

A New Algorithm for the Numerical Solution of Diffusion Problems Related to the Smoluchowski Equation

By Bernhard Nickel

Abteilung Spektroskopie und Photochemische Kinetik, Max-Planck-Institut
für biophysikalische Chemie, Am Fassberg 11, D-37077 Göttingen, Germany

*Dedicated to Prof. Dr. Dr. h. c. mult. H. Gg. Wagner
on the occasion of his 70th birthday*

(Received February 2, 2000; accepted February 8, 2000)

Algorithm / Smoluchowski Equation / Diffusion-Influenced Reactions

Diffusion-influenced reactions can often be described with simple kinetic models, whose basic features are a spherically symmetric potential, a distance-dependent relative diffusion coefficient, and a distance-dependent first-order rate coefficient. A new algorithm for the solution of the corresponding Smoluchowski equation has been developed. Its peculiarities are: (1) A logarithmic increase of the radius; (2) the systematic use of numerical fundamental solutions w of the Smoluchowski equation; (3) the use of polynomials of up to the 8th degree for the definition of the first and second partial derivatives of w with respect to the radius; (4) successive doubling of the total diffusion time. The power of the algorithm is illustrated by examples. In particular its usefulness for the combination of a short-range potential with a large radial range is demonstrated. Some aspects of the algorithm are explained in the context of one-dimensional diffusion. Diffusion in a harmonic potential (Ornstein-Uhlenbeck process) and in a double-minimum potential is treated in detail. It is shown that a detailed balance will in general not lead to the best approximation of the time-dependence of a distribution.

1. Introduction

Diffusion-influenced reactions can often be described with simple kinetic models, whose basic common feature is the relative diffusive motion of a particle B in the spherically symmetric potential of a particle A [1–20]. For an ensemble of A···B pairs, the average relative motion of A and B and the disappearance of A···B pairs by a reaction of A or B can be described with the Smoluchowski Eq. (1):

$$\frac{\partial \rho(r, t)}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 D(r) \left[\frac{\partial \rho(r, t)}{\partial r} + \frac{1}{k_B T} \frac{dU(r)}{dr} \rho(r, t) \right] \right) - k(r) \rho(r, t). \quad (1)$$

ρ is a probability density (often normalized for $t = 0$), r is the distance between the centers of A and B, t is the time, D is the distance-dependent relative diffusion coefficient of A and B, k_B is the Boltzmann constant, T is the temperature, U is a spherically-symmetric potential, and k is a distance-dependent first-order rate coefficient. The partial differential Eq. (1) is to be solved numerically for a given initial condition $\rho(r, 0)$, subject to adequate boundary conditions.

The development of the present algorithm started from a specific problem, which is described in the Appendix. Technically this problem is characterized by the combination of a short-range deep potential with a large radial range. The basic assumption has been that a standard algorithm would not yield sufficiently accurate numerical solutions of problems of this kind within an acceptable computation time. In view of the lot of systematic work on the development of efficient diffusion algorithms [6–8, 13–20], the publication of a new algorithm requires a justification. In particular it should become evident that the new algorithm incorporates features that are of interest beyond the specific kinetic problem, for whose solution it has been designed. This justification is given in part in the following items and in part in the Discussion in section 4.

Dimension. It has been customary to treat one-dimensional diffusion, circularly symmetric two-dimensional diffusion, and spherically symmetric three-dimensional diffusion formally in the same way by using general formulae, in which the dimension is specified by a dimension variable d (cf. for instance ref. [19, 20]). The present algorithm was originally developed only for spherically symmetric diffusion. Later it turned out that some peculiarities of the algorithm are most easily explained and tested in the context of one-dimensional diffusion, and for this reason one-dimensional diffusion has been also included. Circularly symmetric diffusion has not been taken into account, but its treatment by the present algorithm should offer no difficulties.

Discretization of space coordinate. A short-range potential requires small radial steps Δr , which are unsuitable for the coverage of large radial ranges (R_0, R_N), where R_0 is the inner radius and R_N is the outer radius of the spherical shell, to which the motion of the center of B relative to the center of A is confined. Different strategies were developed to solve this problem: (a) Different constant values of Δr are used in sub-intervals of the total range (R_0, R_N) [6, 19]. (b) The discrete radius is logarithmically increased: $r_{i+1} = r_i \times (1 + q)$ with $0 < q = \text{const.} \ll 1$, and $z = \ln r$ is introduced as new independent spatial variable [14]. (c) The independent variable r is transformed by means of an isomorphism onto the *finite* range

of a new variable $z(r)$. By a suitable choice of the function $z(r)$, an equidistant discretization of z gives an optimal representation of the potential [17, 18]. The strategy of the present algorithm is in part similar to strategy (b). It will be shown that very large radial ranges can be covered by using a very small value q_1 in the near zone and a larger value q_2 in the far zone with a smooth transition from q_1 to q_2 . In contrast to the strategy (b), the radius r is retained as independent spatial variable.

Partial derivatives. In all diffusion algorithms, the first and the second partial derivatives of the distribution function ρ with respect to r at discrete radii r_i , $\rho'(r_i, t)$ and $\rho''(r_i, t)$, are needed, directly or indirectly. The usual procedures for calculating these partial derivatives are based on finite differencing, which is related to polynomial interpolation. The present procedure employs *direct* polynomial interpolation [21–23]: The polynomial $Y(r)$ of degree $2J$, centered at r_i and passing through $\rho(r_{i-j}, t), \dots, \rho(r_i, t), \dots, \rho(r_{i+j}, t)$, is calculated and the required partial derivatives are defined by $\rho'(r_i, t) \approx Y'(r_i)$ and $\rho''(r_i, t) \approx Y''(r_i)$. With this procedure also the partial derivatives of a numerical δ -function are obtained, which is defined by

$$w(r_i, 0 | r_n) = V_n^{-1} \delta_{in}, \quad (2)$$

where V_n is the volume of the n -th spherical shell, r_n is its mean radius, and δ_{in} is the Kronecker delta ($\delta_{in} = 1$ if $i = n$ and $\delta_{in} = 0$ if $i \neq n$).

Unsuitability of standard algorithms. A standard algorithm for the solution of Eq. (1) can be characterized as follows [24–26]: By using a finite difference procedure, the partial derivatives $\rho'(r_i, t)$ and $\rho''(r_i, t)$ in Eq. (1) are replaced by linear combinations of $\rho(r_i, t)$ and two or more of its neighboring values $\rho(r_{i \pm j}, t)$ with $j \geq 1$. Thus a single partial differential equation is replaced by a system of N coupled ordinary differential equations, to which Runge-Kutta methods can be applied [21–26]. It can be easily shown that a standard algorithm of this type is unsuitable for the solution of the present problem [23]. For a pure diffusion problem and in the absence of a potential, the total time range of interest is of the order $t_{\text{tot}} \approx (R_N - R_0)^2/D$. In the presence of a short-range potential, the radial step $\Delta r = r_{i+1} - r_i$ in the vicinity of the greatest slope of the potential is limited by the acceptable ratio of neighboring equilibrium values ρ_{eq} ,

$$\sigma^{-1} \leq \rho_{\text{eq}}(r_{i+1})/\rho_{\text{eq}}(r_i) \leq \sigma, \quad (3)$$

with the value of σ depending on the particular algorithm and the desired accuracy of the numerical solution; in general σ will be in the range $1 < \sigma < 2$. If the maximum value of Δr is fixed according to Eq. (3), then the maximum time step Δt in a Runge-Kutta procedure is $\Delta t \approx (\Delta r)^2/D$ (in the case of a δ -function $w(r_i, 0 | r_n)$, $w(r_n, \Delta t | r_n)$ must remain positive). Thus the total number of time steps is of the order of $t_{\text{tot}}/\Delta t \approx (R_N/\Delta r)^2$. The problem described in the Appendix would require $t_{\text{tot}}/\Delta t \geq 4 \times 10^{10}$, which

means that it cannot be easily solved with a standard method. With the present algorithm problems of this kind can be solved without difficulty. In section 3 an extreme example with $(R_N/\Delta r)^2 \approx 2.5 \times 10^{17}$ will be presented.

Conservation of total probability. An acceptable numerical algorithm must conserve the *total probability* in the absence of a reaction. In the case of a numerical δ -function $w(r_i, 0 | r_n)$ that means with respect to the first time step that the weighted loss of probability density at r_n is exactly equal to the sum of the weighted gains of probability density at the neighboring grid points $r_{n \pm j}$ with $j = 1, 2, \dots, J$. The present polynomial interpolation yields an exact conservation of total probability in the following special case: One-dimensional diffusion, constant steps Δx , and a potential defined by a polynomial of a degree $\leq 2J + 1$. In the case of spherically symmetric relative diffusion and logarithmically increasing radii, the total probability is no longer exactly conserved, but the relative error can be made extremely small in the absence of a potential. If q in the radial increment factor $(1 + q)$ is small and constant, the relative error is approximately proportional to q^{2J+2} . For instance, with $q = 0.02$ and $J = 4$, the total probability is virtually exactly conserved within the limits of double-precision calculations. In the presence of a potential, the deviations from the exact probability conservation remain very small, if the ratio σ of neighboring equilibrium values ρ_{eq} is small enough, for instance $\sigma \approx 1.5$ for $J = 4$. The required conservation of total probability can be always achieved by a very small correction.

Detailed balance. In the case of reflecting boundaries and in the absence of reactions, the final stationary numerical solution $\rho_{\text{stat}}(r_i)$ should be *close* to the theoretical equilibrium distribution $\rho_{\text{eq}}(r_i)$. In some algorithms the *exact* equality $\rho_{\text{stat}}(r_i) = \rho_{\text{eq}}(r_i)$ is implemented by a *detailed balance* [15, 16, 19]. Its meaning is most easily understood in the case of a Monte-Carlo algorithm [13, 15]: In the simplest case of one-dimensional diffusion with equidistant grid points, diffusion is characterized by the frequency Ω_{ij} of hopping from x_i to x_j . In this case, *detailed balance* means that the frequencies Ω_{ij} *exactly* satisfy the equilibrium condition $\Omega_{ij} \rho_{\text{eq}}(r_i) = \Omega_{ji} \rho_{\text{eq}}(r_j)$. A detailed balance seems to be regarded by some authors [15, 16, 19] as a *necessary* condition for a good numerical algorithm. The present algorithm does not satisfy this condition, and for this reason alone it might be considered inferior to other algorithms. However, by the example of one-dimensional diffusion in a harmonic potential (Ornstein-Uhlenbeck process [13, 19, 27]) it will be demonstrated that, by requiring a detailed balance, not the best numerical approximation of the *evolution* of a distribution function is obtained.

Time propagation. With respect to time propagation, two types of algorithms can be distinguished. In algorithms of the first type, the time step Δt is small and constant either during the whole time range of interest or at least during rather large time intervals [20]. During each time step, the

transfer of probability density is restricted to neighboring grid points. The upper limit of Δt is roughly given by the requirement that the transfer of probability density from a grid point r_i to its two adjacent grid points must not lead to a negative probability density at r_i itself. In algorithms of the second type, the so-called Chebyshev time propagation is applied, which permits large time steps and involves simultaneous changes of ρ in many or even all grid points [16, 19]. This method is similar to a propagation scheme for the time-dependent Schrödinger equation [28–30]. As far as very long time propagations are of interest, an algorithm of the second type might be the first choice among the known algorithms, for a kinetic problem described like that described in the Appendix. The present algorithm belongs to the first type with respect to the first time step, with the difference that $2J \geq 2$ instead of 2 neighboring points are involved. With respect to the further time propagation, the present algorithm seems to be new. The basic idea is the systematic use of *numerical* fundamental solutions $w(r_i, t | r_n)$, in analogy to the use of *analytical* fundamental solutions $g(r_i, t | r_n)$ in the analytical treatment of diffusion problems [1, 2, 4, 5]. Let the N fundamental solutions at time t , $w(r_i, t | r_n)$ ($n = 1, 2, \dots, N$) and the initial condition $\rho(r_i, 0)$ be known. $\rho(r_i, t)$ is then given by

$$\rho(r_i, t) = \sum_{n=1}^N [\rho(r_n, 0) \times V_n] \times w(r_i, t | r_n). \quad (4)$$

Eq. (4) is applied also to the fundamental solutions themselves. By successive doubling of the total diffusion time t , each fundamental solution at the time $2t$, $w(r_i, 2t | r_n)$, can be expressed as the weighted sum of all N fundamental solutions at the time t , $w(r_i, t | r_l)$ ($l = 1, 2, \dots, N$). Since with this procedure the computation time becomes proportional to the logarithm of the total diffusion time, the computation of $\rho(r_i, t)$ can be always extended to the attainment of the final stationary distribution $\rho_{\text{stat}}(r_i)$ (in the absence of reactions and with reflecting boundaries).

Duration of the first time step. In algorithms of the first type, the total computation time is inversely proportional to the time step Δt . Hence, in general Δt should be as large as possible, and for this reason an algorithm that is of second order with respect to time is preferable to one that is only of first order with respect to time like the present algorithm. However, in connection with successive time-doubling, the limitation to first order is no real disadvantage. For very short times t , most elements of $w(r_i, t | r_n)$ are either equal to zero or can be equated with zero because of their extreme smallness; by limiting all numerical operations to elements of $w(r_i, t | r_n)$ that are not equal to zero, the computation time for the initial time-doubling cycles becomes very short. Therefore Δt can be chosen almost arbitrarily small without appreciably increasing the total computation time. Finally it is to be mentioned that the algorithm is completely stable also with large values of Δt like the algorithms of the first type.

Boundary conditions. The present numerical implementation of reflecting or absorbing boundaries is completely analogous to Smoluchowski's theoretical treatment of boundaries [1, 2, 4, 31]. The first J numerical δ -functions $w(r_i, 0 | r_n)$ ($n = 1, 2, \dots, J$) are supplemented by their mirror images $w(r_i, 0 | r_n)$ ($n = 0, -1, \dots, -J + 1$). In the case of a reflecting boundary, the transfer of probability through the boundary during the first time step is exactly compensated by the inverse transfer of probability from the mirror image. In the case of an absorbing boundary, the same mirror-image numerical δ -functions are used, but now with a negative sign. Other boundary conditions (radiation, constant concentration) are implemented in a similar way.

Reactions. In a standard algorithm, a reaction (the sink term $k(r)$ in Eq. (1)) is taken into account simultaneously with diffusion. In the present algorithm, a reaction is taken into account by a single, very small change of the fundamental solutions at a suitable diffusion time, which is much shorter than $(k_{\max})^{-1}$.

Examples. The efficiency of the present algorithm is demonstrated by five examples with an adequate representation of relative errors: (1) One-dimensional diffusion in a harmonic potential (Ornstein-Uhlenbeck process); (2) one-dimensional diffusion in a double-minimum potential that is defined by a polynomial of the 8th degree; (3) three-dimensional diffusion without potential in a large radial range with constant radial increment factor $(1 + q)$ and a reflecting or absorbing inner boundary; (4) three-dimensional diffusion with a short-range potential barrier, a distance-dependent relative diffusion coefficient, a small radial range with a constant radial increment factor $(1 + q)$, and reflecting boundaries; (5) three-dimensional diffusion with the same short-range potential barrier and the same distance-dependent relative diffusion coefficient and also with reflecting boundaries, but with a very large radial range with a variable radial increment factor $(1 + q)$.

The structure of the article is as follows: In section 2, the main features of the algorithm are treated in the context of one-dimensional diffusion. In particular it will be shown that the new algorithm permits the numerical solution of the Ornstein-Uhlenbeck problem with high accuracy. In section 3, three-dimensional diffusion, reactions, and computational aspects of the algorithm are treated. In section 4, the merits and deficiencies of the algorithm are discussed. In section 5 the basic features of the algorithm are summarized.

2. One-dimensional diffusion

2.1 Definition of intervals and boundaries

Let diffusion be restricted by boundaries at $x = X_0$ and $x = X_N$ and let the interval (X_0, X_N) be divided into N intervals of equal width $h = (X_N - X_0)/$

N . In standard algorithms, the interval boundaries $X_n = X_0 + hn$ ($n = 0, 1, \dots, N$) are the grid points. To each inner grid point X_n belongs a symmetrical interval $(X_n - h/2, X_n + h/2)$. The disadvantage of this definition of grid points is that the associated intervals of the boundaries $(X_0, X_0 + h/2)$ and $(X_N - h/2, X_N)$, are smaller and unsymmetrical. The ensuing conceptual and mathematical difficulties are avoided by choosing the *centers* x_n of the N intervals as grid points: $x_n = X_0 + h(n - \frac{1}{2})$ with $(n = 1, 2, \dots, N)$.

2.2 Initial condition and fundamental solutions

In the case of one-dimensional diffusion with a constant relative diffusion coefficient and in the absence of a reaction, the one-dimensional analogue of the partial differential Eq. (1) is

$$\frac{\partial \rho(x, t)}{\partial t} = D \left(\frac{\partial^2 \rho(x, t)}{\partial x^2} + \frac{du(x)}{dx} \frac{\partial \rho(x, t)}{\partial x} + \frac{d^2 u(x)}{dx^2} \rho(x, t) \right), \quad (5)$$

where $u(x) \equiv U(x)/k_B T$ is a reduced potential. Let the initial condition $\rho(x, 0)$ be numerically given by a normalized set of N values $\rho(x_n, 0)$:

$$\sum_{n=1}^N \rho(x_n, 0) h = 1. \quad (6)$$

The evolution of ρ , $\rho(x_i, t)$, is completely defined by the evolution of the N fundamental solutions, $w(x_i, t | x_n)$. For the moment it is sufficient to calculate $w(x_i, t_a | x_n)$, that is, the result of the first diffusion step of duration $t_a \equiv \Delta t$ on $w(x_i, 0 | x_n)$. The procedure for calculating $w(x_i, t_a | x_n)$ is essentially the same as that, which one would apply in the direct numerical solution of the partial differential Eq. (5). For arbitrary t , $\rho(x, t + \Delta t)$ can be calculated from $\rho(x, t)$ by using Euler's method [24]:

$$\rho(x, t + \Delta t) \approx \rho(x, t) + \frac{\partial \rho(x, t)}{\partial t} \Delta t. \quad (7)$$

For the calculation of $\partial \rho(x, t)/\partial t$ with Eq. (5), one has to know the first and the second partial derivatives $\rho'(x, t) \equiv \partial \rho(x, t)/\partial x$ and $\rho''(x, t) \equiv \partial^2 \rho(x, t)/\partial x^2$. If $\rho(x, t)$ is known only for discrete values x_n with $x_n < x_{n+1}$, one way of calculating $\rho'(x_n, t)$ and $\rho''(x_n, t)$ is to calculate a suitable differentiable function $Y_n(x)$ passing through $\rho(x_n, t)$ and its $2J$ neighboring points $\rho(x_{n\pm 1}, t)$, $\rho(x_{n\pm 2}, t)$, ..., $\rho(x_{n\pm J}, t)$, and having the property $Y'_n(x_n) \approx \rho'(x_n, t)$ and $Y''_n(x_n) \approx \rho''(x_n, t)$ [21–23]. If $Y_n(x)$ is a polynomial of the degree $2J$ ($J = 1, 2, \dots$), then its minimum degree is 4, if inflection points are to be accounted for. The present algorithm is based on the use of polynomials up to the degree $2J = 8$.

2.3 Definition of fundamental solutions by polynomials

Let $\{y_{n,n+j}(x)\}$ be a set of $2J + 1$ polynomials of the degree $2J$, centered at x_n (first subscript), with the property

$$y_{n,n+j}(x_{n+i}) = \delta_{ij} \quad (-J \leq i \leq J), \quad (8)$$

where δ_{ij} is the Kronecker delta. $Y_n(x)$ can be written as the weighted sum of these polynomials:

$$Y_n(x) = \sum_{j=-J}^J \rho(x_{n+j}, t) y_{n,n+j}(x). \quad (9)$$

Eq. (9) is equivalent to Lagrange's interpolation formula [21–23]. The first and the second partial derivatives of ρ with respect to x at x_n are then approximately given by

$$\rho'(x_n, t) \approx Y'_n(x_n) = \sum_{j=-J}^J \rho(x_{n+j}, t) y'_{n,n+j}(x_n), \quad (10)$$

$$\rho''(x_n, t) \approx Y''_n(x_n) = \sum_{j=-J}^J \rho(x_{n+j}, t) y''_{n,n+j}(x_n). \quad (11)$$

The same procedure is applied to the formal calculation of $w'(x_i, 0 | x_n)$ and $w''(x_i, 0 | x_n)$ for $i = n, n \pm 1, \dots, n \pm J$. The only difference is that for each i the sum of $2J + 1$ terms in Eqs. (10) and (11) reduces to a single term:

$$w'(x_{n+j}, 0 | x_n) = h^{-1} y'_{n+j,n}(x_{n+j}), \quad (12)$$

$$w''(x_{n+j}, 0 | x_n) = h^{-1} y''_{n+j,n}(x_{n+j}). \quad (13)$$

The equality signs in Eqs. (12) and (13) have the meaning of definitions depending on the degree of the polynomials.

A necessary condition for any diffusion algorithm is the conservation of total probability in the absence of reactions. With respect to the numerical δ -functions $w(x_i, 0 | x_n)$ that means, the weighted decrease of w in the n -th interval must be equal to the weighted gains of w in the $2J$ neighboring intervals or, in other words, the normalization condition is to be satisfied:

$$h \sum_{j=-J}^J w(x_{n+j}, t_n | x_n) = 1. \quad (14)$$

As a corollary of the definition of $w'(x_i, 0 | x_n)$ and $w''(x_i, 0 | x_n)$ by Eqs. (12) and (13), the condition of probability conservation (14) is exactly satisfied, firstly, for free diffusion for arbitrary J and, secondly, for diffusion in a potential that is given by a polynomial up to the degree $2J + 1$.

It is advantageous to prove the correctness of the first preceding statement first for $J = 2$ (the simplest nontrivial case). For the formal calculation of the partial derivatives $w'(x_i, 0 | x_n)$ and $w''(x_i, 0 | x_n)$, a polynomial of the 4th degree, centered at x_i , is calculated. The following five polynomials are needed:

$$y_{n,n}(x) = + (4h^4)^{-1} (x - x_n + 2h) (x - x_n + h) (x - x_n - h) (x - x_n - 2h) \quad (15)$$

$$y_{n+1,n}(x) = - (6h^4)^{-1} (x - x_n + h) (x - x_n - h) (x - x_n - 2h) (x - x_n - 3h) \quad (16)$$

$$y_{n-1,n}(x) = -(6h^4)^{-1} (x - x_n + 3h) (x - x_n + 2h) (x - x_n + h) (x - x_n - h) \quad (17)$$

$$y_{n+2,n}(x) = +(24h^4)^{-1} (x - x_n - h) (x - x_n - 2h) (x - x_n - 3h) (x - x_n - 4h) \quad (18)$$

$$y_{n-2,n}(x) = +(24h^4)^{-1} (x - x_n + 4h) (x - x_n + 3h) (x - x_n + 2h) (x - x_n + h). \quad (19)$$

The first and second derivatives of these polynomials at their respective centers x_{n+j} are:

$$y'_{n,n}(x_n) = 0 \quad (20a)$$

$$y''_{n,n}(x_n) = -\frac{5}{2h^2} \quad (20b)$$

$$y'_{n\pm 1,n}(x_n \pm h) = \mp \frac{2}{3h} \quad (21a)$$

$$y''_{n\pm 1,n}(x_n \pm h) = +\frac{4}{3h^2} \quad (21b)$$

$$y'_{n\pm 2,n}(x_n \pm 2h) = \pm \frac{1}{12h} \quad (22a)$$

$$y''_{n\pm 2,n}(x_n \pm 2h) = -\frac{1}{12h^2}. \quad (22b)$$

The probability is exactly conserved for arbitrary t_a ,

$$(D h t_a) \times \sum_{j=-J}^J y''_{n+j,n}(x_n + hj) = 0, \quad (23)$$

with $J = 2$ in the present case.

The proof of the validity of Eq. (23) for arbitrary $J \geq 1$ runs as follows. In the special case of $\rho(x_{n+j}, 0) = 1$ for $-J \leq j \leq J$, the analogues of Eqs. (9)–(11) are

$$Y_n(x_n) = \sum_{j=-J}^J y_{n,n+j}(x_n) = 1, \quad (24)$$

$$Y'_n(x_n) = \sum_{j=-J}^J y'_{n,n+j}(x_n) = 0, \quad (24a)$$

$$Y''_n(x_n) = \sum_{j=-J}^J y''_{n,n+j}(x_n) = 0. \quad (24b)$$

The terms of the sum in Eq. (24b) can be written in a different form:

$$y''_{n,n+j}(x_n) = y''_{n-j,n}(x_n - hj). \quad (25)$$

Table 1. First and second derivatives of the polynomials $y_{n+j,n}(x)$ of the degree $2J$ at their centers $x_n + jh$ ($-J \leq j \leq J$). If the terms in the rows 3 to 6 are multiplied by 2, the sum of the terms in each column is equal to zero.

J	1	2	3	4
$y''_{n,n}(x_n)$	$-\frac{2}{h^2}$	$-\frac{5}{2h^2}$	$-\frac{49}{18h^2}$	$-\frac{205}{72h^2}$
$y''_{n\pm 1,n}(x_n \pm h)$	$+\frac{1}{h^2}$	$+\frac{4}{3h^2}$	$+\frac{3}{2h^2}$	$+\frac{8}{5h^2}$
$y''_{n\pm 2,n}(x_n \pm 2h)$		$-\frac{1}{12h^2}$	$-\frac{3}{20h^2}$	$-\frac{1}{5h^2}$
$y''_{n\pm 3,n}(x_n \pm 3h)$			$+\frac{1}{90h^2}$	$+\frac{8}{9 \cdot 35h^2}$
$y''_{n\pm 4,n}(x_n \pm 4h)$				$-\frac{1}{16 \cdot 35h^2}$
$y'_{n,n}(x_n)$	0	0	0	0
$y'_{n\pm 1,n}(x_n \pm h)$	$\mp \frac{1}{2h}$	$\mp \frac{2}{3h}$	$\mp \frac{3}{4h}$	$\mp \frac{4}{5h}$
$y'_{n\pm 2,n}(x_n \pm 2h)$		$\pm \frac{1}{12h}$	$\pm \frac{3}{20h}$	$\pm \frac{1}{5h}$
$y'_{n\pm 3,n}(x_n \pm 3h)$			$\mp \frac{1}{60h}$	$\mp \frac{4}{105h}$
$y'_{n\pm 4,n}(x_n \pm 4h)$				$\pm \frac{1}{280h}$

That means, the set of terms in Eq. (24b) is identical with that in Eq. (23); only the order of the terms in one sum is inverted relative to the order of the terms in the other sum. In Table 1 all terms $y'_{n+j,n}(x_n + hj)$ and $y''_{n+j,n}(x_n + hj)$ are listed for $J = 1, 2, 3, 4$.

The total probability is also exactly conserved, if the reduced potential $u(x)$ is given by a polynomial of the degree $m \leq 2J + 1$ and $1 \leq J \leq 4$:

$$u(x_n + hj) = \sum_{k=0}^m \alpha_k (x_n + hj)^k, \quad (26)$$

$$u'(x_n + hj) = \sum_{k=1}^m k \alpha_k (x_n + hj)^{k-1}, \quad (26a)$$

$$u''(x_n + hj) = \sum_{k=2}^m k(k-1) \alpha_k (x_n + hj)^{k-2}. \quad (26b)$$

The probability conservation is exact, because for each term of the polynomial $u(x)$ the following relation holds:

$$k(k-1)x_n^{k-2}y_{n,n}(x_n) + \sum_{j=-J}^J k(x_n+hj)^{k-1}y'_{n+j,n}(x_n+hj) = 0 \quad (0 \leq k \leq 2J+1). \quad (27)$$

The correctness of Eq. (27) has been proved for $J = 1, 2, 3, 4$. Eq. (27) probably holds for arbitrary $J \geq 1$, but no serious attempt has been made by the author to prove its validity for $J > 4$.

2.4 Fundamental solutions for the first diffusion step

For the calculation of the fundamental solutions $w(x_{n+j}, t_a | x_n)$ with $j = 0, \pm 1, \dots, \pm J$ with Eq. (7), the quantities

$$H_{n,j} \equiv h [\partial w(x, t | x_n) / \partial t]_{x=x_{n+j}, t=0} \quad (28)$$

are needed, which are according to Eqs. (5), (12) and (13):

$$H_{n,j} = D [y''_{n+j,n}(x_{n+j}) + u'(x_{n+j})y'_{n+j,n}(x_{n+j}) + u''(x_{n+j})y_{n+j,n}(x_{n+j})]. \quad (29)$$

For calculating the result of the first diffusion step, it is advantageous to define the changes $b_{n,j}$ of $w(x_{n+j}, 0 | x_n)$ relative to $w(x_n, 0 | x_n)$:

$$b_{n,0} = 1 + H_{n,0} t_a, \quad (30)$$

$$b_{n,j} = H_{n,j} t_a \quad (j = \pm 1, \pm 2, \dots, \pm J). \quad (31)$$

A practical upper limit of t_a is obtained by requiring $0.5 \leq b_{n,0} < 1$ or $0 < |H_{n,0} t_a| \leq 0.5$. If boundary conditions are neglected, then the effect of the first diffusion step on $w(x_i, 0 | x_n)$ is

$$w(x_{n+j}, t_a | x_n) = w(x_n, 0 | x_n) \times b_{n,j} \quad \text{for } j = 0, \pm 1, \dots, \pm J, \quad (32)$$

$$w(x_{n+j}, t_a | x_n) = 0 \quad \text{for } |j| > J. \quad (33)$$

The quantities $b_{n,j}$ contain all the information that is needed for calculating the evolution of the fundamental solutions in the absence of reactions.

2.5 Boundary conditions

Two boundaries are of interest: A *left* boundary at X_0 and a *right* boundary at X_N . In the following, first the two limiting cases of a completely reflecting boundary and a completely absorbing boundary are treated. The left reflecting boundary is defined by [4]

$$\left(\frac{\partial \rho(x,t)}{\partial x} \right)_{x=X_0} = 0, \quad (34)$$

$$\left(\frac{du(x)}{dx} \right)_{x=X_0} = 0, \quad (35)$$

and the left absorbing boundary is defined by

$$\rho(X_0, t) = 0 \quad \text{for } t > 0. \quad (36)$$

These boundary conditions are numerically implemented as follows. First, $2J$ additional intervals are defined with indices $n = 0, -1, \dots, -(2J - 1)$ and center values $x_n = x_1 + h(n - 1)$. Second, the J additional sets of quantities $H_{n,j}, b_{n,j}$ ($n = 0, -1, \dots, -J + 1$) are calculated. Third, in the case of a reflecting left boundary, the loss of probability density in the intervals $1, 2, \dots, J$ due to diffusion from right to left through the left boundary is exactly balanced by the gain of probability density due to diffusion from left to right through the left boundary from the intervals $0, -1, \dots, -(J - 1)$, which are symmetrical to the intervals $1, 2, \dots, J$ with respect to the reflecting boundary at X_0 . The following equations refer to the case $J = 2$:

$$b_{1,-1} + b_{1,-2} = c_0 (b_{0,+1} + b_{0,+2}), \quad (37)$$

$$b_{2,-2} = c_{-1} b_{-1,+2}. \quad (38)$$

The factors c_0 and c_{-1} are equal to unity in the absence of a potential and close to unity in the presence of a potential. The final values of $b_{1,j}$ and $b_{2,j}$, referring to a particular boundary condition (BC), are

$$(b_{1,0})_{\text{BC}} = b_{1,0} + b_{0,+1} \times c_0 \times c_{\text{BC}}, \quad (39)$$

$$(b_{1,+1})_{\text{BC}} = b_{1,+1} + b_{0,+2} \times c_0 \times c_{\text{BC}}, \quad (40)$$

$$(b_{2,-1})_{\text{BC}} = b_{2,-1} + b_{-1,+2} \times c_{-1} \times c_{\text{BC}}, \quad (41)$$

$$(b_{1,-1})_{\text{BC}} = (b_{1,-2})_{\text{BC}} = (b_{2,-2})_{\text{BC}} = 0. \quad (42)$$

The factor c_{BC} specifies the particular boundary condition: c_{BC} is equal to $+1$ for a reflecting boundary and equal to -1 for an absorbing boundary. The present implementation of the two boundary conditions is the complete analogue of the analytical procedure (cf. refs. [1, 2, 4, 31]). A reflecting or absorbing *right* boundary is implemented in the same way by defining the corresponding quantities $(b_{N,0})_{\text{BC}}, (b_{N,-1})_{\text{BC}}, (b_{N-1,+1})_{\text{BC}}$, and $(b_{N,+1})_{\text{BC}} = (b_{N,+2})_{\text{BC}} = (b_{N-1,+2})_{\text{BC}} = 0$. In the following the subscript "BC" will be omitted, and it will be always assumed that the quantities $b_{1,j}, b_{2,j}$ and $b_{N,j}, b_{N-1,j}$ are defined in accord with the chosen boundary condition.

The extension of the definition of the boundary conditions to arbitrary $J \geq 1$ is straightforward. With these redefinitions of the quantities $b_{1,j}, b_{2,j}, \dots, b_{J,j}$ and $b_{N,j}, b_{N-1,j}, \dots, b_{N-J+1,j}$, Eq. (32) can be applied to the whole range $1 \leq n \leq N$, if the range of the index j is appropriately restricted in the vicinity of a boundary:

$$w(x_{n+j}, t_a | x_n) = w(x_n, 0 | x_n) \times b_{n,j} \quad \text{for } j_1 \leq j \leq j_2, \quad (43a)$$

$$w(x_{n+j}, t_a | x_n) = 0 \quad \text{for } j < j_1 \quad \text{or} \quad j > j_2, \quad (43b)$$

where

$$j_1 = -J \quad \text{and} \quad j_2 = +J \quad \text{for } J + 1 \leq n \leq N - J, \quad (44a)$$

$$j_1 = 1 - n \quad \text{and} \quad j_2 = +J \quad \text{for } 1 \leq n \leq J, \quad (44b)$$

$$j_1 = -J \quad \text{and} \quad j_2 = N - n \quad \text{for } N - J + 1 \leq n \leq J. \quad (44c)$$

Two other boundary conditions are of general interest: A *radiation* boundary can be defined by a linear combination of a reflecting boundary with an absorbing boundary. In practice that means, the factor c_{BC} in Eqs. (39)–(41) may have arbitrary values in the interval $-1 \leq c_{BC} \leq +1$. A *constant-concentration* boundary can be implemented as follows. In contrast to the analytical definition, ρ or w is kept constant not at X_0 or X_N but at x_1 or x_N . The definition of the quantities $H_{n,j}$ is the same as before. In the special case $J = 2$, the condition $w(x_1, t_a | x_n) = w(x_1, 0 | x_n)$ is satisfied, if the Eqs. (37) and (38) are replaced by the equations

$$b_{1,0} + s_{1,0} b_{0,+1} = 1, \quad (37a)$$

$$b_{2,-1} + s_{2,-1} b_{0,+1} = 0, \quad (38a)$$

$$b_{3,-2} + s_{3,-2} b_{-1,+2} = 0. \quad (38b)$$

Thus the new definitions of the quantities $b_{n,j}$ for $n = 1, 2, 3$ are:

$$(b_{1,0})_{BC} = b_{1,0} + b_{0,+1} \times s_{1,0}, \quad (39a)$$

$$(b_{1,+1})_{BC} = b_{1,+1} + b_{0,+2} \times s_{1,0}, \quad (40a)$$

$$(b_{2,-1})_{BC} = b_{2,-1} + b_{0,+1} \times s_{2,-1} = 0, \quad (41a)$$

$$(b_{2,0})_{BC} = b_{2,0} + b_{0,+2} \times s_{2,-1}, \quad (41b)$$

$$(b_{3,-2})_{BC} = b_{3,-2} + b_{-1,+2} \times s_{3,-2}, \quad (41c)$$

$$(b_{1,-1})_{BC} = (b_{1,-2})_{BC} = (b_{2,-2})_{BC} = 0. \quad (42a)$$

The extension of this implementation of a constant-concentration boundary to arbitrary J offers no difficulties.

2.6 The second and subsequent diffusion steps:

Successive time doubling

The fundamental solutions $w(x_i, t_a m | x_n)$ and arbitrary distribution functions $\rho(x_i, t_a m)$ with $m \geq 2$ can be calculated with one of three different procedures:

Procedure (I). $w(x_i, t_a m | x_n)$ or $\rho(x_i, t_a m)$ is calculated by repeatedly using the quantities $b_{i,j}$:

$$w(x_i, t_a m | x_n) = \sum_{j=i_1}^{j_2} w(x_{i+j}, t_a(m-1) | x_n) \times b_{i+j,-j}. \quad (45)$$

Procedure (I) corresponds to standard procedures. It is the fastest one if only a single fundamental solution is needed and the number of diffusion steps is not extremely large.

Procedure (II). For very long diffusion times it is advantageous, to calculate $w(x_i, t | x_n)$ not by successive application of the same time step t_a , but by successive *doubling* of the time, and by defining each fundamental solution at the time $t_{m+1} = t_a 2^m$ as a linear combination of all N fundamental solutions at the time $t_m = t_a 2^{m-1}$:

$$w(x_i, t_a \times 2^m | x_n) = \sum_{l=1}^N [w(x_i, t_a \times 2^{m-1} | x_n) \times h] \times w(x_i, t_a \times 2^{m-1} | x_l). \quad (46)$$

The term in brackets is the probability of finding a particle at the time $t_m = t_a \times 2^{m-1}$ in the l -th interval, if it has been in the n -th interval at time $t_0 = 0$ with unit probability. In the following examples always procedure (II) was used.

Procedure (III). In typical applications, the evolution of ρ , $\rho(x_i, t_c | v)$, is of interest, where the constant time step t_c is much longer than the first time step t_a . The calculation of $\rho(x_i, t_c | v)$ by procedure (III) consists of two parts. First, the N fundamental solutions $w(x_i, t_c | x_n)$ are calculated with procedure (II). Second, $\rho(x_i, t_c | v | x_n)$ with $v = 1, 2, \dots$ is calculated by repeatedly using the fundamental solutions $w(x_i, t_c | x_n)$:

$$\rho(x_i, t_c | v) = \sum_{l=1}^N [\rho(x_i, t_c | (v-1) | x_l) \times h] \times w(x_i, t_c | x_l). \quad (46a)$$

2.7 Example 1: Diffusion in a harmonic potential (Ornstein-Uhlenbeck process)

For one-dimensional diffusion in a harmonic potential $u = \alpha x^2$ the analytical solution of Eq. (5) is known [13, 19, 27]:

$$g(x, t | x_n) = \frac{\exp(-\alpha[(x - x_n \exp(-2\alpha Dt))^2 [1 - \exp(-4\alpha Dt)]^{-1}])}{\sqrt{\pi[1 - \exp(-4\alpha Dt)]/\alpha}}. \quad (47)$$

The following computations were performed with $\alpha = 4 \text{ nm}^{-2}$, $x_n = 8 \text{ nm}$, $D = 6 \times 10^9 \text{ nm}^2 \text{ s}^{-1}$, $h = 0.01 \text{ nm}$, $X_0 = -5.005 \text{ nm}$, $X_N = 9.005 \text{ nm}$, $N = 1401$, and reflecting boundaries at X_0 and X_N . The initial speed of computation was strongly enhanced by equating all extremely small values of $w(x, t | x_n)$ with zero and by limiting all summations to terms that differed from zero. The smallest equilibrium value was $g_{\text{eq}}(x_N) \approx 2.19 \times 10^{-141} \text{ nm}^{-1}$. After each time-doubling cycle, $w(x_i, t | x_n)$ was equated with zero if $|w(x_i, t | x_n)|$ was smaller than $10^{-150} \text{ nm}^{-1}$.

In the steady state, $g_{\text{eq}}(x) \equiv g(x, \infty | x_n)$ is a Gaussian centered at $x = 0$. $g_{\text{eq}}(x)$ is shown in Fig. 1a on a logarithmic scale. The stationary numerical solutions $w_{\text{stat}}(x) \equiv w(x, t_{\text{stat}} | x_n)$ obtained with $J = 1, 2, 3, 4$ are compared with the exact solution $g_{\text{eq}}(x)$ (t_{stat} is the time at which all N fundamental functions are equal within a specified accuracy). The relative deviation F of w from g is defined by

$$F = (w/g) - 1. \quad (48)$$

Since F changes sign and $|F|$ extends over 10 orders of magnitude and, the following quantity z is used for the graphical representation of relative deviations:

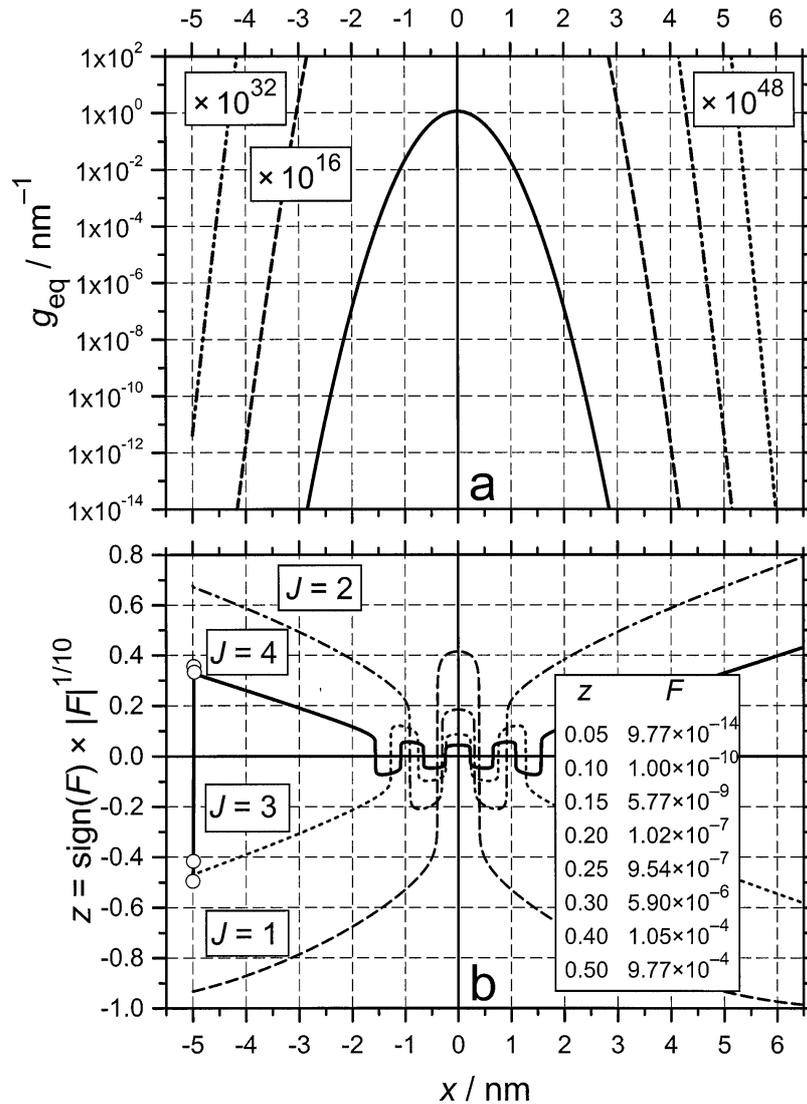


Fig. 1. Example 1: One-dimensional diffusion in a reduced harmonic potential $u(x) = \alpha x^2$ with $\alpha = 4 \text{ nm}^2$ and $D = 6 \times 10^9 \text{ nm}^2 \text{ s}^{-1}$ ($h = 0.01 \text{ nm}$, reflecting boundaries at $X_0 = -5.005 \text{ nm}$ and $X_N = 9.005 \text{ nm}$). (a) Equilibrium distribution $g_{\text{eq}}(x)$, calculated with Eq. (47). (b) Relative deviations F (see Eq. (48)) of numerical stationary distributions $w_{\text{stat}}(x)$ from the equilibrium distribution $g_{\text{eq}}(x)$ in the z -representation (see Eq. (49)); $w_{\text{stat}}(x)$ depends on the degree $2J$ of the interpolation polynomials; for $J = 4$, the four leftmost points z_i ($i = 1, 2, 3, 4$) are represented by open circles. $w_{\text{stat}}(x)$ is independent of the start position x_n and the length of the first time step t_a (as long as t_a is below the stability limit) with an accuracy of 13 digits. $w_{\text{stat}}(x)$ is reached at the time $t_{\text{stat}} = t_a \times 2^{18} \approx 1.12 \text{ ns}$ ($t_a \approx 4.27 \text{ fs}$).

$$z = \text{sign}(F) \times |F|^s, \quad (49)$$

with $s = 0.1$ in the present paper. A table with selected values of z and $F(z)$ is shown as inset in Fig. 1b. For simplicity, the quantity z will be also called *relative deviation*.

In Fig. 1b the quantity $z(x)$ is shown for $J = 1, 2, 3, 4$. At $x = 0$ nm, an increase of J by 1 reduces F roughly by a factor of 10^{-3} and increases the number of intersections of z with the line $z = 0$ by 2. $w_{\text{stat}}(x)$ is independent of the length of the first time step t_a . The present data were obtained with $t_a \approx 4.27$ fs, which corresponds for $w(0, 0 | 0)$ to a decrease from 100 nm^{-1} to $\approx 48.8 \text{ nm}^{-1}$. At the time $t_{\text{stat}} = t_a \times 2^{18} \approx 1.12$ ns the numerical curves $w(x, t_{\text{stat}} | x_n)$ were independent of x_n with an accuracy of 13 digits.

Independently of the neglect of extremely small terms, the accuracy of the computation of $w(x, t | x_n)$ is limited by the accumulation of round-off errors, which entails a deviation of the normalization sums, S_{norm} , from unity. In the present case, at $t_{\text{stat}} \approx 1.12$ ns, the deviation of S_{norm} , from unity, $S_{\text{norm}} - 1 \approx 3 \times 10^{-12}$, is about 150 times larger than the true value of $F(0) \approx 2 \times 10^{-14}$ for $J = 4$. This error is corrected for by a renormalization of the numerical fundamental solutions after every 10 time-doubling cycles.

With $J = 4$, the relative deviation $|F|$ is less than 10^{-4} in a range of g_{eq} of about 62 orders of magnitude ($z(6 \text{ nm}) = 0.4$ corresponds to $F(6 \text{ nm}) \approx 1 \times 10^{-4}$ in Fig. 1b and $g_{\text{eq}}(6 \text{ nm})/g_{\text{eq}}(0 \text{ nm}) \approx 10^{-62}$ in Fig. 1a). The only strongly irregular data points are the four leftmost points z_i ($i = 1, 2, 3, 4$), which are represented by open circles in Fig. 1b. They are a consequence of the fact that g_{eq} does not conform to the boundary condition Eq. (34). Finally, this computation illustrates the numerical stability of the algorithm for a value of t_a close to the upper stability limit.

The present algorithm is only of first order with respect to time. Hence one may expect a strong increase of accuracy by reducing the length of the first time step t_a . Let $(t_a)_m$ be defined by $(t_a)_m = (t_a)_0 \times 2^{-m}$, where $(t_a)_0$ is the time step used in the calculation of $w_{\text{stat}}(x)$ in Fig. 1 and $m = 0, 1, 2, \dots$. In Fig. 2 the results are shown for $J = 4$ and a rather short time $t \approx 1.07$ ps, when the center of $g(x, t | x_n)$ (≈ 7.6 nm) is still close to the start position $x_n = 8.0$ nm. The pattern of the relative deviations $z(x_i)$ initially strongly changes with increasing m and finally becomes virtually constant for $m \geq 24$. $|F|$ is less than 10^{-4} in a range of about 8 orders of magnitude of $g(x, t | x_n)$. Close to the center of $g(x, t | x_n)$, $|F|$ is $< 10^{-8}$.

The analogous results for $t \approx 8.53$ ps are shown in Fig. 3. Relative to the results in Fig. 2, with $m = 36$ a much shorter first time step is needed for the attainment of a constant pattern of z , and at the same time $|F|$ is less than 10^{-4} in a range of about 34 orders of magnitude of $g(x, t | x_n)$. Close to the center of $g(x, t | x_n)$, $|F|$ is $< 10^{-11}$. Finally it should be noted that $g(x, t | x_n)$ in Fig. 3 approximately corresponds to the rightmost curve in Fig. 1 of ref. [19], where the same example is used for the illustration of the gain of accuracy by the so-called virtual gridding technique.

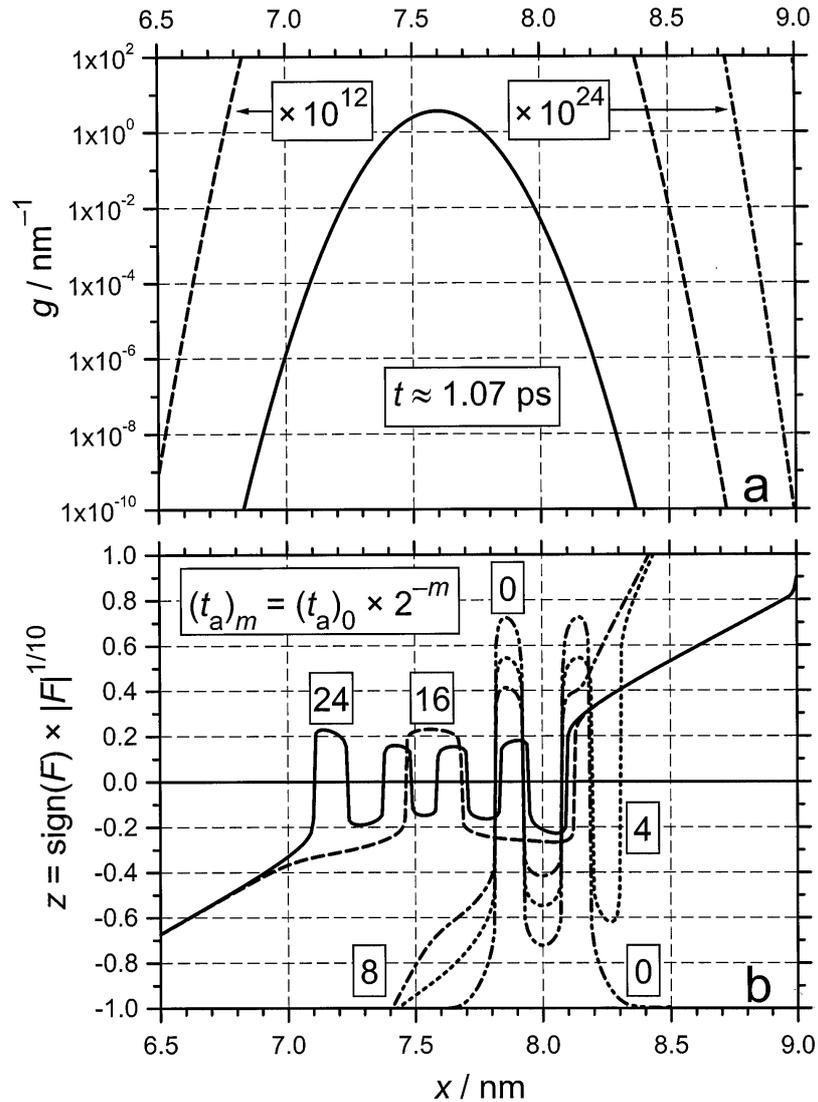


Fig. 2. Example 1: One-dimensional diffusion in a reduced harmonic potential $u(x) = \alpha x^2$ with $\alpha = 4 \text{ nm}^2$ and $D = 6 \times 10^9 \text{ nm}^2 \text{ s}^{-1}$ ($h = 0.01 \text{ nm}$, reflecting boundaries at $X_0 = -5.005 \text{ nm}$ and $X_N = 9.005 \text{ nm}$). (a) Fundamental solution $g(x, t | x_n)$, calculated with Eq. (47) with $x_n = 8 \text{ nm}$ and $t \approx 1.07 \text{ ps}$. (b) Relative deviations of the numerical fundamental solution $w(x_i, t | x_n)$ from $g(x_i, t | x_n)$. $w(x_i, t | x_n)$ was calculated with $J = 4$ and different values of $t_a = (t_a)_0 \times 2^{-m}$, where $(t_a)_0 \approx 4.27 \text{ fs}$. The numbers in frames are the values of m .

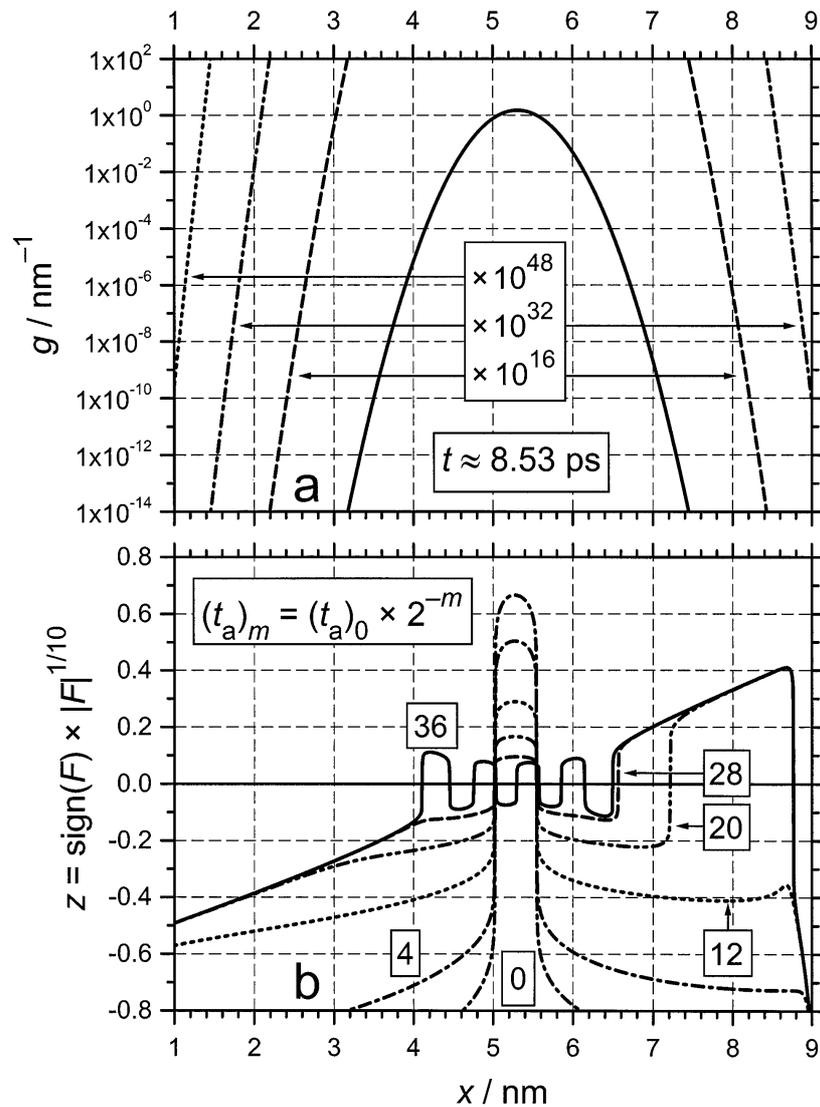


Fig. 3. Example 1: One-dimensional diffusion in a reduced harmonic potential $u(x) = \alpha x^2$ with $\alpha = 4 \text{ nm}^2$ and $D = 6 \times 10^9 \text{ nm}^2 \text{ s}^{-1}$ ($\hbar = 0.01 \text{ nm}$, reflecting boundaries at $X_0 = -5.005 \text{ nm}$ and $X_N = 9.005 \text{ nm}$). (a) Fundamental solution $g(x_i, t | x_n)$, calculated with Eq. (47) with $x_n = 8 \text{ nm}$ and $t \approx 8.53 \text{ ps}$. (b) Relative deviations z of the numerical fundamental solution $w(x_i, t | x_n)$ from $g(x_i, t | x_n)$. $w(x_i, t | x_n)$ was calculated with $J = 4$ and different values of $t_a = (t_a)_0 \times 2^{-m}$, where $(t_a)_0 \approx 4.27 \text{ fs}$. The numbers in frames are the values of m .

2.8 Negative effect of detailed balance on the accuracy of the algorithm

As already mentioned in the Introduction, the *exact* detailed balance is considered by some authors [15, 16, 19] to be a necessary condition for a good diffusion algorithm. In the case of a harmonic potential it can be easily demonstrated that, on the contrary, the enforced detailed balance leads to *completely wrong* numerical solutions, if the potential range of interest is very large like in example 1. Two sets of numerical curves w were computed with $J = 1$. The first set corresponds to the present algorithm. The second set satisfies the additional requirement of detailed balance, which is easily implemented by a redefinition of the quantities $H_{n,j}$. By assigning the index $n = 0$ to $x = 0$, the new set of $H_{n,j}$ is obtained by the identity

$$(H_{0,j})_{\text{new}} = (H_{0,j})_{\text{old}} \quad (j = 0, \pm 1), \quad (50)$$

and the recursion formulae:

$$(H_{n+1,-1})_{\text{new}} \exp(-\alpha x_{n+1}^2) = (H_{n,+1})_{\text{new}} \exp(-\alpha x_n^2) \quad (n \geq 0), \quad (51a)$$

$$(H_{n+1,-1})_{\text{new}} = (H_{n,+1})_{\text{new}} \exp(+ah^2(2n+1)), \quad (51b)$$

$$\eta = (H_{n+1,-1})_{\text{new}} / (H_{n+1,-1})_{\text{old}}, \quad (51c)$$

$$(H_{n+1,j})_{\text{new}} = (H_{n+1,j})_{\text{old}} \times \eta \quad (j = 0, \pm 1). \quad (51d)$$

The calculation of $(H_{n,j})_{\text{new}}$ with $n \leq -1$ is analogous.

In Fig. 4 the results are shown. Quantities referring to detailed balance are labeled with an asterisk. In the logarithmic plot of Fig. 4a, the curves w and g virtually coincide in the upper range of two orders of magnitude and cannot be distinguished from each other for arbitrary $|x_n| \leq 8$ nm and $t \leq 1.07$ ps. The curves w^* differ already significantly from the curves g for $x_n = -4$ nm and are completely wrong for $x_n = 8$ nm. The latter wrong result is not surprising because of the accumulation of a very large systematic error in the quantities $(H_{n,j})_{\text{new}}$. In Fig. 4a the ratio η is plotted (see Eq. (51c)): $\eta \approx 0.75$ for $x_n = -4$ nm and $\eta \approx 10^{-2}$ for $x_n = +8$ nm. In Fig. 4b and 4c the relative deviations z of w from g are shown for $x_n = 8$ nm and $x_n = -4$ nm. The smallest deviations are always found close to the maximum of each curve g (the three open circles correspond to the indices $i_{\text{max}}, i_{\text{max}} \pm 10$).

With $x_n = 0$ nm, the initial differences between w and w^* are very small. Significant differences between z and z^* occur only at times, when w is already very close to the stationary distribution w_{stat} . The final stationary distribution w_{stat}^* is reached at the time $t_{13} \approx 8.74$ ns and agrees with g_{eq} with a relative error $|F| < 10^{-12}$ in the whole range of x_i except for the three rightmost points ($i = N-2, N-1, N$), where $|F| < 10^{-10}$ (note that the range of g covers 140 orders of magnitude!). For comparison, the final stationary distribution w_{stat} is reached already after 1.09 ns (w_{stat} is independent of x_n with an accuracy of 13 digits).

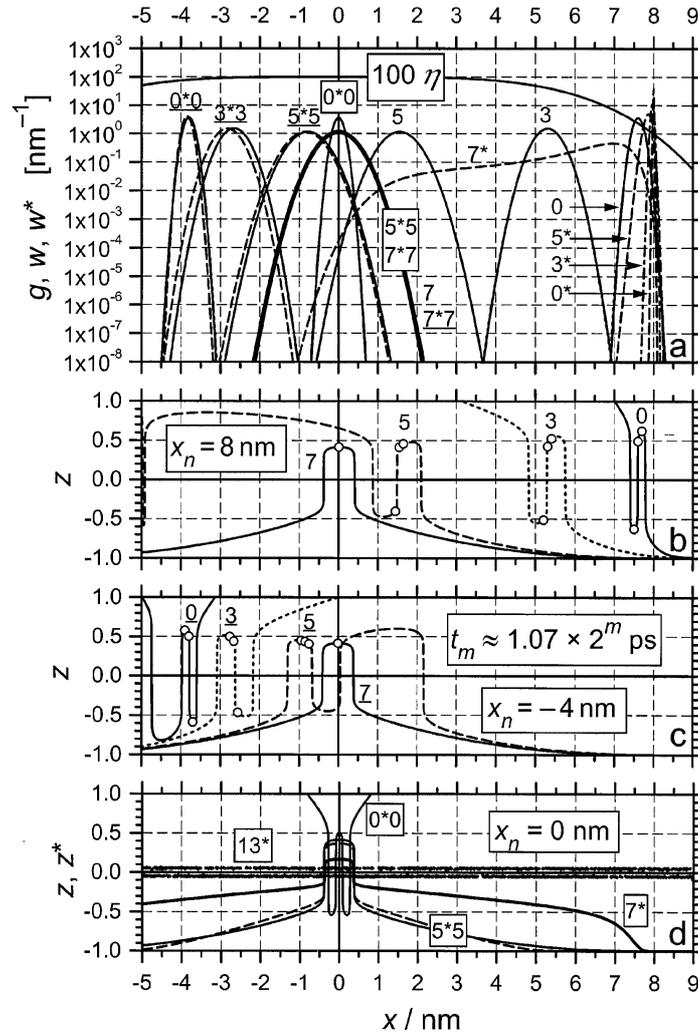


Fig. 4. Example 1: One-dimensional diffusion in a reduced harmonic potential $u(x) = \alpha x^2$ with $\alpha = 4 \text{ nm}^2$ and $D = 6 \times 10^9 \text{ nm}^2 \text{ s}^{-1}$; negative effect of a *detailed balance* on the accuracy of the numerical fundamental solutions calculated with $J = 1$. (a) Three sets of curves are shown, referring to the start positions $x_n = 8 \text{ nm}$, $x_n = -4 \text{ nm}$ (underlined numbers), and $x_n = 0 \text{ nm}$ (numbers in frames). The numbers $m = 0, 3, 5, 7, 13$ refer to the times $t_m \approx 1.07 \times 2^m \text{ ps}$. The curves g (—) and w (⋯) virtually coincide and cannot be distinguished from each other; the curves corresponding to t_7 virtually coincide with the stationary distribution (thick solid curve). The curves w^* (---) correspond to a detailed balance. η is defined by Eq. (51c). (b) Relative deviation z of w from g for $x_n = 8 \text{ nm}$. (c) Relative deviation z of w from g for $x_n = -4 \text{ nm}$. (d) Relative deviations z (---) and z^* (—) of w and w^* from g for $x_n = 0 \text{ nm}$. The stationary distribution w_{stat}^* was reached at $t_{13} \approx 8.74 \text{ ns}$; the corresponding data points z_{stat}^* are statistically distributed relative to $z = 0$ with $|z_{\text{stat}}^*| \approx 0.05$.

2.9 Example 2: Diffusion in a double-minimum potential with reflecting boundaries

A double minimum potential with reflecting boundaries at $x = X_0$ and $x = X_N = -X_0$ is defined by a symmetrical polynomial of the 8th degree:

$$u(x) = u(0) [1 - 8 (x/X_N)^2 + 20 (x/X_N)^4 - 16 (x/X_N)^6 + 4 (x/X_N)^8]. \quad (52)$$

(This polynomial is easily obtained by a transformation of the Chebyshev polynomial $T_8(x)$ [21, 23].) The following computations were performed with $X_0 = -8.005$ nm, $X_N = 8.005$ nm, $N = 1601$, $h = 0.01$ nm, and $u(0) \approx 34.539$, which corresponds to an equilibrium distribution extending over 15 orders of magnitude. The results in Fig. 5 illustrate four aspects of the algorithm:

(a) The probability conservation remains exact if the potential is given by a polynomial of a degree $\leq 2J + 1$.

(b) The systematic errors of the stationary numerical solution close to a reflecting boundary are very small, if the potential at the boundary satisfies Eq. (35).

(c) If the barrier between the two potential minima is high, two very different time scales can be distinguished (as is well known from Kramers's work [31, 32]). If $x_1 \leq x_n < 0$ nm, then in the left part of the potential a quasi-stationary distribution is attained in about 1 ns. The attainment of the final equilibrium distribution is extremely slow: Close to the left potential minimum, the difference $w(x_i, t | x_n) - w_{\text{stat}}(x_i)$ decays exponentially almost exactly, with a time constant $\tau_{\text{stat}} \approx 7.28 \times 10^4$ s. At the time $t_{\text{stat}} \approx 2.46 \times 10^6$ s $\approx 33.8 \tau_{\text{stat}}$, all N fundamental solutions $w(x_i, t | x_n)$ were equal with an accuracy of 13 digits. The last time step of 1.23×10^6 s was by a factor of $2^{108} \approx 3.245 \times 10^{32}$ longer than the first time step, $t_a \approx 3.79 \times 10^{-27}$ s. Note that this problem cannot be solved with a standard method, since even in the most favorable case, where $t_a \approx 4.27 \times 10^{-15}$ s is already close to the stability limit of the algorithm and $t_{\text{stat}} \approx 5\tau_{\text{stat}} \approx 3.64 \times 10^5$ s, about 10^{20} time steps would be needed.

(d) The agreement of the stationary numerical distribution, $w_{\text{stat}}(x)$, with the exact equilibrium distribution, $g_{\text{eq}}(x)$, is very good. $|F|$ is less than 10^{-11} in the vicinity of the maxima of $g_{\text{eq}}(x)$, less than 10^{-10} in a range of more than 5 orders of magnitude of $g_{\text{eq}}(x)$, and less than 10^{-8} everywhere except for x_1 and x_N , where $F \approx 3 \times 10^{-8}$. The lack of complete mirror symmetry of the curve $z(x)$ in Fig. 5b is caused by the insufficient numerical accuracy.

3. Spherically symmetric three-dimensional diffusion

The treatment of spherically symmetric three-dimensional diffusion is essentially analogous to that of one-dimensional diffusion. Therefore, in the following only the peculiar aspects of three-dimensional diffusion are treated in detail.

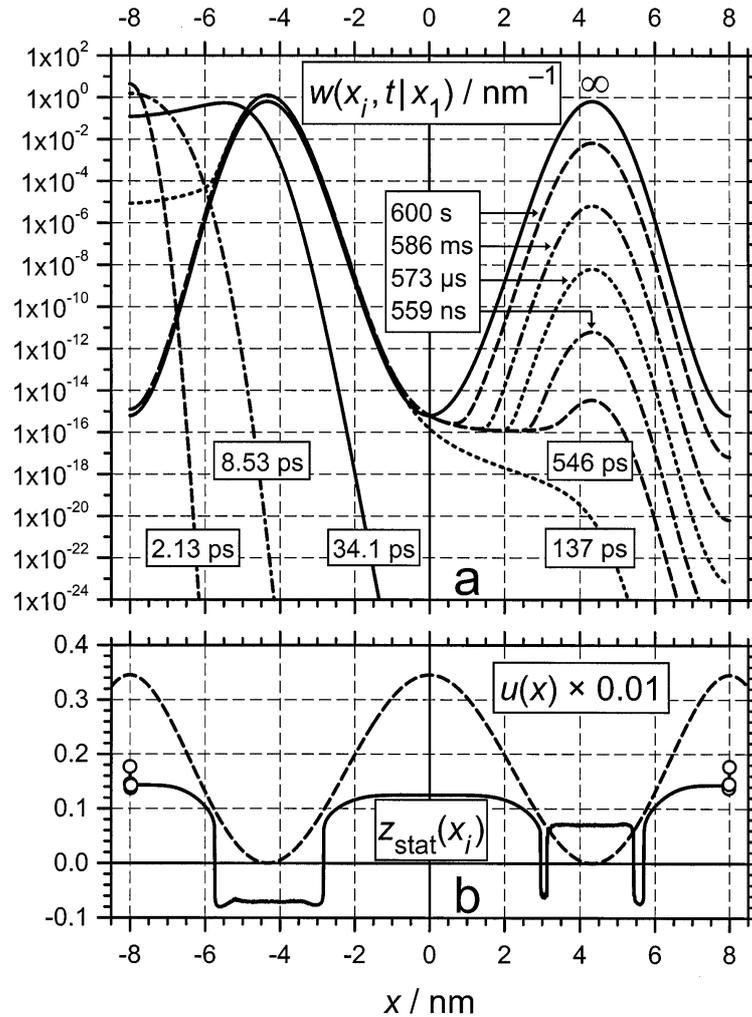


Fig. 5. Example 2: One-dimensional diffusion in a reduced double minimum potential $u(x)$ defined by Eq. (52): $X_0 = -8.005 \text{ nm}$, $X_N = 8.005 \text{ nm}$, $N = 1601$, $h = 0.01 \text{ nm}$, $u(0) \approx 34.539$, $D = 6 \times 10^9 \text{ nm}^2 \text{ s}^{-1}$. (a) Numerical fundamental solutions w (first time step $t_a \approx 3.79 \times 10^{-27} \text{ s}$). (b) Reduced potential $u(x)$ and relative deviation z_{stat} of the stationary distribution w_{stat} from the equilibrium distribution g_{eq} .

3.1 Definition of spherical shells and mean radii

Let a sphere of radius R_0 be surrounded by a series of concentric spherical surfaces with increasing radii $R_1, R_2, \dots, R_i, \dots, R_{N-1}, R_N$ with

$$R_i = R_{i-1} \times f_i. \quad (53)$$

The volume of the i -th spherical shell between the radii R_{i-1} and R_i is

$$V_i = (4\pi/3) (R_i^3 - R_{i-1}^3). \quad (54)$$

As effective radius r_i of the i -th spherical shell the mean radius of the shell is chosen:

$$r_i = \langle r \rangle_{V_i} = \frac{4\pi}{3V_i} \int_{R_{i-1}}^{R_i} r^3 dr. \quad (55)$$

In the important special case of constant f , Eqs. (53) to (55) can be written as follows:

$$R_i = R_0 \times f^i, \quad (53a)$$

$$v_i = (4\pi/3) R_0^3 f^{3i-3} (f^3 - 1) = V_i f^{3i-3}, \quad (54a)$$

$$r_i = R_0 \left(\frac{3(f^4 - 1)}{4(f^3 - 1)} \right) f^{i-1} = \langle r \rangle_{V_i} f^{i-1}. \quad (55a)$$

If the initial radial increments must be very small and the maximum radius r_N is very large, different values of f can be used for the near zone and for the far zone: $f_1 = 1 + q_1$ and $f_2 = 1 + q_2$ with $q_1 < q_2$. Let $f = f(i)$ be a continuous function of the index i . A necessary condition for a smooth transition from f_1 to f_2 is that the second derivative of $f(i)$ with respect to i be continuous. The following function $f(i)$ has this property and turns out to be flexible enough for the present purpose:

$$f(i) = f_1 + (f_2 - f_1) (1 - \exp(-(i/i_{1-2})[1 - \exp(-i/i_{1-2})]^\kappa)). \quad (56)$$

Here the index i_{1-2} roughly defines the beginning transition from f_1 to f_2 and the exponent κ defines the steepness of that transition. An application of Eq. (56) is presented in section 3.6 (the corresponding function $f(i)$ is shown in Fig. 9b). Other functions $f(i)$ were also tested. A single polynomial and a combination of two polynomials (for different ranges of i) turned out to be unsuitable.

3.2 Fundamental solutions and probability conservation

Let a fundamental solution at the time $t = t_a$, $w(r_i, t_a | r_n)$, be defined by Eq. (57):

$$w(r_i, t_a | r_n) \approx w(r_i, 0 | r_n) + \left(\frac{\partial w(r_i, t | r_n)}{\partial t} \right)_{t=0} \times t_a. \quad (57)$$

$w(r_i, 0 | r_n)$ is defined by Eq. (2) and $[\partial w(r_i, t | r_n)/\partial t]_{t=0}$ is calculated with Eq. (1). The required first and second derivatives with respect to r , $w'(r_i, 0 | r_n)$ and $w''(r_i, 0 | r_n)$, are again defined by a set of $2J + 1$ polynomials $y_{n+j,n}(r)$ of the degree $2J$ with $j = 0, \pm 1, \dots, \pm J$:

$$w'(r_{n+j}, 0 | r_n) = V_n^{-1} y'_{n+j,n}(r_{n+j}), \quad (58)$$

$$w''(r_{n+j}, 0 | r_n) = V_n^{-1} y''_{n+j,n}(r_{n+j}). \quad (59)$$

With $J = 2$, the $2J + 1 = 5$ polynomials are

$$y_{n,n}(r) = A_{n,0} (r - r_{n-2}) (r - r_{n-1}) (r - r_{n+1}) (r - r_{n+2}), \quad (60)$$

$$y_{n+1,n}(r) = A_{n,+1} (r - r_{n-1}) (r - r_{n+1}) (r - r_{n+2}) (r - r_{n+3}), \quad (61)$$

$$y_{n-1,n}(r) = A_{n,-1} (r - r_{n-3}) (r - r_{n-2}) (r - r_{n-1}) (r - r_{n+1}), \quad (62)$$

$$y_{n+2,n}(r) = A_{n,+2} (r - r_{n+1}) (r - r_{n+2}) (r - r_{n+3}) (r - r_{n+4}), \quad (63)$$

$$y_{n-2,n}(r) = A_{n,-2} (r - r_{n-4}) (r - r_{n-3}) (r - r_{n-2}) (r - r_{n-1}). \quad (64)$$

The factors $A_{n,j}$ are obtained from the condition that all five polynomials pass through the point $y_{n,n}(r_n)$:

$$y_{n+j,n}(r_n) = y_{n,n}(r_n) = 1 \quad (j = 0, \pm 1, \pm 2). \quad (65)$$

For the calculation of the fundamental solutions $w(r_{n+j}, t_a | r_n)$ with $j = 0, \pm 1, \dots, \pm J$, the quantities

$$H_{n,j} \equiv V_n [\partial w(r, t | r_n) / \partial t]_{r=r_{n+j}, t=0} \quad (66)$$

are needed, which are according to Eq. (1) and with $u \equiv U/k_B T$ and $k = 0$:

$$\begin{aligned} H_{n,j} = & D(r_{n+j}) y''_{n+j,n}(r_{n+j}) + [(2/r_{n+j})D(r_{n+j}) + D'(r_{n+j}) \\ & + D(r_{n+j}) u'(r_{n+j})] y'_{n+j,n}(r_{n+j}) + [(2/r_{n+j})D(r_{n+j}) u'(r_{n+j}) \\ & + D'(r_{n+j}) u'(r_{n+j}) + D(r_{n+j}) u''(r_{n+j})] y_{n+j,n}(r_{n+j}). \end{aligned} \quad (67)$$

In the special case of a constant potential and a constant relative diffusion coefficient, Eq. (67) reduces to

$$H_{n,j} = D[y''_{n+j,n}(r_{n+j}) + (2/r_{n+j}) y'_{n+j,n}(r_{n+j})]. \quad (68)$$

The probability balance is now defined by

$$S_n = \sum_{j=-J}^{+J} H_{n,j} V_{n+j}. \quad (69)$$

In contrast to one-dimensional diffusion with constant intervals Δx , the probability balance S_n is no longer exactly equal to zero. The relative deviation of S_n from zero is defined by

$$\omega_n = S_n / |H_{n,0} V_n|. \quad (70)$$

In the special case of constant u and D (cf. Eq. (68)) and a constant radial increment factor $f \equiv 1 + q$, ω is independent of n . By using the definitions of V_i and r_i in Eqs. (54a) and (55a), the following expressions for ω_n can be derived for $J = 1, 2$:

Table 2. Relative deviation $(\omega_n)_{2J}$ ($J = 1, 2, 3, 4$) of the three-dimensional diffusion balance from zero for different values of q . $(\omega_n)_2$ and $(\omega_n)_4$ were calculated with Eqs. (71) and (72), respectively. The values of $(\omega_n)_6$ and $(\omega_n)_8$ are the results of purely numerical calculations.

q	$(\omega_n)_2$	$(\omega_n)_4$	$(\omega_n)_6$	$(\omega_n)_8$
0.001	-9.98×10^{-13}	$+6.38 \times 10^{-18}$		
0.002	-1.59×10^{-11}	$+4.07 \times 10^{-16}$		
0.003	-8.05×10^{-11}	$+4.62 \times 10^{-15}$		
0.004	-2.54×10^{-10}	$+2.59 \times 10^{-14}$		
0.006	-1.28×10^{-9}	$+2.93 \times 10^{-13}$		
0.008	-4.03×10^{-9}	$+1.64 \times 10^{-12}$		
0.010	-9.80×10^{-9}	$+6.21 \times 10^{-12}$		
0.015	-4.92×10^{-8}	$+6.98 \times 10^{-11}$		
0.020	-1.54×10^{-7}	$+3.86 \times 10^{-10}$	-1.9×10^{-12}	
0.025	-3.72×10^{-7}	$+1.45 \times 10^{-9}$	-1.10×10^{-11}	
0.030	-7.64×10^{-7}	$+4.28 \times 10^{-9}$	-4.64×10^{-11}	$+8. \times 10^{-13}$
0.040	-2.37×10^{-6}	$+2.34 \times 10^{-8}$	-4.48×10^{-10}	$+1.41 \times 10^{-11}$
0.060	-1.16×10^{-5}	$+2.53 \times 10^{-7}$	-1.08×10^{-8}	$+7.55 \times 10^{-10}$
0.080	-3.53×10^{-5}	$+1.35 \times 10^{-6}$	-1.01×10^{-7}	$+1.25 \times 10^{-8}$
0.100	-8.34×10^{-5}	$+4.93 \times 10^{-6}$	-5.70×10^{-7}	$+1.10 \times 10^{-7}$
0.150	-3.90×10^{-4}	$+5.06 \times 10^{-5}$	-1.30×10^{-5}	$+5.63 \times 10^{-6}$
0.200	-1.15×10^{-3}	$+2.60 \times 10^{-4}$	-1.19×10^{-4}	$+9.40 \times 10^{-5}$
0.250	-2.61×10^{-3}	$+9.23 \times 10^{-4}$	-6.70×10^{-4}	$+8.78 \times 10^{-4}$

$$(\omega_n)_{2J=2} = \frac{-q^4}{1 + 2q - q^3}, \tag{71}$$

$$(\omega_n)_{2J=4} = \frac{32q^6(1 + 4q + \frac{22}{3}q^2 + 8q^3 + \frac{271}{48}q^4 + \frac{21}{8}q^5 + \frac{25}{32}q^6 + \frac{13}{96}q^7 + \frac{1}{96}q^8)}{5(1 + q)^3(1 + q + \frac{1}{2}q^2)(1 + q + \frac{1}{3}q^2)(1 + 2q - \frac{6}{5}q^2 - \frac{11}{5}q^3 - \frac{3}{5}q^4)}. \tag{72}$$

In Table 2 values of $(\omega_n)_{2J}$ are listed for $J = 1, 2, 3, 4$ and different values of q . For $J \geq 2$, $(\omega_n)_{2J}$ is extremely small in the whole range of q that is of practical interest. The purely numerically calculated values of $(\omega_n)_6$ and $(\omega_n)_8$ become rather inaccurate and finally meaningless with decreasing q (as a result of insufficient numerical accuracy). Nevertheless one may estimate that $(\omega_n)_6$ and $(\omega_n)_8$ are approximately proportional to q^8 and q^{10} , respectively, for small values of q . Thus, for small values of q , $(\omega_n)_{2J}$ seems to be approximately proportional to q^{2J+2} for arbitrary $J \geq 1$.

In the general case of $u = u(r_i)$, $D = D(r_i)$, and $f = f(i)$, additional systematic errors are introduced in the probability balance. However, by a suitable choice of $f = f(i)$ and with $J = 4$, $|(\omega_n)_8|$ can be kept very small. For instance, in the example in section 3.6, the relation $|(\omega_n)_8| \leq 1.2 \times 10^{-9}$ was always satisfied.

The very small error in the probability balance is corrected for by introducing correction factors φ . Since it is unknown how the deviation from the exact probability conservation is distributed over the central term $H_{n,0}V_n$ on the one hand and the $2J$ other terms $H_{n,j}V_{n+j}$ with $j \neq 0$ on the other hand, the following symmetrical definition of φ is preferred:

$$H_{n,0}V_n \times \varphi_{n,0} + (S_n - H_{n,0}V_n) \times \varphi_{n,j \neq 0} = 0 \quad (73)$$

or

$$\varphi_{n,0}/\varphi_{n,j \neq 0} = 1 - [S_n/(H_{n,0}V_n)]. \quad (74)$$

The corrected values of H are: $(H_{n,0})_{\text{corr}} = H_{n,0} \times \varphi_{n,0}$ and $(H_{n,j \neq 0})_{\text{corr}} = H_{n,j \neq 0} \times \varphi_{n,j \neq 0}$ with $j = \pm 1, \pm 2, \dots, \pm J$. Three possibilities were tested: (a) $\varphi_{n,0} = 1$ and $\varphi_{n,j \neq 0} \neq 1$; (b) $\varphi_{n,0} \neq 1$ and $\varphi_{n,j \neq 0} = 1$; (c) $\varphi_{n,0} = 1/\varphi_{n,j \neq 0}$. In practice it turned out that the final results were almost independent of the specific correction procedure (in the examples 3, 4, and 5, always the correction (c) was applied). Asymmetric corrections, in which either the terms $H_{n,j \geq 1}$ or the terms $H_{n,j \leq -1}$ were multiplied by a correction factor, gave nearly the same results. In the following the subscript ‘‘corr’’ is omitted, since always corrected values of $H_{n,j}$ are implied.

3.3 The first diffusion step and boundary conditions

For calculating the result of the first diffusion step of duration t_a , relative changes $b_{n,j}$ of $w(r_{n+j}, 0 | r_n)$ are again defined by Eqs. (30) and (31), but with the quantities $H_{n,j}$ now given by Eq. (67) and corrected according to Eq. (73). The boundary conditions in Eqs. (34)–(36) remain also valid if x and X_0 are replaced by r and R_0 . The only essential change concerns the numerical implementation of the boundary conditions. Since the constant interval $\Delta x = h$ is replaced with the variable volume V_i of the spherical shells, Eqs. (37) and (38) are to be replaced by the equations

$$b_{1,-1} \times V_0 + b_{1,-2} \times V_{-1} = c_0 (b_{0,+1} \times V_1 + b_{0,+2} \times V_2), \quad (75)$$

$$b_{2,-2} \times V_0 = c_{-1} (b_{-1,+2} \times V_1). \quad (76)$$

The definition of a constant-concentration boundary by Eqs. (37a)–(42a) remains unchanged.

3.4 Example 3: Free diffusion near a reflecting or absorbing spherical surface

Let the inner boundary be again specified (as in section 2.5) by the parameter c_{BC} with $c_{\text{BC}} = +1$ for reflection and $c_{\text{BC}} = -1$ for absorption. Then, with constant potential and constant relative diffusion coefficient, the common theoretical fundamental solution can be written in the form [4, 33]:

$$g(r_i, t | r_n) = \frac{1}{8\pi r_i r_n \sqrt{\pi D t}} \left[\exp\left(-\frac{(r_i - r_n)^2}{4Dt}\right) + c_{\text{BC}} \exp\left(-\frac{(r_i + r_n - 2R_0)^2}{4Dt}\right) \right. \\ \left. \times \left(1 - \frac{2(1 + c_{\text{BC}}) \sqrt{Dt}}{R_0} \int_0^{\infty} \exp\left(-x^2 - 2\left[\frac{\sqrt{Dt}}{R_0} + \frac{(r_i + r_n - 2R_0)}{2\sqrt{Dt}}\right]x\right) dx\right) \right]. \quad (77)$$

Eq. (77) is obtained from Eq. (24) in ref. [4] by equating $k_{\text{act}} = 0$ and inserting c_{BC} before the second term and $(1 + c_{\text{BC}})/2$ before the integral, and finally by transforming the complementary error function into the present integral. (The computational disadvantage of the complementary error function is the simultaneous appearance of an exponential function with a positive exponent, which makes it impossible to calculate $g(r_i, t | r_n)$ for large values of $r_i + r_n$ or Dt . The integral was calculated with a numerical accuracy of 15 digits for arbitrary values of the bracketed term in the integrand.)

For the comparison of theoretical curves with numerically calculated curves it is advantageous to apply to theoretical curves g the same normalization as to numerical curves w . This is achieved by introducing a factor $(1 + \varepsilon)$, which is defined by the equation

$$(1 + \varepsilon(t)) \times \sum_{i=1}^N g(r_i, t | r_n) V_i = 4\pi \int_{R_0}^{\infty} g(r, t | r_n) r^2 dr. \quad (78)$$

In the following the symbol $g^* \equiv (1 + \varepsilon) g$ is used. In the present example, ε is almost independent of time in the time range of interest (from 4 ps to 4 ns): $\varepsilon \approx 1.244 \times 10^{-5}$. The notation for the fundamental solutions is simplified. Since the values of w and g^* always refer to specified values of t and r_n , the simplified notation is $g_i^* \equiv g^*(r_i, t | r_n)$ and $w_i \equiv w(r_i, t | r_n)$.

In Fig. 6a six pairs of theoretical curves g_i^* are shown, which were calculated with $R_0 = 0.8$ nm, $f = 1.005$, $r_n \approx 2.00$ nm, $n = 184$, $r_N \approx 2319$ nm, $N = 1600$, $D = 6 \times 10^9$ nm² s⁻¹. The solid curves refer to a reflecting boundary at R_0 ($c_{\text{BC}} = +1$ in Eq. (77)), and the dashed curves refer to an absorbing boundary at R_0 ($c_{\text{BC}} = -1$ in Eq. (77)). The curves cover the time range from ≈ 4 ps (curves 0) to 4 ns (curves 10) in steps of a factor of 4. The four leftmost data points of the dashed curves 0, 2, and 4 are represented by circles.

The numerical fundamental solutions $w(r_i, t | r_n)$ were computed with $J = 4$. In the case of a reflecting boundary at R_0 , a very small value of t_a ($\approx 4.136 \times 10^{-34}$ s) was used, and the fundamental solutions $w(r_i, t | r_n)$ were renormalized after every 10 time-doubling cycles. The relative deviations F of w from g^* are defined in analogy to Eq. (48) and again given in the z -representation (see Eq. (49)) in Fig. 6b. Curve 0 shows the typical limiting pattern of z for $J = 4$ with nine intersections with the $z = 0$ line; a further reduction of t_a would leave $z(r)$ virtually unchanged. The only irregular points are the four leftmost points of each curve, due to the imperfect nu-

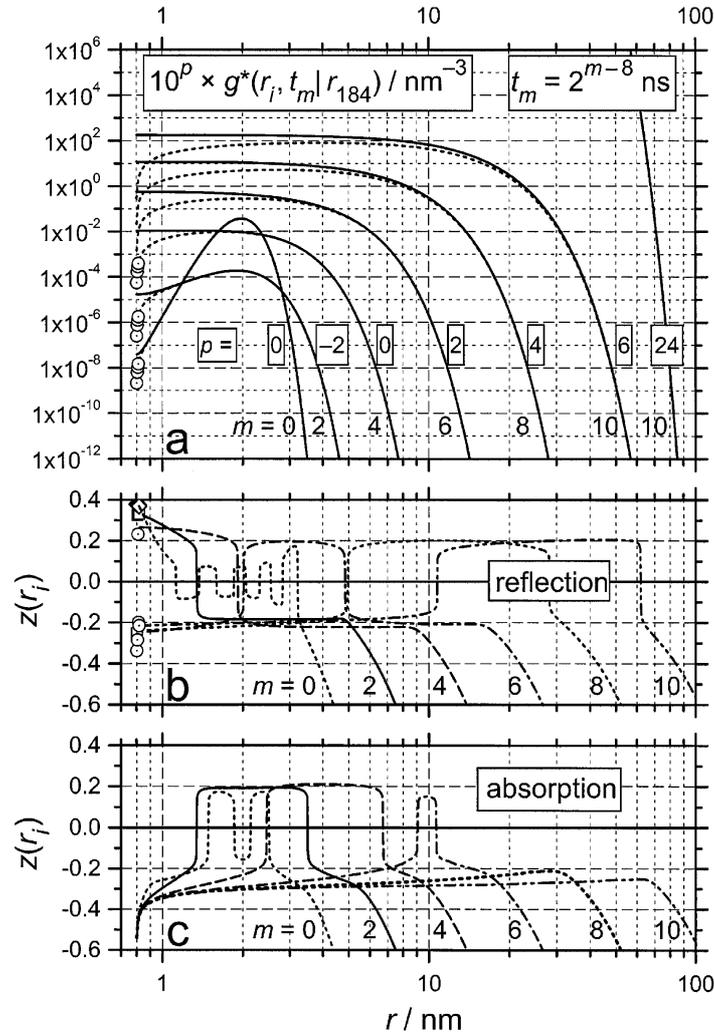


Fig. 6. Example 3: Spherically symmetric three-dimensional diffusion with constant potential and constant relative diffusion coefficient. Parameters: $R_0 = 0.8$ nm, $f = 1.005$, $r_n \approx 2.00$ nm, $n = 184$, $r_N \approx 2319$ nm, $N = 1600$, $D = 6 \times 10^9$ nm² s⁻¹; reflecting boundary at R_N . (a) Theoretical fundamental solutions g^* (see Eqs. (77) and (78)) referring to a reflecting boundary (—) or absorbing boundary (---) at R_0 ; the curves cover the time range from $t_0 \approx 4$ ps (curves 0) to $t_{10} = 4$ ns (curves 10) in steps of a factor of 4. The four leftmost points of the dashed curves 0, 2, and 4 are represented by open circles. For a better distinction of the curves close to the boundary or the coverage of a larger range of g^* , the curves with index $m = 2, 6, 8, 10$ have been multiplied by suitable factors 10^p . (b) Reflecting boundary: Relative deviations z of w from g^* (first time step $t_a \approx 4.136 \times 10^{-34}$ s, renormalization of all w after every 10 time-doubling cycles). (c) Absorbing boundary, relative deviations z of w from g^* ($t_a \approx 1.164 \times 10^{-19}$ s, no renormalization).

merical implementation of the boundary condition. $|F|$ is less than 10^{-4} in a range of 15.7 orders of magnitude for $m = 0$ and in a range of 31.4 orders of magnitude for $m = 10$. At the right side of each curve, $z = -0.4$ or $F \approx -10^{-4}$ corresponds to $g_{i+1}^*/g_i^* \approx 0.5$.

In the case of an absorbing boundary at R_0 , a correction for the accumulation of round-off errors by renormalization can no longer be applied. Therefore, a $2^{48} \approx 2.81 \times 10^{14}$ times larger value of t_a was chosen: $t_a \approx 1.164 \times 10^{-19}$ s. The values of z are shown in Fig. 6c. The only strongly irregular value of z , $z_1 \approx -0.537$, corresponds to $F_1 \approx -2.0 \times 10^{-3}$ and is practically independent of time. This rather large value of F_1 is not surprising in view of the large ratio $g_2^*/g_1^* \approx 3$. The accumulation of round-off errors becomes apparent in the z -curves for $m = 6, 8, 10$.

3.5 Example 4: $u = u(r)$, $D = D(r)$, small radial range with constant ratio r_{i+1}/r_i

For the present test calculations an arbitrary potential function was chosen:

$$u(r) = u(R_0) \times \exp[-\ln(2)|(r - R_0)/\Gamma|^{n_{\text{pot}}}] \times [1 + f_{\text{pot}} \ln(2)|(r - R_0)/\Gamma|^{n_{\text{pot}}}] \quad (79)$$

The second factor is a good approximation to a box potential for high powers n_{pot} . The last factor allows one to introduce additionally a potential barrier. Γ is the halfwidth of the potential if $f_{\text{pot}} = 0$. For the relative diffusion coefficient the simple expression proposed by Northrup and Hynes [11] was used,

$$D(r) = D(\infty) [1 - a_{\text{diff}} \exp(-(r - R_0)/R_0)], \quad (80)$$

with $0 \leq a_{\text{diff}} < 1$. The following computations were performed with $R_0 = 0.8$ nm, $R_N \approx 1.78$ nm, $u(R_0) = -14.39$, $n_{\text{pot}} = 4$, $\Gamma = 0.2$ nm, $f_{\text{pot}} = 2$, $D(\infty) = 6 \times 10^9$ nm² s⁻¹ and $a_{\text{diff}} = 0.5$. The corresponding function $u(r)$ is shown in Fig. 7. The present values of R_0 , $u(R_0)$, and $D(\infty)$ roughly correspond to the formation and dissociation of pyrene excimers in a fluid solvent like hexane at room temperature (cf. the Appendix).

Starting from $f = 1.002$, $N = 400$ and polynomials of the degree $2J = 4$, the numerical accuracy was increased either by a reduction of f (and a corresponding increase of N) or by an increase of J . Two sets of computations were performed, whose main results are presented in Fig. 8. The computations of the first set had $2J = 4$ and $t_a \approx 1.421 \times 10^{-23}$ s in common, and the values of f and N were varied: $f = 1.002$, $N = 400$; $f = 1.002^{1/2} \approx 1.001$, $N = 800$; $f = 1.002^{1/4} \approx 1.0005$, $N = 1600$. The computations of the second set had $f = 1.002$, $N = 400$, and $t_a \approx 3.638 \times 10^{-21}$ s in common, and the degree of the polynomials was varied: $2J = 4, 6, 8$. The final stationary distributions were independent of the value of t_a . $t_{\text{stat}} = 512$ ns was a practically infinitely long time: $w(r_i, t_{\text{stat}} | r_N)$ agreed with $w(r_i, t_{\text{stat}} | r_1)$ with an accuracy of 13 digits.

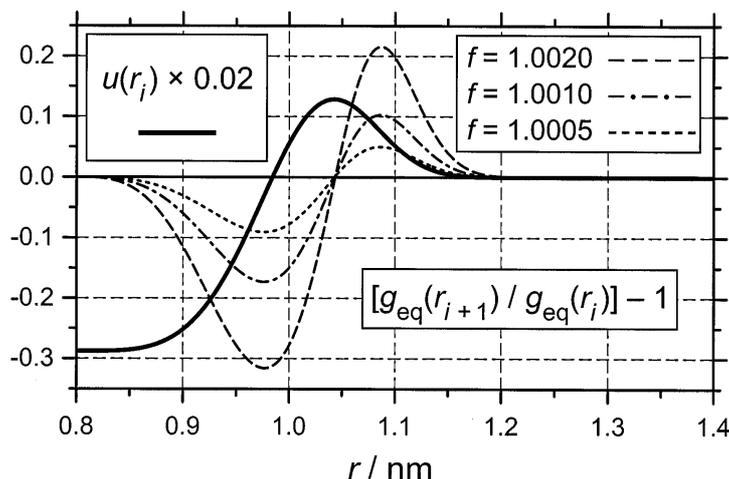


Fig. 7. Reduced potential $u(r)$, calculated with Eq. (79) and the parameters $R_0 = 0.8$ nm, $u(R_0) = -14.39$, $n_{\text{pot}} = 4$, $\Gamma = 0.2$ nm, and $f_{\text{pot}} = 2$. The ratio of successive equilibrium values of g , $g_{\text{eq}}(r_{i+1})/g_{\text{eq}}(r_i) = \exp(-[u(r_{i+1}) - u(r_i)])$, was calculated with the radial increment factors $f = 1.002$, 1.001 , and 1.0005 .

The increase of numerical accuracy by a reduction of $q = f - 1$ is demonstrated in Fig. 8b. The almost constant systematic error F for $r > 1.05$ nm ranges from $\approx 1.83 \times 10^{-5}$ for $q \approx 0.0005$ to $\approx 4.75 \times 10^{-3}$ for $q = 0.002$ and is rather accurately proportional to q^4 . The irregular systematic error F_1 due to the imperfect implementation of the boundary condition ranges from $\approx 5.1 \times 10^{-6}$ for $q \approx 0.0005$ to $\approx 1.99 \times 10^{-4}$ for $q = 0.002$ and is rather accurately proportional to q .

In Fig. 8c the increase of numerical accuracy by the increase of the degree of the interpolation polynomials is demonstrated. In order to eliminate the irregularities due to the boundaries, the difference $(F_{390} - F_{11})$ is taken as a measure for the attainable accuracy: $(F_{390} - F_{11})_{J=2} \approx 4.75 \times 10^{-3}$, $(F_{390} - F_{11})_{J=3} \approx -1.36 \times 10^{-4}$, $(F_{390} - F_{11})_{J=4} \approx 3.72 \times 10^{-6}$. Thus, in the present example, an increase of the degree of the polynomials by 2 entails a reduction of the systematic error due to the potential to $\approx 1/36$.

3.6 Example 5: $u = u(r)$, $D = D(r)$, very large radial range with increasing ratio r_{i+1}/r_i

The last example differs from the preceding one in two respects: First, $q = f - 1$ is gradually increased from $q_1 \approx 0.002$ to the final value $q_N \approx 0.02$ (see Fig. 9b). Second, the radial range is extremely large: $R_0 = 0.8$ nm and $R_N \approx 8 \times 10^5$ nm ($N = 1409$). As a representative example, the evolution of the fundamental solution with $r_n \approx 800$ nm ($n = 1043$) is shown in

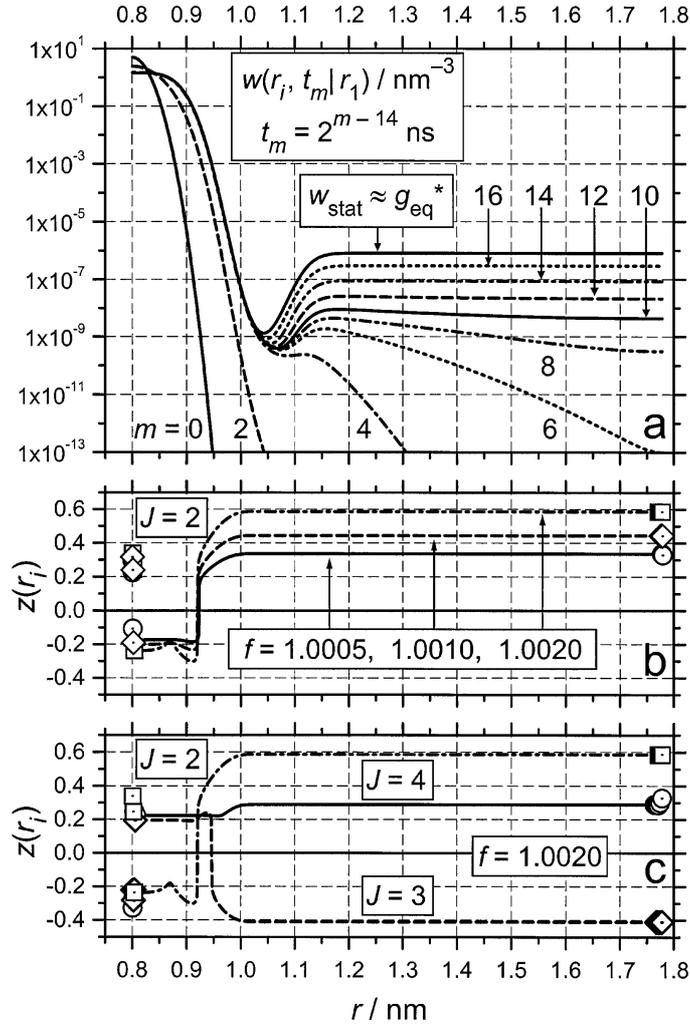


Fig. 8. Example 4: Relative diffusion in a small radial range with a reduced potential $u = u(r)$ (cf. Fig. 7), $D = D(r)$ ($a_{\text{diff}} = 0.5$, $D(\infty) = 6 \times 10^9 \text{ nm}^2 \text{ s}^{-1}$; cf. Eq. (80)) and reflecting boundaries at $R_0 = 0.8 \text{ nm}$ and $R_N \approx 1.78 \text{ nm}$. (a) w was calculated with $f = 1.002$, $N = 400$, $t_a \approx 1.084 \times 10^{-28} \text{ s}$, and $J = 4$. The times $t_m = 2^{m-14} \text{ ns}$ for the curves in the figure range from $t_0 \approx 61 \text{ fs}$ to $t_{16} = 4 \text{ ns}$. The solid curve with the label ∞ is the stationary distribution, which is reached at $t_{\text{stat}} = t_{23} = 512 \text{ ns}$ (w_{stat} is independent of r_n with an accuracy of 13 digits). (b) $J = 2$: Effect of radial increment factor f on the relative deviation z of w_{stat} from the equilibrium distribution g_{eq}^* ($f = 1.002$, $N = 400$; $f = 1.002^{1/2} \approx 1.001$, $N = 800$; $f = 1.002^{1/4} \approx 1.0005$, $N = 1600$). The three leftmost and the three rightmost points of each z -curve are represented by separate symbols. (c) $f = 1.002$, $N = 400$: Effect of the degree $2J$ of the interpolation polynomials on the relative deviation z of w_{stat} from the equilibrium distribution g_{eq}^* . The $J + 1$ leftmost and $J + 1$ rightmost points of each z -curve are represented by separate symbols.

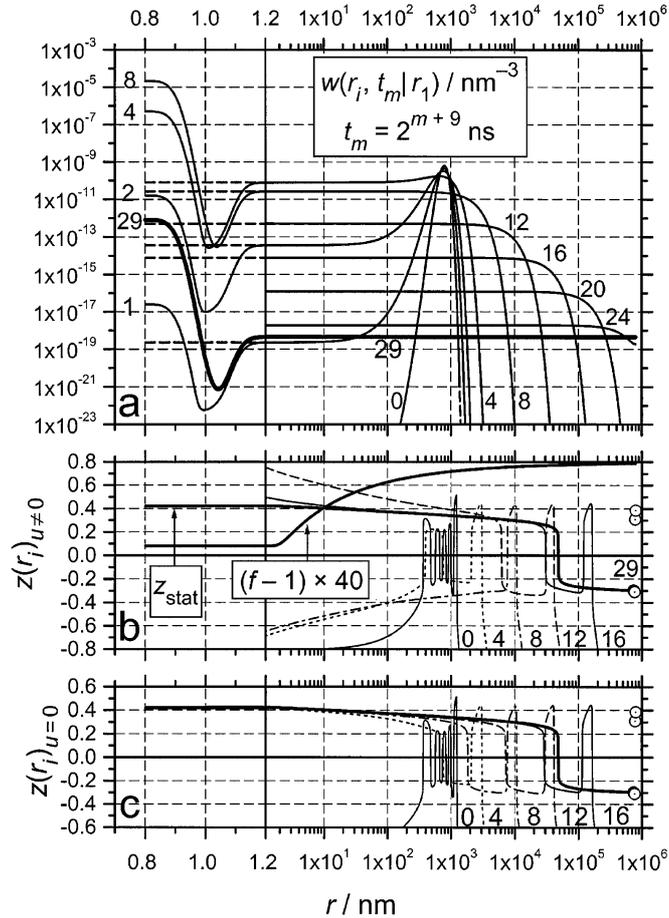


Fig. 9. Example 5: Relative diffusion with the same reduced potential $u(r)$ and relative diffusion coefficient as in Example 4 (Fig. 8), but with a very large radial range: $R_0 = 0.8 \text{ nm}$ and $R_N \approx 8 \times 10^5 \text{ nm}$ ($f_1 = 1.002$, $f_2 = 1.02$, $N = 1409$, transition from f_1 to f_2 with Eq. (56) and $\kappa = 16$ and $i_{1-2} = 300$, $r_1 \approx 0.8008 \text{ nm}$, $r_N \approx 8.005 \times 10^5 \text{ nm}$, $t_a \approx 1.084 \times 10^{-28} \text{ s}$). All curves refer to a start position $r_n \approx 800 \text{ nm}$ ($n = 1043$). The abscissa is linear from 0.7 to 1.2 nm and logarithmic from 1.2 to 10^6 nm . (a) Numerical fundamental solutions w (—): The times $t_m = 2^{m+9} \text{ ns}$ for the curves in the figure range from $t_0 = 512 \text{ ns}$ to $t_{29} = t_{\text{stat}} \approx 275 \text{ s}$ (thick solid curve, $m = 29$). Since for times $t_8 < t_m < t_{24}$ the shape of the curves $w(r_i, t_m | r_n)$ remains virtually constant in the range $0.8 \text{ nm} < r_i < 1.2 \text{ nm}$ (only the amplitude decreases), these curves are not shown. For comparison the theoretical curves g^* (---) were calculated for the times t_0 to t_{16} with Eqs. (77) and (78), $D = D(\infty)$, and $u = 0$. (b) Relative deviation z of w from g^* ($m \leq 16$, $r_i > 1.2 \text{ nm}$, thin curves) and of w_{stat} from the equilibrium distribution g_{eq}^* (thick solid curve); the four rightmost points of z_{stat} are represented by open circles. The curve $q = f - 1$ was calculated with Eq. (56) and the parameter values given above. (c) Relative deviation z of w (computed also with $D = D(\infty)$ and $u = 0$) from g^* ; the thick solid curve is identical with curve 29 in Fig. 9b.

Fig. 9a. The scale for the radius is linear from 0.7 to 1.2 nm and logarithmic from 1.2 to 10^6 nm. Since the initial evolution takes place in the radial range where u and D are virtually constant, $w(r_i, t | r_n)$ can be compared with the theoretical curves $g^*(r_i, t | r_n)$, calculated with Eqs. (77) and (78) (the value of ε ranged from 1.62×10^{-4} with curve 0 to 1.84×10^{-4} with curve 16). The main results are:

(a) The N numerical fundamental solutions $w(r_i, t | r_n)$ were equal within 13 digits at the time $t_{\text{stat}} = t_a \times 2^{101} \approx 2.749 \times 10^2$ s ($t_a \approx 1.084 \times 10^{-28}$ s). Thus the last time step ($\approx 1.374 \times 10^2$ s) was by 30 orders of magnitude longer than the first time step.

(b) The maximum relative error F of about 1.8×10^{-4} appears now in the near zone at ≈ 1 nm. This is a consequence of the large radial range and the increasing factor f . Therefore, although the equilibrium probability density in the first shell is by a factor of 1.8×10^6 greater than that in the N -th shell, the equilibrium probability $V_N g_{\text{eq}}(r_N) \approx 5.67 \times 10^{-2}$ is about 5.5×10^{12} times larger than the equilibrium probability $V_1 g_{\text{eq}}(r_1) \approx 1.04 \times 10^{-14}$.

(c) For large radii, where the potential u and the relative diffusion coefficient D are virtually constant, the relative deviation F of w from the theoretical curves g^* is small (see Fig. 9b).

(d) The analogous computations were performed with constant potential u and constant relative diffusion coefficient $D = D(\infty)$. The corresponding relative deviations z of w from the theoretical curves g^* are shown in Fig. 9c. Between 0.8 and 2×10^4 nm, curve 16 in Fig. 9c virtually coincides with curve z_{stat} of Fig. 9b. That means, if q is virtually constant in the near zone and the degree of the polynomials is high enough as in this example, the contribution of the near zone to the total systematic error remains small.

3.7 Implementation of reactions

The effect of a distance-dependent rate coefficient k on a distribution $\rho(r_i, t)$ can be taken into account in different ways. The simplest way is its implementation in the fundamental solutions. Let $1/k$ be in the range $\tau_{\text{min}} \leq 1/k \leq \tau_{\text{max}}$ and let the times t_b, t_c, t_d satisfy the relation

$$t_a \leq t_b \leq t_c \ll \tau_{\text{min}} \leq \tau_{\text{max}} < t_d \approx 7\tau_{\text{max}}. \quad (81)$$

t_b is the time at which the reaction is taken into account for the first time. t_c is the adequate time step for calculating the course of the reaction, and t_d is the longest time of interest for the given kinetic problem. If $w(r_i, t_b | r_n)_{\text{diff}}$ is a fundamental solution implying pure diffusion, then the corresponding fundamental solution implying also reaction is given by

$$w(r_i, t_b | r_n)_{\text{react}} \approx w(r_i, t_b | r_n)_{\text{diff}} \times \exp[-k(r_i)t_b]. \quad (82)$$

The further evolution of the fundamental solutions is calculated as before up to the time t_c . Finally, when the time $t_c \approx 0.01 \tau_{\min}$ is reached, the evolution of the distribution function ρ is calculated with procedure (III) with constant time steps t_c up to a time t_d (cf. section 2.6):

$$\rho(r_i, (v+1) \times t_c) = \sum_{n=1}^N [\rho(r_n, v t_c) \times V_n] \times w_n(r_i, t_c | r_n). \quad (83)$$

Note that the computation time needed for each time step of length t_c is only $1/N$ of the time needed for computing the N fundamental solutions in a time-doubling cycle (if all elements of all fundamental solutions significantly differ from zero).

Within the frame of double-precision computations, the practical choice of t_b is governed by two conditions. First, t_b/τ_{\min} should be very small, and in this respect $t_b = t_a$ would be the best choice, if the effect of round-off errors is neglected. Second, t_b/τ_{\max} must be large enough, e.g. $t_b/\tau_{\max} \geq 10^{-10}$, in order that the round-off error in the computation of $w(r_i, t_b | r_n)_{\text{react}}$ remain small enough for arbitrary r_i . Obviously both conditions can be simultaneously satisfied as long as the ratio τ_{\max}/τ_{\min} is not extremely large (e.g. $\leq 10^6$).

3.8 Computational aspects

The numerical calculations were performed on a personal computer with 128 MB random-access memory and a 366 MHz Intel Pentium II processor. The computer programs were written in FORTRAN 90 (compiler: Absoft Pro Fortran, version 6.2). The computer memory requirements are almost completely given by two two-dimensional arrays of N^2 double-precision floating-point numbers for the storage of the N fundamental solutions and a buffer for the intermediate storage of a new set of them. Thus, for $N = 1000$, roughly 16 MB of random-access memory are required. The length of the executable programs was kept below 1 MB by defining most large arrays as allocatable arrays. No special efforts were made to optimize the computer programs. With $N = 1000$, one cycle of time-doubling took $t_{\text{cycle}} \approx 4$ min, when *all* N elements of *all* N fundamental solutions were significantly different from zero.

The interpolation polynomials were calculated with Newton's method [21, 22], which yields the polynomials in a form that is very suitable for calculating their first and second derivatives. Most of the mathematical operations can be formulated as matrix operations. From a practical point of view the present use of explicit summations seems to be preferable, because it is better suited for limiting all summations to those terms that are significantly different from zero. At least in the initial time-doubling cycles, the speed of the computation of $w(r_i, t | r_n)$ is enhanced by several orders of

magnitude, if all extremely small elements $w(r_i, t | r_n)$ are equated with zero and only the remaining elements are taken into account. The computation of the curves in Fig. 8a may serve as an example ($N = 1600$ and $t_{\text{cycle}} \approx 16$ min). The initial speed of computation was strongly enhanced by equating all $w(r_i, t | r_n) < 10^{-50} g_{\text{eq}}(r_i)$ with zero and by limiting all summations to terms that differed from zero. Curve 2 in Fig. 8a is the last one ($t \approx 244$ fs), in which part of the values $w(r_i, t | r_i)$ are equal to zero. The 56 cycles needed for computing curve 2 required 38 min. The remaining 21 cycles for computing the stationary distribution ($t_{\text{stat}} = 512$ ns) required 340 min.

4. Discussion

4.1 Merits and deficiencies of the algorithm

The present algorithm was criticized in several respects. Part of the criticism has been already answered in the Introduction and by the numerical solution of the Ornstein-Uhlenbeck problem in section 2.7. Some other points of the criticism, however, deserve an explicit discussion and are treated in the following items.

Comparison with other algorithms. One may ask whether the present algorithm permits one to solve diffusion problems of interest with higher accuracy and in shorter time than any other known algorithm. This question cannot be definitively answered by the author for two reasons. First, the algorithms of most interest [16–19] are mathematically complex, and their implementation would have been difficult for the author. Moreover, the published details are not always sufficient for writing the pertinent program code [16, 19]. Second, although general statements on the accuracy of algorithms can be found in the literature, there is a lack of illustrative examples that demonstrate the range of application and achievable accuracy of a particular algorithm. For these reasons an alternative procedure is suggested. A reader may apply the algorithms at his disposal to the diffusion problems in examples 1 to 5. If he will be able to surpass the present results with respect to numerical accuracy and speed of computation, then the present algorithm will be of no interest to him. To facilitate this comparison, in particular the relative deviations of numerical solutions from exact solutions are shown in a suitable form.

Time-doubling. The repeated *time-doubling* in the present algorithm may look similar to the *doubling-time* in the *boundary-doubling method* developed by Kim *et al.* [20]. The meaning of the term *doubling-time* is, however, completely different from the meaning of the term *time-doubling* in the present algorithm. The doubling-time in ref. [20] is the time at which the *outer boundary* is doubled. If simultaneously the radial step width Δr is

doubled, then the time step Δt can be quadrupled (*adaptive step size control by step doubling*). After each adaptation of step size, however, Δr and Δt remain constant for many successive steps.

Interpolating functions. In certain diffusion or heat-conduction problems, an exponential approximation closely corresponds to the local concentration or temperature profile in a space layer and seems to give the best results [34]. For this reason it was argued that a polynomial interpolation cannot be advantageous for the present purpose. However, close to minima or maxima of a distribution, an exponential approximation can be no longer optimal and is likely to be inferior to the present polynomial approximation, apart from the loss of mathematical simplicity.

Degree of the polynomials. The very small relative error over many orders of magnitude of the numerical fundamental solutions w is perhaps the most surprising result of the present investigation and not yet completely understood. The strong increase of numerical accuracy with increasing degree of the polynomials can be easily understood (cf. pp. 198–200 in ref. [21]). One may ask which gain in accuracy is expected by using polynomials of higher degree than 8. Probably, within the limitations of double-precision calculations, it will not be possible to calculate the required first and second derivatives of the polynomials with sufficient accuracy, if the polynomials are calculated with Newton's algorithm [21] as in the present computer program. The calculation of the data in Table 1 revealed that the smallest terms in a column are sums of large terms with alternating signs. In conclusion, within the limitations of double-precision accuracy, polynomials of the 8th degree seem to be an optimum. Finally it should be noted that the total computation time is virtually independent of the degree of the polynomials.

Boundaries. Another limitation of the attainable accuracy results from the present implementation of a reflecting or absorbing boundary. The systematic errors resulting from this imperfect implementation are not reduced by increasing the degree of polynomials. One might expect that these systematic errors are smaller, when $(\Delta r)_i = r - r_{i-1}$ is constant near the boundaries. In this connection a smooth transition from $(\Delta r)_2$ to $(\Delta r)_N$ with $(\Delta r)_2 \ll (\Delta r)_N$ has been of interest. The pertinent computations yielded, however, no significant reduction of the systematic error.

Detailed balance. Nadler and Schulten [15] have shown that the approximate calculation of partial derivatives with a finite-difference method guarantees neither probability conservation nor a detailed balance. They and some of their successors [16, 19] emphasized the importance of a detailed balance for a good diffusion algorithm. The author has not been able to share this view. On the contrary, it seems to him more likely that, by implementing a detailed balance, in general one cannot obtain an optimal algorithm for the *evolution* of a distribution function. Apart from the illustration of this statement in section 2.8 and Fig. 4, the following consider-

ation may be helpful. The numerical solution of a pure diffusion problem can be compared with the fitting of a function to a set of data points. Equating the stationary distribution with the equilibrium distribution by the enforced detailed balance corresponds to forcing the fitted function to pass exactly through one of the data points. In general the fitted function thus obtained will not be the optimal function for the whole data set. In the case of a detailed balance, the equality of the stationary distribution with the equilibrium distribution is only a necessary condition for a correct program code, but contains no information on the quality of the algorithm. Without an exact detailed balance, the deviation of the stationary distribution from the equilibrium distribution already contains information on the quality of the algorithm.

Nonnegativity. Physically, negative probability densities have no meaning and should not occur in a good diffusion algorithm that involves only transfer of probability density to *adjacent* grid points. This is also true for the present algorithm with polynomials of the second degree. With higher polynomials ($J \geq 4$), negative probability densities occur at the relative positions $j = \pm 2, \pm 4, \dots$ (see Table 1). In the actual calculation of the numerical fundamental solutions w , negative values occur only in those ranges of x or r , where exact values g_i and g_{i+1} differ by an order of magnitude or more.

Restriction to first order with respect to time. The restriction to first order with respect to time would indeed be a disadvantage of an algorithm working with constant time steps t_a . The distinctive feature of the present algorithm is, however, that the time needed for the computation of $\rho(r_i, t)$ in the time range of interest is almost independent of the length t_a of the first time step. Hence t_a can be chosen almost arbitrarily short, and nothing would be gained by a second-order procedure with respect to time.

Stability of first-order scheme. The fact that in general an extremely short first time step t_a has been used was misinterpreted as evidence for a numerical instability of the algorithm in the case of larger values of t_a . The opposite is true, as is illustrated by example 1 and Fig. 1: In the absence of reactions, the stationary numerical distribution w_{stat} is independent of t_a as long as t_a is below the sharp stability limit. The unavoidable accumulation of round-off errors is not a peculiarity of the present algorithm but results from the limitations imposed by double-precision accuracy (cf. e.g. section 1.3.4 in ref. [24]).

Renormalization of the numerical fundamental solutions. Any numerical procedure will in the end be limited by the accumulation of round-off errors. This limitation will be the more obvious the more accurate the given procedure is. Moreover, example 3 in section 3.4 shows that even without renormalization a very high numerical accuracy is achieved. An unavoidable numerical error results from the combination of a large radial range with a very short first time step t_a . In example 5, for instance, $w(r_N, t | r_N)$

remains constant in the whole time range from $t = 0$ to $t = t_a \times 2^{39} \approx 2.98 \times 10^{-17}$ s, whereas the ratio $w(r_{N-1}, t | r_N) / w(r_N, t | r_N)$ steadily increases from 0 to $\approx 10^{-15}$ in the same time range.

Renormalization of the theoretical fundamental solutions. The present use of renormalized theoretical fundamental solutions g^* is consistent with the normalization of the theoretical equilibrium distributions, which is tacitly applied in other algorithms.

Corrections. The correction of the quantities $H_{n,j}$ (see Eq. (67)) according to Eq. (73) is to some degree arbitrary and therefore unsatisfactory. However, this correction can easily be kept very small, and it is completely negligible in comparison to the large systematic errors that would be introduced by using the recommended detailed balance [15, 16, 19]. Moreover, the very weak dependence of the numerical fundamental solution on the particular correction of the quantities H_n shows that the correction for probability conservation is not a limiting factor in the application of the algorithm (as long as this correction is very small).

Control parameters. An ideal algorithm should contain control parameters that limit the maximum deviation of the numerical solution from the true solution of a partial differential equation. Admittedly, for the present algorithm no control parameters are given. However, this should not seriously limit its usefulness. The practical application of the algorithm to three-dimensional diffusion is governed by a few rules:

Rule 1. If possible, a constant radial increment factor $f = 1 + q$ with $q \lesssim 0.02$ should be used.

Rule 2. The maximum acceptable value of q is obtained from the limits σ^{-1} and σ for the acceptable ratio of neighboring equilibrium values ρ_{eq} , (cf. Eq. (3)).

Rule 3. The ratio σ can be the greater the higher the degree $2J$ of the interpolation polynomials. For this reason, $2J = 8$ will be in general the best choice.

Rule 4. If a variable radial increment factor $f_i = 1 + q_i$ is needed, the best results are obtained, if q is virtually constant in the effective range of the potential. The relative error ω_n in the probability conservation during the first time step (cf. Eq. (70)) should be extremely small in the whole radial range, e.g. $|\omega_n| < 10^{-8}$.

Rule 5. In the case of reflecting boundaries, the first time step t_a can be made extremely short, for instance by a factor 10^{-12} shorter than the maximum $(t_a)_{\text{max}}$ corresponding to the limit of numerical stability. The concomitant accumulation of round-off errors is corrected for by a renormalization of all N fundamental solutions after every ten time-doubling cycles up the diffusion time t_b , when the reaction term $k(r)$ is taken into account (cf. section 3.7).

Rule 6. In the case of absorbing or constant-concentration boundaries, the accumulation of round-off errors can no longer be corrected for by a

renormalization. Therefore the first time step t_a should not be made extremely short, but still much shorter than $(t_a)_{\max}$.

Rule 7. The convergence of the algorithm with decreasing q should be checked.

Waste of computer resources. This objection would have been justified in 1980, when the present computations would have required a large computer, shared by a community of researchers. At the end of the year 1999, when the typical random-access memory of a new personal computer was 64 MB, this objection did not make sense anymore.

Soft-sphere approximations. The relative ease, with which the combination of a short-range potential with a large radial range can be treated, offers the possibility to overcome the limitations of hard-sphere models. A suitable repulsive potential is introduced, and the inner reflecting boundary is shifted to a smaller radius, where the repulsive term of the potential already strongly dominates (and where the equilibrium probability density is already much smaller than in the potential minimum). Example 1 in section 2.7 has shown that the present implementation of a reflecting boundary yields acceptable results even if the potential does not satisfy Eq. (35).

4.2 On the most efficient algorithm for long time propagations

The alleged *slowness* of the present algorithm was perhaps the main objection against it. The argument is simple: If only *one* distribution function at a *single* time t is needed, it is an enormous waste of computation time to compute N distribution functions instead of one. In practice, however, nearly always the *evolution* of a distribution function is of interest and, for instance, 100 to 1000 distributions with constant time steps t_c are needed (cf. section 3.7).

The computation time needed for a time-doubling cycle, t_{cycle} , is proportional to the third power of the number of the fundamental solutions, N , if *all* elements of *each* fundamental solution significantly differ from zero. Obviously, due to this fact, the present algorithm is very slow if N is large. For instance, with $N = 1600$ (cf. section 3.8) and without parallel computing, it would be difficult to fit the parameters of a kinetic model to experimental data within a reasonable time. However, in many cases, values of N in the range $100 \leq N \leq 200$ will be sufficient and the time t_d , up to which the distribution $\rho(r, t)$ is to be calculated, may be much shorter than the time needed for complete equilibration in the absence of a reaction. In such cases the time for the computation of the whole *evolution* of ρ reduces to about 1 min or less, and the present algorithm becomes applicable as part of a curve-fitting procedure.

In a comparison of different algorithms that allow the application of very large constant time steps t_c , the distinction between two different computation times is important. The first computation time is the time needed

for computing $\rho(r_i, t_c)$, which requires in the present algorithm the computation of N fundamental solutions $w(r_i, t_c | r_n)$. The second computation time is the time needed for computing v_d distribution functions $\rho(r_i, t_c v)$ with $v = 1, 2, \dots, v_d$ and $t_d = t_c v_d$. With respect to the second computation time, the present algorithm seems to be the most efficient one, because there is no simpler solution for this problem than the summation of N weighted fundamental solutions. The first computation time will be in general longer than the second computation time. For instance, if $v_d = N$ and $w(r_i, t_c | r_n) \neq 0$ for arbitrary n and i , then the time needed for computing the v_d distribution functions $\rho(r_i, t_c v)$ would be equal to be the time needed for computing a new set of N fundamental solutions $w(r_i, 2t_c | r_n)$.

With the algorithm in refs. [16, 19] the situation is reversed. Due to the direct computation of the distribution $\rho(r_i, t_c)$, the first computation time can be rather short. On the other hand, the same procedure is to be repeated $(v_d - 1)$ times; hence, even in the most favorable case, the second computation time will be at least $(v_d - 1)$ times longer than the first computations time. An advantage of that algorithm is that the time step need not be constant.

It is not a priori evident, which of the two algorithms is more efficient in computing the *evolution* of a distribution function. Finally, if for some problems of interest the present algorithm will turn out to be the only practicable one, slowness will be no longer an argument.

5. Summary

With the present algorithm, the Smoluchowski equation of a spherically symmetric three-dimensional diffusion problem can be numerically solved with high accuracy even in the unfavorable case of a short-range potential barrier in combination with a large radial range. The basic features of the algorithm are:

(a) A series of N spherical shells of volume V_i is defined by the radii $R_0, R_1, \dots, R_i, \dots, R_N$ with $R_{i+1}/R_i = 1 + q$, where q is either constant or a slowly increasing function of i . The mean radii r_i of the spherical shells are taken as grid points.

(b) Numerical fundamental solutions $w(r_i, t | r_n)$ are defined. $w(r_i, 0 | r_n) = V_n^{-1} \delta_{in}$ is the numerical analogue of a δ -function. All that is needed for the calculation of $w(r_i, t | r_n)$ is the set of fundamental solutions $w(r_i, t_a | r_n)$, where t_a is the duration of the first diffusion step.

(c) The first and second derivatives of the numeric δ -function $w(r_i, 0 | r_n)$ with respect to r at r_i are defined by the respective derivative of a polynomial of the degree $2J$ ($J = 1, 2, 3, 4$), centered at r_i . In the calculation of $w(r_i, t_a | r_n)$, the error in probability conservation becomes very small in the range of q that is of practical interest ($q \lesssim 0.02$). For very small q , the probability error is approximately proportional to q^{2J+2} .

(d) A reflecting inner boundary is implemented by compensating the loss due to diffusion from the shells $1, 2, \dots, J$ through the boundary by the gain due to the opposite diffusion from the auxiliary shells $0, -1, \dots, -J + 1$ through the boundary. An absorbing inner boundary is implemented by the same terms, but with opposite signs. Radiation and constant-concentration boundaries are implemented in a similar way. The treatment of outer boundaries is analogous.

(e) By successive doubling of the total diffusion time t , each fundamental function $w(r_i, 2t | r_n)$ can be expressed as the weighted sum of all N fundamental solutions $w(r_i, t | r_l)$ ($l = 1, 2, \dots, N$). Since the total computation time is roughly proportional to the logarithm of the total diffusion time, the stationary distribution w_{stat} can be always computed (in the case reflecting boundaries and in the absence of reactions).

(f) The evolution of an arbitrary distribution ρ can be calculated, starting from an initial condition $\rho(r_i, 0)$, in time steps of the length t_c by repeatedly applying the set of fundamental solutions $w(r_i, t_c | r_l)$.

(g) A distance-depending rate coefficient k can be taken into account by multiplying at a time $t_2 \ll k_{\text{max}}^{-1}$ all N^2 values $w(r_i, t_2 | r_n)$ with $\exp[-k(r_i) t_2]$.

(h) Within the limits of double-precision computations, the best results were obtained with polynomials of the degree 8 (the highest degree tested).

(i) The exact equality of the stationary distribution with the equilibrium distribution (enforced by a detailed balance) will in general not lead to the best approximation of the *evolution* of a distribution.

(j) With increasing q_i ($q_N \approx 10 q_1$), a radial range of six orders of magnitude can be covered with a maximum relative error of w of less than 2×10^{-4} .

Acknowledgements

The author thanks Professor J. Troe for support of this work and Dr. W. Naumann for long discussions on diffusion-influenced reactions and for the theoretical treatment of some problems in the kinetics of triplet-triplet annihilation. The author also thanks Professor N. Ernsting, Dr. G. Käß, and Professor R. Schinke for explaining to him the Chebyshev time propagation. Finally, the detailed comments by a referee were helpful in writing the final version of this paper.

Appendix

The need of a new algorithm appeared in connection with a well-known special case of reversible excimer formation [10, 35–37]: The lowest excited singlet state $^1A^*$ of an aromatic hydrocarbon A is populated by triplet-

triplet annihilation ${}^3\text{A}^* + {}^3\text{A}^* \rightarrow {}^1\text{A}^* + {}^1\text{A}$. The subsequent reversible formation of excimers, ${}^1\text{A}^* + {}^1\text{A} \rightleftharpoons {}^1(\text{AA})^*$, and the observable intensity ratio of delayed monomer fluorescence and delayed excimer fluorescence depend on the initial spatial distribution $\rho(r, 0)$ of ${}^1\text{A}^*$ relative to ${}^1\text{A}$, on an effective potential $U(r)$ for the attraction between ${}^1\text{A}^*$ and ${}^1\text{A}$, and on the rate constants for the radiative and nonradiative decays of ${}^1\text{A}^*$ and ${}^1(\text{AA})^*$ to the respective ground state or triplet state. In the case of pyrene (the classic example of delayed excimer fluorescence [35]), the numerical solution of the corresponding Smoluchowski Eq. (1) has to satisfy the following requirements:

- (a) $U(r)$ is a short-range potential. Hence the radial steps Δr must be small, e.g. $\Delta r \approx 10^{-3} R_0$ or less, where R_0 is the contact distance between A and B.
- (b) The radial range is large with a maximum $r_N \approx 100 R$.
- (c) The potential minimum is $U(R_0) \approx -hc \times 3000 \text{ cm}^{-1}$.
- (d) The temperature range of interest ($T_{\text{max}} \approx 300 \text{ K}$, $T_{\text{min}} \approx 120 \text{ K}$) corresponds to the range $14 \leq |U(R)|/k_{\text{B}}T \leq 36$.
- (e) The relative diffusion coefficient D ranges from $\approx 5 \times 10^{-8} \text{ cm}^2 \text{ s}^{-1}$ at the lowest temperature to $\approx 5 \times 10^{-5} \text{ cm}^2 \text{ s}^{-1}$ at the highest temperature.
- (f) The constant time step is $\leq 10^{-16} \text{ s}$; the longest time of interest is $\approx 5 \times 10^{-6} \text{ s}$.
- (g) The deviation of the numerical solution from the exact solution of Eq. (1) must be less than 1% in the time range of interest for all temperatures.

References

1. M. v. Smoluchowski, *Physik. Zeitschr.* **17** (1916) 585 [p. 588, Eq. (50)].
2. M. v. Smoluchowski, *Z. Physik. Chem.* **92** (1917) 129.
3. P. Debye, *Transact. Electrochem. Soc.* **82** (1942) 265.
4. S. A. Rice, *Diffusion-Limited Reactions*, vol. 25 of the series *Comprehensive Chemical Kinetics*, ed. C. H. Bamford, C. F. H. Tipper and R. G. Compton (Elsevier, Amsterdam, 1985).
5. A. A. Ovchinnikov, S. F. Timashev and A. A. Belyi, *Kinetics of Diffusion-Controlled Chemical Processes* (Nova, Commack, 1989; Russian original: Moscow, Khimiya, 1986).
6. J. B. Pedersen and J. H. Freed, *J. Chem. Phys.* **58** (1973) 2746.
7. J. B. Pedersen and J. H. Freed, *J. Chem. Phys.* **59** (1973) 2869.
8. Z. Schulten and K. Schulten, *J. Chem. Phys.* **66** (1977) 4616.
9. M. J. Pilling and S. A. Rice, *J. Chem. Soc. Faraday Trans. II*, **72** (1976) 792.
10. P. R. Butler and M. J. Pilling, *J. Chem. Soc. Faraday II*, **73** (1977) 886.
11. S. H. Northrup and J. T. Hynes, *J. Chem. Phys.* **69** (1978) 5246.
12. S. H. Northrup and J. T. Hynes, *J. Chem. Phys.* **71** (1979) 871.
13. G. Lamm and K. Schulten, *J. Chem. Phys.* **78** (1983) 2713.
14. P. Clifford, N. J. B. Green and M. J. Pilling, *J. Phys. Chem.* **88** (1984) 4171.
15. W. Nadler and K. Schulten, *J. Chem. Phys.* **84** (1986) 4015.
16. E. Pines, D. Huppert and N. Agmon, *J. Chem. Phys.* **88** (1988) 5620.
17. J. B. Pedersen, L. I. Lolle and J. S. Jørgensen, *Chem. Phys.* **165** (1992) 339.

18. K. L. Aminov and J. B. Pedersen, *Chem. Phys.* **193** (1995) 297.
19. E. B. Krissinel[†] and N. Agmon, *J. Comput. Chem.* **17** (1996) 1085.
20. H. Kim, S. Shin and K. J. Shin, *J. Chem. Phys.* **108** (1998) 5861.
21. R. Zurmühl, *Praktische Mathematik für Ingenieure und Physiker*, 4th Edition (Springer-Verlag, Berlin, 1963).
22. L. F. Shampine and M. K. Gordon, *Computer solution of ordinary differential equations – The initial value problem* (W. H. Freeman, San Francisco, 1975), chapter 2.
23. W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes – The Art of Scientific Computing (FORTRAN Version)* (Cambridge University Press, Cambridge, 1989).
24. C. W. Gear, *Numerical initial values problems in ordinary differential equations* (Prentice-Hall, Englewood Cliffs, 1971).
25. A. R. Mitchell and D. F. Griffiths, *The finite difference method in partial differential equations* (J. Wiley & Sons, Chichester, 1980).
26. K. Dekker and J. G. Verwer, *Stability of Runge-Kutta methods for stiff nonlinear differential equations* (North-Holland, Amsterdam, 1984), chapter 10.
27. G. E. Uhlenbeck and L. S. Ornstein, *Phys. Rev.* **36** (1930) 823.
28. H. Tal-Ezer and R. Kosloff, *J. Chem. Phys.* **81** (1984) 3967.
29. C. Leforestier, R. H. Bisseling, C. Cerjan, M. D. Feit, R. Friesner, A. Guldberg, A. Hammerich, G. Jolicard, W. Karlein, H.-D. Meyer, N. Lipkin, O. Roncero and R. Kosloff, *J. Comput. Phys.* **94** (1991) 59.
30. J. Broeckhove and L. Lathouwers (Editors), *Time-dependent quantum molecular dynamics* (Plenum Press, New York, 1992).
31. S. Chandrasekhar, *Rev. Mod. Phys.* **15** (1943) 1.
32. H. A. Kramers, *Physica* **7** (1940) 284.
33. K. Razi Naqvi, K. J. Mork and S. Waldenström, *J. Phys. Chem.* **84** (1980) 1315.
34. S. V. Patankar, *Numerical heat transfer and fluid flow* (Taylor & Francis, 1980).
35. C. A. Parker and C. G. Hatchard, *Trans. Faraday Soc.* **59** (1963) 284.
36. J. B. Birks, B. N. Srinivasan and P. S. McGlynn, *J. Mol. Spectr.* **27** (1968) 266.
37. B. Stevens and M. I. Ban, *Mol. Cryst.* **4** (1968) 173.