$ ). Jülicher *et al.* [81] have obtained identical results for a molecular motor driven by a difference in chemical potential rather than temperature.

Using the exact results obtained in nonzero external load subsection, we differentiate  $v$  and  $\Phi$  with respect to  $f$  and  $\gamma$  to evaluate the elements of the matrix  $\mathbf{M}$  for our model<sup>18</sup>:

$$\mathbf{M} = \begin{pmatrix} 3\beta^2 d^2 C(3 + 4\zeta) & \alpha\beta dC(1 - \zeta) \\ \alpha\beta dC(1 - \zeta) & \alpha^2 C[4 + \zeta(29 + \zeta)]/(4 + 3\zeta) \end{pmatrix}, \quad (201)$$

where

$$\zeta = e^{-\alpha\beta} \quad (0 < \zeta < 1) \quad , \quad C = \frac{3\zeta\Gamma}{(16 + 5\zeta)[1 + \zeta(4 + \zeta)]} > 0. \quad (202)$$

The variable  $\zeta$  is akin to  $\mu$  and  $\nu$ : it is the “rescaled temperature” of the equilibrium state with respect to which the linear response behavior is defined. The determinant of this matrix is:

$$\det \mathbf{M} = \frac{9(\alpha\beta\zeta d\Gamma)^2(2 + 19\zeta + 21\zeta^2)}{(4 + 3\zeta)(16 + 5\zeta)[1 + \zeta(4 + \zeta)]^2}. \quad (203)$$

We see by inspection that  $M_{11}, M_{22} > 0$ ; that  $M_{12} = M_{21}$ , as mandated by Onsager reciprocity (Eq.200); and that  $\det \mathbf{M} > 0$ , in agreement with the Second Law (Eq.199).

It is interesting to consider the operation of our system as a heat engine and refrigerator, in the linear response regime. In this regime, the conditions for these two behaviors are:  $vf > 0$  for a heat engine (as before), and  $\gamma\Phi < 0$  for a refrigerator (since  $\Phi = \dot{Q}_A = \dot{Q}_B$ , to leading order in  $f, \gamma$ ). In Fig.A.7, the shaded regions indicate values of  $(f, \gamma)$  for which the system acts as a heat engine or refrigerator (compare with Fig.[2] of Ref. [81].) The two diagonal lines which form boundaries of these regions are given by  $v = 0$  and  $\Phi = 0$ . The slopes of these lines are:

$$\lambda_{v=0} = -\frac{M_{11}}{M_{12}} \quad , \quad \lambda_{\Phi=0} = -\frac{M_{21}}{M_{22}}. \quad (204)$$

The Second Law, by requiring that  $\det \mathbf{M} \geq 0$ , guarantees that these shaded regions do not overlap:  $|\lambda_{v=0}| \geq |\lambda_{\Phi=0}|$ .

The *efficiency* of our system, when operating as a heat engine, is given by:

$$\eta_{\text{eng}} = \frac{\dot{W}}{|\Phi|} \quad (205)$$

---

<sup>18</sup>Recall that the expressions for  $v(\mathbf{x})$ ,  $\dot{Q}_A(\mathbf{x})$ , and  $\dot{Q}_B(\mathbf{x})$  differ according to the sign of  $f$ . We have verified that, regardless of whether we use the results valid for range (b) ( $f < 0$ ), or those for range (c) ( $f > 0$ ), we obtain the same results for the elements of  $\mathbf{M}$ .

(again using  $\Phi = \dot{Q}_A = \dot{Q}_B$  to leading order). The *Carnot efficiency* defined for the temperatures  $T_A$  and  $T_B$  is:

$$\eta_{\text{eng}}^C = \frac{|T_A - T_B|}{T_{>}} = \frac{|\gamma|}{\beta} + O(\gamma^2). \quad (206)$$

Then we can get an explicit expression for the relative efficiency ( $y = \eta/\eta^C$ ) of our system, in the linear response regime:

$$y_{\text{eng}} \equiv \frac{\eta_{\text{eng}}}{\eta_{\text{eng}}^C} = -\frac{1}{\lambda} \frac{M_{11} + M_{12}\lambda}{M_{21} + M_{22}\lambda}, \quad (207)$$

where  $\lambda = \gamma/f$ . The relative efficiency is the same for all points along any straight line through the origin, and Eq.207 gives that relative efficiency as a function of the slope  $\lambda$  of the line. A similar analysis holds for the case of refrigeration:

$$y_{\text{ref}} \equiv \frac{\eta_{\text{ref}}}{\eta_{\text{ref}}^C} = -\lambda \frac{M_{21} + M_{22}\lambda}{M_{11} + M_{12}\lambda}. \quad (208)$$

Note that  $y_{\text{ref}}$  happens to be the inverse of  $y_{\text{eng}}$ , although the two expressions are valid for different ranges of  $\lambda$  values, corresponding to the shaded regions in Fig.A.7.

The results of the previous two paragraphs were derived with the implicit assumption that  $M_{12}, M_{21} > 0$ . This happens to be true for our model, but in general these elements can be either positive or negative (or zero), so long as they are equal. If  $M_{12}$  and  $M_{21}$  were negative, then the shaded regions would occur in the the first and third quadrants of the  $(f, \gamma)$  plane, and the negative signs would not appear in Eqs.207 and 208.

The above results imply that, near equilibrium, a microscopic device operating between two reservoirs either can act *both* as a heat engine and as a refrigerator (if  $M_{12} = M_{21} \neq 0$ ), or will exhibit neither behavior (if  $M_{12} = M_{21} = 0$ ). For instance, in a microscopic ratchet-and-pawl device, if the sawteeth on the wheel have a symmetric shape, then the system cannot behave as a heat engine:  $M_{12} = 0$ , as is obvious by symmetry. What is not so immediately obvious, but follows from the condition  $M_{12} = M_{21}$ , is that it is equally impossible for a system with symmetric teeth to operate as a refrigerator, in the linear response regime.

Finally, for a given inverse temperature  $\beta$  of the equilibrium state, we can solve for the *maximal* relative efficiency achievable in the linear response regime, by maximizing  $y_{\text{eng}}$  and  $y_{\text{ref}}$  with respect to  $\lambda$ . It turns out that the maximal efficiencies are equal in the two cases, and depend only on a single parameter  $r$  characterizing the behavior of the system near equilibrium:

$$y^{\text{MAX}} \equiv y_{\text{eng}}^{\text{MAX}} = y_{\text{ref}}^{\text{MAX}} = \frac{r}{(1 + \sqrt{1 - r})^2}, \quad r(\beta) = \frac{M_{12}M_{21}}{M_{11}M_{22}} \quad (0 \leq r \leq 1). \quad (209)$$

The value of  $y^{\text{MAX}}$  increases monotonically as  $r$  goes from 0 to 1.

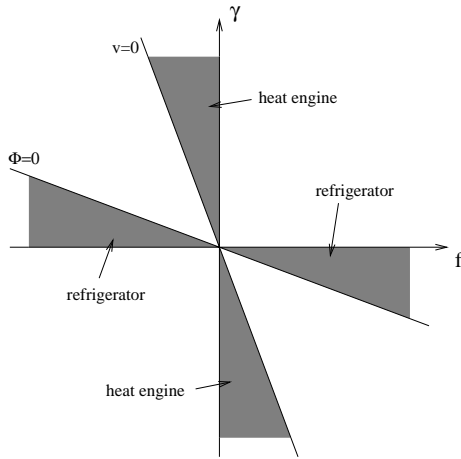
For our discrete model, Eq.201 gives the matrix  $\mathbf{M}$  – which completely describes the near-equilibrium behavior – in terms of  $\beta$ , along with the internal parameters  $\alpha$ ,  $d$ , and  $\Gamma$ . From this result, we find that our model gives:

$$r = \frac{(1 - \zeta)^2(4 + 3\zeta)}{3(3 + 4\zeta)[4 + \zeta(29 + 9\zeta)]}, \quad (210)$$

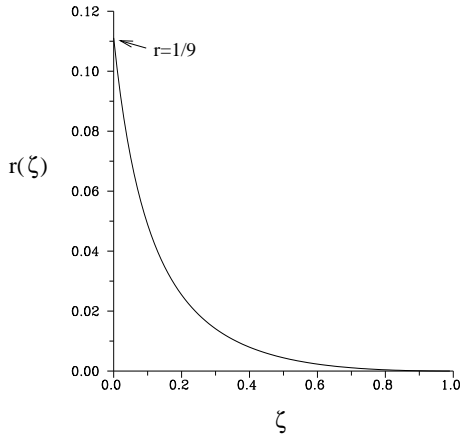
which depends only on the rescaled temperature  $\zeta = \exp(-\alpha\beta)$  of the equilibrium state. In Fig.A.8 we plot  $r(\zeta)$ . We see that  $r$  approaches a limiting value of  $1/9$  as the equilibrium temperature  $T = \beta^{-1}$  goes to zero (i.e.  $\zeta \rightarrow 0$ ), and decreases to zero as  $T \rightarrow \infty$  (i.e.  $\zeta \rightarrow 1$ ). For the limiting value  $r = 1/9$ , our model gives a maximal relative efficiency

$$y^{\text{MAX}}(T \rightarrow 0) = \frac{1}{17 + 12\sqrt{2}} \approx 0.0294. \quad (211)$$

This is the best relative efficiency which our system can achieve in the near-equilibrium regime.



**Fig A.7** General predictions based on linear response. The shaded regions adjacent to the vertical axis indicate the values of  $(f, \gamma)$  for which the system behaves as a heat engine; those adjacent to the horizontal axis indicate that the system is a refrigerator. The diagonal lines bounding these regions are the contours  $v(f, \gamma) = 0$  and  $\Phi(f, \gamma) = 0$ . This figure essentially combines Figs.5.5 and 5.6, for near-equilibrium values of  $f, \gamma \approx 0$ .



**Fig A.8** The quantity  $r(\zeta)$  is plotted for the entire range of values of the (rescaled) equilibrium state temperature ( $0 < \zeta < 1$ ).

We have carried out a cursory numerical search –  $10^8$  points sampled randomly in  $(f, \mu, \nu)$ -space – and have found, *away from equilibrium*, relative efficiencies as high as  $y \approx 0.0432$  (heat engine) and  $y \approx 0.0647$  (refrigerator). These are greater than the near-equilibrium value quoted in Eq.211, but still far short of unity. This suggests that the efficiency of our model is always considerably lower than the corresponding Carnot efficiency. Such a conclusion is in agreement with Parrondo and Espanol [85], and Sekimoto [86], who have argued that Feynman’s analysis on the point of efficiency – in which he concluded that the ratchet and pawl would operate at Carnot efficiency – was in error.



**Summary.** Our aim in this section has been to introduce a discrete model of Feynman's ratchet and pawl, and to solve exactly for the behavior of that model as a function of external load ( $f$ ) and reservoir temperatures ( $T_A, T_B$ ). The central result, Eqs.190, gives the average directed motion ( $v$ ) and heat flows ( $\dot{Q}_A, \dot{Q}_B$ ), in the steady state. We have shown that our model can act both as a heat engine and as a refrigerator, and we have investigated its behavior in the near-equilibrium, linear response regime.

**Explicit expressions.** Here we present explicit expressions for the polynomials  $P_n^j, X^j$ , and  $Y^j$ . The  $P_n^j$ 's follow from Jordan elimination, performed on the matrices  $\mathcal{R}^j$ ,  $j = b, c$ ; and  $X_j$  and  $Y_j$  then follow from Eq.184.

$$\begin{aligned}
P_1^b &= \nu^3\sigma(1 + \sigma + 3\sigma^2) + \nu\sigma^2(4 + 3\mu + (4 + \mu)(1 + 2\mu)\sigma + (4 + 3\mu(3 + 2\mu))\sigma^2) \\
&\quad + 4\mu\sigma(3 + \sigma(5 + 3\mu + 5\sigma + 4\mu\sigma)) + \nu^2(4 + \sigma(4 + 2\mu + \sigma(6 + \sigma + \sigma^2 + 3\sigma^3))) \\
P_2^b &= \nu\sigma^2(4 + 3\mu + (12 + 7\mu)\sigma + (12 + \mu(15 + 2\mu))\sigma^2) + \nu^2(4 + 2(6 + \mu)\sigma + 6(1 + \mu)\sigma^2 \\
&\quad + \sigma^3 + 3\sigma^4 + \sigma^5) + \nu^3\sigma(1 + \sigma(3 + \sigma)) + 4\mu\sigma(3 + \sigma(3 + \mu + 3(1 + \mu)\sigma)) \\
P_3^b &= 4\mu\sigma(3 + \sigma(5 + 3\sigma)) + \nu^3\sigma(1 + \sigma(5 + 9\sigma)) + \nu\sigma^2(4(1 + \sigma + 3\sigma^2) \\
&\quad + \mu(3 + \sigma(9 + 7\sigma))) + \nu^2(4 + \sigma(12 + \sigma(18 + \sigma + 5\sigma^2 + 9\sigma^3))) \\
P_4^b &= \sigma(16 + 4(6 + \nu)\sigma + 6(4 + \nu)\sigma^2 + 10\nu\sigma^3 + \mu(4 + \sigma(20 + \nu + 24\sigma + 5\nu\sigma + 9\nu\sigma^2))) \\
P_5^b &= \nu^2(4 + \mu + 5(2 + \mu)\sigma + 3(2 + 3\mu)\sigma^2) + 2\mu\sigma(4 + \mu + (4 + 3\mu)\sigma + (4 + 5\mu)\sigma^2) \\
&\quad + 2\nu\sigma^2(4 + \mu + (8 + 5\mu)\sigma + (8 + 9\mu)\sigma^2) \\
P_6^b &= 2\mu\nu\sigma^2(2 + \mu + (2 + 3\mu)\sigma + (4 + 3\mu)\sigma^2) + 2\mu^2\sigma(3 + \sigma(5 + 3\sigma)) \\
&\quad + \nu^3\sigma(2 + 4\sigma(1 + \sigma) + \mu(1 + \sigma(5 + 9\sigma))) + \nu^2(4 + 6\sigma + 2\sigma^2(3 + \sigma(1 + 2\sigma(1 + \sigma)))) \\
&\quad + \mu(3 + \sigma(11 + \sigma(15 + \sigma + 5\sigma^2 + 9\sigma^3)))) \tag{212}
\end{aligned}$$

$$\begin{aligned}
X^b &= \nu^3\sigma(-1 - 2\sigma + 3\sigma^3) + \nu\sigma^2(-4 - 3\mu - 2\mu\sigma + 4\mu\sigma^2 + 3(4 + \mu(5 + 2\mu))\sigma^3) \\
&\quad + \nu^2(-4 - 3(4 + \mu)\sigma - 3(4 + 3\mu)\sigma^2 - (1 + 9\mu)\sigma^3 + (2 + \mu)\sigma^4 + (6 + 5\mu)\sigma^5 \\
&\quad + (13 + 9\mu)\sigma^6) + 2\mu\sigma(-6 + \sigma(-4 + 6\sigma^2 + 3\mu(-1 + (-1 + \sigma)\sigma))) \\
Y^b &= -(\nu^3\sigma(3 + \sigma(5 + 7\sigma)) + \nu\sigma^2(-2\mu^2(1 + \sigma)(1 + 2\sigma) + \mu(-1 + (3 - 5\sigma)\sigma) \\
&\quad + 4(1 + \sigma + \sigma^2)) - 2\mu\sigma(2(1 + \sigma + \sigma^2) + \mu(5 + \sigma(9 + 7\sigma))) \\
&\quad + \nu^2(8 + \mu(-1 + \sigma - 3\sigma^2) + \sigma(10 + \sigma(12 + \sigma(3 + \sigma(5 + 7\sigma)))))) \tag{213}
\end{aligned}$$

$$\begin{aligned}
P_1^c &= \nu^3\sigma(1 + \sigma(3 + \sigma)) + 4\mu\sigma(5 + \mu(5 + 2\sigma) + \sigma(5 + 3\sigma)) + \nu\sigma^2(2\mu^2(3 + \sigma) \\
&\quad + 4(1 + \sigma + \sigma^2) + 3\mu(3 + \sigma(3 + \sigma))) + \nu^2(4 + 2\mu + \sigma(6 + \sigma(4 + \sigma + 3\sigma^2 + \sigma^3))) \\
P_2^c &= \nu^3\sigma(3 + \sigma + \sigma^2) + 4\mu\sigma(\mu + 3\mu\sigma + 3(1 + \sigma + \sigma^2)) + \nu\sigma^2(2\mu^2\sigma + 4(3 + \sigma(3 + \sigma)) \\
&\quad + \mu(15 + \sigma(7 + 3\sigma))) + \nu^2(12 + 2\mu(3 + \sigma) + \sigma(6 + \sigma(4 + \sigma(3 + \sigma + \sigma^2)))) \\
P_3^c &= 4\mu\sigma(5 + 3\sigma(1 + \sigma)) + \nu^3\sigma(9 + \sigma(5 + \sigma)) + \nu\sigma^2(4(1 + \sigma(3 + \sigma)) \\
&\quad + \mu(9 + \sigma(7 + 3\sigma))) + \nu^2(16 + \sigma(14 + \sigma(4 + \sigma(9 + \sigma(5 + \sigma)))))) \\
P_4^c &= \sigma(24 + 2\sigma(3 + 2\sigma)(4 + \nu + \nu\sigma) + \mu(28 + \sigma(4(4 + \sigma) + \nu(9 + \sigma(5 + \sigma)))))) \\
P_5^c &= 2\mu\sigma(4(1 + \sigma + \sigma^2) + \mu(3 + \sigma(5 + \sigma))) + \nu^2(10 + 6\sigma + 4\sigma^2 + \mu(9 + \sigma(5 + \sigma)))
\end{aligned}$$

$$\begin{aligned}
& +2\nu\sigma^2(4(2+\sigma(2+\sigma))+\mu(9+\sigma(5+\sigma))) \\
P_6^c = & 2\mu^2\sigma(5+3\sigma(1+\sigma))+2\mu\nu\sigma^2(2(1+\sigma)^2+\mu(3+\sigma(3+\sigma)))+\nu^3\sigma(2(2+\sigma(2+\sigma)) \\
& +\mu(9+\sigma(5+\sigma)))+\nu^2(2(3+\sigma(3+\sigma(2+\sigma(2+\sigma(2+\sigma)))))) \\
& +\mu(15+\sigma(11+\sigma(3+\sigma(9+\sigma(5+\sigma))))))
\end{aligned} \tag{214}$$

$$\begin{aligned}
X^c = & \sigma(\nu\sigma^2(-4-5\mu+4\mu(2+\mu)\sigma+2(2+\mu)^2\sigma^2+(4+3\mu)\sigma^3)+\nu^3(-3\sigma+2\sigma^3+\sigma^4) \\
& +2\mu\sigma(-10+6\sigma^3+\mu(-5+\sigma(-1+3\sigma))) \\
& +\nu^2(-22+\mu(-15+\sigma(-5+\sigma(-1+\sigma(9+\sigma(5+\sigma)))))) \\
& +\sigma(-8+\sigma(-2+\sigma(7+\sigma(10+\sigma(6+\sigma)))))) \\
Y^c = & -(\nu^3\sigma(5+\sigma(7+3\sigma))+\nu\sigma^2(4(1+\sigma+\sigma^2)-2\mu^2(3+\sigma(2+\sigma)) \\
& -\mu(-3+\sigma(5+\sigma)))-2\mu\sigma(2(1+\sigma+\sigma^2)+\mu(9+\sigma(7+5\sigma))) \\
& +\nu^2(10-\mu(-1+\sigma(3+\sigma))+\sigma(12+\sigma(8+\sigma(5+\sigma(7+3\sigma))))))
\end{aligned} \tag{215}$$

### 3. The piston model

**Introduction.** In Ref. [95] a *detailed fluctuation theorem* (DFT) was derived. This result pertains to the evolution of a system over a finite time interval  $\tau$ , and has a structure similar to that of the usual FT, but in addition exhibits a dependence on the initial and final microstates of the system. Specifically, consider a process during which a system of interest evolves in contact with one or more heat reservoirs, while – possibly – some *work parameter*  $\lambda$  (e.g. an applied field) is being manipulated externally. Let  $\lambda^+(t)$  be the externally imposed time-dependence of this parameter, from  $t = 0$  to  $t = \tau$ , and let  $\Delta s$  denote the entropy produced during a particular realization of this process. This entropy production is defined in terms of heat exchanged with the reservoir(s); see Ref. [95] or Eq.247 below. Let  $\Pi^+$  denote the process just described. Now consider the “reverse” process,  $\Pi^-$ , defined exactly as  $\Pi^+$ , but with the time-dependence of the work parameter reversed:  $\lambda^-(t) = \lambda^+(\tau - t)$ ,  $t \in [0, \tau]$ . For instance, if  $\Pi^+$  entails turning on an external field, then  $\Pi^-$  entails turning it off. Finally, for the forward process  $\Pi^+$ , let  $P_+(\mathbf{z}_f, \Delta s | \mathbf{z}_i)$  denote the joint probability that the entropy produced over the interval of observation will be  $\Delta s$ , and the final microstate of the system will be  $\mathbf{z}_f$ , *given* an initial microstate  $\mathbf{z}_i$ ; and let  $P_-(\mathbf{z}_f, \Delta s | \mathbf{z}_i)$  denote the same for the reverse process,  $\Pi^-$ . Then the DFT states that these joint, conditional probability distributions satisfy the following relation:

$$\frac{P_+(\mathbf{z}_B, +\Delta S | \mathbf{z}_A)}{P_-(\mathbf{z}_A^*, -\Delta S | \mathbf{z}_B^*)} = \exp(\Delta S/k_B), \tag{216}$$

where the asterisk (\*) denotes a reversal of momenta:  $(\mathbf{q}, \mathbf{p})^* = (\mathbf{q}, -\mathbf{p})$ . Note that, if no external parameter is being changed, then the subscripts on  $P_+$  and  $P_-$  are dropped, and then the DFT makes a prediction about a single function  $P(\mathbf{z}_f, \Delta S | \mathbf{z}_i)$ , corresponding to some static process  $\Pi$ .

**The model.** We would like to present the following model. Consider a piston moving freely in a tube. Let the piston be constrained by two boundaries in a way that

it reflects elastically from the boundaries conserving its kinetic energy. Now let the tube be filled with ideal gas with temperature  $T_l$  from the left side of the piston and with temperature  $T_r$  from the right side. We will consider that the tube extends infinitely to the both sides of the piston. So the temperatures of gases are held fixed and do not depend on the process of interaction with the piston. Now let us take the limit when the boundaries approaching to each other. It is clear that the average rate of collision between particles and the piston converges to a finite value. In this case we can treat the piston having the fixed coordinate, but possible to have different kinetic energies. Thus we abstract from the motion of the piston and consider only the changing its kinetic energy under the influence of particles. We consider the particles of gases having the kinetic energy distributed according to the canonical distribution. We consider also that the average rates of collisions from the left side and from the right side are constant in time, and, generally, can be different.

Our observation of the given model starts at some moment of time and lasts for a period when  $M$  collisions between particles and the piston happened. We will watch the system evolution as an evolution which depends on number of collisions rather than on time. And we will define the final state by a number of been collisions, not by the elapsed time. Each collision redistributes energy between the piston and gases or, say, heat reservoirs. Let us define the entropy changing for one collision as:  $\Delta S = -(u_B - u_A)/T$ , where  $u_B$  is the kinetic energy of the piston after the collision,  $u_A$  - before, and  $T$  is the temperature of the gas, which a hitting particle belongs to. Actually  $u_B - u_A$  is the heat transferred from gas to the piston. After  $M$  collisions we can find the piston having the energy  $u_B$  and total change of entropy equal to  $\Delta S$ . Here we would like to present an explicit expression for  $P_M(u_B, \Delta S | u_A)$  — joint probability distribution to find the piston with kinetic energy  $u_B$  and that entropy of the system has changed on value  $\Delta S$ , conditional on the initial kinetic energy of the piston being  $u_A$ . This expression (Eq.220 and the following ingredients) is the main result of the current subsection. We also stress here (and prove it below) that this expression satisfies the relation:

$$\frac{P_M(u_B, \Delta S | u_A)}{P_M(u_A, -\Delta S | u_B)} = e^{\Delta S} \quad (217)$$

**The main result.** Let  $\gamma$  denote a sequence of collision. We will be interested only a side which collision was from. For instance,  $\gamma = (llrrrrlr)$  means that the first collision was from the left side then again from the left, then 4 particles hit the piston from the right side, then 1 from left and 1 from right. Now let  $P^\gamma(u_B, \Delta S | u_A)$  denote the probability to find the quantities  $\Delta S$  and  $u_B$  at the end given  $u_A$  at the beginning after the exact sequence  $\gamma$ . Then  $P_M(u_B, \Delta S | u_A)$  one can present as:

$$P_M(u_B, \Delta S | u_A) = \sum_{\gamma; M} P^\gamma(u_B, \Delta S | u_A) p_\gamma, \quad (218)$$

where  $p_\gamma$  is the probability to find the exact sequence  $\gamma$  and the sum is taken over all possible sequences  $\gamma$  of length  $M$ .

We state here (and this will be proven below) that  $P^\gamma(u_B, \Delta S | u_A)$  depends only on the number of *new collisions* and the fact of which side was the first collision from. By a *new*

*collision* we mean either the first collision, or a collision from one side if the previous one was from the opposite side. For example, if  $\gamma = (lll)$  then there was 1 *new collision* — the first one; if  $\gamma = (rrlllr)$  then there were 3 *new collisions* — 1st, 3rd and 6th. Thus the minimal  $N$  can be 1 when a sequence consist of only one type of events, and the maximal  $N$  can be equal to  $M$  when each collision is a *new collision*. So we can write:

$$P^\gamma(u_B, \Delta S|u_A) = P_{N_\gamma}(u_B, \Delta S|u_A, \pi_\gamma), \quad (219)$$

where  $P_N(u_B, \Delta S|u_A, \pi)$  is the probability to find  $u_B, \Delta S$  with given  $u_A$  after  $N$  *new collisions* with given the first collision  $\pi$ , which can be  $l$  or  $r$ . Indexes on  $N$  and  $\pi$  in Eq.219 says that  $\gamma$  defines uniquely  $N$  and  $\pi$ . Grouping the terms with equal  $N$  in the sum of Eq.218 we find (using Eq.219) that the probability  $P_M(u_B, \Delta S|u_A)$  can be presented as:

$$P_M(u_B, \Delta S|u_A) = \sum_{N=1}^M P_N(u_B, \Delta S|u_A, l)P(N, l|M) + P_N(u_B, \Delta S|u_A, r)P(N, r|M), \quad (220)$$

where  $P(N, \pi|M)$  is the probability that in a sequence of length  $M$  one finds  $N$  *new collisions* with the first collision  $\pi$ . In the last subsection of this section we solve explicitly for  $P_N(u_B, \Delta S|u_A, \pi)$ , obtaining:

$$P_N(u_B, \Delta S|u_A, \pi) = \mathcal{N}' e^D \min[1, e^{\beta a}] \frac{1}{N_c!} \sum_{k=0}^{N_c} \binom{N_c}{k} |a|^{N_c-k} \frac{(N_c+k)!}{\beta^{N_c+k+1}}; \quad \text{odd } N \quad (221)$$

$$P_N(u_B, \Delta S|u_A, \pi) = \mathcal{N}' \frac{e^D}{N_c!(N_c-1)!} \sum_{k=0}^{N_c} \binom{N_c}{k} a^{N_c-k} \frac{(N_c+k-1)!}{\beta^{N_c+k}}; \quad \text{even } N; a > 0 \quad (222)$$

$$\begin{aligned} P_N(u_B, \Delta S|u_A, \pi) &= \\ &= \mathcal{N}' \frac{e^D e^{\beta a}}{N_c!(N_c-1)!} \sum_{k=0}^{N_c-1} \binom{N_c-1}{k} |a|^{N_c-k-1} \frac{(N_c+k)!}{\beta^{N_c+k+1}}; \quad \text{even } N; a < 0 \end{aligned} \quad (223)$$

The parameters  $a$ ,  $D$ ,  $N_c$  and  $\mathcal{N}'$  depend on the arguments  $u_B, \Delta S, u_A, \pi$  and parity of  $N$ . They are presented explicitly in the last subsection of this section (Eqs.228, 229 and 230). The parameter  $\beta = \frac{1}{T_l} + \frac{1}{T_r}$ .

$$P(N, l|M) = \sum_{\gamma; N, l} p_l^{n_l} p_r^{n_r}, \quad \text{and} \quad P(N, r|M) = \sum_{\gamma; N, r} p_l^{n_l} p_r^{n_r}, \quad (224)$$

where  $p_l$  and  $p_r$  are probabilities that collision was exactly from left or right side, and  $n_l$  and  $n_r$  are numbers of collisions from left and right side ( $n_l + n_r = M$ ) — they defined by  $\gamma$ , and the sum is taken over all  $\gamma$  with fixed  $N$  and fixed the first collision  $l$  or  $r$ . Exact expression for  $P(N, \pi|M)$  can be found solving the combinatorial problem. This is not necessary to prove Eq.217.

Now we will argue the validity of Eq.219. Let us first consider a single collision between a particle and the piston. Because the particle and piston masses are equal, the kinetic energy of the piston after the collision will equal the kinetic energy of the particle before the collision, associated with the direction normal to the plain of piston. That means that the probability distribution of the piston energy after the collision will be the canonical distribution corresponding to temperatures  $T_l$  or  $T_r$ , depending on whether the colliding particle come from the left or the right. In this case the joint probability is:

$$P^{\gamma=(\pi)}(u_B, \Delta S | u_A) = P_{N=1}(u_B, \Delta S | u_A, \pi) = \frac{1}{T_\pi} e^{-u_B/T_\pi} \delta\left(\Delta S + \frac{u_B - u_A}{T_\pi}\right). \quad (225)$$

For two collisions from the same side,  $\gamma = (\pi\pi)$ , we find that:

$$P^{(\pi\pi)}(u_B, \Delta S | u_A) = \int_0^\infty du_1 \int_{-\infty}^\infty dS_1 \int_{-\infty}^\infty dS_2 \delta(\Delta S - S_1 - S_2) \cdot P_{N=1}(u_B, S_2 | u_1, \pi) P_{N=1}(u_1, S_1 | u_A, \pi) = P_{N=1}(u_B, \Delta S | u_A, \pi) \quad (226)$$

This shows us that the probability distribution does not change after a collision from the same side. It is not hard to see that this property extends to any sequence  $\gamma$ , validating Eq.219.

**Fluctuation theorem relation.** To prove Eq.217 it is sufficient to show that the three following conditions are satisfied.

$$1) P_N(u_B, \Delta S | u_A, \pi) = P_N(u_A, -\Delta S | u_B, \pi) e^{\Delta S} \quad \text{when } N \text{ is odd}; \quad (C1)$$

$$2) P_N(u_B, \Delta S | u_A, \pi) = P_N(u_A, -\Delta S | u_B, \pi^*) e^{\Delta S} \quad \text{when } N \text{ is even}; \quad (C2)$$

$$3) P(N, l | M) = P(N, r | M) \quad \text{when } N \text{ is even}. \quad (C3)$$

Here  $\pi^*$  is the opposite of  $\pi$ . The fact that these three conditions are sufficient to prove Eq.217 can be easily verified by substituting these relations into Eq.220.

Let us prove C1. For the case when  $N$  is odd the quantity  $a$  is antisymmetric under exchange of  $u_A$  and  $u_B$ , and  $\Delta S$  to  $-\Delta S$  (see Eq.229):

$$a(u_B, u_A, \Delta S, \pi) = -a(u_A, u_B, -\Delta S, \pi).$$

That means that the sum in Eq.221 remains the same for  $P_N(u_A, -\Delta S | u_B, \pi)$ . Therefore we get:

$$\ln \frac{P_N(u_B, \Delta S | u_A, \pi)}{P_N(u_A, -\Delta S | u_B, \pi)} = \beta a(u_B, u_A, \Delta S, \pi) + D(u_B, u_A, \Delta S, \pi) - D(u_A, u_B, -\Delta S, \pi) = \Delta S$$

The explicit expressions for  $a$  and  $D$  are given in the last subsection of this section.

To prove C2 we use the following identity (see Eq.230):

$$a(u_B, u_A, \Delta S, \pi) = a(u_A, u_B, -\Delta S, \pi^*).$$

This implies that  $P_N(u_B, \Delta S | u_A, \pi)$  and  $P_N(u_A, -\Delta S | u_B, \pi^*)$  will differ only by term  $e^D$ . We then get

$$\ln \frac{P_N(u_B, \Delta S | u_A, \pi)}{P_N(u_A, -\Delta S | u_B, \pi^*)} = D(u_B, u_A, \Delta S, \pi) - D(u_A, u_B, -\Delta S, \pi^*) = \Delta S$$

To show that  $P(N, l|M) = P(N, r|M)$  **when  $N$  is even** we note that the sums in Eqs.224 can be rewritten as sums over sequences  $\gamma^*$  which are exactly reversed to  $\gamma$ . For example, if  $\gamma = (llrrrlrrll)$  then  $\gamma^* = (rllrlrrlll)$ . And it is clear that the first symbol changes to the opposite. The transformation  $\gamma$  to  $\gamma^*$  is unique and does not change  $n_r$ ,  $n_l$  and  $N$ . So we can write:

$$\sum_{\gamma; N, l} p_l^{n_l} p_r^{n_r} = \sum_{\gamma^*; N, r} p_l^{n_l} p_r^{n_r} = \sum_{\gamma; N, r} p_l^{n_l} p_r^{n_r},$$

which is nothing else but the condition C3 (see Eq.224).

**Derivation.** Here we sketch the derivation of Eqs. 221,222 and 223.

The general expression for any  $N$  is obtained from the expression for  $N = 1$  by multiplying a sequence of factors  $P_{N=1}$  and then integrating over internal  $u_i$ ,  $0 < i < N$ . Then it takes the form:

$$P_N(u_N, \Delta S | u_0, \pi) = \mathcal{N} \left[ \prod_{i=1}^{N-1} \int_0^\infty du_i \right] e^{-\sum_{i=1}^N u_i / T_i} \delta \left( \Delta S + \sum_{i=1}^N \frac{u_i - u_{i-1}}{T_i} \right), \quad (227)$$

where the normalizing term  $\mathcal{N}^{-1} = (T_l T_r)^{N/2}$  when  $N$  is even and  $\mathcal{N}^{-1} = T_\pi (T_l T_r)^{(N-1)/2}$  when  $N$  is odd; and  $T_i = T_\pi$  if  $i$  is odd and  $T_i = T_{\pi^*}$  if  $i$  is even. Here we take  $u_0 = u_A$  and  $u_N = u_B$ .

Consider first the case when  $N$  is odd. To evaluate the integral in Eq.227 we first integrate over  $u_1$  and then introduce new variables:  $x_{(i-1)/2} = u_i / T_\pi$  for  $i$  odd, and  $y_{i/2} = u_i / T_{\pi^*}$  for  $i$  even. The right side of Eq.227 can then be represented as:

$$\mathcal{N}' e^D \left[ \prod_{i=1}^{N_c+1} \int_0^\infty dx_i \right] e^{-\beta \sum_{i=1}^{N_c+1} x_i} \left[ \prod_{j=1}^{N_c} \int_0^\infty dy_j \right] \theta(a + \sum_{i=1}^{N_c+1} x_i - \sum_{j=1}^{N_c} y_j),$$

where the following definitions are introduced

$$\mathcal{N}' = \frac{\mathcal{N} T_R T_L}{|T_R - T_L|}; \quad N_c = \frac{N-3}{2}; \quad (228)$$

$$a = \frac{-\Delta S T_\pi T_{\pi^*} + (u_0 - u_N) T_{\pi^*}}{T_{\pi^*} - T_\pi}; \quad D = \frac{\Delta S T_\pi T_{\pi^*} + u_N T_\pi - u_0 T_{\pi^*}}{(T_{\pi^*} - T_\pi) T_\pi}. \quad (229)$$

$\theta$  is Heaviside's function. The term  $\theta(\cdot)$  appears after the integration over  $u_1$ . Now the integral can be rewritten as:

$$\mathcal{N}' e^D (N_c!)^{-2} \int_0^\infty dz e^{-\beta z} z^{N_c} (a+z)^{N_c} \theta(a+z).$$

Here we used the identity:

$$\left[ \prod_{i=1}^N \int_0^\infty dx_i \right] f\left(\sum_{i=1}^N x_i\right) = \frac{1}{(N-1)!} \int_0^\infty z^{N-1} f(z) dz.$$

And finally we obtain Eq.221 by expanding  $(a + z)^{N_c}$  and evaluating the  $\Gamma$ -integrals.

A similar procedure can be used when  $N$  is even. In this case after integrating Eq.227 over  $u_1$  and substituting  $x_i$  and  $y_i$  we obtain:

$$\mathcal{N}' e^D \left[ \prod_{i=1}^{N_c} \int_0^\infty dx_i \right] e^{-\beta \sum_{i=1}^{N_c} x_i} \left[ \prod_{j=1}^{N_c} \int_0^\infty dy_j \right] \theta(a + \sum_{i=1}^{N_c} x_i - \sum_{j=1}^{N_c} y_j),$$

where now  $N_c = (N - 1)/2$  and

$$a = \frac{-\Delta S T_\pi T_{\pi^*} + u_0 T_{\pi^*} - u_N T_\pi}{T_{\pi^*} - T_\pi}; \quad D = \frac{\Delta S T_\pi T_{\pi^*}^2 + u_N T_\pi^2 - u_0 T_{\pi^*}^2}{(T_{\pi^*} - T_\pi) T_\pi T_{\pi^*}}. \quad (230)$$

This multidimensional integral can again be cast into a one-dimensional form:

$$\mathcal{N}' e^D (N_c! (N_c - 1))^{-1} \int_0^\infty dz e^{-\beta z} z^{N_c-1} (a + z)^{N_c} \theta(a + z).$$

And finally we get Eq.222 and Eq.223.

#### 4. Dragged harmonic oscillator

**Introduction.** Statements of the FT which have appeared in the literature can be expressed (after appropriate definitions of the quantities appearing) by the following equation, pertaining to a system evolving away from thermal equilibrium:

$$\left[ \lim_{\tau \rightarrow \infty} \right] \frac{1}{\tau} \ln \frac{p_\tau(+\bar{\sigma})}{p_\tau(-\bar{\sigma})} = \bar{\sigma}. \quad (231)$$

Here,  $\bar{\sigma}$  is the average rate at which entropy is generated during a given time interval – or *segment* – of duration  $\tau$ ;  $p_\tau(\bar{\sigma})$  is the distribution of values of  $\bar{\sigma}$  over a statistical ensemble of such segments; and Boltzmann's constant  $k_B = 1$ . Note that the words “average rate” denote the time average over a given segment, *not* an average over the ensemble of segments. Following the literature, we will distinguish between two versions of the FT, *steady-state* and *transient*, which differ in the definition of the ensemble of segments considered. In the steady-state case, we imagine observing the system as it evolves in a nonequilibrium steady state for an “infinite” length of time. We divide this time of observation into infinitely many consecutive segments of duration  $\tau$ , and compute the average entropy generation rate,  $\bar{\sigma}$ , for each segment;  $p_\tau(\bar{\sigma})$  is then the distribution of values of  $\bar{\sigma}$  over this ensemble of segments. In the transient case, by contrast, we imagine that the system of interest begins in an equilibrium state, but then is driven *away* from equilibrium, for instance by the sudden application of an external force. We observe the response of the system for a segment of duration  $\tau$ , starting from the moment the external perturbation is applied, and compute the average entropy generation rate. Then  $p_\tau(\bar{\sigma})$  is the distribution of values of  $\bar{\sigma}$  over infinitely many repetitions of this process, always starting from equilibrium. The transient FT is valid for any duration  $\tau$ , whereas the steady-state FT becomes valid as  $\tau \rightarrow \infty$ ; hence the parenthetical appearance of that limit in Eq.231.

In this section we consider a one-dimensional model of a particle dragged through a thermal medium. We first introduce the model and then we solve it exactly: given the particle's initial location, we compute the joint probability distribution of the final location and the net external work performed on the particle, after an arbitrary time of evolution. We then use this result to confirm both the steady-state and transient fluctuation theorems, as well as the detailed fluctuation theorem and the nonequilibrium work relation for free energy differences.

In our model, the particle under consideration obeys Langevin dynamics. This model can therefore be viewed as a special case of the situation considered by Kurchan [98] (and later generalized by Crooks [94], and Lebowitz and Spohn [100]), who showed that the FT is satisfied for this class of stochastic dynamics. In essence, we are illustrating Kurchan's results with a specific model. However, the definition of entropy generated which we choose differs from that of Ref. [98], and so we are obliged to express the steady-state and transient FT's differently, in terms of *power delivered* rather than *entropy generation rate*. Nevertheless, as discussed in greater detail below, the results which we present are indeed equivalent to the predictions of Kurchan, despite the difference in terminology.

**Introduce and solve the model.** Consider the following situation. A particle, in contact with a thermal medium at temperature  $\beta^{-1}$ , is dragged through that medium by a time-dependent external harmonic force. Assuming a single degree of freedom, let  $x$  denote the location of the particle, and let

$$U(x, t) = \frac{k}{2}(x - ut)^2 \quad (232)$$

be the potential well, moving with constant speed  $u$ , which drags the particle. Assume furthermore that the thermal forces can be modeled as the sum of linear friction and white noise, and that the motion of the particle is overdamped. Then the equation of motion for the position of the particle is:

$$\dot{x} = -\frac{k}{\gamma}(x - ut) + \tilde{v}, \quad (233)$$

where  $\dot{x} \equiv dx/dt$ ,  $\gamma$  is the coefficient of friction, and  $\tilde{v}(t)$  represents delta-correlated white noise with variance  $2/\beta\gamma$  (as mandated by the fluctuation-dissipation relation):

$$\langle \tilde{v}(t_1)\tilde{v}(t_2) \rangle = (2/\beta\gamma) \cdot \delta(t_2 - t_1). \quad (234)$$

Imagine that we observe the evolution of such a particle over a time interval from  $t = 0$  to  $t = \tau$ , and from the observed trajectory we compute the total work  $W$  performed by the external potential over this interval. Then the central result of this section will be an answer to the following question. Given an initial location  $x_0$ , what is the joint probability distribution for the final location and value of work performed  $(x, W)$ ? We will then make use of the answer to show that various forms of the FT are satisfied for this simple model.

To answer the above question, we first introduce a “work accumulated” function,  $w(t)$ , which gives the work performed on the particle up to time  $t$ ; hence,  $W = w(\tau)$ . This function satisfies



$$\dot{w}(t) = \frac{\partial U}{\partial t}(x(t), t) = -uk(x(t) - ut), \quad (235)$$

along with the initial condition  $w(0) = 0$ . (To motivate Eq.235, see e.g. Refs. [96,97].)

It will furthermore prove convenient to specify the location of the particle by a variable  $y = x - ut$  (i.e. in the reference frame of the moving well), rather than  $x$ . Under this change of variables, Eqs.233 and 235 become:

$$\dot{y} = -\frac{k}{\gamma}y - u + \tilde{v} \quad (236a)$$

$$\dot{w} = -uky. \quad (236b)$$

Now imagine a statistical ensemble of such particles, represented by an evolving probability distribution  $f(y, w, t)$ . Eq.236 then translates into the Fokker-Planck equation

$$\frac{\partial f}{\partial t} = \frac{k}{\gamma} \frac{\partial}{\partial y}(yf) + u \frac{\partial f}{\partial y} + uk y \frac{\partial f}{\partial w} + \frac{1}{\beta\gamma} \frac{\partial^2 f}{\partial y^2}. \quad (237)$$

What we now want is an expression for  $f(y, w, t|y_0)$ , by which we mean the solution to Eq.237 satisfying the initial conditions

$$f(y, w, 0|y_0) = \delta(y - y_0)\delta(w). \quad (238)$$

The function  $f(y, w, t|y_0)$  is the joint probability distribution for achieving a location  $y$  and a value of work accumulated  $w$ , at time  $t$ , given  $y_0$  at time 0. By evaluating this solution at  $t = \tau$ , and making the change of variables from  $y$  back to  $x$ , we have the answer to the question posed earlier.

To solve for  $f(y, w, t|y_0)$ , we first note that Eq.237 has the following property: if at one instant in time the distribution happens to be Gaussian, then it will remain Gaussian for all subsequent times. This follows from the fact that the drift and diffusion coefficients in Eq.237 are either constant or linear in  $y$  and  $w$ . Now, a normalized Gaussian distribution  $f^G(y, w)$  is uniquely defined by the values of the following moments:

$$\hat{y} \equiv \int f^G(y, w) y \quad (239a)$$

$$\hat{w} \equiv \int f^G(y, w) w \quad (239b)$$

$$\sigma_y^2 \equiv \int f^G(y, w) (y^2 - \hat{y}^2) \quad (239c)$$

$$\sigma_w^2 \equiv \int f^G(y, w) (w^2 - \hat{w}^2) \quad (239d)$$

$$c_{yw} \equiv \int f^G(y, w) (yw - \hat{y}\hat{w}), \quad (239e)$$

where the integrals are over  $(y, w)$ -space. An explicit expression for  $f^G(y, w)$  in terms of these moments is:

$$f^G(y, w) = \frac{\sqrt{C}}{2\pi} \exp(-\mathbf{z}^T \mathbf{C} \mathbf{z} / 2), \quad (240a)$$

where

$$\mathbf{z} = \begin{pmatrix} y - \hat{y} \\ w - \hat{w} \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} C\sigma_w^2 & -Cc_{yw} \\ -Cc_{yw} & C\sigma_y^2 \end{pmatrix}, \quad (240b)$$

$C = (\sigma_y^2\sigma_w^2 - c_{yw}^2)^{-1} = \det \mathbf{C}$ , and  $\mathbf{z}^T$  denotes the transpose of  $\mathbf{z}$ .

The evolution of a time-dependent Gaussian distribution  $f^G(y, w, t)$  is thus uniquely specified by the evolution of the moments  $\hat{y}, \dots, c_{yw}$ . Given a distribution evolving under Eq.237, we get from Eq.239 the following set of coupled equations of motion for these moments:

$$\frac{d}{dt}\hat{y} = -\frac{k}{\gamma}\hat{y} - u, \quad \frac{d}{dt}\hat{w} = -uk\hat{y} \quad (241a)$$

$$\frac{d}{dt}\sigma_y^2 = -\frac{2k}{\gamma}\left(\sigma_y^2 - \frac{1}{\beta k}\right), \quad \frac{d}{dt}c_{yw} = -\frac{k}{\gamma}(c_{yw} + u\gamma\sigma_y^2), \quad \frac{d}{dt}\sigma_w^2 = -2ukc_{yw}. \quad (241b)$$

Note that Eq.241a forms one closed set of equations, and Eq.241b forms another closed set. Also, the equations for  $\hat{y}$  and  $\sigma_y^2$  are individually closed, which makes sense: the evolution of the position of the particle, from time  $t$  onward, does not depend on how much work has been performed up to time  $t$ .

The distribution  $f(y, w, t|y_0)$  will then be a Gaussian whose moments evolve according to Eq.241, and satisfy the initial conditions:  $\hat{y} = y_0$ ,  $\hat{w} = \sigma_y^2 = c_{yw} = \sigma_w^2 = 0$ . The solution, as easily verified by substitution, is:

$$\hat{y}(t|y_0) = -l + (y_0 + l)e^{-kt/\gamma} \quad (242a)$$

$$\hat{w}(t|y_0) = uklt + u\gamma(y_0 + l)(e^{-kt/\gamma} - 1) \quad (242b)$$

$$\sigma_y^2(t|y_0) = \frac{1}{\beta k}(1 - e^{-2kt/\gamma}) \quad (242c)$$

$$c_{yw}(t|y_0) = -\frac{u\gamma}{\beta k}(e^{-kt/\gamma} - 1)^2 \quad (242d)$$

$$\sigma_w^2(t|y_0) = \frac{u^2\gamma^2}{\beta k}\left(\frac{2kt}{\gamma} - e^{-2kt/\gamma} + 4e^{-kt/\gamma} - 3\right), \quad (242e)$$

where

$$l = \frac{\gamma u}{k}. \quad (243)$$

The combination of Eqs.242 and 240 gives the solution for  $f(y, w, t|y_0)$  for all times  $t \geq 0$  (and for all values of  $y$ ,  $w$ , and  $y_0$ ). An explicit expression for this solution is somewhat lengthy, and is given in the last subsection of this section. Note that, by projecting out either of the two independent variables  $y$  and  $w$ , we can obtain the marginal probability distributions for the other variable:

$$\rho(y, t|y_0) \equiv \int dw f(y, w, t|y_0) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp[-(y - \hat{y})^2/2\sigma_y^2] \quad (244a)$$

$$\eta(w, t|y_0) \equiv \int dy f(y, w, t|y_0) = \frac{1}{\sqrt{2\pi\sigma_w^2}} \exp[-(w - \hat{w})^2/2\sigma_w^2]. \quad (244b)$$

It is instructive to consider the limit of asymptotically long times,  $kt/\gamma \gg 1$ . In this limit,

$$\hat{y}(t|y_0) \rightarrow -l \quad (245a)$$

$$\hat{w}(t|y_0) \rightarrow uklt + O(1) \quad (245b)$$

$$\sigma_y^2(t|y_0) \rightarrow 1/\beta k \quad (245c)$$

$$c_{yw}(t|y_0) \rightarrow -u\gamma/\beta k \quad (245d)$$

$$\sigma_w^2(t|y_0) \rightarrow 2u^2\gamma t/\beta + O(1), \quad (245e)$$

where the “order unity” corrections to  $\hat{w}(t|y_0)$  and  $\sigma_w^2(t|y_0)$  represent terms which converge to a constant as  $t \rightarrow \infty$ . We see that the distribution of positions settles into a steady-state Gaussian of variance  $1/\beta k$ , centered at a displacement  $-l$  from the minimum of the confining potential  $U$  (Eqs.245a, 245c). Note that a *canonical* distribution of positions would have the same variance, but would be centered at the minimum. Hence,  $l$  represents the extent to which the steady-state distribution of positions “lags behind” the canonical distribution, in the long-time limit. This lag could have been predicted with an educated guess, by ignoring fluctuations and simply balancing the frictional and harmonic forces,  $-\gamma u$  and  $-ky$ . One can similarly understand the leading behavior of  $\hat{w}(t)$ : if the position of the particle is (on average) a distance  $l$  behind the instantaneous minimum of the well, then the harmonic spring is pulling the particle with a force  $+kl$ , at a velocity  $u$ , hence delivering a power  $ukl$ .

**Energy conservation and entropy generation.** In our Langevin model, energy conservation translates into the following balance equation between the external work  $W$  performed on the particle, the heat  $Q$  absorbed from the thermal surroundings, and the net change in the internal energy of the particle [96]:

$$W + Q = \Delta U, \quad (246)$$

where  $\Delta U = U(x(\tau), \tau) - U(x(0), 0)$ , and  $W = w(\tau)$  as above. Eq.246 in essence *defines* the net heat absorbed by the particle from the thermal environment,  $Q$ .

In addition to these quantities ( $W$ ,  $Q$ , and  $\Delta U$ ), we would like to have a microscopic definition of the *entropy generated* over one realization of the process, from  $t = 0$  to  $t = \tau$ . There is necessarily some arbitrariness in such a definition, but we will use the following one:

$$\Delta S \equiv -\beta Q. \quad (247)$$

Thus, we identify the entropy generated with the amount of heat dumped into the environment ( $-Q$ ), divided by the temperature. This definition – motivated by macroscopic thermodynamics (see e.g. Ref. [95]) – differs from that of Kurchan [98], who identifies the entropy production rate with the external power delivered to the system, divided by the reservoir temperature. Both definitions seem to be reasonable, but we will stick with Eq.247.

**Transient and steady-state fluctuation theorems.** In this subsection we use the results derived above to show that our model obeys both a transient and a steady-state FT. Ordinarily the FT is expressed in terms of the average entropy production rate (Eq.231). By contrast, the relations which we will show to be satisfied by our model (Eq.249) are expressed in terms of the average *power delivered* to our particle as it is dragged through its thermal environment. This difference is a consequence of our choice of definition of average entropy generation rate,  $\bar{\sigma} = \Delta S/\tau = -\beta Q/\tau$ . (Kurchan, by contrast, defines entropy generation in terms of power delivered,  $\bar{\sigma}_{\text{Kurchan}} = \beta W/\tau$ . [98]) In the calculations to follow, the reader should bear in mind that the transient and steady-state relations which we establish *are identical to those of Kurchan*; only our definition of entropy production prevents us from presenting them as such.

For a time interval of duration  $\tau$ , let

$$X = W/\tau \quad (248)$$

denote the time-averaged rate at which work is performed on the particle – i.e. the average power delivered – by the moving harmonic potential. Now imagine that we observe the particle for an “infinitely” long time as it evolves in the steady state; we divide this time of observation into infinitely many segments of duration  $\tau$ ; we compute the average power delivered,  $X$ , over each segment; and we construct the statistical distribution of these values,  $p_\tau^S(X)$ . We will obtain an explicit expression for  $p_\tau^S(X)$  for our model, and will show that it satisfies the following *steady-state* FT:

$$\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \ln \frac{p_\tau^S(+X)}{p_\tau^S(-X)} = \beta X. \quad (249a)$$

Similarly, imagine that we begin with the particle in thermal equilibrium, and then we drag it for a time  $\tau$  with the harmonic confining potential. Again defining  $X$  to be the average power delivered over this interval, let  $p_\tau^C(X)$  denote the statistical distribution of values of  $X$ , over infinitely many repetitions of this process, always starting from equilibrium.<sup>19</sup> We will solve for  $p_\tau^C(X)$  and show that it obeys the following *transient* FT:

$$\frac{1}{\tau} \ln \frac{p_\tau^C(+X)}{p_\tau^C(-X)} = \beta X, \quad (249b)$$

whose validity does *not* require the limit  $\tau \rightarrow \infty$ .

Let us begin by considering the steady-state case. Since the right sides of Eqs.245a and 245c are independent of  $y_0$ , an arbitrary initial distribution of particle positions  $\rho(y, 0)$  will settle into a Gaussian:

$$\lim_{t \rightarrow \infty} \rho(y, t) = \rho^S(y) \equiv \sqrt{\frac{\beta k}{2\pi}} \exp[-\beta k(y + l)^2/2], \quad (250)$$

---

<sup>19</sup>The superscript  $C$  indicates that the particle’s initial conditions are sampled from a canonical ensemble.

which defines the instantaneous distribution of particle positions in the steady state. Then  $p_\tau^S(X)$  – the distribution of values of average power delivered,  $X = W/\tau$ , over time intervals of duration  $\tau$  sampled during the steady state – can be constructed by folding together  $\rho^S(y_0)$  and  $\eta(w, \tau|y_0)$  (Eq.244b):

$$p_\tau^S(X) = \int dw \delta(X - w/\tau) \int dy_0 \rho^S(y_0) \eta(w, \tau|y_0). \quad (251)$$

Here, the integral over  $dy_0$  produces the distribution of values of work, after time  $\tau$ , given initial conditions sampled from the steady state. The integral over  $dw$  converts that distribution of values of  $w$  into one of values of  $X$ . Examining Eqs.242b, 242e, 244b, and 250, we see that, in the product  $\rho^S(y_0)\eta(w, \tau|y_0)$ , the variable  $y_0$  appears only in powers up to the quadratic inside an exponent; hence this product is a Gaussian in  $y_0$ , and the integral can be carried out explicitly. Without going through the (modestly tedious) details, we present the result:

$$p_\tau^S(X) = \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left[-\frac{(X - \bar{X})^2}{2\sigma_X^2}\right], \quad (252)$$

where

$$\bar{X} = ukl \quad , \quad \sigma_X^2 = \frac{2\mu\bar{X}}{\beta\tau} \quad , \quad \mu(\tau) = 1 + \frac{\gamma}{k\tau}(e^{-k\tau/\gamma} - 1). \quad (253)$$

( $\bar{X}$  represents the average instantaneous power delivered to the particle, in the steady state.) From this we obtain an explicit expression for the left side of Eq.249a:

$$\frac{1}{\tau} \ln \frac{p_\tau^S(+X)}{p_\tau^S(-X)} = \frac{1}{\tau} \cdot \frac{2\bar{X}X}{\sigma_X^2} = \frac{\beta X}{\mu}. \quad (254)$$

Since  $\lim_{\tau \rightarrow \infty} \mu = 1$ , we conclude that the FT for power delivered in the steady state (Eq.249a) is indeed satisfied for our model. Note, however, that the limit  $\tau \rightarrow \infty$  is necessary.

In the case of the *transient* FT, pertaining to a system driven away from an initial state of canonical equilibrium, we can construct  $p_\tau^C(X)$  in much the same way as  $p_\tau^S(X)$ , only now we fold  $\eta$  in with a canonical distribution,  $\rho^C(y) \propto \exp(-\beta ky^2/2)$ , rather than the steady-state distribution:

$$p_\tau^C(X) = \int dw \delta(X - w/\tau) \int dy_0 \rho^C(y_0) \eta(w, \tau|y_0) \quad (255)$$

$$= \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left[-\frac{(X - \mu\bar{X})^2}{2\sigma_X^2}\right], \quad (256)$$

where  $\sigma_X^2$ ,  $\bar{X}$ , and  $\mu(\tau)$  are exactly as above. The only difference between the steady-state and the transient distribution of values of  $X$  is, we see, the factor  $\mu$  appearing inside the exponent in the latter. This small difference has the effect that the FT for power delivered in the transient case is satisfied for all positive values of  $\tau$ :

$$\frac{1}{\tau} \ln \frac{p_\tau^C(+X)}{p_\tau^C(-X)} = \frac{1}{\tau} \cdot \frac{2\mu\bar{X}X}{\sigma_X^2} = \beta X. \quad (257)$$

**Detailed fluctuation theorem.** We now show that our model satisfies the detailed fluctuation theorem (DFT). As mentioned in the Introduction, the DFT is stated in terms of two processes,  $\Pi^+$  (“forward”) and  $\Pi^-$  (“reverse”), related by time-reversal. In the present context, we take  $\Pi^+$  to be the process studied above: a particle is dragged through a thermal medium by a time-dependent potential well,

$$U^+(x, t) = \frac{k}{2}(x - ut)^2. \quad (258)$$

The reverse process,  $\Pi^-$ , is then obtained by moving the well in the opposite direction:

$$U^-(x, t) = \frac{k}{2}(x + ut - u\tau)^2. \quad (259)$$

Formally, we can think of the minimum of the potential as being given by  $\lambda u\tau$ , where  $\lambda$  is our externally controlled parameter. During the process  $\Pi^+$ ,  $\lambda$  is changed uniformly from 0 to 1; during  $\Pi^-$ , from 1 to 0.

Again taking  $y$  to be the displacement of the particle relative to the instantaneous minimum of the potential, we define  $f_+(y, w, t|y_0)$  and  $f_-(y, w, t|y_0)$  to be the joint probability distributions for attaining  $(y, w)$  at time  $t$ , given  $y_0$  at time  $t$ , for the two processes. We then solve for these two distributions exactly as we solved for  $f(y, w, t|y_0)$  in the first subsection. The solutions are:

$$f_+(y, w, t|y_0) = f(y, w, t|y_0) \quad (260)$$

$$f_-(y, w, t|y_0) = f(y, w, t|y_0)_{u \rightarrow -u}, \quad (261)$$

where  $f$  is just the solution obtained in the first subsection. Thus, the solution for the forward process is identical to the solution of the first subsection (as it must be, since  $\Pi^+$  is exactly the process studied there!), whereas the solution for the reverse process is obtained from  $f$  by replacing  $u$  by  $-u$  everywhere, *including in the definition of  $l$*  (hence,  $l \rightarrow -l$ ). This replacement is easily understood: in terms of the variable  $y$ , the process  $\Pi^-$  is no different than that obtained by starting with the minimum of the potential at  $x = 0$ , and moving it with velocity  $-u$  for a time  $\tau$ .

Let us now put these solutions to good use.

The DFT, in the context of this problem, claims the following:

$$\frac{P_+(y_B, +\Delta S|y_A)}{P_-(y_A, -\Delta S|y_B)} = \exp \Delta S. \quad (262)$$

Here,  $P_\pm(y_f, \Delta s|y_i)$  denote the joint probability distributions of finding the particle at a final point  $y_f$ , and a value of entropy generated  $\Delta s$ , given an initial location  $y_i$ , for the forward and reverse processes.<sup>20</sup> Now consider a single realization of either process.

---

<sup>20</sup>Strictly speaking, these should be defined in terms of absolute locations  $x_A$  and  $x_B$ , but the change of variables to relative displacements  $y_A$  and  $y_B$  is immediate.

Energy conservation, along with our definition of entropy produced (Eqs.246 and 247), give us the following relation between  $y_i$ ,  $y_f$ ,  $\Delta s$ , and the work  $w$  performed:

$$w - \beta^{-1}\Delta s = \frac{k}{2}(y_f^2 - y_i^2) \equiv \Delta U(y_i, y_f). \quad (263)$$

We can then re-express  $P_{\pm}$  in terms of  $f_{\pm}$ :

$$P_{\pm}(y_f, \Delta s | y_i) = \int dw f_{\pm}(y_f, w, \tau | y_i) \delta(\Delta s + \beta \Delta U - \beta w) \quad (264)$$

$$= \beta^{-1} f_{\pm}(y_f, \frac{k}{2}(y_f^2 - y_i^2) + \beta^{-1}\Delta s, \tau | y_i), \quad (265)$$

from which we obtain the following explicit expressions for the numerator and denominator in Eq.262:

$$P_+(y_B, +\Delta S | y_A) = \beta^{-1} f_+(y_B, +\Omega, \tau | y_A) \quad (266)$$

$$P_-(y_A, -\Delta S | y_B) = \beta^{-1} f_-(y_A, -\Omega, \tau | y_B), \quad (267)$$

where

$$\Omega \equiv \frac{k}{2}(y_B^2 - y_A^2) + \beta^{-1}\Delta S. \quad (268)$$

Eq.262 is thus equivalent to the following relation:

$$\frac{f_+(y_B, +\Omega, \tau | y_A)}{f_-(y_A, -\Omega, \tau | y_B)} = \exp\left[\beta\Omega - \frac{\beta k}{2}(y_B^2 - y_A^2)\right], \quad (269)$$

where, if Eq.262 is to be valid for all real  $\Delta S$ , then Eq.269 must hold for all real  $\Omega$ . Using the expression for  $f$  given in the last subsection of this section, one can verify that, indeed, Eq.269 is valid for all values of  $\Omega$ , hence the DFT is satisfied by our model.

**Generalization.** Before proceeding to the nonequilibrium work relation, we point out that our model easily generalized to include an additional uniform, constant external force. Namely, suppose we modify the time-dependent potentials  $U^{\pm}(x, t)$  for the forward and reverse processes, as follows:

$$U^{\pm}(x, t) \rightarrow U^{\pm}(x, t) + \alpha x \quad , \quad \alpha > 0. \quad (270)$$

This corresponds to subjecting the particle to an additional leftward-pushing force, of magnitude  $\alpha$ . Thus, assuming  $u > 0$ , the particle is dragged “up” the potential energy slope  $\alpha x$  (i.e. *against* the additional force) during  $\Pi^+$ , and “down” the slope during  $\Pi^-$ .

The solution in this case, for the forward process  $\Pi^+$ , is the same as that in the previous subsection, except that  $l$  is everywhere replaced by

$$l_{\alpha} = l + \frac{\alpha}{k} = \frac{\alpha + \gamma u}{k}. \quad (271)$$

Thus, in the steady state, the average position of the particle is displaced by an amount  $\alpha/k$  to the left, relative to the case with no additional force (Eq.245a). This implies that

additional work is performed on the particle at an average rate  $u\alpha$  (Eq.245b); this is simply the average rate at which we drag the particle up the slope.

For the reverse process, the solution is obtained (as in the previous subsection) by the further replacement  $u \rightarrow -u$ , including in the definition of  $l_\alpha$ .

It is straightforward to show that, with these replacements, the DFT remains valid.

**Nonequilibrium work relation for free energy differences.** In Ref. [99] the following *nonequilibrium work relation* for free energy differences was established:

$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta F}. \quad (272)$$

This pertains to a situation in which a system of interest, coupled to a heat reservoir at temperature  $\beta^{-1}$ , evolves in time as an external parameter  $\lambda$  is switched *at a finite rate* from an initial to a final value, say, from 0 to 1.  $W$  is the work performed during one realization of this process; the angular brackets denote an average over a statistical ensemble of such realizations, and  $\Delta F$  is the free energy difference between the equilibrium states (at temperature  $\beta^{-1}$ ) corresponding to the values  $\lambda = 0$  and  $\lambda = 1$ . It is assumed that, at the start (though not necessarily at the end) of each realization, the system is in equilibrium with the reservoir. Let us use the results derived above to show explicitly that Eq.272 is satisfied by our model. We will consider the forward process defined by  $U^+(x, t)$  in the previous subsection (Eq.270). Formally, as before, we let  $\lambda u\tau$  define the minimum of the confining potential itself, so that we move that minimum from 0 to  $u\tau$  by changing  $\lambda$  from 0 to 1:

$$U_\lambda(x) = \frac{k}{2}(x - \lambda u\tau)^2 + \alpha x \quad (273)$$

$$U^+(x, t) = U_{\lambda(t)}(x) \quad , \quad \lambda(t) = t/\tau. \quad (274)$$

For any value of  $\lambda$ , and at fixed temperature  $\beta^{-1}$ , there exists an equilibrium state of the system, defined microscopically by a canonical distribution. The associated free energy  $F_\lambda$  is then defined in terms of the logarithm of the corresponding partition function:

$$F_\lambda = -\beta^{-1} \ln \int dx e^{-\beta U_\lambda(x)} \quad (275)$$

$$= \alpha \lambda u\tau - \frac{\alpha^2}{2k} + \frac{1}{2\beta} \ln \frac{\beta k}{2\pi}. \quad (276)$$

Hence the free energy difference is simply

$$\Delta F = F_1 - F_0 = \alpha u\tau. \quad (277)$$

Physically, this is the work required to *reversibly* change  $\lambda$  from 0 to 1, at constant temperature.

Now consider a statistical ensemble of realizations of our finite-time, irreversible process, with initial conditions sampled from the canonical ensemble corresponding to  $\lambda = 0$ . The associated distribution of values of work,  $W = w(\tau)$ , can be written as:

$$\eta(W) = \int dy_0 \rho_\alpha^C(y_0) \eta(W, \tau | y_0), \quad (278)$$



where  $\rho_\alpha^C(y_0)$  is the canonical distribution of initial conditions (for finite  $\alpha$ ), and  $\eta(W, \tau|y_0)$  is the distribution of work values at time  $\tau$ , given initial condition  $y_0$ . We can use the results of the first subsection, with the substitution  $l \rightarrow l_\alpha$ , to compute this integral explicitly. Once again skipping the algebra, we present the result:  $\eta(W)$  is a Gaussian distribution of mean and variance

$$\langle W \rangle = \alpha u \tau + \mu \gamma u^2 \tau \quad (279)$$

$$\sigma_W^2 = 2\mu \gamma u^2 \tau / \beta, \quad (280)$$

with  $\mu(\tau)$  as defined by Eq.253. Thus,  $\eta(W)$  is a Gaussian whose mean and variance are related by

$$\langle W \rangle = \Delta F + \beta \sigma_W^2 / 2. \quad (281)$$

This immediately implies (see, e.g. Eq.[12] of Ref. [99]) that the nonequilibrium work relation for free energy differences, Eq.272 above, is satisfied.

**Explicit expressions.** Here we give an explicit expression for the function  $f(y, w, t|y_0)$  introduced in the first subsection. This solution is essentially the combination of Eqs.240 and 242:

$$f(y, w, t|y_0) = \frac{\sqrt{C}}{2\pi} \exp(AC/2), \quad (282)$$

where

$$C = \det \mathbf{C} = -\frac{\beta^2 k^2}{2\gamma u^2 \nu_- [2\gamma \nu_- + kt \nu_+]} \quad (283)$$

$$A = (1/\beta k) \left\{ \nu_+ \nu_- (w - kltu)^2 + \gamma^2 u^2 \nu_- (y - y_0) [4l \nu_- + (3\nu - 1)y_0 + (\nu - 3)y] \right. \quad (284)$$

$$\left. + 2\gamma u \left( l^2 kut \nu_-^2 - l \nu_- [2w \nu_- + ktu \nu_+ (y_0 - y)] - w \nu_-^2 (y_0 + y) - ktu (y - \nu y_0)^2 \right) \right\}, \quad (285)$$

and

$$\nu = \exp(-kt/\gamma) \quad , \quad \nu_\pm = \nu \pm 1. \quad (286)$$

## REFERENCES

- [1] Y.Arimoto et al, Phys. Rev. C55, R1011 (1997)
- [2] W.J.Świątecki, Physica Scripta **24**, 113 (1981).
- [3] J.Łocki, H.Feldmeier and W.J.Swiątecki, Nucl. Phys. A459, 145 (1986)
- [4] S.Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943);
- [5] C.Dellago, P.G.Bolhuis, F.S.Csajka, and D. Chandler, J.Chem. Phys. 108, 1964 (1998)
- [6] J.Łocki, R.Łaneta, J.Brzychczyk, and K.Grotowski, Z. Phys. A-Hadrons 341, 307 (1992)
- [7] J.Łocki and W.J.Swiątecki, report LBL-12811 (1982)
- [8] W.D.Myers, W.J.Swiątecki, Nucl. Phys 81, 1 (1966)
- [9] Świątecki, W.J., J.Phys. (Paris) 33,C5-45, (1972)
- [10] W.J.Swiątecki, S.Björnholm, Phis. Rep. 4, 325 (1972)
- [11] J.Łocki, Y.Boneh, J.R.Nix, J.Randrup, M.Robel, A.J.Sierk, W.J.Swiątecki, Ann. Phys. (NY) 113, 330 (1978)
- [12] P.Bonche, S.E.Koonin, J.W.Negele, Phys. Rev. C13, 1226 (1976)
- [13] R.Bass, Nuclear Reactions with Heavy Ions, Springer-Verlag, 1980.
- [14] R.Y.Cusson, J.Maruhn, Phys. Lett. 62B, 134 (1976)
- [15] P.Bonche, J.Phys. 37, C5-213 (1976)
- [16] R.Y.Cusson, R.K.Smith, J.Maruhn, Phys. Rev. Lett. 36, 1166 (1976)
- [17] S.E.Koonin et al., Phys. Rev. C15, 1539 (1977)
- [18] V.Maruhn-Rezwani, K.T.R.Davis, S.E.Koonin, Phys. Lett. 67B, 134 (1977)
- [19] W.Nörenberg, Z.Phys. A274, 241 (1975)
- [20] S.Ayik, B.Schürmann, W.Nörenberg, Z.Phys. A277, 299 (1976)
- [21] W.Nörenberg, Z.Phys. A276, 84 (1976)
- [22] S.Ayik, B.Schürmann, W.Nörenberg, Z.Phys. A279, 145 (1976)
- [23] H.Hofmann, C.Ngô, Phys. Lett. 65B, 97 (1976)
- [24] W.Nörenberg, J.Phys. (Paris) 37, C5-141 (1976)
- [25] H.Hofmann, P.J.Siemens, Nucl. Phys. A275, 464 (1977)
- [26] C.M.Ko, H.J.Pirner, H.A.Weidenmüller, Phys. Lett. 62B, 248 (1976)
- [27] S.Ayik, B.Schürmann, W.Nörenberg, Z.Phys. A286, 271 (1978)
- [28] B.Schürmann, W.Nörenberg, M.Simbel, Z.Phys. A286, 263 (1978)
- [29] H.Hofmann, P.J.Siemens, Nucl. Phys. A257, 165 (1976)
- [30] D.Agassi, C.M.Ko, H.A.Weidenmüller, Ann. Phys. (NY) 107, 140 (1977)
- [31] D.Agassi, C.M.Ko, H.A.Weidenmüller, Phys. Rev. C18, 223 (1978)
- [32] J.McIver and A.Komornicki, J. Am. Chem. Soc. 94, 2625 (1972)
- [33] C.J.Cerjan and W.H.Miller, J. Chem. Phys. 92, 5580 (1990)
- [34] R.Elber, D.P.Chen, D.Rojewska, and R.Eisenberg, Biophys. J. 68, 906 (1995)
- [35] L.R.Pratt, J. Chem. Phys. 85, 5045 (1986)
- [36] R.Elber and M.Karplus, Chem. Phys. Lett 139, 375 (1987)
- [37] E.M.Sevick, A.T.Bell, and D.N.Theodorou, J. Chem. Phys. 98, 3196 (1993)
- [38] R.Czerminski and R.Elber, Int. J. Quant. Chem. 24, 167 (1990)
- [39] G.Mills and H.Jónsson, Phys. Rev. Lett. 72, 1124 (1994)
- [40] R.E.Gillian and K.R.Wilson, J. Chem. Phys. 97, 1757 (1992)
- [41] E.A.Carter, G.Ciccotti, J.T.Hynes, and R.Kapral, Chem. Phys. Lett. 156, 472 (1989)
- [42] N.Metropolis, A.W.Rosenbluth, M.N.Rosenbluth, A.H.Teller, and E.Teller, Equation of state calculations for fast computing machines, J.Chem.Phys. 21, 1087-92 (1953)

- [43] L.D.Landau, E.M.Lifshitz, Statistical Physics
- [44] J.G.Kirkwood J.Chem. Phys. **3**, 300 (1935)
- [45] C.Jarzynski, Phys. Rev. Lett., **78**, 2690 (1997)
- [46] C.Jarzynski, private communications.
- [47] C.Jarzynski, Phys. Rev. E **V56 N5**, 5018 (1997)
- [48] C.Jarzynski, Phys. Rev. Lett. **V74 N15**, 2937 (1995)
- [49] C.Jarzynski, Phys. Rev. A **46**, 7498 (1992)
- [50] C.Jarzynski, Phys. Rev. E **V48**, 4340 (1993)
- [51] C.Jarzynski, Ph.D. Thesis, Lawrence Berkeley Lab. (1995)
- [52] R.P.Feynman, A.R.Hibbs, Quantum Mechanics and Path Integrals
- [53] 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 and S.Machlup, Phys. Rev. **91**, 1505 (1953) and Phys.Rev.**91**, 1512 (1953).
- [54] A.Bohr, B.R.Mottelson, Nuclear Structure, Vol. 1, W.A.Benjamin, Inc. 1969
- [55] R.Kubo, M.Toda, N.Hashitsume, Statistical Physics II, Nonequilibrium Statistical Mechanics, The Fluctuation-Dissipation Theorem, Springer Series in Solid-State Science **V31** (1998)
- [56] R.Kubo, M.Toda, N.Hashitsume, Statistical Physics II, Nonequilibrium Statistical Mechanics, Fokker-Plank Equation, Springer Series in Solid-State Science **V31** (1998)
- [57] P.Fröbrich, I.I.Gontchar, Phys. Rep. **292**, 131 (1998)
- [58] Yu.L.Klimontovich, Statistical Teory of Open Systems, **V1**, Kluwer Academic Publishers **V67** (1995)
- [59] I.Kelson, Phys.Rev. B **136**, 1667 (1964)
- [60] H.Krappe, J.R.Nix, A.J.Sierk, Phys. Rev. C **20**, 992 (1979)
- [61] W.D.Myers, W.J.Swiatecki, Art. Fys **36**, 343 (1967)
- [62] Y.Aritomo, T.Wada, M.Ohta and Y.Abe, Phys. Rev. C **V52 (2)** Feb (1999)
- [63] J.Randrup, W.J.Swiatecki, Nucl. Phys. A **429**, 105 (1984)
- [64] H.Feldmeier, H.Spangenberg, Nucl. Phys. A **428**, 223 (1984)
- [65] H.Feldmeier, Rep. Prog. Phys. **50**, 915 (1987)
- [66] J.Randrup, Nucl. Phys. A **307**, 319 (1978)
- [67] C.E.Aguiar, V.C.Barbosa, and R.Donangelo, Nucl.Phys.A **514**, 205 (1990).
- [68] C.W.Gardiner, Handbook of Stochastic Methods, Springer-Verlag, 1983.
- [69] A.Wieloch, Z.Sosin and J.Blocki, Acta Phys. Pol. B **V30** (1999) 1087
- [70] R.P.Feynman, R.B.Leighton, and M.Sands, *The Feynman Lectures on Physics* (Addison-Wesley, Reading, MA, 1966), Vol.I, Chap.46.
- [71] M.Smoluchowski, Phys.Z **13**, 1069 (1912); “Vortgage uber die Kinetische Theorie der Materie und der Elektrizitat”, ed. M.Planck (Teubner und Leipzig, Berlin, 1914), pp. 89-121.
- [72] M.O.Magnasco, Phys.Rev.Lett. **71**, 1477 (1993).
- [73] R.D.Astumian and M.Bier, Phys.Rev.Lett. **72**, 1766 (1994).
- [74] J.Prost, J.-F.Chauwin, L.Peliti, and A.Ajdari, Phys.Rev.Lett. **72**, 2652 (1994).
- [75] J.-F.Chauwin, A.Ajdari, and J.Prost, Europhys.Lett. **27**, 421 (1994).
- [76] J.Rousselet, L.Salome, A.Ajdari, and J.Prost, Nature (London) **370** 446 (1994).

- [77] L.P.Faucheux, L.S.Bourdieu, P.D.Kaplan, and A.J.Libchaber, Phys.Rev.Lett. **74**, 1504 (1995).
- [78] C.R.Doering, Nuovo Cimento **D17**, 685 (1995).
- [79] R.D.Astumian, Science **276**, 917 (1997).
- [80] M.Bier, Contemporary Physics **38**, 371 (1997).
- [81] F.Jülicher, A.Ajdari, and J.Prost, Rev.Mod.Phys. **69**, 1269 (1997), Section II.
- [82] P.Hänggi and R.Bartussek, in *Nonlinear Physics of Complex Systems: Current Status and Future Trends*, p.294. Edited by J.Parisi, S.C.Müller, and W.Zimmermann (Springer-Verlag, Berlin, 1996).
- [83] R.D.Astumian and M.Bier, Biophys.J. **70**, 637 (1996).
- [84] L.Schimansky-Geier, M.Kschischo, and T.Fricke, Phys.Rev.Lett. **79**, 3335 (1997).
- [85] J.M.R.Parrondo and P.Espanol, Am.J.Phys. **64**, 1125 (1996).
- [86] K.Sekimoto, J.Phys.Soc.Jap. **66**, 1234 (1997).
- [87] N.Metropolis *et al*, J.Chem.Phys.**21**, 1087 (1953).
- [88] L.Onsager, Phys.Rev. **37**, 405 (1931); Phys.Rev. **38**, 2265 (1931).
- [89] D.J. Evans, E.G.D. Cohen, and G.P. Morriss, Phys. Rev. Lett. **71**, 2401 (1993).
- [90] D.J. Evans and D.J. Searles, Phys. Rev. E **50**, 1645 (1994).
- [91] G. Gallavotti and E.G.D. Cohen, J. Stat. Phys. **80**, 931 (1995).
- [92] J.L. Lebowitz, H. Spohn, J. Stat. Phys. **95**, 333 (1999).
- [93] C. Maes, J. Stat. Phys. **95**, 367 (1999).
- [94] G.E. Crooks, Phys. Rev. E **60**, 2721 (1999).
- [95] C. Jarzynski, to appear in J. Stat. Phys.
- [96] K. Sekimoto, J. Phys. Soc. Jpn. **66**, 1234 (1997).
- [97] To motivate this definition of external work, see C. Jarzynski, Acta Phys. Pol. B **29**, 1609 (1998).
- [98] J. Kurchan, J. Phys. A **31**, 3719 (1998).
- [99] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
- [100] J.L. Lebowitz, H. Spohn, J. Stat. Phys. **95**, 333 (1999).
- [101] D.H.E.Gross, H.Kalinowski, Phys. Rep. C 45 (1978) 175
- [102] P.Fröbrich, Phys. Rep. C 116 (1984) 337
- [103] J.Birkelund, L.E.Tubbs, J.R.Huizenga, J.N.De, D.Sperber, Phys. Rep. C 56 (1979) 107
- [104] Christelle Stodel, Ph.D. dissertation, LPC-GANIL
- [105] W.Reisdorf, F.P.Hessberger, K.D.Hildenbrand, S.Hofmann, G.Münzenberg, K.H.Schmidt, W.F.W.Schneider, K.Sümmerer, G.Wirth, J.V.Kratz, K.Schlitt, C.C.Sahm, Nucl. Phys. A 444 (1985) 154
- [106] P.Armbruster, private communication

**The material presented in this thesis is partly contained in the following publications:**

- [107] O.Mazonka and C.Jarzynski, J.Stat.Phys., submitted for publication
- [108] C.Jarzynski and O.Mazonka, Phys.Rev.E, V59 (1999) 6448
- [109] O.Mazonka, C.Jarzynski and J.Blocki, Nucl.Phys.A 641 (1998) 335.
- [110] O.Mazonka, J.Blocki and J.Wilczyński Acta Phys.Pol. B, V30, 469 (1999).
- [111] O.Mazonka, J.Blocki and C.Jarzynski Acta Phys.Pol. B, V30, 1577 (1999).

# Acknowledgments

I was extremely lucky to get a possibility to come to the Institute for Nuclear Studies, Warsaw for Ph.D. studies. I would like to express my gratitude to the Institute for hospitality and especially to the director prof. Z.Sujkowski and a leader of Ph.D. studies prof. L.Lukaszuk. This period of time I spent with people who widened my horizons, taught and inspired interests in me to many wonderful things.

I am happy that I have met my supervisor Jan Blocki, because Jan is not only talented scientist and skillful leader but a sympathetic and kind person. By his own example Jan has taught me to work. I will value forever his trust and sincerity, his friendly attitude he showed towards me and my family, as well his helpful advises and deeds in hard for us moments.

With no doubt I can call Chris Jarzynski my second advisor. Though separated by geography we spent a lot of time together discussing a wide range of topics. Chris has introduced me to statistical physics, and I have learned from him many beautiful ideas in this field.

I have benefited from the seminars, discussions and the creative atmosphere at our Institute. I would like to thank all the members and specially to Janusz Wilczynski and Janusz Skalski.

I am grateful to my colleagues in Krakow, prof. Stanislaw Drozd, Andrzej Wieloch and Zbigniew Sosin for interesting discussion and their hospitality during my visits in Krakow.

Special thanks to my friends Oleg Utyuzh, Andrej and Mirosława Pasternak, Alexandr Undynko, Igor Muntian, Boria Sidorenko, Vladimir Pascalutsa.

My thanks go also to my parents and my family. I owe my wife a lot for her love and support. She accepted my work with understanding and often helped me with wise woman's advice.

The financial support of Maria Skłodowska-Curie Fund (PAA/NSF-96-253 and PAA/DOE-98-34) is greatly acknowledged.

