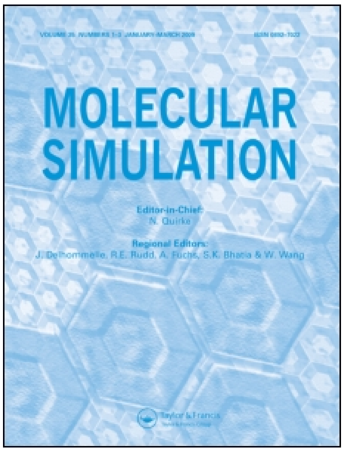

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

System Relaxation and Thermodynamic Integration

Jürgen Schlitter[a]; Dirk Husmeier[a]
[a] Institut für Biophysik, Ruhr-Universität, Bochum, Germany

To cite this Article Schlitter, Jürgen and Husmeier, Dirk(1992) 'System Relaxation and Thermodynamic Integration', Molecular Simulation, 8: 3, 285 — 295

To link to this Article: DOI: 10.1080/08927029208022483

URL: http://dx.doi.org/10.1080/08927029208022483

Molecular Simulation, 1992, Vol. 8, pp. 285–295
Reprints available directly from the publisher
Photocopying permitted by license only

SYSTEM RELAXATION AND THERMODYNAMIC INTEGRATION

JÜRGEN SCHLITTER AND DIRK HUSMEIER

Institut für Biophysik, Ruhr-Universität, D-4630 Bochum, Germany

(Received March 1991, accepted March 1991)

The thermodynamic integration (TI) method for calculating free energy differences has three inherent problems: statistics, numerical integration, and relaxation. In this paper the latter is analyzed for the nonequilibrium TI method introduced by Postma which combines molecular dynamics simulation and TI in a very effective way. A nontrivial extrapolation technique is presented to remove the relaxation error and to calculate the underlying relaxation time. It is shown that the optimal choice of grid points, derived in a previous paper, for minimizing statistical errors not only removes integration errors, but also minimizes relaxation errors. The methods are applied in a calculation of the free energy of cavity formation in water.

KEY WORDS: Thermodynamic integration, perturbation methods, free energy

1 INTRODUCTION

In the last few years numerical methods for calculating free energy differences have become an important tool in several fields from the theory of liquids to biochemistry and drug design. Among the various methods [1], thermodynamic integration (TI) as introduced by Kirkwood [2] has recently generated great interest. One reason is that it is easily combined with the powerful method of molecular dynamics computer simulation (MD) which enables one to simulate a system in time at constant temperature and/or pressure. A useful detailed introduction to methods and applications is found in two recent review articles [3, 4].

For calculating a free energy difference by TI one has to perform the transition from the initial to the final state by a continuous change of the energy function, and obtain the desired result by the integration of equilibrium averages calculated during the transition. As discussed in previous work [5], three types of errors arise when TI is performed in practice: (i) statistical errors, (ii) integration errors due to singularities of the integrand, and (iii) relaxation errors of the averages. It was shown previously that in the framework of linear TI the statistical error can be minimized by a unique presciption for the distribution of grid points, and the same prescription removes the integration error. The present paper deals with the relaxation error which naturally occurs when the calculation of equilibrium averages is combined with a change of the system Hamiltonian.

In some previous applications of the method, see e.g. [6–8], equilibration and sampling periods of sufficient length are repeated in order to compute reliable estimates for the required averages. Postma [9] was the first to try the extreme case of minimum length, i.e. by smoothly changing the Hamiltonian after each time step during the MD simulation and replacing the mean by the actual value of the TI integrand. Numerical investigations have shown that convergence is achieved with increasing total length of the simulation [10]. We shall discuss the latter version of TI on a general level in order to gain some insight into the underlying mechanism. The resulting methods are applied in a numerical calculation of the free energy for cavity formation in water.

2 NONEQUILIBRIUM THERMODYNAMIC INTEGRATION

2.1 Basic relations

The method of thermodynamic integration was introduced by Kirkwood [2] for calculating the free energy difference between two states of a system which can be defined by their respective Hamiltonians H_0 and H_1. In its simplest and most widely used form called linear TI, one starts from an intermediate Hamiltonian $H(\lambda) = H_0 + \lambda \Delta H$, where $\Delta H = H_1 - H_0$ is the difference Hamiltonian, calculates mean values $f(\lambda) = \langle \Delta H \rangle_\lambda$ for an ensemble governed by $H(\lambda)$, and obtains the free energy difference ΔA by the integral

$$\Delta A = \int_0^1 f(\lambda)\, \mathrm{d}\lambda \tag{2.1}$$

To distinguish this approach to that described below it will henceforth be denoted as equilibrium TI (ETI) since $f(\lambda)$ must be the true equilibrium mean of the perturbation ΔH. This makes the calculation costly in all cases where averages are computed by simulation methods like Monte Carlo or molecular dynamics. After each step on the way from H_0 to H_1 the system has to relax towards equilibrium before an average can be taken. On the other hand ETI has proven to be very successful in treating even singular terms arising in ΔH and f if the integration is performed in the right way [6].

Postma [9] proposed and first applied a method which abandons the concept of equilibrium and integrates TI in the MD simulation of a system in time. By analogy with the above method this nonequilibrium TI (NETI) can be described as follows. The transition between the states is performed during a simulation time T in a MD run. The perturbation parameter is therefore defined by a monotonous function $\lambda(t)$ with $\lambda(0) = 0$ and $\lambda(T) = 1$. It determines the time-dependent Hamiltonian

$$H(t) = H_0 + \lambda(t)\Delta H \tag{2.2}$$

Instead of the above equilibrium average, a time-dependent nonequilibrium ensemble average

$$\phi(t) = \langle \Delta H \rangle_{\lambda,t} \tag{2.3}$$

is inserted in the perturbation integral which now reads

$$\Delta A_T = \int_0^T \phi(t)\, \mathrm{d}\lambda/\mathrm{d}t\, \mathrm{d}t \tag{2.4}$$

A rough justification for this proceeding is given by the argument that during each of the many (10^3–10^4) small time steps of a typical MD run the Hamiltonian is changed only little so that the system may be considered as always being near to equilibrium. More precisely, when the simulation period T is gradually enhanced with a simultaneous redefinition of $\lambda(t)$, one expects ΔA_T to converge to the desired value $\Delta A_\infty = \Delta A$.

In order to stay close to the numerical practice, we will rewrite the integrals as sums and neglect completely integration errors. We consider MD trajectories with fixed time step Δt where all values with index i are taken at $t_i = i\Delta t$, $i = 0, 1, \ldots, N$. In particular one has $T = N\Delta t$ and

$$\begin{aligned} \Delta\lambda_i &= \lambda_i - \lambda_{i-1} \qquad \text{for} \qquad i = 1, \ldots, N \\ \phi_i &= \phi(t_i) \\ f_i &= f(\lambda_i) \end{aligned} \tag{2.5}$$

At $t = 0$ and $\lambda = 0$, i.e. before the perturbation is switched on, one has of course $\phi_0 = f_0$. The exact free energy difference and the NETI approximation thus read

$$\Delta A = \sum f_i \Delta\lambda_i \qquad \text{and} \qquad \Delta A_T = \sum \phi_i \Delta\lambda_i \tag{2.6}$$

In the theoretical analysis of relaxation we shall ignore the fact that $\phi(t)$ is usually approximated by the instantaneous values of the perturbation, $(\Delta H)_{\lambda(t)}$, which estimates the ensemble mean by just one sample value and bears a statistical error like any other estimate. The present work aims at discussing and eliminating the relaxation error originating from the systematic deviation from equilibrium which is obviously inherent in the NETI approach. We note, however, the important result of our previous investigation of ETI [5] that by an appropriate choice for the density of grid points, $dn/d\lambda$, the statistical error $\sigma_{\Delta A}$ is minimized and the integration error is eliminated. By analogy the best results are expected in NETI if $\lambda(t)$ is chosen according to the prescription

$$dt/d\lambda \propto dn/d\lambda \propto |d\langle \Delta H \rangle_\lambda / d\lambda|^{1/2} \tag{2.7}$$

2.2 *Relaxation error*

The difference between the expressions for the free energy difference from NETI and ETI,

$$\delta_T = \Delta A_T - \Delta A = \sum (\phi_i - f_i)\Delta\lambda_i \tag{2.8}$$

depends on the difference between the time-dependent means ϕ_i and the equilibrium means f_i. It would vanish if after each perturbation step the system could relax over a sufficiently long period instead of the short interval Δt, usually about 1 fs, before the mean is calculated, and hence is a pure relaxation error. This idea may be expressed by an equation

$$\phi_i = r(\Delta t)\phi_{i-1} + (1 - r(\Delta t))f_i \tag{2.9}$$

where $r(t')$ is a relaxation function with $r(0) = 1$ and $r \to 0$ for $t' \to \infty$. Immediately after a perturbation step, ϕ_i has its former value ϕ_{i-1} and then tends towards the equilibrium value f_i. In order to simplify the following analysis we will make the assumption that the relaxation is homogeneous in time, i.e. the relaxation function is the same for each perturbation step and hence independent on the time t and the index i. The assumption will be discussed later, and it will be shown that it can be checked to some extent by the information obtainable from the simulations. Next we expand r as

$$r(\Delta t) = 1 - \Delta t/r + O_2(\Delta t/\tau) \tag{2.10}$$

and retain only the linear term. This is generally valid in practice as τ, which has the meaning of an effective relaxation time of the perturbation ΔH, is certainly much larger than the usual time step $\Delta t \sim 1$ fs. By multiplying (2.9) by $\Delta\lambda_i$ and summing over all time steps one obtains, with the abbreviation $\Delta\phi_i = \phi_i - \phi_{i-1}$ and again neglecting terms nonlinear in $(\tau/\Delta t)$, a simple relation between the free energy expressions (2.6)

$$\Delta A_T = \Delta A - (\tau/\Delta t) \sum \Delta\phi_i \Delta\lambda_i \tag{2.11}$$

which implies a relaxation error

$$\delta_T = -(\tau/\Delta t) \sum \Delta\phi_i \Delta\lambda_i \tag{2.12}$$

Let us now consider the case of equidistant perturbation steps with $\Delta\lambda_i = 1/N$ where

$$\delta_T = (f_0 - \phi_N)\tau/T = (f_0 - f_N)\tau/T - O_2(\tau/T) \tag{2.13}$$

It can be shown [1] that $f(\lambda)$ and hence also $\phi(t)$ are monotonously decreasing functions, and it follows from (2.9) that $\phi_N \geqslant f_N$. Consequently the relaxation error is always positive and the free energy difference is overestimated by using ΔA_T. Reliable results are expected only if the simulation is performed over a period much longer than the relaxation time, which is a serious, but not unexpected limitation of the method.

For most applications equidistant steps $\Delta\lambda_i$ are not the best choice or even not applicable [5]. In general, therefore, the correct equation (2.12) containing a complicated dependence on T in all terms yields a more complicated behaviour, but it is still possible to give an upper bound for the relaxation error, namely

$$\delta_T = (\tau/\Delta t) \sum (-\Delta\phi_i/\Delta\lambda_i)(\Delta\lambda_i)^2 \lesssim \beta N \sigma_{\Delta A}^2 \tau/T \tag{2.14}$$

This very informative relation connects the relaxation error (2.12) with the statistical error of the free energy ΔA and is readily derived from the fact that ϕ decreases more slowly than f, $-\Delta\phi/\Delta\lambda \leqslant -\Delta f/\Delta\lambda$, and the equation $\sigma_f^2 = -\beta^{-1}\, \mathrm{d}f/\mathrm{d}\lambda$ [1]. It shows that minimizing the statistical error is the essential means of minimizing all kinds of errors arising in linear TI when τ and T are fixed quantities. As shown in previous work [5] this is achieved by choosing the unique optimal distribution of grid points given in (2.7). In a recent publication a very promising attempt was presented to reduce the relaxation time itself by weakening the relevant activation barriers [11].

3 IMPROVED METHODS

3.1 Forward-backward integration

The fact that the relaxation error is always positive can be used to partially remove it and to give an estimate of ΔA. This was recognized early on and is the common way of coping with the problem of the inherent relaxation. Indeed, if the transition 0 → 1 yields $\Delta A_T = \Delta A + \delta$, and the reverse transition (after inversion of the sign) $\Delta A'_T = \Delta A - \delta'$, then the arithmetic mean of both values is $\Delta A + (\delta - \delta')/2$. The error is partially compensated and of course smaller than $(\delta + \delta')/2 = (\Delta A_T - \Delta A'_T)/2$, which hence is considered as an upper limit of the relaxation error. Forward-backward integration thus yields the improved result

$$\Delta A = (\Delta A_T + \Delta A'_T)/2 \pm (\Delta A_T - \Delta A'_T)/2 \tag{3.1}$$

The argument is based on the assumption that both the forward and backward integration start from equilibrium. If this assumption is seriously violated, which may occur at the second step, then the interpretation of results can be very misleading [3]. Small error limits in (3.1) can be shown to occur not only when the simulation time is sufficiently long, but also when it is too short to allow the system to reach equilibrium.

The main shortcoming of this method arises from the fact that two simulations of the same length are used, which gives an idea of the relaxation error but no insight into the underlying relaxation behaviour. This is only achieved by repeated simulations of different length as done by Straatsma and Berendsen [10].

3.2 Extrapolation methods

The fact that according to (2.13) the relaxation error will vanish for $T \rightarrow \infty$ suggests an expansion in powers of T^{-1}

$$\Delta A_T = \Delta A + bT^{-1} + cT^{-2} + \ldots \tag{3.2}$$

and to calculate ΔA from the results of several simulations by extrapolation to $T^{-1} = 0$. It may even be possible to estimate the relaxation time as $\tau = b/(f_0 - f_N)$, but in general the values f_0 and f_N have a large statistical error or can be shown to be infinite. Hence this naive extrapolation method is not of much use in practice.

The extrapolation method we propose starts from (2.11) which can be written as

$$y_T = \Delta A + (\tau/\Delta t)x_T \tag{3.3}$$

with the abbreviations

$$y_T = \Delta A_T = \sum \phi_i \Delta\lambda_i, \qquad x_T = -\sum \Delta\phi_i \Delta\lambda_i \tag{3.4}$$

(3.3) predicts a linear relationship between the quantities x_T and y_T which can be obtained by MD runs with NETI. According to the above investigations of the relaxation error, x_T vanishes from $T \rightarrow \infty$ and y_T then tends to the desired ΔA. Obviously only two NETI runs with different simulation periods T and their respective values of x_T and y_T are sufficient to solve (3.3) for the interesting quantities ΔA and $\tau/\Delta t$, without any restriction of τ/T.

In practice not the exact means ϕ_i and the corresponding x_T and y_T but only numerical estimates contaminated by statistical errors are available. We hence propose the following scheme consisting of three steps.

(i) calculate x_T and y_T for at least three $T = T_1, T_2, T_3, \ldots$

(ii) perform a regression analysis with the quadratic ansatz

$$y = a + bx + cx^2$$

to determine a, b, and c with their respective statistical errors

(iii) if the results are compatible with the assumption $c = 0$, i.e. $|\sigma_c/c| > 1$, repeat the regression with the linear ansatz $y = a + bx$

The step (ii) allows one to test the assumption of a homogeneous relaxation. If one does not find linearity, the assumption must be discarded and one has to decide in detail whether the extrapolation is well described by the quadratic form. If the linear approximation can be used, step (iii) will yield $a = \Delta A$ and $b = \tau/\Delta t$ with higher accuracy.

The method aims at removing the relaxation error and minimizing the actual statistical error by the regression analysis. The remaining statistical error of ΔA is obtained by considering the error propagation as usual. Comparison with the result from the independent extrapolation (3.2) provides a check of consistency.

We note that the statistical error of y_T depends on T, i.e. on the number of integration steps used. More precisely, when a special time dependence $\lambda = \Lambda(t/T)$ is used with different values of T, then the variance of y_T, σ_T^2, should be proportional to $N^{-1} \propto T^{-1}$ [5]. Hence regression should be performed by giving each y_T a weight factor proportional to T.

4 CAVITY FORMATION IN WATER

In order to illustrate the extrapolation technique we applied NETI to the formation of a cavity in liquid SPC water. The same problem was treated before by Postma, Berendsen and Haak [12] using the windowing technique. The example has the advantage that ΔA can also be compared with scaled-particle theory which yields a rigorous expression for the reversible work done in the process, and that the relaxation behaviour of the system is well known.

We carried out isothermal-isobaric simulations of 216 water molecules in a periodic rectangular box using GROMOS programs [13]. For means of conformity with Postma *et al.* we used the same parameters: a temperature of 305 K, a pressure of 40 atm, time constants of 0.4 and 0.5 ps for the coupling of temperature and pressure, resp., a cut-off distance of 0.9 nm, and a time step of 0.001 ps. After an equilibration period of 40 ps for the pure water, the soft cavity-water potential of the repulsive van der Waals type, $\lambda \Delta V = V_{cw} = \lambda (B/R)^{12}$, was introduced. B was chosen equal to the parameter of the repulsive water–water interaction ($(0.3428\,\mathrm{nm\,mm\,kJ\,mol})^{12}$) which corresponds to a thermal radius of 0.317 nm at 305 K. The transition from the pure solvent state, $\lambda = 0$, to the final state with cavity, $\lambda = 1$, was made using a time dependence $\lambda = (t/T)^8$. This choice yields the minimum statistical error for the repulsive Lennard-Jones potential [5]. Four simulations were performed with

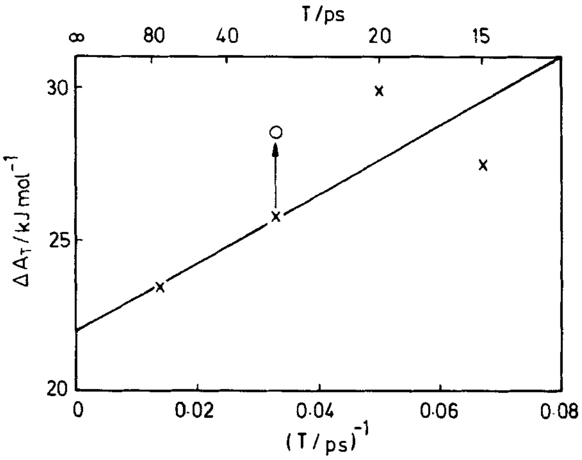

Figure 1 Free energy of cavity formation in water, ΔA_T, as a function of the inverse simulation time, T^{-1}. The straight line was obtained by linear regression from values (x) computed with the optimal distribution of grid points. The arrow points at a value (O) computed with a different distribution, see text.

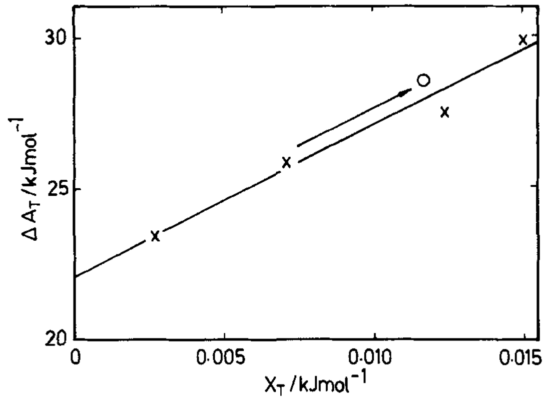

Figure 2 Free energy of cavity formation in water, ΔA_T, as a function of $x_T = -\Sigma\, \Delta\phi_i \Delta\lambda_i$. The straight line was obtained by linear regression from values (x) computed with the optimal distribution of grid points. The equilibrium ΔA is found at $x_T = 0$ by extrapolation, the slope gives the relaxation time as $\tau/\Delta t$. The shift to the larger value (o) at equal simulation time 30 ps, but nonoptimal distribution is marked by the arrow.

different simulation periods (15, 20, 30, 70 ps). Figures 1 and 2 show the resulting ΔA_T vs. T^{-1} and x_T, resp., marked by crosses together with the regression lines.

For the evaluation of (3.3) it is important to recognize that both x_T and y_T are prone to statistical errors. At the linear regression (iii) we therefore consider two extreme cases, the usual form $y = y(x)$ with errors only in y and the alternative form $x = x(y)$ with errors only in x. When (3.2) is used, errors occur only in y. The numerical results are summarized in Table 1.

All types of numerical evaluation yield consistent values for the free energy of cavity formation with satisfactory small errors. As expected, ΔA_T in Figure 1 is not linear in T^{-1} since $|\sigma_c/c| < 1$ which may explain the slightly different values of ΔA from the linear and quadratic regression. By way of contrast the plot ΔA_T vs. x_T in Figure 2 which illustrates our novel extrapolation method, exhibits a clearly linear relation. Evidently the quadratic term can be omitted as its relative error is much greater than unity. The good fit to a straight line can be interpreted as a justification of the assumption that in the present case the relaxation of the perturbation potential ΔV is indeed homogeneous in time.

We find that the relaxation time of ΔV, $\tau = 0.5$ ps, is smaller than the correlation time of SPC water [9] for rotation, 2 ps, and larger than for translation, 0.1 ps, which

Table 1 Evaluation of the results for cavity formation in water shown in Figures 1 and 2 by extrapolation techniques.

Fit function	*Weights*	σ_c/c	τ/ps	$\Delta A/kJ\,mol^{-1}$
$y(T^{-1}) = a + bT^{-1} + cT^{-2}$	$\propto T$	-0.4		19.2 ± 1.1
$y(T^{-1}) = a + bT^{-1}$	$\propto T$			21.9 ± 0.5
$y(x) = a + bx + cx^2$	$\propto T$	$+9.6$	0.47 ± 0.33	22.1 ± 1.0
$y(x) = a + bx$	$\propto T$		0.50 ± 0.06	22.0 ± 0.5
$x(y) = a + by$	const		0.51 ± 0.07	21.9 ± 0.7
Scaled particle theory				23.25
Postma *et al.* [12]				26.3 ± 1.1

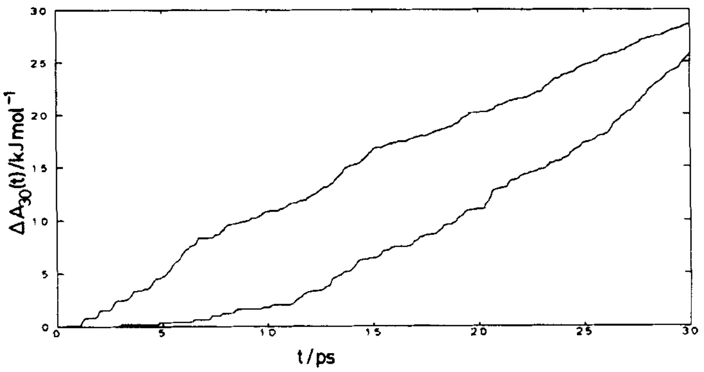

Figure 3 Increase of free energy during two 30 ps runs. The quadratic curve (lower line) is characteristic of the optimal choice of grid points, $\lambda = (t/T)^8$. A linear increase by $\lambda = (t/T)^4$ (upper line) yields a larger value, i.e. a larger relaxation error, see also Figures 1 and 2.

are the microscopic processes induced in the solvent during formation of the cavity. The coincidence with the time constants for coupling to an external bath is merely hazardous as we know from other applications [14]. It is usually argued in literature that these pure system times are the crucial quantities for relaxation in free energy calculations. In the appendix we show at a simple example that instead τ of ΔV must depend on both the system and the particular perturbation applied to it. As τ decreases with increasing strength and steepness of the perturbation potential a short relaxation time for the present Lennard-Jones potential is not surprising.

The inequality (2.14) emphasizes the close relation between the relaxation error and the statistical error which both can be minimized by the choice of the function $\Lambda(t/T)$ determing the distribution of grid points. By theoretical arguments $\lambda = (t/T)^8$, which gives a quadratically increasing $\Delta A_T(t)$, should give minimum errors in the present case [5]. In order to get a numerical confirmation we repeated the 30 ps run with a dependence $\lambda = (t/T)^4$ which produces a linearly increasing $\Delta A_T(t)$, see Figure 3. The resulting $\Delta A_T = \Delta A_T(T)$ is marked by an open circle in Figures 1 and 2 and exhibits an error which is 3 kJ/mol larger than before. This drastic increase is not hazardous in the statistical sense as it might seem in Figure 1 since the additional point fits very well with the straight line in Figure 2. Obviously a lot of computer time can be saved by the right choice of $\Lambda(t/T)$ (about one half in this case) without loss of accuracy, and a linear increase of the free energy is not characteristic of an optimal numerical procedure.

Finally, our ΔA for cavity formation is 1 kJ/mol (4 percent) smaller than the value of scaled particle theory and 4 kJ/mol smaller than calculated by Postma *et al.* [12]. Their result is comparable with our ΔA obtained in the 30 ps run, which may be due to uncomplete relaxation, but a simple comparison of simulation times is not possible because of the different perturbation method used.

5 SUMMARY AND DISCUSSION

The present study is part of a systematic investigation of errors arising in applications of TI which was begun in a previous paper and aims at optimizing the performance and evaluation of the numerical calculations. Of course, all types of errors due to statistics, numerical integration, and relaxation can be made small by sufficiently long computer runs, but it is a theoretical as well as a practical challenge to find optimal strategies for obtaining reliable results at a minimum expense of computer time.

Among the various forms of TI we have studied the method of Postma [9] which in a systematic way incorporates the continuous change of the Hamiltonian and sampling in the MD simulation. It is known that this technique also called slow-growth or continuous coupling technique, works successfully and is superior e.g. to the windowing technique [3]. For the sake of simplicity we analyzed the relaxation error δ_T on the assumption that the relaxation of the interesting perturbation *Hamiltonian* or potential is homogeneous in time, i.e. the relaxation function should be the same at each step. This also implies the definition of an effective relaxation time τ. For δ_T we derived an expression which holds to first order in τ/T, where T is the simulation time, and an upper bound showing how it can be minimized. A new method was developed to calculate the true free energy change ΔA, δ_T, and τ by linear regression and extrapolation. The trivial and less informative extrapolation on time axis for $T \to \infty$ can be used as a check of consistency.

Despite of the fact that the extrapolation method works very well in the numerical example of cavity formation in water, the basic assumption seems questionable for two reasons. Firstly, it certainly works only if the rearrangement of a system after perturbation is dominated by one or several processes of similar correlation time. Secondly, as demonstrated in the appendix different parts of a complicated perturbation potential may have different relaxations even in a homogeneous system. Indeed we found different values of τ at different steps of an internal reaction in a protein [14]. Thus, separate treatment of *a priori* different processes is generally required. In practice the applicability of the assumption can be checked to some extent by the goodness of fit in linear regression.

In any case estimation of the relaxation error requires more than one single simulation. Forward-backward integration is useful when T is much greater than τ (which in most cases will be unknown), but makes no distinction between relaxation and statistical errors. At least three simulations are required for the new extrapolation method which in principle removes the relaxation error and ends up with a residual statistical error of ΔA.

Finally we like to stress the important role of the grid points $\lambda_i = \Lambda(t_i/T)$. In a previous work we developed a prescription for an optimal distribution of grid points in ETI which minimizes the statistical error. Surprisingly it turned out that also the integration error is removed in the same way. Now we can add the new result that also the relaxation error in NETI is minimized by the same choice of grid points which minimizes the statistical error in ETI. This theoretical finding is confirmed by numerical calculations which were discussed above in detail.

Acknowledgment

The computer simulations were performed on the CRAY Y-MP of the HLRZ Jülich and the Cyber 205 of the Rechenzentrum der Ruhr-Universität Bochum. We like to thank these institutions for generous support.

References

[1] M. Mezi and D.L. Beveridge, "Free energy simulations", *Ann. N.Y. Acad. Sci.*, **482**, 1 (1986).

[2] J.G. Kirkwood, "Statistical mechanics of fluid mixtures", *J. Chem. Phys.*, **3**, 300 (1935).

[3] W.F. van Gunsteren, "Methods for calculation of free energies and binding constants: Successes and problems" in: Computation of free energy for biomolecular systems, W.F. van Gunsteren and P.K. Weiner, eds., Escom Science Publishers, Leiden, 1989.
[4] W.F. van Gunsteren and H.J.C. Berendsen, "Computer simulation of molecular dynamics: Methodology, applications, and perspectives in chemistry", *Angew. Chem. Int. Ed. Engl.*, **29**, 992 (1990).
[5] J. Schlitter, "Methods for minimizing errors in linear thermodynamic integration", *Molecular Simulation* (1990), in press.
[6] D.R. Squire and W.G. Hoover, "Monte Carlo simulation of vacancies in rare-gas crystals", *J. Chem. Phys.*, **50**, 701 (1969).
[7] M.R. Mruzik, F.F. Abraham, D.E. Schreiber and G.M. Pound, "A Monte Carlo study of ion-water clusters", *J. Chem. Phys.*, **64**, 481 (1976).
[8] J. Gao, K. Kuczera, B. Tidor and M. Karplus, "Hidden thermodynamics of mutant proteins: A molecular dynamics analysis", *Science*, **244**, 1069 (1989).
[9] J.P.M. Postma, "MD of H_2O: A molecular dynamics study of water", Ph.D. thesis, University of Groningen, Groningen, The Netherlands, 1985.
[10] T.P. Straatsma and H.J.C. Berendsen, "Free energy of ionic hydration: Analysis of a thermodynamic integration technique to evaluate free energy differences by molecular dynamics", *J. Chem. Phys.*, **89**, 5876 (1988).
[11] A.E. Mark, W.F. van Gunsteren and H.J.C. Berendsen, "Calculation of relative free energy via indirect pathways", *J. Chem. Phys.*, (1990) in press.
[12] J.P.M. Postma, H.J.C. Berensen and J.R. Haak, "Thermodynamics of cavity formation in water. A molecular dynamics study". *Faraday Symp. Chem. Soc.*, **17**, 55 (1982).
[13] W.F. van Gunsteren and H.J.C. Berendsen, BIOMAS, Groningen 1987.
[14] J. Schlitter, to be published.
[15] G.E. Uhlenbeck and L.S. Ornstein, "On the theory of brownian motion", *Phys. Rev.*, **36**, 823 (1930).

APPENDIX

The following consideration aims at clarifying the meaning of the relaxation time τ introduced above and to get some insight into its dependence on the system and the particular perturbation applied. As a one-dimensional example, the diffusive motion of a particle is investigated which for $t \leqslant 0$ moves freely and becomes subject to an external force $F(x)$ exerted by a potential $\Delta V(x)$ which is switched on at $t = 0$. The time evolution of the probability distribution can be described by a Fokker-Planck equation which for a harmonical potential $\Delta V(x) = cx^2$ yields a mean position $\langle x \rangle_t = x_0 \exp(-t/\tau_x)$ with $\tau_x = m\gamma/2c$, m being the particle mass and γ the friction coefficient of the system [15]. The relaxation of the perturbation potential is given by $\langle x^2 \rangle_t$ and becomes

$$\langle \Delta V \rangle_t = \exp(-2t/\tau_x)\Delta V_0 + (1 - \exp(-2t/\tau_x))\langle \Delta V \rangle \qquad \text{(A1)}$$

which has the functional form used in equation (2.9). The same physical process obviously leads to different relations for different variables, in particular the relaxation time of ΔV, $\tau = \tau_x/2$, is shorter than that of x.

A generalization of this finding is obtained by using the Langevin equation for the stationary case

$$\mathrm{d}x/\mathrm{d}t = (F(x) + R(t)/m\gamma \qquad \text{(A2)}$$

Multiplying (A2) with kx^{k-1} and taking the ensemble average gives

$$d\langle x^k \rangle_t/\mathrm{d}t = k\langle x^{k-1}F(x)\rangle_t/m\gamma \qquad \text{(A3)}$$

on the assumption that x and R are statistically independent. Let us now consider a perturbation $\Delta V = cx^{-n}$ with a force $F = ncx^{-n-1}$, and a δ-shaped initial distribution at x_0 for $t = 0$. To calculate the time derivatives at $t = 0$ from (A3) one can replace the average on the right hand side by the value at x_0. With $k = 1$ and $k = -n$ and finds

$$d\langle x \rangle_t / \mathrm{d}t = nc/m\gamma x_0^{n+1} \quad \text{and} \quad d\langle \Delta V \rangle_t / \mathrm{d}t = -(nc)^2/m\gamma x_0^{2n+2} \tag{A4}$$

or, by expanding the averages according to

$$\langle x \rangle_t = x_0(1 + t/\tau_x + \ldots) \quad \text{and} \quad \langle \Delta V \rangle_t = \Delta V_0(1 - t/\tau + \ldots) \tag{A5}$$

with relaxation times

$$\tau_x = x_0^{n+2} m\gamma / nc \quad \text{and} \quad \tau = \tau_x / n \tag{A6}$$

The simplified Langevin equation (A2) is, of course, not fully equivalent to the Fokker–Planck equation and hence does not give the correct asymptotic behaviour. The time constants (A6), however, are in agreement with the above exact solutions for the case $n = -2$, which justifys the present approximate treatment. The important result is that τ not only depends on the system variables m and γ, but also on the parameters of the perturbation potential. Relaxation becomes fast with increasing mobility of the system ($\propto 1/\gamma$). In the same system environment relaxation times decrease with increasing steepness n and strength c of ΔV. For a given system, relatively short relaxation times are hence expected when it is perturbed by Lennard-Jones type repulsion potentials with $n = 12$.