INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Université d'Ottawa • University of Ottawa

CALCULATING THE DIFFUSION COEFFICIENT OF A RANDOM WALKER AMONG IMMOBILE OBSTACLES: NEW NUMERICALLY EXACT THEORY AND APPLICATIONS.

By

© Jean-François Mercier

A thesis submitted to
the School of Graduate Studies and Research
in partial fulfillment of the requirements
for the degree of Master of Science in Physics

Department of Physics

University of Ottawa
Ottawa, Ontario
7 June 1999

Summary

Many biological, chemical and physical problems can be reduced to that of the diffusion of a particle in a quenched system of obstacles. For example, the diffusion of proteins in the plane of a biomembrane, the migration of analytes in low field gel electrophoresis and the diffusion of particles in porous media can all be reduced to such systems. The standard method to study such problems is to use Monte Carlo simulations on finite-size lattices with periodic boundary conditions. This approach is very simple and one can obtain diffusion coefficients with error bars as small as 0.1%. We developed a new algebraic method to calculate the exact diffusion coefficients for those systems. This method can treat systems similar (in size and complexity) to the ones currently studied with Monte Carlo simulations, but gives exact results and usually requires less CPU time. Furthermore, this new approach can easily be adapted to any type of lattice in any dimensionality greater than one.

After explaining the new method and its numerical implementation, some examples will be given to demonstrate both its viability and its power. Finally, a standard mean-field theory of gel electrophoresis, which predicts that the mobility of charged particles is directly related to the fractional gel volume available to it, will be tested for three-dimensional gels.

Sommaire

Plusieurs problèmes biologiques, chimiques et physiques peuvent être réduits à la diffusion aléatoire de particule(s) sur un réseau ponctué d'obstacles fixes et impénétrables. Parmi les exemples les plus connus on compte la diffusion de protéines dans le plan d'une membrane biologique, la diffusion de particules dans une solution de polymères et la migration d'un analyte dans un gel d'électrophorèse. La méthode couramment utilisée pour étudier ce type de problème est la simulation Monte-Carlo (SMC) sur un réseau de taille finie avec des conditions aux frontières périodiques. Cette approche, extrêmement simple, permet de suivre la diffusion de millions de particules évoluant sur le réseau et d'en déduire le coefficient de diffusion avec des incertitudes de l'ordre de 0.1%. Nous avons développé une technique permettant de calculer, pour une distribution d'obstacles donnée, le coefficient de diffusion d'une particule de façon exacte. Cette technique de calcul permet d'étudier des problèmes similaires (en taille et en complexité) à ceux étudiés à l'aide de simulations Monte Carlo, mais elle donne des résultats exacts et ce, en moins de temps que les SMC. De plus, elle peut facilement être adaptée à tout type de réseau (carré, triangulaire, ...) de dimension supérieure à un.

Après avoir décrit la méthode et notre approche numérique, nous donnerons des exemples afin de démontrer son efficacité. Finalement, une théorie de champ moyen, couramment utilisée dans le domaine de l'éléctrophorèse sur gel et prédisant que la mobilité d'une particule est directement proportionnelle au volume qui lui est accessible, sera verifiée pour des gels à trois dimensions.

Acknowledgements

I would like to deeply thank my supervisor Dr. Gary W. Slater for his constant support, his more than generous availability and the great confidence that he has put in me throughout my Master Degree. His contagious love for science and his wonderful passion for it were great sources of inspiration for me. I also really appreciate that he gave me the opportunity to participate to many conferences.

I would like to give a special thank to Hong Guo, who was the first student to work on this project. His superb work has greatly facilitated mine.

I would also like to thank the entire Department of Physics: professors, support staff and students. Together, they build a very dynamic and friendly research environment.

I would like to express my sincere gratitude to my girlfriend Pascale Sévigny. For her understanding and support without which this work would have been a lot harder, but also for correcting most of my texts (which are usually a disaster) and for listening repetitively to my boring talks.

Finally, thanks to the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (FCAR, Québec) and the University of Ottawa for scholarships.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The diffusion of a probe particle in a fluid medium is a stochastic phenomenon in which a small particle moves randomly (this is called the Brownian motion). Because of its numerous applications in many fields of science (physics, chemistry, biology, ect.), this process has been studied extensively.

If the medium is homogenous, diffusion dynamics is relatively easy to understand and accurate predictions can be made. In such media the diffusion depends on parameters like the temperature, the viscosity of the medium, and the particle characteristics (shape, mass ...). In many important areas, however, the medium is not homogenous, and some obstacles obstruct the diffusion of the particle and slow its migration. The sieving effect of the obstacles is not well defined, but may presumably include excluded volume effects, percolation constraints (the particle cannot move over macroscopic distances), hydrodynamic interactions and specific particle-obstacle interactions (attractive or repulsive). Many widely used separation techniques, such as chromatography and electrophoresis, exploit the sieving properties of inhomogeneous media for practical applications. This is actually where our initial interest came from.

Understanding how the obstacles will affect the diffusion is often very important. For example, some features (e.g., the concentration of fixed proteins) of a biomembrane can be deduced by studying the diffusion of a traceable mobile protein, and one needs a good understanding of restricted diffusion in order to correctly interpret

the observations. In gel electrophoresis, a field-driven biased diffusion is used to separate particles of different sizes. In order to design new gels with better separation properties, one must first understand the effect of the gel fibers on the diffusion of the solute particle. Among the other examples in which obstructed diffusion plays a major role, we count the diffusion of particles in polymer solutions and in porous rocks.

1.1 Free diffusion in homogenous systems

If the medium is homogenous, the motion of the particle can be predicted to high accuracy. In such systems, it has been shown that the mean square displacement $\langle r^2 \rangle$ of a particle is completely defined by the equation[1]:

$$\langle r^2 \rangle = 2dD_0t, \tag{1.1}$$

where d is the space dimensionality, t the time and D_0 the diffusion coefficient (also called the diffusion constant). Thus the diffusion coefficient is a parameter which measures the facility of a particle to diffuse in a given medium and it depends on the properties of both the fluid (viscosity) and the particle (size, shape). For example, the diffusion coefficient of a spherical particle of radius R in a three-dimensional fluid is given by[1]:

$$D_0 = \frac{k_BT}{6\pi\eta R}, \tag{1.2}$$

where k_B is the Boltzmann constant and η is the viscosity of the fluid.

1.2 Diffusion in the presence of obstacles

When a particle moves in a medium containing obstacles, the dynamics is more subtle than in the homogenous case. The system now has two main regimes: the short time dynamics (anomalous diffusion), where the mean square displacement is not linearly proportional to time, and the steady-state regime where the mean

square displacement is strictly proportional to time. For the steady-state regime we can define the diffusion coefficient as:

$$D = \lim_{t \to \infty} \frac{\langle r^2 \rangle}{2dt}. \tag{1.3}$$

An interesting parameter, which allows one to isolate the effect of the obstacles on the diffusion, is the scaled diffusion coefficient D* defined as:

$$D^* = \frac{D}{D_0}, \tag{1.4}$$

where D_0 is the diffusion coefficient for a similar system (same liquid, temperature and diffusing particle) without obstacles. Note that since the obstacles slow down the particle, the scaled diffusion coefficient is generally smaller than unity.

1.3 The Nernst-Einstein relation

In many physical systems the diffusion of the probe particle is biased (a move in one of the directions is more probable than in the other directions). For example a charged particle undergoes a biased diffusion when an electric field is applied. Chemical potentials and gravity can also bias the diffusion of particles. In such systems the net dynamics can be characterized by the mobility μ defined as:

$$\mu = \frac{\langle v \rangle}{\varepsilon}, \tag{1.5}$$

where $\langle v \rangle$ is the mean velocity of the particle in the field direction and ε is the field intensity.

For systems at equilibrium, the diffusion coefficient D and the zero field ($\varepsilon \to 0$) mobility μ are related by the Nernst-Einstein relation. For a charged particle in an electric field the relation is given by[1]:

$$D = \lim_{\varepsilon \to 0} \frac{k_B T}{q} \mu, \tag{1.6}$$

where q is the electric charge of the particle.

Since k_BT/q is constant, Eq.(1.6) implies that (we further assume that the obstacle-free mobility μ_0 is field-independent):

$$\frac{D}{D_0} = \lim_{\varepsilon \to 0} \frac{\mu}{\mu_0}. \tag{1.7}$$

In the following, we will use the standard notation $D^* \equiv D/D_0$ and $\mu^*(\varepsilon) \equiv \mu(\varepsilon)/\mu_0$ for the scaled diffusion coefficient and mobility, respectively. Thus for a vanishing field ($\varepsilon \to 0$), the scaled mobility is equal to the scaled diffusion coefficient. Note that this equation actually gives the diffusion coefficient in the field direction.

1.4 Electrophoresis and the OMRC model

As briefly mentioned above, gel electrophoresis exploits the sieving properties of inhomogeneous media, such as agarose gels, to separate charged particles of different sizes. Such a gel is formed of cross-linked uncharged fibers, which can be considered immobile for the migrating particle. When an electric field is applied to the system, the charged particles drift in a direction which depends on their charge. The larger particles tend to hit the gel fibers more often than the smaller ones, and are therefore more retarded. This method is commonly used to separate a variety of biological molecules such as DNA and proteins.

In 1967, Morris[2] published a study of the sieving effect of gel electrophoresis. In this paper he showed that his experimental data (proteins in polycrylamide gels) were compatible with a theoretical model developed by Giddings and Boyack[3], which predicted that the scaled mobility for this specific system was given by[2]:

$$\mu^* \approx 1.1 f(C), \tag{1.8}$$

where f (C) is the fractional gel volume available to the particle, and C is the concentration of obstacles.

Given the simplicity of Eq.(1.8), Rodbard and Chrambach[4] then suggested to change the $\approx$ for an $=$ and to forget the 1.1 factor. They also used a result, first obtained by Ogston[5], which shows that the volume accessible to a spherical particle in a random array of infinitely long fibers is given by:

$$f(C) = e^{-KC}, \tag{1.9}$$

where K is a constant that depends on the radius of the probe particle. Therefore, they conjectured that:

$$\mu^* = f(C) = e^{-KC}, \tag{1.10}$$

which we will call the Ogston-Morris-Rodbard-Chrambach (OMRC)[2-7] model. This model is probably the most used model to predict the sieving effect of obstacles, especially in the field of gel electrophoresis. In particular, this model lead to the Ferguson plot[6] (widely used for gel electrophoresis) where the logarithm of the mobility is plotted as a function of the gel concentration to obtain the so-called retardation coefficient K (the slope). Over the years, a lot of efforts have been spent computing the available volume $f(C)$ for a variety of media and probe particles in order to better fit experimental data. Surprisingly, very few attempts, if any, have focused on verifying the validity of Eq.(1.10). To verify such a model, we must rely on either very accurate experimental measurements or computer simulations. However direct measurements have failed to validate the OMRC model because μ^* and $f(C)$ cannot generally be measured independently. We must therefore rely on computer simulations.

In 1995, Slater and Guo[8] proposed a lattice model in order to explain gel electrophoretic sieving at a more microscopic level. In a series of papers[9-11], they used this approach to test the predictions of the OMRC model, and they found that it was actually equivalent to a mean-field model (the only "gel" for which the OMRC model prediction are accurate is a uniform annealed gel where the obstacles are reset at random at each time step[9]). The work presented in this thesis is in part the continuity of the early work of Slater and Guo.

1.5 Presentation of the thesis

The lattice model, frequently used to study diffusion and mobility, and the corresponding lattice Monte Carlo (LMC) simulation method, are presented in Chapter 2. Chapter 3 explains our approach to the lattice model which allows us to obtain exact diffusion coefficients or mobilities. Chapter 4 describes how our method can be implemented numerically and gives some examples. A study of three-dimensional gel electrophoresis is presented in Chapter 5, while Chapter 6 explains the limitations of our approach and proposes possible solutions or alternatives. My conclusions are presented in Chapter 7.

1.6 Statement of originality

The lattice model and the numerical method to solve the lattice system exactly, presented in Section 3.2 and 3.3, have been developed by Slater and Guo [8–11]. The rest of the work presented in this thesis, including the algebraic simplifications (Section 3.5), the numerical implementation (Section 4.2) and the corresponding Fortran program, and all the applications, are the product of my work during the last two years. Chapter 3, 4 and 5 are based on published articles[20,21,26] written conjointly with my supervisor Dr. Gary W. Slater (I am the first author of the three papers).

Chapter 2

The lattice model of diffusion

Many computational approaches can be used to simulate the migration of particles in obstructed media (e.g. lattice Monte Carlo (LMC), Brownian Dynamics, Molecular Dynamics, etc). To study the long time behavior, LMC is often the best algorithm because it is quite efficient. In lattice models, the space is discrete, which means that the particle, and the obstacles, can only occupy certain locations (the lattice sites, like on a chess board). The sites occupied by the obstacles are forbidden to the particle and the latter is moving on the lattice by making discrete jumps of fixed length. Figure 2.1 shows an example of a two-dimensional square lattice where the obstacles (the black sites) are placed periodically. The next two sections will explain how unbiased and biased diffusion are modeled on a lattice.

2.1 Random walk on a lattice

Consider a particle (a 1×1 lattice plaquette) which undergoes unbiased, free (no obstacles) random jumps on a two-dimensional square lattice. Its continuous diffusion process is thus modeled by random Brownian jumps of fixed length $a = 1$ (the lattice parameter in scaled units) and mean duration time τ=1 (also in scaled units). At each step, the particle is simply jumping to one of the four adjacent sites (noted $+x$, $-x$, $+y$ and $-y$). For unbiased migration, the jumping probabilities are uniform, and on a two-dimensional square lattice we have:

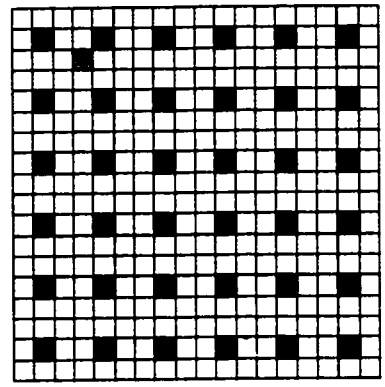

Figure 2.1: Example of a simple two-dimensional square lattice. The dark squares represent the obstacles, while the gray one represents the migrating particle. Periodic boundary conditions are imposed.

$$p_{\pm x} = p_{\pm y} = \frac{1}{4}. \tag{2.1}$$

The free diffusion coefficient (for a system without obstacles) is then given by $D_0 = D_{0x}/2 = D_{0y}/2 = 1/4$. In the presence of obstacles, we have an additional rule: if a move leads to a particle overlapping with an obstacle, the move is rejected and the particle returns to its previous position. However, the time step counts regardless of the acceptance of such moves. These hard core interactions between the particle and the obstacles will obviously reduce the scaled diffusion coefficient D^* of the particle. Note that, in such models, the hydrodynamic effects are not taken into account.

2.2 Diffusion with a weak field

2.2.1 The Metropolis algorithm

Let us now consider what happens when a weak external field is applied to the particle. The particle has a larger probability of jumping in the direction of the field than in the opposite direction. This problem can be treated using various strategies. One of them is the Metropolis algorithm, where a move is always accepted

if it decreases the energy of the particle, but not always if it increases this energy. For a 1D model with a field direction $(+x)$, the jumping probabilities become:

$$p_{+x} = \frac{1}{2}, \quad \text{and} \quad p_{-x} = \frac{e^{-\Delta E/k_B T}}{2} = \frac{e^{-\varepsilon}}{2}, \tag{2.2}$$

where ΔE is the energy difference between the initial and final states of the particle and ε is the scaled field intensity. Note that the time step is kept untouched at $\tau = 1$ while there is a probability $(1/2 - e^{-\varepsilon}/2)$ of not moving.

2.2.2 The Glauber approach

Another approach has been developed for lattice systems by Slater and Rousseau[12]. Although originally based on the theory of first-passage problems, it can be easily understood in terms of Glauber[13] dynamics (as often used for lattice spin systems). This approach gives the following jumping probabilities for a one-dimensional system with a scaled field ε pointing in the $+x$ direction:

$$p_{\pm x} = \frac{e^{\pm\varepsilon}}{e^{+\varepsilon} + e^{-\varepsilon}}. \tag{2.3}$$

To define the field-dependent time step, we first choose the field-free time step τ_0 and the lattice spacing a such that the scaled velocity of a free particle is equal to ε for any arbitrary value of ε. Thus:

$$v = \varepsilon = \frac{p_{+x} - p_{-x}}{\tau} = \frac{e^{+\varepsilon} - e^{-\varepsilon}}{\tau(e^{+\varepsilon} + e^{-\varepsilon})} = \frac{tanh(\varepsilon)}{\tau} \tag{2.4}$$

giving a field-dependent time step:

$$\tau = \frac{tanh(\varepsilon)}{\varepsilon}. \tag{2.5}$$

For a low field intensity $\varepsilon \ll 1$ we have:

$$\tau = \frac{tanh(\varepsilon)}{\varepsilon} \approx 1 - \frac{\varepsilon^2}{3} + O(\varepsilon^4). \tag{2.6}$$

Note that this series expansion does not have a first order term (ε^1). In other words, the time durations of the jumps are not perturbed, to first order in ε, by the

external field. This greatly simplifies the problem since it means that the jumps will be evenly distributed in time between different spatial directions (when $d > 1$) in the $\varepsilon \to 0$ limit.

The Metropolis and the Glauber approaches will give exactly the same results for D^* and μ^* in the $\varepsilon \to 0$ limits (see Appendix A). The work presented in this thesis will use the Glauber approach for two reasons. First, this approach is conceptually more intuitive (when a field is applied, a particle has a greater chance to jump in the field direction, and less to jump in the opposite direction). Second, although it has not yet been investigated in details, the Glauber approach could potentially give useful results for high field intensities (this is because we can now adjust the time steps which can be non-uniform [Hong Guo, personal communication]).

2.2.3 Higher dimensionality

For a hypercubic system of dimension d, with the field direction $(+x)$, the probability for the next jump to be in direction i is:

$$p_{\pm i} = \frac{1}{2(d-1) + e^{\varepsilon} + e^{-\varepsilon}} \simeq \frac{1}{2d} + O(\varepsilon^3) \qquad (i \neq x), \tag{2.7}$$

while the probability for the next jump to be in the $\pm x$ direction is

$$p_{\pm x} = \frac{e^{\pm\varepsilon}}{2(d-1) + e^{\varepsilon} + e^{-\varepsilon}} \simeq \frac{1 \pm \varepsilon}{2d} + O(\varepsilon^3). \tag{2.8}$$

Note that $p_{+i} + p_{-i} = p_{+x} + p_{-x} = 1/d$, as it should. These probabilities do depend on the field intensity ε to the first order; this is indeed why the mobility is finite in the zero-field limit. In the absence of obstacles, the mean velocity is given by $V_0 = (p_{+x} - p_{-x})/\tau = \varepsilon/d$, and the free-solution mobility by:

$$\mu_0 \equiv \frac{V_0}{\varepsilon} = \frac{1}{d}. \tag{2.9}$$

As in the unbiased case discussed previously, the obstacles will lead to jump rejections and smaller mobilities.

2.3 Lattice Monte Carlo simulations

The standard method to study lattice models is to perform lattice Monte Carlo simulations (LMC). For example, the diffusion of proteins on the surface of biomembranes in the presence of fractal and random aggregates of obstacles has been studied extensively by M.J. Saxton[14–19] using LMC simulations. The approach is quite simple and straightforward. One simply puts the particle at a random position on the lattice and let the system evolve given the overlapping rules (the particle cannot overlap an obstacle) and the jumping probabilities (Eq.(2.1) for a system with no field and Eqs.(2.7), (2.8) for a system with a field). By recording the mean square displacement r^2 of the particle for many runs and many starting points, and using Eq.(1.3) to analyze the data, one can obtain diffusion coefficients D^* (or mobilities μ^*) with error bars as small as 0.1%[14]. Researchers frequently use periodic boundary conditions (PBC) with LMC simulations (if a particle reaches the end of the lattice in one direction, it is reinjected at the other end of the lattice), although one can in principle avoid this by using various numerical tricks.

This thesis proposes a novel method to compute D^* (or μ^*) for such lattice models. Our method provides exact results and does not normally require more computational resources than LMC methods.

Chapter 3

Exact results: Theory

3.1 Introduction

In this chapter, we derive an algebraically exact method to calculate the scaled diffusion coefficient D^* of a particle in a quenched system of obstacles with periodic boundary conditions (PBC)[1]. This method uses the Nernst-Einstein relation (Eq.(1.7)) to transform the problem into a field-driven drift problem where D^* is related to the zero-field mobility (μ^*). It provides exact values for D^* which can be useful to test theoretical models and understand subtle geometrical effects. The general idea is quite simple. A LMC simulation can be seen as a Markov chain for which one can write coupled linear master equations, one per available lattice site. These equations can be solved exactly for systems that are not too large, and the diffusion coefficient D^* can be obtained from the resulting probability distribution functions. Since our approach can provide exact D^* values for two-dimensional systems of size larger than 256×256 (which is the typical size used for LMC simulations) in a few minutes of CPU on a standard workstation, we believe that our method offers a powerful alternative to LMC simulations. Calculating the exact diffusion coefficient D^* in the presence of closed volumes and/or multiple unconnected channels is a further challenge, even when using computer simulations. Such cases will be discussed

[1]This work has been published in the *Journal of Chemical Physics*[20].

after some simpler examples are treated.

3.2 General description of the method

In order to compute the mobility, we simply calculate the probability of presence on each lattice site in the steady-state, and multiply these probabilities by the local velocities. This procedure can be generalized to complicated situations (e.g., to random fields).

The method presented here requires one to use a finite-size lattice with periodic boundary conditions (PBC). Let the system size be L^d, where d is the dimensionality and L is the linear size of the box, and let N be the number of sites occupied by obstacles. Briefly, the method is as follows. We first calculate the steady-state probability of presence of the particle on each of the $L^d - N$ vacant sites of the lattice, using PBC's. This probability is not uniform if $\varepsilon \neq 0$. Once this is done, we calculate the mean (global) velocity v by multiplying the probability of presence on each site by the local mean velocity. Finally, the mobility is calculated and, from Eq.(1.7), the diffusion coefficient is obtained. This approach reduces the problem to the solution of a (large) system of linear equations.

Various algorithms can be used to solve systems of linear equations. If an undefined value of ε is used, an algebraic solution, in terms of ε, can be obtained using, e.g., Mathematica. This gives an exact solution in terms of a rational fraction, but such an approach takes too much memory resources to be applied to large and complex systems. It is often easier to use a very small value such as $\varepsilon = 10^{-y}$, and to solve the equations numerically (using, e.g., an iterative procedure) to obtain the diffusion coefficient with a precision of about y decimal places. Slater and Guo[10] have shown how these arbitrary precise results can be transformed into exact results (in terms of rational fractions) using the numerical capabilities of Mathematica. Section 3.5 presents another approach that allows us to solve our linear equations numerically without having to choose a specific value for the field intensity ε.

Note that this procedure actually gives the diffusion coefficient in the direction

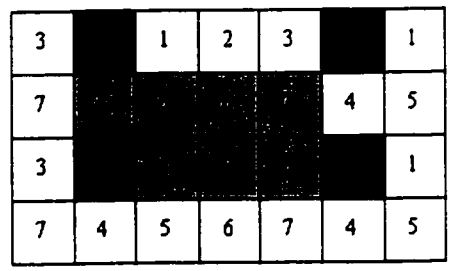

Figure 3.1: The model system used in Section 3.3. The periodic cell is of size 4×2 (in grey).

of the applied field. To better illustrate the method, we calculate in the following section the diffusion coefficient D^* of a 1×1 particle in the system shown in Fig. 3.1, for both the x and y directions. Systems with multiple independent zones are discussed in Section 3.4, while the algebraic simplifications and a powerful computational approach are presented in Section 3.5 and Chapter 4 respectively.

3.3 D^* via the zero-field mobility μ^*

3.3.1 Diffusion in the x-direction

Figure 3.1 shows the site labeling scheme for our model system. Because of the periodic boundary conditions, there are only 7 independent sites. The probability $n(i, t+1)$ for the particle to be on site i ($i = 1, 2, 3, 4, 5, 6, 7$) at time $t+1$ can easily be computed from the probabilities $n(i, t)$ and the jumping probabilities given by Eqs.(2.7)-(2.8). Indeed, $n(i, t+1)$ is related to the flow from adjacent sites towards site i (the particle moves once per unit time). For example, the one-step rate equation for site #3 reads:

$$n(3, t+1) = p_{+x} n(3, t) + \frac{1}{2} n(7, t) + p_{+x} n(2, t). \tag{3.1}$$

The first term on the r.h.s. of Eq.(3.1) is due to the reflection on the obstacle in front of site #3, while the next two terms calculate the probability for the particles on the adjacent sites (#7 and #2) to jump onto site #3. In the steady-state, we have $n(i,t+1) = n(i,t) = n(i,t \to \infty) \equiv n(i)$. We must therefore solve the following system of equations (we use the Dirac notation where $|t\rangle$ and $\langle t|$ represent column and row vectors, respectively, while matrices are denoted by capital letters and scalars by lower case letters):

$$T|n\rangle = |n\rangle,$$

where:

$$T = \frac{1}{4}\begin{pmatrix} 1-\varepsilon & 1-\varepsilon & 0 & 0 & 2 & 0 & 0 \\ 1+\varepsilon & 0 & 1-\varepsilon & 0 & 0 & 2 & 0 \\ 0 & 1+\varepsilon & 1+\varepsilon & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 & 1-\varepsilon & 0 & 1+\varepsilon \\ 2 & 0 & 0 & 1+\varepsilon & 0 & 1-\varepsilon & 0 \\ 0 & 2 & 0 & 0 & 1+\varepsilon & 0 & 1-\varepsilon \\ 0 & 0 & 2 & 1-\varepsilon & 0 & 1+\varepsilon & 0 \end{pmatrix} \tag{3.2}$$

and

$$|n\rangle \equiv \begin{pmatrix} n(1) \\ n(2) \\ n(3) \\ n(4) \\ n(5) \\ n(6) \\ n(7) \end{pmatrix}.$$

Only 6 of the 7 equations are independent. Besides these rate equations, we also have the normalization condition:

$$\sum_{i=1}^{7} n(i) = 1. \tag{3.3}$$

This system of linear equations (which can also be seen as an eigenvector problem $T|n\rangle = \lambda|n\rangle$ with an eigenvalue $\lambda = 1$) can be solved either symbolically or numerically. To first order in ε, we find that the normalized probability vector $|n\rangle$ is given by:

$$|n\rangle \equiv \begin{pmatrix} n(1) \\ n(2) \\ n(3) \\ n(4) \\ n(5) \\ n(6) \\ n(7) \end{pmatrix} = \frac{1}{7}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \frac{\varepsilon}{14}\begin{pmatrix} -2 \\ 0 \\ 2 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}. \tag{3.4}$$

The mean velocity on each site is $v(i) = p_{+x}L_+(i) - p_{-x}L_-(i)$, where the displacements $L = 1$ if there is no obstacle in the given direction, and zero otherwise. Here, we thus have:

$$|v\rangle \equiv \begin{pmatrix} v(1) \\ v(2) \\ v(3) \\ v(4) \\ v(5) \\ v(6) \\ v(7) \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1+\varepsilon \\ 2\varepsilon \\ -1+\varepsilon \\ 2\varepsilon \\ 2\varepsilon \\ 2\varepsilon \\ 2\varepsilon \end{pmatrix}. \tag{3.5}$$

Therefore, the reduced diffusion coefficient is given by:

$$D_x^* = \mu^* = \frac{\mu}{\mu_0} = \frac{v/\varepsilon}{1/2} = \frac{\left[\left(\sum_{i=1}^{7} v(i)n(i)\right)/\varepsilon\right]}{1/2} = \frac{\langle v|n\rangle/\varepsilon}{1/2} = 5/7. \tag{3.6}$$

3.3.2 Diffusion in the y-direction

A similar approach for a field pointing in the y-direction gives

$$|n\rangle = \frac{1}{7}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{and} \quad |v\rangle = \frac{\varepsilon}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \tag{3.7}$$

for all values of ε. We thus obtain $v = \langle v|n\rangle = 3\varepsilon/7$ and $D_y^* = \mu^* = \mu/\mu_0 = 6/7$.

This method reduces the problem to solving a system of J linear equations with J unknowns, where $J \leq L^d - N$ is the number of unique empty sites on the lattice of size L (we can sometimes use the symmetries of the lattice to minimize J). Solving such a system of equations using a direct technique takes a time $T \sim J^3$. Therefore, the CPU time increases like $T \sim L^{3d}$ for a system of linear size L with PBCs; this represents the most important limitation of this approach, especially if large precision (or a rational fraction) is seeked. More efficient iterative methods will be discussed in Chapter 4.

3.4 Closed volumes and multiple pathways

When the obstacles are placed randomly, our finite-size systems can contain closed volumes (lakes) or multiple independent pathways (channels). Figure 3.2 shows a simple example of a two-dimensional system containing a closed lake and two unconnected channels (the field is assumed to point in the $+x$ direction). Of course, it would be unrealistic to simply reject these cases in the calculation of the mean diffusion coefficient. Therefore, we must be able to treat such cases in a systematic way. Note that $D_y^* = 0$ for this example.

The lattice sites inside a lake or a secondary channel are not related, through rate equations, to the sites of the "main" channel. In other words, each lake or channel results in an independent set of probability equations. Therefore, we must treat each lake or channel as an independent "system" or "zone"; for instance, we must

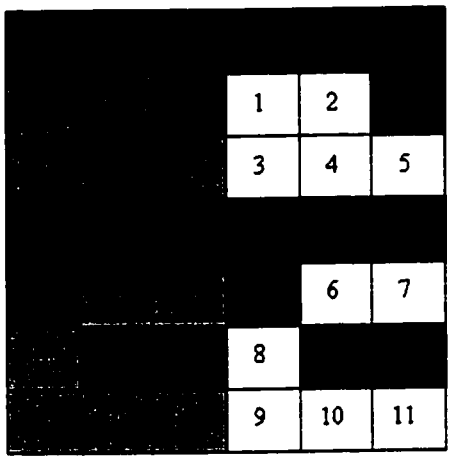

Figure 3.2: Two periodic cells (the primitive cell is in grey) of a simple system with periodicities $L_x = 3$ and $L_y = 7$. Sites # 6 and #7 form a closed lake. Sites #1, 2, 3, 4 and 5 form an independent channel, as do sites #8, 9, 10 and 11. We thus have three unconnected zones.

normalize each set of solutions independently. This will lead to a different diffusion coefficient for each independent zone. How these different diffusion coefficients are used to obtain an average value depends on the problem that is being modeled, as will be discussed later.

Our example (Fig. 3.2) has one lake (zone b, formed by sites #6 and #7) and two channels (zone a, formed by sites #1 to #5, and zone c, formed by sites #8 to #11) , i.e. three independent "zones". If we follow the method described in Section 3.3, the probability matrix T is block-diagonal. These blocks can be solved (and normalized) separately in order to determine the relative probability for the particle to be on each site if it solely diffuses in the given zone. To first order in ε, the normalized solutions are given by:

$$
\begin{aligned}
|n_a\rangle &\equiv \begin{pmatrix} n(1) \\ n(2) \\ n(3) \\ n(4) \\ n(5) \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 1-\frac{8}{11}\varepsilon \\ 1+\frac{8}{11}\varepsilon \\ 1-\frac{2}{11}\varepsilon \\ 1+\frac{2}{11}\varepsilon \\ 1 \end{pmatrix}, \\
|n_b\rangle &\equiv \begin{pmatrix} n(6) \\ n(7) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1-\varepsilon \\ 1+\varepsilon \end{pmatrix}, \\
|n_c\rangle &\equiv \begin{pmatrix} n(8) \\ n(9) \\ n(10) \\ n(11) \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix},
\end{aligned}
\tag{3.8}
$$

while the mean velocity on each site are

$$
\begin{aligned}
|v_a\rangle &\equiv \begin{pmatrix} v(1) \\ v(2) \\ v(3) \\ v(4) \\ v(5) \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1+\varepsilon \\ -1+\varepsilon \\ 2\varepsilon \\ 2\varepsilon \\ 2\varepsilon \end{pmatrix}, \\
|v_b\rangle &\equiv \begin{pmatrix} v(6) \\ v(7) \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1+\varepsilon \\ -1+\varepsilon \end{pmatrix}, \\
|v_c\rangle &\equiv \begin{pmatrix} v(8) \\ v(9) \\ v(10) \\ v(11) \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 0 \\ 2\varepsilon \\ 2\varepsilon \\ 2\varepsilon \end{pmatrix}.
\end{aligned}
\tag{3.9}
$$

Finally, the three different diffusion coefficients can be calculated:

$$
\begin{aligned}
D_a^* &= \lim_{\varepsilon \to 0} \frac{\langle v_a | n_a \rangle}{\mu_0 \varepsilon} = \frac{36}{55}, \\
D_b^* &= \lim_{\varepsilon \to 0} \frac{\langle v_b | n_b \rangle}{\mu_0 \varepsilon} = 0, \\
D_c^* &= \lim_{\varepsilon \to 0} \frac{\langle v_c | n_c \rangle}{\mu_0 \varepsilon} = \frac{3}{4}.
\end{aligned}
\tag{3.10}
$$

Note that the diffusion coefficient in the closed lake is zero, as it should, since a particle trapped in a lake cannot move over macroscopic distances. One can thus save time by attributing, e.g., an *a priori* zero diffusion coefficient to all zones containing iess sites than the length of the cell in the field direction (in this example, the lengths are: $L_x = 3$ and $L_y = 6$).

When the values of D^* are known for each independent zone, a global (or weighted) value of D^* can be calculated. However, this step is very problem-dependent. For example, if particles are injected randomly and the measured diffusion coefficient takes all of them into account, one must weight the various values using the number of empty sites in each "zone"; here, this procedure would give $D^* = (5D_a^* + 2D_b^* + 4D_c^*)/11 = 69/121$. One could also disregard closed lakes (e.g., if it is impossible to follow these molecules), or consider only the largest "zone". Clearly, this depends on the problem being studied. In our studies of random systems, we will weight all channels using their size (number of sites), but we will not count the lakes (equivalently, we can say that we fill the lakes with obstacles).

3.5 Algebraic simplifications

In order to save one mathematical step, the normalization condition can be introduced directly into the matrix equation. For each separate zone, one can replace one of the equations of the matrix block (which is linearly dependent upon the others anyway) by the normalization equation $\sum_{\{i \in zone\}} n_i = 1$. In this way, the solution is automatically normalized. In practice, we first rewrite the matrix equation for a given zone $T|n\rangle = |n\rangle$, as $(T-I)|n\rangle = |0\rangle$, I being the identity matrix ($I_{ij} = \delta_{ij}$) and

$|0\rangle$ the zero-vector. To normalize the solution, we then replace the last line of the $T - I$ matrix and the last element of the $|0\rangle$ vector by one's. The resulting matrix equation is $A|n\rangle = |b\rangle$, where $|b\rangle = \{0, 0, ..., 1\}$ and A is the new (i.e., modified) transition matrix.

We can further simplify the numerical problem by eliminating the need for an arbitrary (finite) value of ε. First, let's separate each term in the $A|n\rangle = |b\rangle$ matrix equation (and the velocity vector) into a constant term (with a subscript I) and a ε^1-dependent term (with a ε subscript):

$$\begin{aligned} \text{Transition matrix:} \quad A &= A_I + \varepsilon A_\varepsilon, \\ \text{Probability vector:} \quad |n\rangle &= |n_I\rangle + \varepsilon|n_\varepsilon\rangle, \\ \text{b vector:} \quad |b\rangle &= |b_I\rangle, \\ \text{Velocity vector:} \quad |v\rangle &= |v_I\rangle + \varepsilon|v_\varepsilon\rangle. \end{aligned} \tag{3.11}$$

Note that the b-vector is always independent of ε. With Eq.(3.11), the matrix equation $A|n\rangle = |b\rangle$ becomes

$$A_I|n_I\rangle + \varepsilon\left(A_I|n_\varepsilon\rangle + A_\varepsilon|n_I\rangle\right) + \varepsilon^2 A_\varepsilon|n_\varepsilon\rangle = |b_I\rangle. \tag{3.12}$$

To first order in ε, this relation can be satisfied for an arbitrary field intensity ε only if

$$A_I|n_I\rangle = |b_I\rangle \qquad \text{and} \qquad A_I|n_\varepsilon\rangle = -A_\varepsilon|n_I\rangle. \tag{3.13}$$

The first relation corresponds to the field-free case and leads to the trivial solution $n_I(i) = 1/J$ (where J is the number of connected sites in the given zone). It is in fact the second relation that contains the non-trivial effects of the fixed obstacles on the diffusion coefficient (or mobility).

Once the total probability vector $|n\rangle = |n_I\rangle + \varepsilon|n_\varepsilon\rangle$ is calculated, the mean velocity v is given by (to first order in ε):

$$\begin{aligned} v = \langle v|n\rangle &= \left(\langle v_I| + \varepsilon\langle v_\varepsilon|\right) \cdot \left(|n_I\rangle + \varepsilon|n_\varepsilon\rangle\right) \\ &= \langle v_I|n_I\rangle + \varepsilon\left(\langle v_I|n_\varepsilon\rangle + \langle v_\varepsilon|n_I\rangle\right). \end{aligned} \tag{3.14}$$

The term $\langle v_I|n_I\rangle = 0$ since it gives the velocity without an external field. The reduced diffusion coefficient is thus given by:

$$D^* = \frac{\langle v_I|n_\varepsilon\rangle + \langle v_\varepsilon|n_I\rangle}{\mu_0}. \tag{3.15}$$

This approach does not require us to use any specific value for the field intensity ε. The only step here is the evaluation of the field-dependent correction term $|n_\varepsilon\rangle = -(A_I)^{-1}A_\varepsilon|n_I\rangle$ of the total probability density vector $|n\rangle = |n_I\rangle + \varepsilon|n_\varepsilon\rangle$. Therefore, exact results in term of rational fractions can be obtained without using symbolic calculations.

Note that we can rewrite Eq.(3.15) as

$$D^* = \frac{(\langle v_\varepsilon| - \langle v_I|A_I^{-1}A_\varepsilon)|n_I\rangle}{D_0} = \frac{\langle d|n_I\rangle}{D_0}, \tag{3.16}$$

where $\langle d| = \langle v_\varepsilon| - \langle v_I|A_I^{-1}A_\varepsilon$ can be seen as a "local diffusion coefficient" vector. While it is not clear how to interpret $\langle d|$, it is the first time, to our knowledge, that someone has actually defined a local diffusion coefficient for a given lattice site.

3.6 Isotropy considerations

When the system is isotropic, the diffusion coefficient D^* (or mobility μ^*) should be identical in all directions. To verify that our method satisfies this requirement, the diffusion coefficient will be calculated for an arbitrary direction θ in this section.

Our approach calculates D^* (or μ^*) in the direction of the external field ε. In order to obtain D^* in a given direction we simply apply the external field in that direction. Starting from Eq.(3.15) we get:

$$D = D_0 \times D^* = \langle v_I|n_\varepsilon\rangle + \langle v_\varepsilon|n_I\rangle = \lim_{\varepsilon\to 0}\frac{v}{\varepsilon}. \tag{3.17}$$

We will now get through each of those terms and see how they change when the field is at an angle θ with the x axis.

3.6.1 The local velocity

This velocity will now have two components, one for the x direction and one for the y direction. Since it is by definition independent of the field, $|v_I\rangle$ will also be independent of θ. Thus the x component of $|v_I\rangle$ will be identical to the case where the field was in the x direction and the y component will be identical to the case where the field was in the y direction:

$$|v_{I\theta}\rangle = (|v_{Ix}\rangle, |v_{Iy}\rangle), \tag{3.18}$$

where the subscripts denote velocities for the cases where the field is in the θ, x and y directions, respectively.

It can easily be shown that the field-dependent term $|v_{\epsilon\theta}\rangle$, will be given by:

$$|v_{\epsilon\theta}\rangle = (cos(\theta)|v_{\epsilon x}\rangle, sin(\theta)|v_{\epsilon y}\rangle). \tag{3.19}$$

3.6.2 The local probability

Similarly, the $|n_I\rangle$ term will be the same for any field direction:

$$|n_{I\theta}\rangle = |n_{Ix}\rangle = |n_{Iy}\rangle = |n_I\rangle = \frac{1}{J}, \tag{3.20}$$

where J is the number of empty sites. The $|n_{\epsilon\theta}\rangle$ term is found from Eq.(3.13). Note that the A_I matrix is also field-independent. The $A_{\epsilon\theta}$ matrix is given by:

$$A_{\epsilon\theta} = cos(\theta)A_{\epsilon x} + sin(\theta)A_{\epsilon y}. \tag{3.21}$$

Thus we have:

$$\begin{aligned} |n_{\epsilon\theta}\rangle &= -A_I^{-1}A_{\epsilon\theta}|n_I\rangle \\ &= -A_I^{-1}(cos(\theta)A_{\epsilon x} + sin(\theta)A_{\epsilon y})|n_I\rangle \\ &= -cos(\theta)A_I^{-1}A_{\epsilon x}|n_I\rangle - sin(\theta)A_I^{-1}A_{\epsilon y}|n_I\rangle) \\ &= cos(\theta)|n_{\epsilon x}\rangle + sin(\theta)|n_{\epsilon y}\rangle. \end{aligned} \tag{3.22}$$

3.6.3 The diffusion coefficient

Using Eqs.(3.18), (3.19), (3.20), and (3.22), we can find the mean velocity of the particle,

$$\begin{aligned} v_\theta &= \varepsilon(\langle v_{\varepsilon\theta}|n_{I\theta}\rangle + \langle v_I|n_{\varepsilon\theta}\rangle) \\ &= (v_x cos(\theta) + \varepsilon sin(\theta)\langle v_{Ix}|n_{\varepsilon y}\rangle, v_y sin(\theta) + \varepsilon cos(\theta)\langle v_{Iy}|n_{\varepsilon x}\rangle), \end{aligned} \tag{3.23}$$

where we used:

$$\begin{aligned} v_x &= \varepsilon(\langle v_{\varepsilon x}|n_{Ix}\rangle + \langle v_{Ix}|n_{\varepsilon x}\rangle), \\ v_y &= \varepsilon(\langle v_{\varepsilon y}|n_{Iy}\rangle + \langle v_{Iy}|n_{\varepsilon y}\rangle). \end{aligned} \tag{3.24}$$

For isotropic systems the $\langle v_{Ix}|n_{\varepsilon y}\rangle$ and the $\langle v_{Iy}|n_{\varepsilon x}\rangle$ terms will vanish if the lattice is sufficiently large (each configuration leading to a negative term being counterbalanced by another one leading to a positive term). Therefore, the "global" mean velocity of an isotropic system is:

$$v_\theta = \sqrt{{v_x}^2 cos^2(\theta) + {v_y}^2 sin^2(\theta)}. \tag{3.25}$$

giving a diffusion coefficient:

$$D_\theta = \frac{v_\theta}{\varepsilon} = \sqrt{{D_x}^2 cos^2(\theta) + {D_y}^2 sin^2(\theta)}. \tag{3.26}$$

Thus, if the system is isotropic ($D_x = D_y$), the diffusion coefficient is independent of the direction for sufficiently large primitive cells.

Since we work with finite lattice sizes, we should always calculate the diffusion coefficient along one of the axis, to avoid the complications of the cross term in Eq.(3.23). Eq.(3.26) shows that there is no loss of generality in doing so. Anisotropic cases are more tricky to treat and have not been considered in this work.

3.7 Conclusion

In this chapter, we have presented an algebraically exact method to calculate the diffusion coefficient D^* of a particle in a system with immobile obstacles and periodic

boundary conditions. The method uses the Nernst-Einstein relation to transform the problem into a field-driven drift problem where D^* is related to the zero-field scaled mobility μ^*. As an example, we calculated the diffusion coefficients for a particle in a simple anisotropic system. In fact, our approach reduces the diffusion problem to the solution of a (large) system of linear equations.

When the obstacles are placed randomly, a system can contain closed volumes (lakes) and/or multiple independent pathways (channels). We have shown how one can treat these cases in a systematic way.

In the next chapter, a powerful numerical implementation will be described and tested and several applications will be given.

Chapter 4

Exact results: Implementation and examples

In the previous chapter, an exact method to calculate the scaled diffusion coefficient D^* of a particle in a lattice system with immobile obstacles and periodic boundary conditions was derived. In this chapter, a numerical implementation of this method will be described and tested. Several applications will be treated using this implementation in order to demonstrate both its validity and its power[1].

4.1 Introduction

Our method is based on the Nernst-Einstein relation and reduces the problem to the solution of a (large) system of linear equations. For small systems, one can use the numerical capabilities (more precisely, the exact arithmetics functions) of, e.g., Mathematica, to solve the linear system exactly and obtain exact results in the form of rational fractions. For example, the diffusion coefficient for the system shown in Fig. 4.1a is exactly:

$D^*(C) =$

$$\frac{1567385513021793366294054168681619296405823446186520108216313595962173294107386267262338478168372096513834291437243198852598921227261066686874813301079743821053753024084429854593459701267220336621340470004480987498437370922904615432651955128922627776}{}$$

[1]This work has been published in the *Journal of Chemical Physics*[21].

Such exact results, given as rational fractions, can provide useful information about model systems. As is demonstrated in Section 4.3.2, the precise value of some series expansion coefficients can be studied using such calculations.

However such a huge rational fraction is often useless and the computer memory and the number of mathematical operations (and thus the CPU time) needed to obtain it increase extremely rapidly with the size of the system. Therefore such calculations are, in practice, limited to relatively small systems. A numerical approach, capable of treating larger periodic cells, is introduced in Section 4.2. This procedure gives D^* values with an arbitrary precision and requires far less memory and mathematical operations than an exact calculation. We have used this numerical method in our studies of randomly distributed obstacles (Section 4.3.3) and to observe subtle effects related to the geometrical structure of deterministic cross fractal obstacles (Section 4.3.4).

If random distributions of obstacles are studied, the exact value of D^* can only be calculated for specific realizations of the disorder; the true thermodynamic value of D^* is then an average over many realizations in the limit of infinite systems. The convergence (and CPU efficiency) of the procedure for such problems will be examined in Section 4.3.3.

4.2 Numerical implementation

As described in Chapter 3, exact calculations using our approach are reduced to solving a (large) system of linear equations of the form $A|n\rangle = |b\rangle$. When the size of the modified transition matrix A is not too large, any standard method, like LU decomposition, should be adequate. However, the memory ($\sim L^{2d}$) and the number of mathematical operations ($\sim L^{3d}$) needed for standard methods increase rapidly with the linear size L of the lattice and the dimensionality d of the space. It is possible to solve much larger systems of equations, while keeping both the memory requirements and CPU time at a reasonable level, by using special index storage methods developed for sparse matrices and iterative solving algorithms.

Large systems actually result in extremely sparse transition matrices. For example, a maximum of 5 elements per row are not equal to zero in the case of a two-dimensional square lattice (each site can only be coupled to its immediate neighbors!). The main idea of sparse storage technique is to store only the non-zero elements and their location. Iterative methods designed to use such storage have been developed. Unlike LU decomposition and other standard direct solving methods, these iterative approaches do not need the expansion of the matrix, and thus need far less limited memory resources.

With the row-indexed compact storage technique[22] (the storage technique) and the biconjugate gradient method[22] (the matrix solving technique), we have been able to solve matrices as large as 250000×250000 (which correspond roughly to two-dimensional 500×500 cells with PBC's) on a 200 MHz IBM PowerPC (running AIX 4.1.5), with 96 MB of RAM. Such a calculation typically takes about 30 minutes for diffusion coefficients that are precise to within one part in 10^8. Mathematica is limited to about 2500×2500 matrices, while lattice Monte Carlo simulation methods typically use 256×256 lattices and can hardly give error bars that are less than one part in 10^3.

4.3 Applications

In this section, we will look at a few systems that make use of a two-dimensional square lattice. The efficiency of the numerical approach is studied for randomly distributed obstacles in Section 4.3.3. Section 4.3.5 shows how our method can be adapted to treat triangular lattices, while Section 4.3.6 investigates higher spatial dimensionalities.

4.3.1 The diffusion coefficient at low concentrations

For any dimensionality or lattice structure, we expect that, in the zero-concentration ($C \rightarrow 0$) limit, the diffusion coefficient dependence upon obstacle concentration C (the fraction of lattice sites occupied by immobile obstacles) will be, to first order

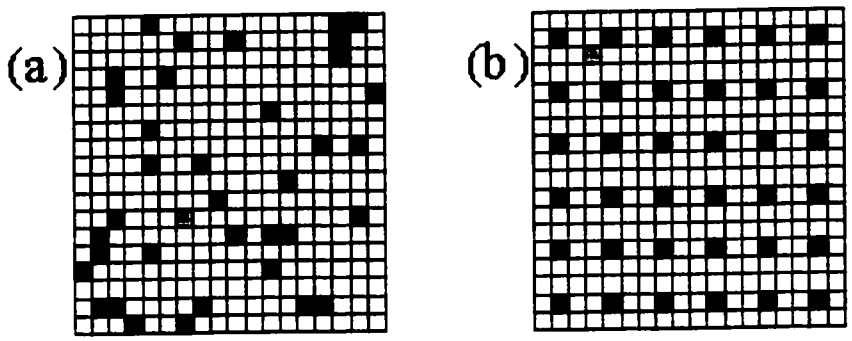

Figure 4.1: (a) An example of a random distribution of obstacles. The lattice size is 18×18, and $D^* = 0.7383...$ (b) The obstacles are now distributed periodically and form a sublattice with a lattice parameter $L = 3$ giving $D^* = 45/56 \approx 0.803571...$ In both cases the concentration of obstacles is $C = 1/9 \approx 11.1\%$ and the diffusing 1×1 particle is in grey.

in C, identical for all types of distribution of obstacles if the basic obstacle shape is the same. For example, if the obstacles are all single site obstructions, random and periodic distributions of obstacles should look the same when $C \to 0$ because in this limit the obstacles are not correlated. Let us write the concentration dependence of the diffusion coefficient as a series expansion:

$$D^*(C) = 1 + a_1 C + a_2 C^2 + a_3 C^3 + \ldots . \tag{4.1}$$

For single site diffusing particles and single site obstacles randomly occupying the vertices of a two-dimensional square lattice (see Fig. 4.1a), Nieuwenhuizen et al.[23] have shown that $a_1 = -(\pi - 1)$ in the thermodynamic $L \to \infty$ limit. Slater and Guo[9] have demonstrated that this is indeed also valid for periodically distributed obstacles (Fig. 4.1b). The second coefficient (a_2) depends on the type of obstacle distribution because it represents the correlation between two obstacles; for a random distribution, we have approximately $a_2 = -0.8558$ (this will be studied in Section 4.3.3), while Slater and Guo[10] have shown that if the obstacles themselves

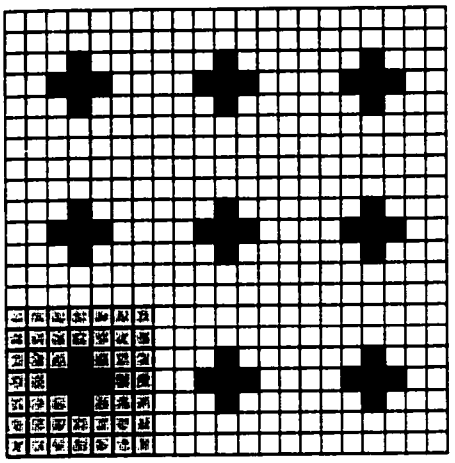

Figure 4.2: A system with periodically distributed cross-shaped obstacles. The grey area shows the size of the repeating unit cell. The cross is formed by 5 obstacles. Each $L \times L$ cell contains one such cross. Here, we have $L = 7$, hence the concentration is $C = 5/49 \approx 10.2\%$ and $D^* = 179585/221892$ for a 1×1 probe particle.

form a square sub-lattice (either simple (Fig. 4.1b) or body-centered), they obtain exactly $a_2 = ((\pi - 1)^2 + 1)/2 \approx 2.7932$. In the latter case, it is the third coefficient that differs between the two types of sub-lattices. Slater and Treurniet[11] also studied systems with larger (square) obstacles and migrating particles (without rotation) and obtained very precise values for the coefficients a_1, a_2 and a_3 even in the limit of infinitely large particles and obstacles (the limit of the continuum). We refer the reader to Refs. $[9 - 11]$ for further information. Appendix B shows that it is sometimes possible to find analytically an exact first-order coefficient for very simple systems.

4.3.2 Periodically distributed cross-shaped obstacles

As an example of exact results one can obtain for periodic systems using our approach, and as an introduction to the deterministic fractals studied in Section 4.3.4, we now present an investigation of the simple system shown in Fig. 4.2. Each obstacle is a 3×3 cross, and the latter are distributed periodically on the lattice such

Table 4.1: Some exact diffusion coefficients D^* for different system sizes L and obstacle concentrations C for the system shown in Fig. 4.2.

L	C	D^*
4	5/16	16/33
5	1/5	55/84
6	5/36	1788/2387
7	5/49	179585/221892
8	5/64	122784/144373
10	1/200	.901326097...
20	1/800	.974207745...
30	1/1800	.988436218...
40	1/3200	.993475221...
50	1/5000	.995818131...
60	1/7200	.9970936500...
75	1/11250	.9981387424...

that they form a square sub-lattice of periodicity $L = 7$. The concentration is thus given by $C = 5/L^2 = 5/49$.

We have solved this problem using the method described in Chapter 3, hence reducing the numerical effort to finding the correction vector $|n_\epsilon\rangle$. We have used Mathematica to find the exact numerical values of D^* (represented by rational fractions) for the smaller systems ($L \leq 18$). Table 4.1 presents some of our results. As we can see, the reduced diffusion coefficient D^* increases very quickly with the lattice size (and sub-lattice parameter) L. When plotted as a function of the concentration C (Fig. 4.3), the data shows a simple monotonic decrease and a positive curvature. The solid line gives the best fit using a fourth-order polynomial:

$$D^*(C) = 1 - 2.09630831C + 2.69726C^2 - 0.51274C^3 + 7.06C^4 + \ldots . \qquad (4.2)$$

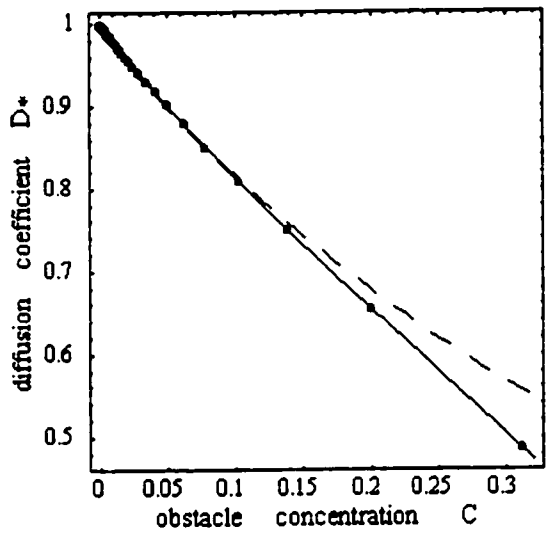

Figure 4.3: Diffusion coefficient D^* vs obstacle concentration C for the system shown in Fig. 4.2 (see also Table 4.1). The solid line shows the best polynomial fit given by Eq.(4.2). The dashed line shows the fit to the stretched exponential $D^* = \exp\left(-(C/0.4075)^{1.046}\right)$.

As one can verify, there is no inflection point in the allowed range $0 \leq C \leq 5/16$. The accuracy here for periodic systems cannot be matched by LMC simulations. It is interesting to note that stretched exponential fits (suggested by some authors for related problems[24]) do not provide better results than the simple polynomial expansion (see Fig. 4.3).

4.3.3 The pseudo-random gel

General principles

While one can obtain the exact diffusion coefficient D^* (or, at least, an arbitrarily precise estimate of it) if the obstacles form a periodic pattern, the situation is different if there is any degree of randomness in the obstacle distribution. Although our approach still outperforms LMC simulations, one must be careful in order to obtain reliable values (of course, care is also required if one uses LMCs). In order

to explain the main issues, we now present a study of the problem of randomly distributed obstacles on a square lattice. The results presented here were all obtained using the numerical implementation described in Section 4.2.

For each concentration C, we first choose a system lattice size (typically $30 < L < 500$) and calculate the number $N = C \times L \times L$ of obstacles needed to obtain the right concentration. We then place these obstacles randomly on the L^2 lattice sites and calculate the exact diffusion coefficient D^* using PBCs. For example, Fig. 4.1a shows a $L = 18$ system with $N = 36$ obstacles (giving a concentration $C = 1/9$). After repeating this procedure for $\Omega \gg 1$ different random distributions of the N obstacles, we average over these exact solutions to yield the average diffusion coefficient $D^*(C, L)$, as well as its uncertainty $\Delta D^*(C, L) = \sigma_D(C, L)/\Omega^{1/2}$, where $\sigma_D(C, L)$ is the standard-deviation of the distribution of exact diffusion coefficients $D^*(C, L)$. Next, in order to obtain the asymptotic diffusion coefficient $D^*(C, L \to \infty)$, we calculate $D^*(C, L)$ for increasing system sizes L until the result becomes independent of L (within computational uncertainties) or until a satisfying extrapolation value can be estimated for $D^*(C, \infty)$. In the next few sections, we will study the convergence properties of D^* and σ_D as well as the efficiency of our computational approach when the lattice size $L \to \infty$.

The convergence of $D^*(L)$

For small obstacle concentrations C, the diffusion coefficient $D^*(C, L)$ becomes essentially independent of the cell dimension L for relatively small values of L. For example, Table 4.2 gives the diffusion coefficient $D^*(C = 4\%, L)$, the standard-deviation $\sigma_D(C = 4\%, L)$, the number Ω of systems generated, and finally the CPU time required per system for increasing system sizes L (all calculations were done using our own Fortran program). The convergence of D^* for large L's is shown in Fig. 4.4; as we can see, the diffusion coefficient reaches an asymptotic value $D^* = 0.91292(3)$ for (approximately) $L \geq 70$. Beyond $L = 70$, the uncertainty on the points are too large to provide any improvement in our estimate of $D^*(C, L \to \infty)$. Therefore, much larger ensemble sizes Ω, not larger system sizes L,

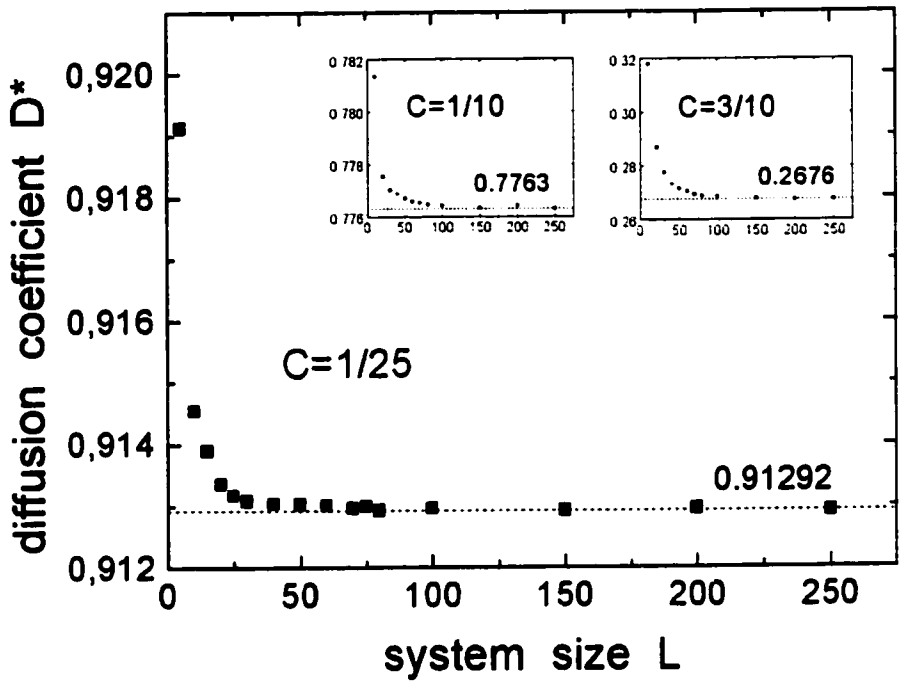

Figure 4.4: Diffusion coefficient $D^*(C, L)$ vs system size L for randomly distributed obstacles at $C = 4\%$ (main figure), and $C = 10\%$ and $C = 30\%$ (insets). The raw data are given in Tables 4.2, 4.3 and 4.4 respectively. The horizontal dashed lines give our best estimates of the asymptotic values $D^*(C, L \to \infty)$. The error bars are comparable to (or smaller than) the size of the points in all cases.

would be necessary to increase the accuracy of $D^*(C, L \to \infty)$.

Tables 4.3 and 4.4, as well as Fig. 4.4, show some data for $C = 10\%$ and $C = 30\%$, respectively. The asymptotic values are found to be $D^*(10\%) = 0.7763(3)$ and $D^*(30\%) = 0.2676(3)$. We note that $D^*(C, L)$ is already within 0.5% of its asymptotic value for approximate system sizes $L = 5$, 15, and 75, for $C = 4\%$, 10% and 33% respectively. Therefore, convergence is slower when the concentration increases. This can be explained in terms of the correlation length $\xi(C)$ since the system "looks" infinite only if $L \gg \xi$. For a 2D random distribution of obstacles,

Table 4.2: Diffusion coefficient $D^*(C = 4\%, L)$ for random two-dimensional systems of size L. Also given are the ensemble size Ω, the standard-deviation σ_D, and the CPU time required per system. Note that the result for $L = 5$ is exact.

L	Ω	D^*	σ_D	CPU(s)
5	1	25/136	0	0
10	10000	0.9145(1)	0.013	≤ 0.01
20	10000	0.91336(8)	0.008	0.1
30	5000	0.91307(8)	0.0056	0.3
50	1000	0.9130(1)	0.0033	1.3
100	1000	0.91295(5)	0.0016	11.7
150	1000	0.91292(3)	0.00109	38.2
200	750	0.91294(3)	0.0008	95.0
250	500	0.91290(3)	0.0006	206.5

the correlation length is known[25] to increase like $\xi \sim (C - C^*)^{-4/3}$ when C is close to C^*, where $C^* = 40.7254\%$ is the percolation threshold of this system (i.e., we have $D^* = 0$ for $C > C^*$ in the $L \to \infty$ limit). Therefore, both the correlation length ξ and the system size L required to reach the asymptotic limit will diverge when we get close to $C = C^*$. We will see in the next section that the correlation length also affects the standard-deviation σ_D.

The standard-deviation $\sigma_D(C, L)$

The shape and width of the distribution function for the diffusion coefficients D^* also depends on the system size L. For example, Fig. 4.5 shows distributions of D^* for many system sizes at $C = 30\%$. For small system sizes, there is a finite probability that the system will not percolate (note that this requires $L < 1/C$): this leads to a peak for $D^* = 0$. The distribution is extremely wide when L is small, and the main peak is asymmetric. More importantly, however, we note that the

Table 4.3: Diffusion coefficient $D^*(C = 10\%, L)$ for random two-dimensional systems of size L. Also given are the ensemble size Ω, the standard-deviation σ_D, and the CPU time required per system.

L	Ω	D^*	σ_D	CPU(s)
10	10000	0.7814(4)	0.04	≤ 0.01
20	10000	0.7775(2)	0.021	0.1
30	1000	0.7770(5)	0.015	0.3
50	1000	0.7767(3)	0.009	0.7
100	1000	0.7764(1)	0.0044	12
150	1000	0.77634(9)	0.003	40.7
200	750	0.77642(8)	0.0021	97.8
250	500	0.77630(8)	0.0018	199.35

Table 4.4: Diffusion coefficient $D^*(C = 30\%, L)$ for random two-dimensional systems of size L. Also given are the ensemble size Ω, the standard-deviation σ_D, and the CPU time required per system.

L	Ω	D^*	σ_D	CPU(s)
10	10000	0.328(1)	0.11	≤ 0.01
20	10000	0.2871(7)	0.07	0.1
30	10000	0.2776(5)	0.052	4.3
50	2000	0.2714(8)	0.034	22
100	1000	0.2685(5)	0.017	23.9
150	1000	0.2680(3)	0.012	88.1
200	750	0.2676(3)	0.0088	223.1
250	500	0.2676(3)	0.0066	450.1

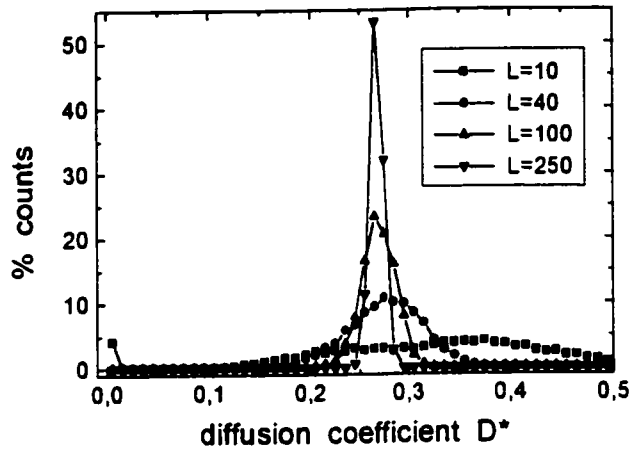

Figure 4.5: Probability distribution function for the diffusion coefficient $D^*(C = 30\%, L)$ for various system sizes L. A total of $\Omega = 1000$ systems were generated in each case.

peaks shift to the left and become much narrower for larger system sizes. This last result is not surprising since the width should vanish for an infinite system size (this is the thermodynamic limit).

Figure 4.6 shows the standard deviation σ_D as a function of the system size L for the three concentrations studied in this section. For sufficiently large system sizes, we observe the scaling law $\sigma_D \sim 1/L$ (which is actually quite useful in practice). When $L > \xi$, the standard-deviation σ_D is related to the variation in diffusion coefficients between independent patches of surface area $\approx \xi^2$. Since there are $(L/\xi)^2$ such patches in a two-dimensional system of size L, the $\sigma_D \sim 1/L$ scaling law is a natural result. For smaller sizes $L < \xi$, however, σ_D is related to the structure of the system, which is then a very correlated "patch", and the scaling law is modified. Our curves indicate that these deviations start for system sizes smaller than approximately $L = 10$, 20 and 50, respectively, for $C = 4\%$, 10% and 30%, in rough agreement with the system sizes found in the previous section for the convergence of D^*. Therefore, we conclude that the correlation lengths are given roughly by $\xi(4\%) \approx 10$, $\xi(10\%) \approx 20$, and $\xi(30\%) \approx 50$. Although a systematic

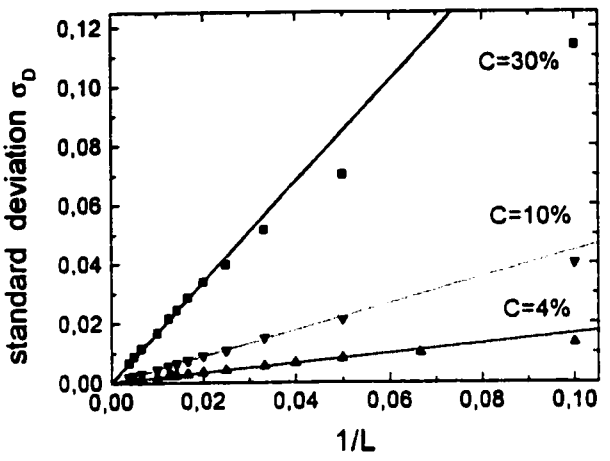

Figure 4.6: Standard-deviation σ_D vs. inverse system size $1/L$ for three different concentrations C. We note that $\sigma_D^* \sim 1/L$ for large system sizes.

study of the correlation length is beyond the scope of our work, this concept is important when one wants to evaluate the error made when using finite-size lattices with PBC's.

Efficiency

As we saw in the two previous sections, a choice of system size $L > \xi(C)$ is important in order to obtain reliable estimates of the thermodynamic value $D^*(C)$. A larger system size means a smaller finite-size error $\delta_\infty = |D^*(C, L \to \infty) - D^*(C, L)|$, but the CPU time increases very quickly with system size. Figure 4.7 shows that the CPU time (per system) increases approximately like L^3 with our iterative computational method. Since the standard-deviation decreases only like $\sigma_D \sim 1/L$, choosing too large a size L is not necessarily advantageous. In practice, the finite-size error $\delta_\infty(C, L)$ quickly becomes much smaller than the achievable ensemble errors $\Delta D^*(C, L)$ for systems $L > \xi$. Consequently, it is better to invest CPU time in reducing $\Delta D^*(C, L) = \sigma_D/\sqrt{\Omega}$, which means increasing the ensemble size Ω, than in increasing the system size L.

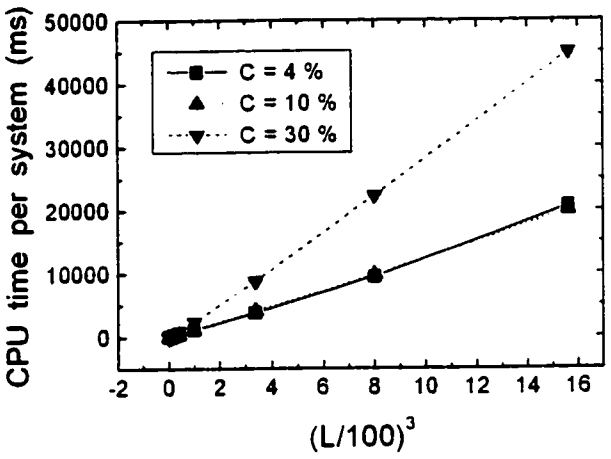

Figure 4.7: Mean CPU time (in ms) necessary to calculate the diffusion coefficient $D^*(C, L)$ for systems of size $L \times L$ with the iterative technique. The values of D^* were calculated to a precision of 10^{-8}.

The same conclusions apply to LMC simulations. In the latter case, however, the uncertainty on the evaluation of the diffusion coefficient for a given system is not zero like here. Hence, the uncertainty on the mean value $D^*(C, N)$ is larger and the convergence is slower.

Percolation and diffusion

Figure 4.8 shows the diffusion coefficient D^* as a function of C for low obstacle concentrations. These data points have all been obtained using a lattice size $L = 150$ and an ensemble size $\Omega = 1000$. The error bars are all smaller than 0.05%. The polynomial fit (solid line), whose coefficient a_1 has been chosen to be the exact value $a_1 = -(\pi - 1)$, results in a coefficient $a_2 = -0.85(3)$, in agreement with the expected value[23] ($a_2 = -0.8558$). The fit thus predicts a percolation concentration $C^* = 40.2(2)\%$, also in fair agreement with the exact result $C^* = 40.7254\%$. More precise values would require larger ensembles and larger system sizes. Again the stretched exponential fit does not provide a better result than the simple polynomial

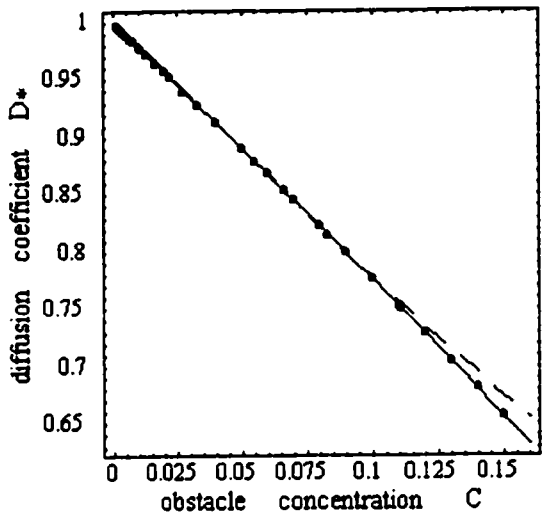

Figure 4.8: Diffusion coefficient D^* vs. obstacle concentration C for single site particles diffusing through randomly placed single site obstacles. The system size is $L = 150$, while the ensemble size is $\Omega = 1000$. The solid line gives the best polynomial fit $D^*(C) = 1 - (\pi - 1)C - 0.85(3)C^2 + ...$, while the dashed line gives the stretched exponential fit $D^* = exp(-(C/0.231)^{1.145})$

expansion (see dashed line in Fig. 4.8) for low concentrations, while such a fit does not predict any percolation threshold C^*.

4.3.4 Deterministic fractals

An interesting application of our computational approach is that of fractal systems. Michael J. Saxton[14] has studied the diffusion of particles of different sizes $M \times M$ in two-dimensional systems with a triangular lattice and fractal aggregates of obstacles formed using either the cluster-cluster aggregation (CCA) or the multicentre diffusion-limited aggregation (MDLA) algorithms. Because of the self-similar nature of the fractal aggregates, he observed that the diffusion coefficient $D^*(M)$ was almost independent of M over a fair range of molecular sizes M. This remarkable and counter-intuitive result can be explained by the lack of a characteristic length scale in a fractal (or self-similar) system.

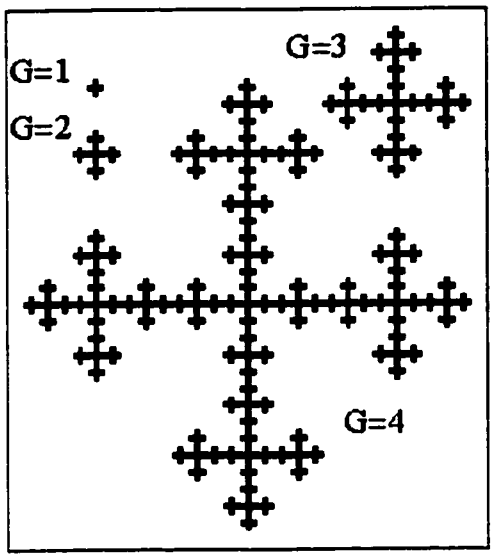

Figure 4.9: Building deterministic fractals by the iterative process.

In order to exploit the fact that our method provides exact values of D^* when the system is periodic, we chose to use periodically distributed deterministic fractals instead of CCA (or MDLA) fractals. In all cases, $D^*(M)$ was calculated to a precision of 10^{-10} using the iterative method.

The deterministic cross-shaped fractal

Figure 4.9 shows that one can transform the cross-shaped obstacle studied previously into a fractal structure by a simple iterative procedure where a given generation can be obtained by substituting each element of the fractal of the previous generation by the original cross pattern. We will thus say that a generation G fractal is obtained after G iterations. The simple cross obstacle for example, is a $G = 1$ object. Figure 4.9 shows the first four generations. For this system the fractal dimension is $d_f = ln(5)/ln(3) \approx 1.46497$. Since we have one such structure per periodic cell, the system is actually a periodic sublattice of fractal obstacles (Fig. 4.10).

Two important length scales exist for our finite-size fractals: the system size L and the fractal linear size $L_G = 3^G$. Note that we must have $3^G < L$ otherwise $D^* = 0$ (see Fig. 4.10). The obstacle concentration $C(G) = 5^G/L^2$ is useless for

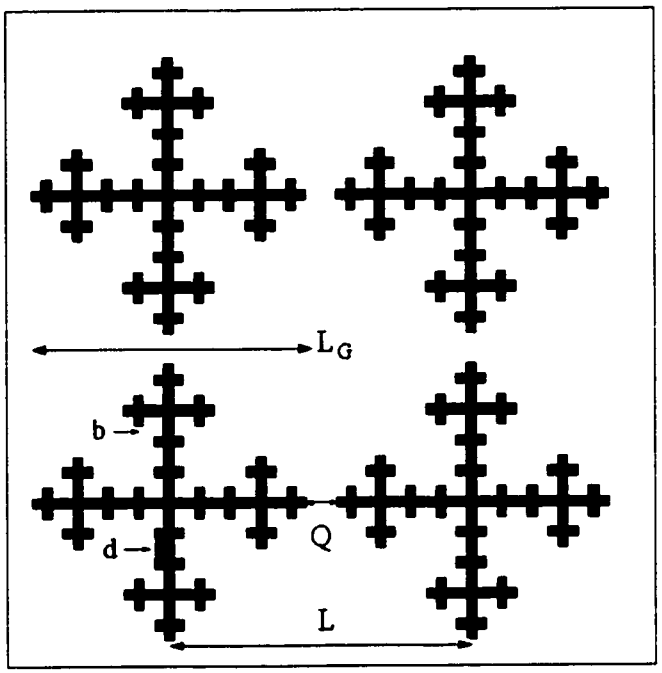

Figure 4.10: A periodic distribution of fractals of generation $G = 3$. The obstacles have a linear size $L_G = 27$ and the periodicity is $L = 30$. The channel width is $Q = L - L_G = 3$. The light grey zone (b) shows a bay, while the dark grey zone (d) is a driveway.

comparing different generations since fractals become infinitely sparse for very large generations:

$$\lim_{G \to \infty} C(G) = 0. \tag{4.3}$$

A more useful parameter is thus the minimum separation between two fractals, $Q \equiv L - L_G$, which is the size of the channel separating two consecutive fractal obstacles (see Fig. 4.10).

A large particle has always access to less free volume than a smaller one. In a system where the obstacles are distributed randomly, less free volume will automatically result in slow long range diffusion and smaller diffusion coefficients. In a fractal environment, however, the situation is quite different. The smaller effective channel width Q will tend to slow down the larger particles. On the other hand, the diffusing

particles can now get trapped into many dead-ends. The loss of useless free volume associated with these dead-ends will actually result in increased diffusion for the larger particles. It is the balance between these two effects, narrower channels but fewer dead-ends for larger particles, that determines D^*. It turns out that because the system "looks" the same over many different length scales (self-similarity), these two effects can cancel each other so the diffusion coefficient becomes insensitive to the size of the particle over a certain range of probe sizes. In the following, we will only study the diffusion of square particles of linear size M and surface area M^2.

While the dead-ends of a non deterministic fractal have a wide range of sizes and shapes, our deterministic fractal contains only two types of dead-ends (Fig. 4.10), which we will name the driveways (d) and the bays (b). The driveways are simple linear roads which end abruptly; these dead-ends do not appear to have a strong impact on D^*, probably because they form linear structures. The bays, on the other hand, strongly affect D^*, as we will see later. For the G-th generation fractal, there are $G-1$ different bay sizes. The largest square particle size that can enter in the i-th type of bay ($i = 1, 2, ..., G-1$), M_i, and the total number m_i of bays whose entry is of linear size M_i, are given by:

$$M_i = \frac{(3^i - 1)}{2} \; ; \qquad m_i = 4 \times 5^{G-1-i}. \tag{4.4}$$

For example, the $G = 4$ fractal has 100 bays that are only accessible to 1×1 particles, 20 bays accessible to particles 4×4 or smaller, and finally 4 bays accessible to particles 13×13 and smaller. One thus expects that the $D^*(M)$ vs M curves might change somewhat abruptly for square particles of sizes $M \approx 1, 4$, and 13.

The generation 4 cross-fractal

Figure 4.11 shows the exact diffusion coefficient D^* vs. the separation $Q = L - 81$ between the $G = 4$ fractals for three different particle sizes M. We first note that $D^*(Q < M) = 0$, as expected. For small separations Q, the diffusion coefficient is larger for smaller particles because it is the length scale Q that dominates the problem. For larger separations, however, the effect of Q becomes less important

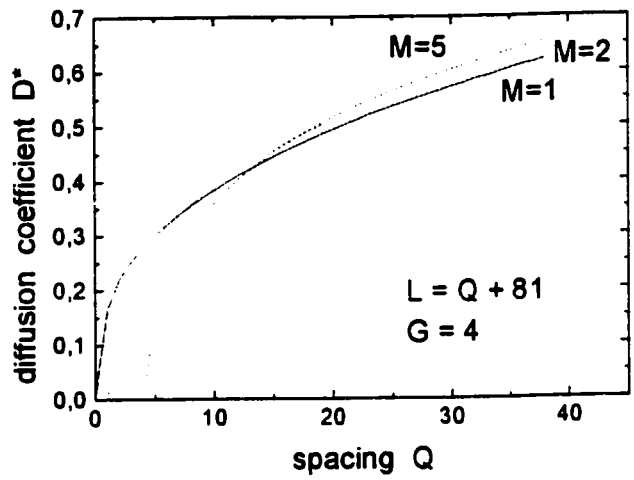

Figure 4.11: Diffusion coefficient D^* vs. free channel width Q for various particle sizes M. Note that the curves are very close to each other for $Q > 12$, and that they cross. A periodic sublattice of $G = 4$ fractal obstacles was used.

and the collisions with the fractal obstacles play a major role. The $M = 1$ and $M = 2$ particles possess the same value of D^* for a channel width $Q \approx 7$, while the $M = 2$ and $M = 5$ values are identical for $Q \approx 17$. The curves remain very close to one another for large separations Q.

Figure 4.12 shows a plot of D^* vs. particle size M for $Q = 16$. For comparison, we have also repeated the calculation for a simple cross-shaped obstacle of size 81×81. While $D^*(M > 16) = 0$ in both cases, the situation is quite different for other particle sizes M. Particles of size $M \leq 4$ can enter all bays of size 1 and 4, and hence are more retarded by the fractals than by the crosses. However, the $M \geq 5$ particles do not have access to the large number of smaller bays and display a higher diffusion coefficient in the fractal system. We also note that the $D^*(M)$ vs. M curve is not monotonic for a fractal system. Indeed, D^* increases from $M = 1$ to $M = 2$, and then stays essentially constant for $M = 2$ to $M = 5$, then decreases steadily towards zero, with a noticeable kink at $M = 14$. The sizes M where the changes in behaviour appear coincide nicely with the sizes $M_i = 1, 4, 13$ of the three

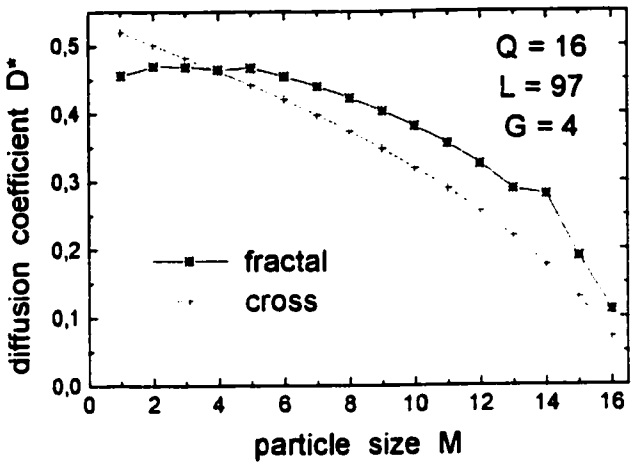

Figure 4.12: Diffusion coefficient $D^*(M)$ vs. particle size M. The system has a periodicity $L = 97$ with either $G = 4$ fractal obstacles ($*$) or 81×81 crosses ($+$).

different types of bays for this system. We thus see that the bays play a major role here. Table 4.5 gives some exact values of $D^*(M)$; note the almost perfectly flat $D^*(M)$ curve for $2 \leq M \leq 5$. Note also that $D^*(M = 2)$ and $D^*(M = 3)$ differ by less than a part in a 400. This small kink would have been almost impossible to identify with lattice Monte Carlo simulations.

A similar calculation for inter-fractal spacings $Q = 8$ (Fig. 4.13) shows that the fractal nature of the system becomes less important when Q is comparable to the size of the particles (clear fractal effects are found only for $M = 1$ and $M = 2$). When $Q = 38$, on the other hand, D^* varies very slowly with size M (by merely a few % while M increases by more than one order of magnitude). In the latter case, small jumps are observed for $M = 2, 5, 14$, as expected, and the curve is again non-monotonic. However, it must be noted that the flatness of the $D^*(M)$ vs. M curve is largely due to the very wide channels in this case. For instance, the exact results for a 81×81 cross-shapped obstacle as well as for a 81×81 square-shaped obstacle are comparable (Fig. 4.13). If the channel width Q was itself a random variable, the channels would sieve the particles and the diffusion coefficient $D^*(M)$

Table 4.5: Diffusion coefficient $D^*(M)$ for a few particle sizes M. The 4-th generation cross fractal system is described in Fig. 4.10; here we have Q=16.

M	D^*
1	0.4556519059
2	0.4698415722
3	0.4686156540
4	0.4646067627
5	0.4671258112
6	0.4540882956
8	0.4225820890
10	0.3807625836
12	0.3243734429
13	0.2879963663
14	0.2503859072
15	0.1891439434
16	0.1106087620
≥ 17	0

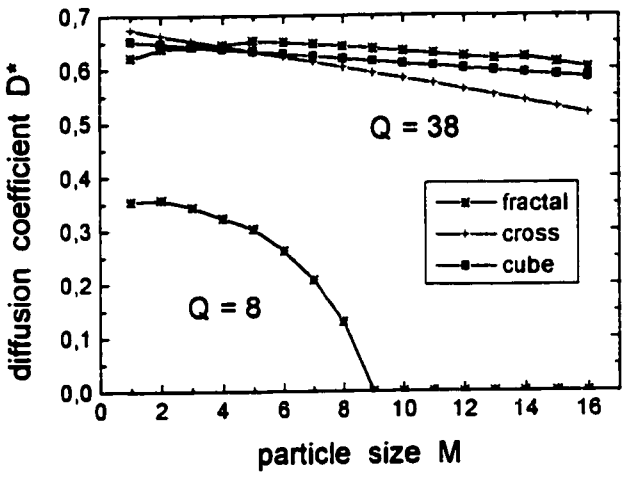

Figure 4.13: Diffusion coefficient $D^*(M)$ vs. particle size M for a $G = 4$ fractal system of periodicity $L = 81 + Q$ $(*)$, with $Q = 8$ and $Q = 38$. Also shown are the diffusion coefficients for a cross and a square both of dimension 81×81 with $Q = 38$.

would be a stronger function of M. This is, in our opinion, why Saxton[14] found that the diffusion coefficient was more dependent upon M for system with multiple fractals (MDLAs) than for a system with a unique fractal (CCAs). The contribution of the fractal nature of the obstacles is thus quite subtle and one requires extremely precise data to sort it out.

In conclusion, fractal structures do lead to qualitatively different results. Our exact data for the fourth generation cross fractals indicate that the particles having access to a large number of small bays show almost identical diffusion properties if the separation between fractals is large in comparison with the size of the particles. This is in agreement with Saxton's simulation results.

4.3.5 Triangular lattices

For some problems, a triangular lattice may actually be more appropriate than a square lattice. Our method can easily be modified for this type of lattice (or any

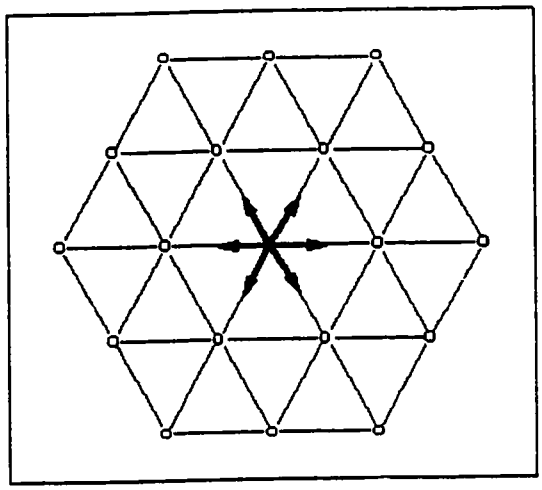

Figure 4.14: A particle on a triangular lattice has six directions to choose from.

other one embedded in a space of any spatial dimensionality $d > 1$; see for example Section 4.3.6). On a triangular lattice, a particle can move in six different directions at each step (Fig. 4.14). When a scaled external field ε is pointing in the positive x direction ($\rightarrow$), the probability for the particle to move in one of these six directions is (to first order in ε):

$$\begin{aligned} P_{\nearrow} = P_{\nwarrow} &\approx \frac{1+\varepsilon/2}{6}, \quad P_{\rightarrow} \approx \frac{1+\varepsilon}{6} \\ P_{\swarrow} = P_{\searrow} &\approx \frac{1-\varepsilon/2}{6}, \quad P_{\leftarrow} \approx \frac{1-\varepsilon}{6} \end{aligned} \tag{4.5}$$

where the arrows specify the direction. Note that, unlike for the square lattice, all the probabilities now depend on the field intensity ε. These relations allow us to build the transition matrix using the method described in the previous chapter. The mean velocity on each site for a triangular lattice is given by:

$$v_i = \frac{1}{2} \times (L_{\nearrow} P_{\nearrow} + 2L_{\rightarrow} P_{\rightarrow} + L_{\searrow} P_{\searrow} - L_{\swarrow} P_{\swarrow} - 2L_{\leftarrow} P_{\leftarrow} - L_{\nwarrow} P_{\nwarrow}) \tag{4.6}$$

where $L = 1$ if there is no obstacle in a given direction and zero otherwise. In absence of obstacles, the mean velocity is thus given by

Table 4.6: Diffusion coefficient $D^*(C)$ vs. obstacle concentration C for one-site particle diffusing on a triangular lattice. Saxton results are taken from ref.[14]. Our calculations used a 250×250 lattice and an ensemble of $\Omega = 250$ randomly generated obstacle systems, while Saxton used 256×256 lattices and $\Omega = 100$.

obstacle concentration C	D^* (this work)	D^* (Saxton results)
0.1	0.81822(7)	0.819
0.2	0.6279(2)	0.629
0.3	0.4248(3)	0.425
0.4	0.2076(4)	0.210

$$v_0 = \left(\frac{P_{\swarrow}}{2} + P_{\rightarrow} + \frac{P_{\searrow}}{2} - \frac{P_{\swarrow}}{2} - P_{\leftarrow} - \frac{P_{\nwarrow}}{2}\right) = \frac{\varepsilon}{2} \tag{4.7}$$

and the corresponding free-solution mobility by:

$$\mu_0 \equiv \frac{v_0}{\varepsilon} = \frac{1}{2}. \tag{4.8}$$

Once μ_0, the transition matrix and the local velocities are known, one can calculate the mobility and the diffusion coefficient in the presence of obstacles as described in Chapter 3. Some results obtained using a triangular lattice with randomly distributed single-site obstacles are shown in Table 4.6. Our results are in good agreement with the LMC results obtained by Saxton[15] (also shown in the Table) under similar conditions, but provide much improved accuracies. A fit of the data with a second-order polynomial gives $D^* = 1 - 1.74(1)C - 0.6(1)C^2$, which predicts a percolation concentration $C^* = 0.49(1)$, in good agreement with the exact result $C^* = 1/2$.

Table 4.7: Diffusion coefficient $D^*(C)$ for a few spatial dimensionalities d. For all these systems, the obstacle concentration is set at $C = 1/4096$. The linear size L of the lattice is also given. Note that $1 - C = 0.999755859...$

d	L	D^*
2	64	0.999477317 ...
3	16	0.999626225 ...
4	8	0.999671940 ...
6	4	0.999706986 ...
12	2	0.999755859 ...

4.3.6 Higher spatial dimensionalities

In order to demonstrate that our method can easily be generalized to any spatial dimensionality $d > 1$, we now investigate the case of a particle diffusing on a hypercubic lattice of dimensionality d in the presence of a periodic sub-lattice of linear size L composed of a single obstacle (see Fig. 4.1b for a $d = 2$ and $L = 3$ example), such that the concentration is $C = 1/L^d$.

Higher dimensions are treated with mathematical steps very similar to the ones described for the triangular lattice. First, we need to adjust the probability for the particles to make one of the many possible moves (with a d-dimensional hypercubic lattice, the particle can move in $2d$ different directions). Given those probabilities, we can build the transition matrix and the local velocity vector and calculate the free-solution mobility $\mu_0 = 1/d$. After that, the calculation of the mobility and the diffusion coefficient can be done as described in Chapter 3.

Using this procedure and the numerical implementation described in Section 4.2, we have calculated the scaled diffusion coefficient D^* of a particle in periodic systems of dimensionalities $d = 2, 3, 4, 6$ and 12. The obstacle concentration was set at $C = 1/4096 = (1/2)^{12}$ in all cases. Therefore, our system unit cells all had one obstacle and the volumes were $64^2, 16^3, 8^6, 4^6$ and 2^{12}, respectively. The results

are shown in Table 4.7. For the same obstacle concentration, the scaled diffusion coefficient D^* is larger for higher spatial dimensionalities. This can be explained by the smaller probability for the particle to hit the same obstacle many times in a row (the particle has more directions to move away from the obstacle in higher dimensionalities). Again, we should note that the variations in D^* are very small (0.03% between $d = 2$ and $d = 12$) and would have been nearly impossible to observe with lattice Monte Carlo simulations.

4.4 Conclusion

In Chapter 3, we presented a new method for calculating the EXACT scaled diffusion coefficient D^* for single particles diffusing in quenched lattice systems with PBCs. The method is based on the Nernst-Einstein relation and reduces the problem to the solution of a system of linear equations, thus offering an attractive alternative to lattice Monte Carlo simulations.

When these linear equations can be solved exactly, the exact diffusion coefficient is given in terms of a rational fraction. Such exact diffusion coefficients can be used to test theoretical models and analytical predictions, and to understand subtle geometrical effects. For instance, we obtained very precise values for the coefficient of the series expansion in the case of periodic cross shaped obstacles. However, the computer memory and the number of mathematical operations needed to obtain exact diffusion coefficients increase rapidly with the size of the lattice and, in practice, one is limited to systems of size 50×50 or less with our currently available computational resources.

If rational fractions are not required, it is possible to solve much larger systems (up to at least 250 000 equations on a standard workstation), while keeping both the memory needs and the CPU time at a reasonable level, by using special indexed storage methods developed for sparse matrices and iterative algorithms. In principle, this approach gives the diffusion coefficient D^* with an arbitrary precision. Moreover, it is substantially faster than exact calculations. We tested this

approach (specifically, the row-indexed compact storage and the biconjugate gradient algorithm) in our brief study of randomly distributed obstacles. We also used this numerical method to observe subtle effects related to the geometrical structure of deterministic cross fractal obstacles.

While one can easily obtain the exact diffusion coefficient D^* if the obstacles form a periodic pattern, the situation is quite different if there is any degree of randomness in the obstacle distribution. In these cases, exact values of D^* can only be calculated for specific realizations of the disorder; the true thermodynamic value of D^* is then an average over many realizations in the limit of infinite systems.

Our method can easily be adapted to any type of lattice of dimensionality $d > 1$. As an example, we studied the triangular lattice and easily obtained results that were one to two orders of magnitude more precise than what was previously obtained by Saxton using similar conditions. We also showed the effect of spatial dimensionality for a low concentration case on a hypercubic lattice.

Since our method can provide arbitrary exact diffusion coefficients for two-dimensional systems of sizes larger than 256×256 within a few minutes of CPU on a standard workstation, we believe it can easily replace lattice Monte Carlo simulations for a wide range of applications. Chapter 5 presents such an application.

Chapter 5

A three-dimensional model of gel electrophoresis

In this chapter, we use the lattice model developed in Chapter 3 to study the low-field electrophoretic sieving of particles in periodic three-dimensional gels made of either isolated obstacles or infinitely long fibers[1]. Exact mobilities are calculated using the numerical method described in Chapter 4. A comparison between the exact mobilities and the free available volumes indicates that the main assumption of the Ogston-Morris-Rodbard-Chrambach (OMRC)[2,6–7] model, which postulates that the mobility (μ) of charged particles is directly related to the fractional gel volume available to them, is not valid. However, a study of the gel concentration and analyte size dependence of the zero-field mobility indicates that the OMRC model and the so-called Ferguson plots ($\ln(\mu)$ vs. gel concentration C) can indeed be used to obtain useful, semi-quantitative information about the gel properties.

5.1 Introduction

The low field electrophoretic mobility μ of a particle (the analyte) in a gel matrix of (volume/volume) concentration C is determined by the average electric charge of the

[1]This work has been published in *Electrophoresis*[26]

particle and the sieving effect of the gel. The latter may include excluded volume and constrictive effects, percolation constraints, hydrodynamic interactions (one of which is the electro-osmotic flow), and even specific analyte-gel interactions (attractive or repulsive). Our understanding of the relationship between the resolution achieved by electrophoresis and the structure, concentration and composition of the sieving matrix is still incomplete. The standard OMRC model assumes that the low-field mobility μ of a rigid particle in a gel is given by:

$$\frac{\mu(C)}{\mu_0} \equiv \mu^*(C) = f(C), \tag{5.1}$$

where μ_0 is the free-solution mobility, $\mu^*(C)$ is the scaled (or relative) mobility, and $f(C)$ is the fractional gel volume available to the molecule. In other words, the OMRC model uses $f(C)$ to model the approximate excluded volume (or crowding) effects. This model thus predicts that if a molecule has access to the same free-volume $f(C)$ in two gels of otherwise entirely different architectures, it will have the same mobility μ^* in both. This leaves little room for gel design and cannot be used to understand the subtle effects of the gel architecture. Other non-excluded volume effects are not specifically described by this model.

Slater and Guo[9] have demonstrated that the OMRC model is actually equivalent to a mean-field model where the gel constraints are spread uniformly in space and time (no memory effect). In other words, this model neglects the existence of percolating pathways as well as the correlations between consecutive analyte-gel fibre collisions. Mean-field models are very useful starting points in many different scientific investigations, but detailed models are necessary in order to elucidate the physical mechanisms responsible for the more subtle effects.

The lattice model developed in Chapter 3 can be applied to gel electrophoresis and allows one to study arbitrary gel structures in $d \geq 2$ dimensions and to compute exact mobilities μ^* and free available volumes $f(C)$. It has been used in a series of papers[9–11] (published by the group of Gary Slater over the last several years) whose aim was to develop a more microscopic model of gel electrophoresis. In Part I[9] of the series, two-dimensional periodic, random and fuzzy gels were studied, and it was shown that the curvature of the Ferguson plot was related to the degree of

disorder in these gels. Part II[10] studied sieving in two-dimensional periodic gels and showed that a modified Ferguson plot can indeed provide useful semi-quantitative information about the gel properties (such as a realistic value of the mean pore size). Part III[11] looked at sieving and percolation in random two-dimensional gels and demonstrated that certain gel architectures have very special sieving properties. The more efficient numerical approach described in Chapter 3[20] allows one to study much larger (hence, more realistic) systems, such as three-dimentional gels.

In this Chapter, we thus present a study of two types of periodic three-dimensional gels. In Section 5.2, the exact mobility is calculated for a three-dimensional periodic lattices containing isolated obstacles while the next section presents a detailed study of a more realistic three-dimensional gel made of infinitely long fibers.

5.2 Periodic isolated obstacles

Let us first study the gel system shown in Fig. 5.1 with a field pointing in the $+x$ direction. The isolated obstacles of volume $1 \times 1 \times 1 = 1$ form a cubic lattice of periodicity $s > 1$. We can treat this problem exactly using a box of size $s \times s \times s$, a single obstacle at position $\{1, 1, 1\}$, and periodic boundary conditions.

The case of an analyte of linear size $R = 1$ (or mass $M = R^3 = 1$) is quite interesting. In this case, we have $S = s^3 - 1$ available sites. With the field pointing in the $+x$ direction, the mean velocity on each site is given by $v(i) = p_{+x}L_+(i) - p_{-x}L_-(i)$, where the displacements $L = 1$ if there is no obstacle in the given direction, and zero otherwise. Thus $v(\{2, 1, 1\}) = (1+\varepsilon)/6$, $v(\{0, 1, 1\}) = (-1+\varepsilon)/6$, and $v(i) = \varepsilon/3$ for all the other sites. As we can see, $v_I(i) = 0$ everywhere, except for the two sites next to the obstacle in the field direction. Furthermore, $v_\varepsilon(i) = 1/6$ for these two special sites, and $v_\varepsilon(i) = 1/3$ for the $S - 2 = s^3 - 3$ other sites. According to Eq.(3.15), we then need only the probabilities $n_\varepsilon(i)$ for the two sites next to the obstacle in order to compute the exact reduced mobility μ^*. Let us write these probability factors as $n_\varepsilon(\{0, 1, 1\}) = -n_\varepsilon(\{2, 1, 1\}) = P_0 > 0$, where we used the symmetry of the lattice to reduce the problem to a single parameter P_0. The scaled mobility is then given

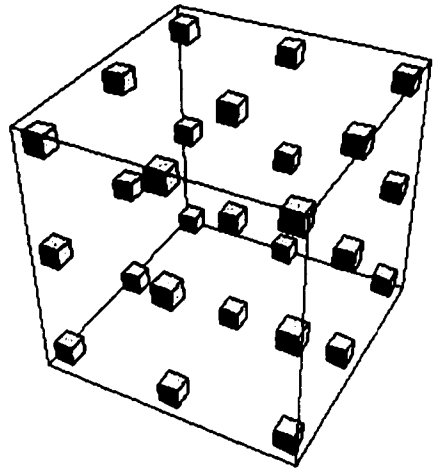

Figure 5.1: A simple three-dimensional periodic gel made of single obstacles repeated periodically in the three dimensions.

by (using $\mu_0 = 1/3$):

$$\mu^* = 1 - \frac{1}{S} - P_0 = 1 - C - \left(\frac{C^2}{1-C} + P_0\right), \tag{5.2}$$

where $C = 1/(S+1)$ is the concentration of obstacles. We note that $1 - C$ is the OMRC model free fractional available volume $f(C)$ for this system. Since $P_0(C) > 0$ and $0 < C < 1$, the last term of this expression is always negative, and hence the exact electrophoretic mobility is always smaller than the value $f(C) = 1 - C$ predicted by the OMRC model. The last term in Eq.(5.2) contains the non-mean field effects; more precisely, it takes into account both the fact that a particle can collide many consecutive times with the same obstacle and the precise architecture of the gel.

Table 5.1 shows some exact electrophoretic mobilities for gel periodicities between $s = 50$ and $s = 2$, which correspond to concentrations between $C = 0.0008\%$ and $C = 12.5\%$, respectively. The results are shown in Fig. 5.2. Fitting the low concentration points with a high order polynomial, we obtain:

$$\mu^* = 1 - 1.53114515C + 0.6482254C^2 + 67.974C^3 + 1377C^4 \ldots . \tag{5.3}$$

Table 5.1: Exact reduced mobilities μ^* for selected values of the obstacle concentration C. The analyte is a cube of linear size $R = 1$ and mass $M = R^3 = 1$. The gel is made of isolated obstacles (of size $1 \times 1 \times 1$) which are distributed periodically on a cubic sub-lattice of parameter s, such that $C = 1/s^3$ (see Fig. 5.1)

s	C	μ^*
2	1/8	6/7
3	1/27	837/884
4	1/64	7936/8127
5	1/125	67125/67952
6	1/216	6039576/6082565
10	1/1000	0.99846957
20	1/8000	0.99980861
30	1/27000	0.99994329
50	1/125000	0.99998911

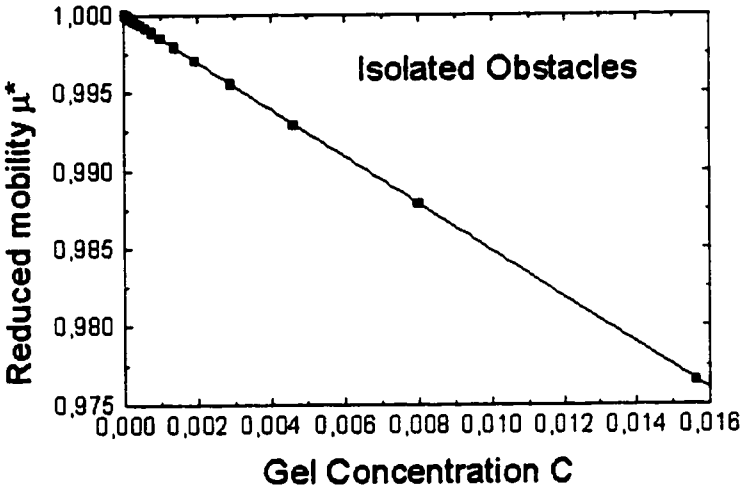

Figure 5.2: Reduced mobility μ^* vs. gel concentration $C = 1/s^3$ for the "gel" shown in Fig. 5.1. The solid line represents the best-fit curve provided by Eq.(5.3). Only concentrations $C \leq 1/64$ are shown.

The plot shows an upward curvature and no inflection point for allowed concentrations $0 \leq C \leq 1/8$. The OMRC model prediction of $\mu^* = 1 - C$ is clearly not appropriate, which demonstrates that this conclusion, first reached by Slater and Guo[9] when studying two-dimensional systems, is not an artifact of the dimensionality of space. Although this model system (Fig. 5.1) is not a good approximation of a real gel, it does represent an interesting generalization of previous work[9–11] on two-dimensional periodic systems. As far as we know, nobody has generalized Nieuwenhuizen's work[23] to calculate analytically the coefficient a_1 for this case.

5.3 A more realistic periodic gel

5.3.1 Small analytes

We now investigate the more realistic three-dimensional gel shown in Fig. 5.3. The fibres have a unit diameter, are infinitely long, and are orthogonal to each other. If

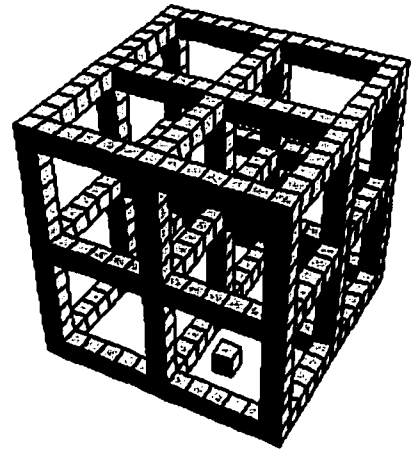

Figure 5.3: A three-dimensional periodic gel with $s = 6$. The $1 \times 1 \times 1$ particle moves between infinitely long fibres oriented along the x, y and z axis of the system. Three fibers share a common unit block at the crosslink location.

the periodicity is s along each axis, we can obtain exact mobilities using a system of size s^3 with periodic boundary conditions. The concentration of obstacles is given by $C = (3s - 2)/s^3$. Obviously, the pore size (diameter) is given roughly by the (constant) fibre spacing s. Note that this gel structure, unlike the one shown in Fig. 5.1, can sustain its own weight in the presence of gravity. We believe that a good model system should always have this property because this constraint will affect the percolation properties of the sieve, and hence the separation one can achieve with it.

For a $R = 1$ and $M = 1$ analyte, the fractional available volume is simply $f(C) = S/s^3 = (s^3 - 3s + 2)/s^3 = 1 - C$. Using the approach described in Section 5.2, we then obtain:

$$\mu^* = 1 - \frac{2s-1}{S} + \frac{\langle n_\varepsilon | v_I \rangle}{1/3} = 1 - C - (g(C) + 3\langle n_\varepsilon | v_I \rangle)), \tag{5.4}$$

where $g(C) = (2s-1)/(s^3(1-C)) - C$, and $s = s(C)$ is determined by the relation $C = (3s-2)/s^3$. Again, the OMRC model fails to predict the correct mobility of the analyte in the gel since the last term of Eq. 5.4 is non-zero.

Table 5.2: Exact reduced mobilities μ^* for selected values of the obstacle concentration C and analyte size R. The gel is made of infinitely long fibres (of unit diameter) oriented along the x, y and z axis of the system (see Fig. 5.3).

s	C	$\mu^*(R=1)$	$\mu^*(R=2)$	$\mu^*(R=5)$
2	1/2	1/2	0	0
3	7/27	27/40	3/7	0
4	5/32	679/864	5/9	0
5	13/125	.85202459	.65906499	0
6	2/27	.89275772	.73590260	3/8
10	7/250	.95882810	.88658339	.60485775
20	29/4000	.98940550	.96917940	.85902134
30	11/3375	.99526464	.98607598	.93209156
50	37/31250	.99829011	.99494491	.97443212

5.3.2 Mobility vs. gel concentration and particle size

Table 5.2 gives some selected exact electrophoretic mobilities μ^* for different gel periodicities s and particle sizes R. The mobility is plotted vs. gel concentration C in Fig. 5.4, and vs. R^2 in Fig. 5.5. An excellent description of the dependence of the mobility μ^* upon gel concentration C is given by the polynomial form:

$$\mu^* = 1 + \sum a_i(R) C^i, \tag{5.5}$$

where the coefficients $a_i(R)$ are functions of the particle size R. Table 5.2 lists the value of the ratios a_1/R^2 and a_2/R^4 as a function of R. The size dependence of the coefficients $a_i(R)$ indicates that these periodic structures can be used to sieve particles according to size. This can also be observed in Fig. 5.6. It is interesting to note that the asymptotic dependence $a_i(R) \sim R^{2i}$, which has been previously established for two-dimensional systems[10], is also valid in three dimensions. This is consistent with Ogston's original work[5], and it also explains why the Ferguson

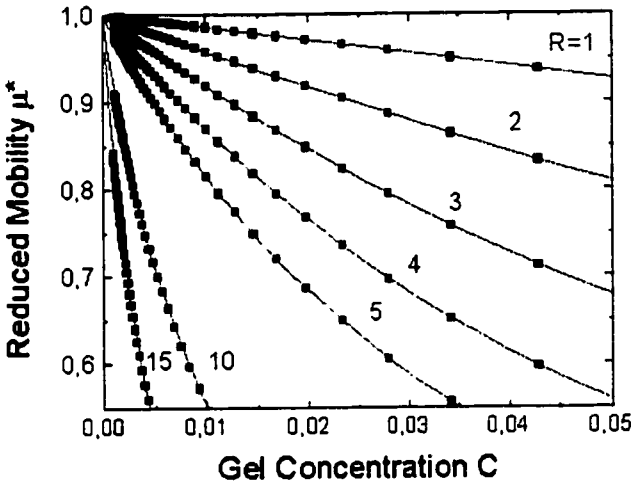

Figure 5.4: Reduced mobility μ^* vs. gel concentration $C = (3s-2)/s^3$ for different particle size R migrating in a periodic gel made of infinitely long fibres oriented along the x, y and z axis of the system. The solid lines show the best polynomial fits suggested by Eq.(5.5). The coefficients a_i of these fits are given in Table 5.3.

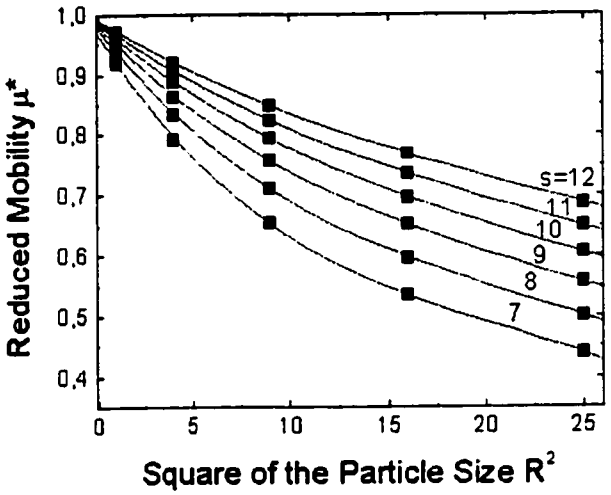

Figure 5.5: Reduced mobility μ^* vs. the square of the particle size R^2 for various gel periodicities s (note that the gel concentration is given by $C = (3s-2)/s^3$). The solid lines show the best polynomial fits.

plot approach provides meaningful information (see next section).

5.3.3 Ferguson plots and gel parameters

Following the work of Slater and Guo[10], the data is now analysed in terms of the standard OMRC model method based mostly on the so-called Ferguson plot. In the framework of the OMRC model, the free available volume $f(C)$, and hence the reduced mobility $\mu^*(C)$, is actually a function of the dimensionless exclusion parameter $\xi = [(R+r)/a]^2$, where R, r and $a(C)$ are the diameter of the analyte, gel fibre and gel pore, respectively. For example, Ogston obtained $\mu^* = exp(-\xi)$ for a random array of non-crosslinked fibres. We first notice that the coefficients $a_i(R)$ given in Table 5.3 scale like R^{2i} when $R \gg r = 1$ (Fig. 5.6), an indication that ξ is indeed the relevant dimensionless parameter in our problem too. We now define the effective retardation coefficient:

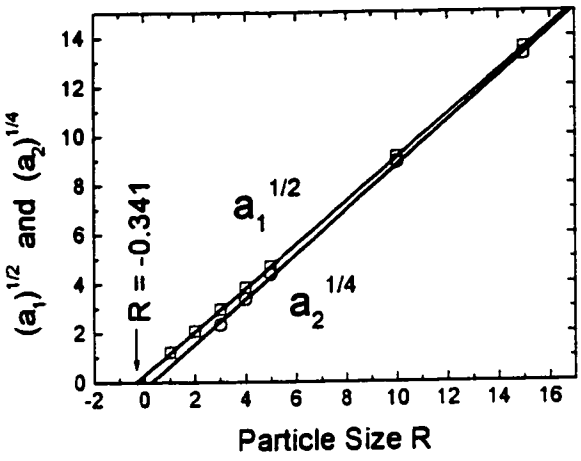

Figure 5.6: Coefficients $(-a_1)^{1/2}$ and $a_2^{1/4}$ as a function of the particle size R. The lines show linear fits.

$$K(R) = -\lim_{C->0} \frac{\partial \mu^*(C)}{\partial C} = -a_1(R). \tag{5.6}$$

A plot of $\sqrt{K} = \sqrt{-a_1}$ vs. particle size R yields a straight line, as was seen in Fig. 5.6. This is the type of plot normally used by experimentalists to find the effective fibre size r_K[27]: the value of R where $K(R) = 0$ (found by extrapolating $K^{1/2}$ to negative values of R) gives $-r_K$, which is an estimate of the fibre size r. Here we obtain $r_K = 0.341$ while $r = 1$.

An effective mean pore size a_K can also be estimated using the OMRC model procedure described in Ref. [10]. Briefly, the idea here is to extrapolate the $\mu^* = 1 - KC$ line (i.e., the best linear fit for low gel concentrations) to find the critical concentration $C = C^* = 1/K$ at which $\mu^* = 0$. At that point, the effective pore size $a_K(C^* = 1/K)$ is necessarily equal to the particle size R (the particle is trapped in the cage-like pore). Since we already know that $K^{1/2} \sim R$ for large particles, we conclude that $a_K(C^* = 1/K \sim 1/R^2) \approx R \sim 1/\sqrt{(C^*)}$, in agreement with the Ogston model. Therefore, except for very small analytes, the procedure normally used in the framework of the OMRC model does actually lead to a realistic estimate

Table 5.3: Coefficients $a_i(R)$ (with $i = 1$ and 2) of the polynomial fits given by Eq.(5.5).

R	$-a_1(R)/R^2$	$a_2(R)/R^4$
1	1.4421	-3.9
2	1.0682	-0.15
3	0.9583	0.36
4	0.9083	0.51
5	0.8813	0.60
10	0.8301	0.63
15	0.8128	0.61

of the pore size for these three-dimensional model gels.

5.4 Discussion

As we have shown in this chapter, the algebraic simplifications and the numerical approach that uses sparse storage techniques and iterative solving algorithms described respectively in Chapters 3 and 4 make three-dimensional systems tractable. Two types of periodic three-dimensional gels have been studied, and our results are qualitatively similar to those found previously for two-dimensional periodic gels. The OMRC model prediction that $\mu^* = f(C) = 1 - C$ was shown to be wrong for both the isolated obstacle and infinitely long fibre gels. In the former case, we could actually demonstrate analytically that the OMRC model systematically overestimates the true electrophoretic mobility of the particles. Despite this, like in two-dimensions, a modified Ferguson plot was found to provide a realistic estimate of the pore size and fibre radius.

For two-dimensional periodic gels made of isolated obstacles, previous work[23] showed that $a_1(R = 1) = -(\pi - 1)$. In this chapter, we showed that for the analogous three-dimensional gel, the first coefficient was $a_1 = -1.5311...$ We can explain

the stronger concentration dependence for the two-dimensional case by the fact that consecutive collisions with the same obstacle are more probable in lower dimensional spaces. As the dimension d increases, the correlation between the obstacles (related to higher order terms in the series expansion) and the probability of multiple collisions with the same obstacle (related to the first order term, or a_1) become less important. In fact the mean-field OMRC model prediction $\mu^* = 1 - C$ should be correct in the limit where $d \to \infty$.

The next logical step in the development of a more microscopic model of gel electrophoresis is the study of "random" three-dimensional gels. However, it is not clear what this actually means. A simple generalization of the two-dimensional random gel, where obstacles of unit size are distributed randomly on the lattice, would lead to an unphysical gel that could not sustain its own weight in the presence of gravity. More realistic gels can be obtained using various approaches. For example, a random three-dimensional gel can be generated using the cluster-cluster aggregation (CCA) technique. In this approach, randomly distributed obstacles perform random walks and stick to each other when colliding; the process continues until only one aggregate (the "gel") remains. This technique produces large connected aggregates, and with periodic boundary conditions, the gel may be quite dilute indeed. It is also possible to use multiple fixed points as seeds for the aggregation process.

A three-dimensional random gel can also be generated using a simple random walk. Starting from an arbitrary site on the lattice, a random walk is performed and an obstacle is created at each step. This procedure will lead to a very long and tortuous fibre that can be seen as a complex gel with periodic boundary conditions (the points where the random walk crossed itself act like crosslinks). The method can be modified to allow a certain number of obstacles to start new branches. Yet another way to produce realistic three-dimensional gels is to randomly place rods (or even planes) of a given size, along the x, y and z directions until a connected aggregate forms. Finally, one can combine one or more of these approaches to form more complicated structures.

Chapter 6

Limitations and possible extensions

Our approach has been developed to study the long time behavior of a single isolated particle migrating in a system with immobile obstacles. In this chapter, some of the limits and unexplored possibilities of the method will be briefly examined.

6.1 Short time dynamics

Unlike lattice Monte Carlo simulations, our method does not provide any information about the short time dynamics of the particle. However, LMC simulations are not the only way to obtain this type of information; other theoretical methods can be used. For example, in the exact enumeration method[28–29], one site is assigned a unit probability of occupancy at $t = 0$, and at each time step the probability density is propagated to nearest-neighbor unobstructed sites. This procedure reduces the noise dramatically in comparison with lattice Monte Carlo simulations. It is also possible to reduce the problem to the numerical analysis of the eigenvalues of the transition matrix[30–32]. This allows one to find very precise values for the dynamical exponents, but does not normally give information about the steady-state regime (the diffusion coefficient D^*). Combining one of those methods with our approach

could provide a global understanding (of both the transient and the steady-state regime) of the dynamics unmatched by current LMC simulations.

6.2 Local binding energies

Until now, we have limited our study to the hard core (steric) interactions between the analytes and the gel fibers. However, our method can be extended to study other types of interactions. For example, we could study a model where the analyte sticks to the gel fiber for a while before continuing its migration. The principle is quite simple: the probability of not moving (p_0) during the next time step is not zero when the particle touches an obstacle. For a two-dimensional square lattice system we then have (with the field pointing in the $+x$ direction):

$$\begin{aligned} p_{\pm x} &= \frac{(1-W)e^{\pm\varepsilon}}{2+e^{\varepsilon}+e^{-\varepsilon}} = \frac{(1-W)(1\pm\varepsilon)}{4} \\ p_{\pm y} &= \frac{1-W}{2+e^{\varepsilon}+e^{-\varepsilon}} = \frac{1-W}{4} \\ p_0 &= W \end{aligned} \tag{6.1}$$

where W in the "sticking" parameter. When the particle does not touch any obstacle, the jumping probabilities are identical to the non-sticky case, namely:

$$\begin{aligned} p_{\pm x} &= \frac{e^{\pm\varepsilon}}{2+e^{\varepsilon}+e^{-\varepsilon}} = \frac{1\pm\varepsilon}{4} \\ p_{\pm y} &= \frac{1}{2+e^{\varepsilon}+e^{-\varepsilon}} = \frac{1}{4} \\ p_0 &= 0 \end{aligned} \tag{6.2}$$

Note that if the particle is touching more than one obstacle, the probability for not moving may be considered to increase proportionally.

Using those equations, we can build the transition matrix and solve the linear system exactly like in the non-sticky case (Section 3.3). Preliminary results (J. Labrie, J.-F. Mercier and G.-W. Slater, manuscript in preparation) indicate that such interactions lead to Ferguson plots with inflection points; such plots have been reported experimentally but have remained without an explanation. The relative

importance of geometry and direct interactions in gel electrophoresis can now be studied using a well-characterized model system.

6.3 Complex particles and finite concentrations

While our approach can easily match lattice Monte Carlo simulations when a single particle diffuses via simple translations, the situation is different when the particle can rotate, is flexible (e.g., a polymer), or when many particles are involved. The main difficulty comes from the fact that it is not possible to define a unique local probability on each site for such cases. This section will investigate this issue.

6.3.1 Finite concentrations

If two particles are diffusing at the same time, the definition of a local probability of presence of the particle on each site must be changed. It is now necessary to define the "joint probability" for the two particles. In other words, we must define the probability $n(i,j)$ for the first particle to be on site i and for the other one to be on site j. Note that $n(i,j) = n(j,i)$ if the two particles are identical. Therefore, the linear system will have $\approx J^2/2$ unknowns, where J is the number of available sites (for comparison, the number of unknowns for a single particle was J).

As one could expect, the situation would be even more complicated if we had more diffusing particles. It is easy to show that the number of unknowns (and equations) will be $\approx J^{N_p}/N_p!$, where N_p is the number of diffusing particles. The complexity of the problem thus increases very quickly with the number of particles involved in the problem. Therefore, our approach can only be used to study very small and simple problems.

6.3.2 Rotating particles

Another interesting case for which our current approach fails is the rotating particle. When the particle is not symmetric (e.g., a rigid rod), it can "migrate" not only

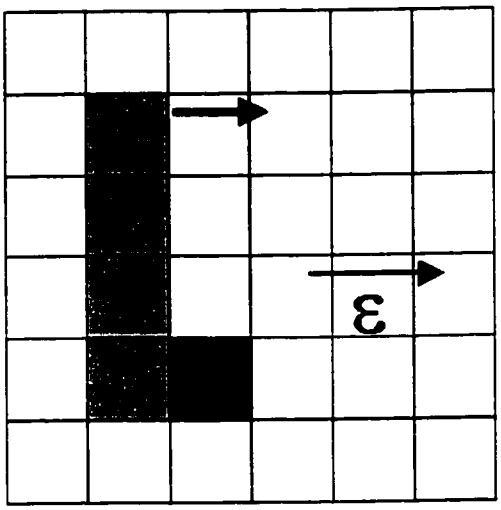

Figure 6.1: One end of the migrating rod (in gray) touches an obstacle (in dark). When a field is applied, the net force makes the rod rotate.

by translation but also by rotation. Like the previous case, the definition of the local probability must be changed when rotations are allowed: we need to define the probability $n(i,j)$ for the particle to be on site i with a given orientation j. Unlike the case with multiple particles, the complexity of the problem will not increase drastically. The number of unknowns, and equations, will be $\approx J \times N_o$, where N_o is the number of possible orientations of the particle (typically 2 or 4 on a two-dimensional square lattice). Therefore, our approach should, in principle, provide good results.

Until now, unfortunately, we have been unable to define satisfying rotation probabilities. To be realistic, such probabilities should depend on the field intensity since a particle having one of its end in contact with an obstacle should naturally rotate in the direction of the field in response to the torque (Fig. 6.1 shows this effect for a rod).

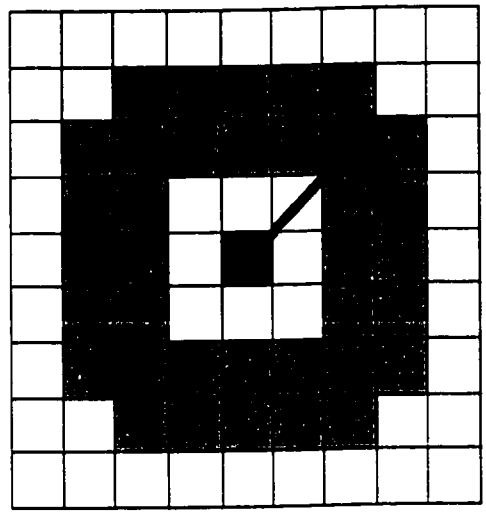

Figure 6.2: The dimer model on a square lattice using the bond fluctuation algorithm. If the first monomer is fixed at the center of the cell, the other one can only be located on one of the 40 grey-shaded sites.

6.3.3 Small polymers (oligomers)

On a square lattice, a dimer can be seen as two particles (the monomers) attached to each other by an elastic link. Many algorithms can be used to describe this system. One of the most useful is the bond fluctuation algorithm, where the centers of the two monomers cannot be closer than 2 and farther apart than $\sqrt{13}$. Any move leading to a conformation violating these rules is rejected. Figure 6.2 shows the sites accessible to one monomer when the other one is fixed at a certain site. In this situation, the monomer could choose from 4 different jumps.

Like for the many particle case, we need to define a local "joint probability" $n(i,j)$ as the probability for the first monomer to be on the site i and for the other one to be on site j. However, if the first particle is on a site i, the second one must be rather close because of the link constraints. Therefore, the linear system will have $\approx J \times N_C/2$ unknowns, where N_C is the number of possible conformations of the dimer ($N_C = 40$ in the bond fluctuation of the square lattice). Of course, the situation will be more complicated for longer polymers.

As we can see, analyzing complex particles or multiple particles with our approach greatly increases the mathematical complexity of the problem. However, for some very simple problems, such as a small polymer in a narrow tube, our approach can still be used in order to obtain exact scaled mobilities or diffusion coefficients.

Chapter 7

Conclusion

This thesis presented an algebraically exact method to calculate the scaled diffusion coefficient D^* (or mobility μ^*) of a probe particle in a system with immobile obstacles and periodic boundary conditions. Our starting point is the fact that it is possible to relate the scaled diffusion coefficient D^* of a particle to the zero-field scaled mobility μ^* with the Nernst-Einstein relation. In fact, our approach reduces the diffusion problem to the solution of a (large) system of linear equations which can be greatly simplified in the zero-field limit. In effect, we are solving the LMC problem exactly, thus eliminating the need for a stochastic simulation method. Exact results can be used to understand subtle geometrical effects or to test theoretical models.

When these linear equations can be solved exactly using a scientific software (such as Mathematica or Maple), the diffusion coefficient is given in terms of a rational fraction. If such rational fractions are not required, it is possible to solve much larger systems, while keeping both the memory needs and the CPU time at a reasonable level, by using special indexed storage methods and iterative algorithms. This approach can give the diffusion coefficient D^* (or mobility μ^*) with an arbitrary precision and is substantially faster. In fact, this approach allows us to treat cases similar (in size and complexity) to the ones currently treated by lattice Monte Carlo simulations, without requiring more CPU resources, with, as a bonus, at least 7 extra decimal places. We tested this approach (specifically, the row-indexed compact storage and the biconjugate gradient algorithm) in our brief study of randomly

distributed obstacles. We also used this numerical method to observe the subtle geometrical effects of deterministic cross fractal obstacles (such effects would have been difficult to see using LMC simulation) and in our study of periodic three-dimensional gel electrophoresis.

For a periodic pattern of obstacles, one can easily obtain the exact diffusion coefficient D^* (or the mobility μ^*). The situation is not the same if there is any randomness in the distribution of obstacles. In these cases, the true thermodynamic value of D^* (or μ^*) is an average over many realizations in the limit of infinite systems. We have shown how one can treat systems containing closed volumes and/or multiple independent pathways (this situation may occur when the obstacles are set randomly on the lattice).

The algebraic simplifications and our numerical approach have made three-dimensional systems tractable. As an example, two types of periodic electrophoresis gels, the isolated obstacles and the infinitely long fibers, have been studied in details. Our results are qualitatively similar to those found previously for two-dimensional periodic gels. In particular, the OMRC model's prediction that $\mu^* = f(C) = 1 - C$ was shown to be wrong for both gels and was even shown to systematically overestimates the true electrophoretic mobility of the particles in the isolated obstacle case. The next step in the study of gel electrophoresis will be to treat more realistic gels where the fibers are not straight and have random directions and crosslinks. Again, we believe that our approach will be far more successful for studying such systems than lattice Monte Carlo simulations. This work is in progress.

Our numerical method does not allow the study of the short-time dynamics of the migrating particle and does not allow (until now) particle rotation. However, it is quite flexible and can easily be adapted to various problems of increasing complexity.

First, it can be adapted to any type of lattice of dimensionality $d > 1$. As an example, we generalized the method to higher dimensionality in a study of a low obstacle concentration on hypercubic lattices. We also studied a random distribution of obstacles on a triangular lattice and easily obtained results that were one to two orders of magnitude more precise than what was previously obtained by LMC simulations with similar conditions.

It is also possible to have non-similar bonds (non-uniform jumping probabilities). For example, it would be easy to study the case where the jumping probabilities are set at random for each site. We could also investigate cases where the particle has a finite probability to make arbitrarily long jumps or to overlap with an obstacle. We could also include local binding energies, or any other effect transforming the jumping probabilities. In fact, the only limitation is that the probability of jumping has to be time-independent and has to sum to unity for each site.

Furthermore, the diffusing particle can take various sizes and shapes without increasing the problem's mathematical complexity. We also showed how we can use our method to treat small flexible dimers or trimers and even finite concentrations (many simultaneously diffusing particles) for simple systems.

Clearly our approach is very versatile and efficient. We believe that it could easily replace lattice Monte Carlo simulations for a wide range of applications, especially those requiring great precision. Also, with the constant increase in computing power, and the democratization of parallel computing (powerful softwares already exist to solve sparse linear systems on parallel computers), the complexity of the systems which would be possible to treat exactly will continue to increase, making this method more and more attractive to researchers.

Appendix A

The mobility with the Metropolis algorithm

In this Appendix, we show that the Metropolis algorithm and the Glauber approach give the same results for μ^* in the $\varepsilon \to 0$ limit for a two-dimensional square lattice.

We start with Eq.(3.15) for the scaled mobility μ^*:

$$\mu^* = \frac{\langle v_I | n_\varepsilon \rangle + \langle v_\varepsilon | n_I \rangle}{\mu_0}. \tag{A.1}$$

We will get through each of those terms and see how they change when using the Metropolis algorithm instead of the Glauber approach.

A.1 The free mobility

In the Metropolis algorithm, the jumping probabilities (for a two-dimensional square system) are:

$$p_{+yM} = p_{-yM} = p_{+xM} = \frac{1}{4}, \quad \text{and} \quad p_{-xM} = \frac{e^{-\varepsilon}}{4}. \tag{A.2}$$

Thus the free mobility is given by:

$$\mu_{0M} = \frac{v}{\varepsilon} = \frac{p_{+xM} - p_{-xM}}{\varepsilon} = \frac{1 - e^{-\varepsilon}}{4\varepsilon} \approx \frac{\varepsilon}{4\varepsilon} = \frac{1}{4} = \frac{\mu_{0G}}{2}, \tag{A.3}$$

where the subscripts M and G denote the values obtained using the Metropolis and the Glauber approach, respectively (recall that $\mu_0 = 1/2$ in the Glauber aproach).

A.2 The velocity

The velocity of a particle is defined as: $v(i) = v_I(i) + \varepsilon v_\varepsilon(i) = p_{+x}L_+(i) - p_{-x}L_-(i)$, where the displacements $L = 1$ if there is no obstacle in the given direction, and zero otherwise. Since the jumping probabilities are identical when $\varepsilon = 0$, the field independent term ($\langle v_I|$) will be identical for the two approaches ($\langle v_{IM}| = \langle v_{IG}| = \langle v_I|$) , namely:

$$\begin{aligned} v_I(i) &= \tfrac{1}{4} \text{ if there is an obstacle in the direction opposite to the field,} \\ &= -\tfrac{1}{4} \text{ if there is an obstacle in the field direction,} \\ &= 0 \text{ if these two sites are both empty or both occupied.} \end{aligned} \tag{A.4}$$

The field correction will depend on the method used; we have:

$$\begin{aligned} v_{\varepsilon M}(i) &= 0 \text{ if there is an obstacle in the direction opposite to the field,} \\ &= \tfrac{1}{4} \text{ if there is an obstacle in the field direction,} \\ &= \tfrac{1}{4} \text{ if these two sites are empty,} \\ &= 0 \text{ if these two sites are both occupied by obstacles,} \end{aligned} \tag{A.5}$$

when using the Metropolis algorithm and

$$\begin{aligned} v_{\varepsilon G}(i) &= \tfrac{1}{4} \text{ if there is an obstacle in the direction opposite to the field,} \\ &= \tfrac{1}{4} \text{ if there is an obstacle in the field direction,} \\ &= \tfrac{1}{2} \text{ if these two sites are empty,} \\ &= 0 \text{ if these two sites are both occupied by obstacles,} \end{aligned} \tag{A.6}$$

when using the Glauber approach.

A.3 The local probabilities

Again, because both approaches have the same jumping probabilities when $\varepsilon = 0$, the field-independent term of the local probability ($|n_I\rangle$) will be identical for the two methods, giving:

$$n_{IM}(i) = n_{IG}(i) = 1/J, \qquad i = 1, 2, ..., J, \tag{A.7}$$

where J is the number of available sites.

The field correction term $|n_\varepsilon\rangle$ to the local probability is found by solving the linear system (3.13):

$$A_I|n_\varepsilon\rangle = -A_\varepsilon|n_I\rangle. \tag{A.8}$$

The A_I matrix will be the same for both methods (again because of the identical jumping probabilities when $\varepsilon = 0$). Since $n_I(i) = 1/J$ for all sites, the $A_\varepsilon|n_I\rangle$ term is in fact the sum of all the first order (ε) terms in the probability rate equations. Again we can break the problem into simple cases. Using the Metropolis approach we have:

$$\begin{aligned} A_\varepsilon|n_I\rangle_M(i) &= -\tfrac{1}{J} && \text{if there is an obstacle in the direction opposite to the field,} \\ &= \tfrac{1}{J} && \text{if there is an obstacle in the field direction,} \\ &= 0 && \text{if these two sites are both empty or both occupied} \end{aligned} \tag{A.9}$$

and with the Glauber approach:

$$\begin{aligned} A_\varepsilon|n_I\rangle_G(i) &= -\tfrac{2}{J} && \text{if there is an obstacle in the direction opposite to the field,} \\ &= \tfrac{2}{J} && \text{if there is an obstacle in the field direction,} \\ &= 0 && \text{if these two sites are both empty or both occupied.} \end{aligned} \tag{A.10}$$

Thus:

$$A_\varepsilon|n_I\rangle_M(i) = \frac{A_\varepsilon|n_I\rangle_G(i)}{2}. \tag{A.11}$$

Substituting Eq.(A.11) in Eq.(A.8) we finally get:

$$|n_\varepsilon\rangle_M = \frac{|n_\varepsilon\rangle_G}{2}. \tag{A.12}$$

A.4 The mobilities

Since for each site facing an obstacle in the field direction there must be another one with an obstacle behind it, we have (using Eqs.(A.5), (A.6) and (A.7)):

$$\langle v_\varepsilon|n_I\rangle_M = \frac{\langle v_\varepsilon|n_I\rangle_G}{2}. \tag{A.13}$$

Combining Eqs.(A.4), (A.11) and (A.13) into Eq(A.1) we get:

$$\mu_M = \frac{1}{2}\left(\langle v_{IG}|n_{\varepsilon G}\rangle + \langle v_{\varepsilon G}|n_{IG}\rangle\right) = \frac{\mu_G}{2}. \tag{A.14}$$

Substituting this result in Eq.(A.13) and using Eq.(A.3) we finally obtain:

$$\mu_M^* = \frac{\mu_M}{\mu_{0M}} = \mu_G^*. \tag{A.15}$$

Thus in the $\varepsilon \to 0$ limit, the Metropolis algorithm and the Glauber approach give exactly the same results.

Appendix B

An exact first-order term

In general, the concentration (C) dependence of the mobility μ^* (or diffusion coefficient D^*) can be expressed as a series expansion:

$$\mu^* = D^* = 1 + a_1 C + a_2 C^2 + a_3 C^3 + a_4 C^4 \ldots \,. \tag{B.1}$$

The standard procedure to obtain such series expansions is to calculate μ^* or D^* for a given (large) number of obstacle concentrations and to fit the data to obtain the parameters a_i. This procedure has been used in Chapter 4 for the study of the cross shaped obstacles and random two-dimensional gels, and in Chapter 5 for the study of gel electrophoresis in three-dimensional gels.

While this procedure is robust and works rather well, it is sometimes possible to analytically find an exact first-order coefficient for very simple systems. As an example, the case of an infinitely long tube with periodic obstacles is analyzed in this Appendix.

B.1 The infinite tube of width two

Consider the system shown in Fig. B.1. It is an infinite tube of diameter $\Phi = 2$ (scaled units), with periodic obstacles separated by a distance L. The system can be considered as a unique cell of length L with periodic boundary conditions.

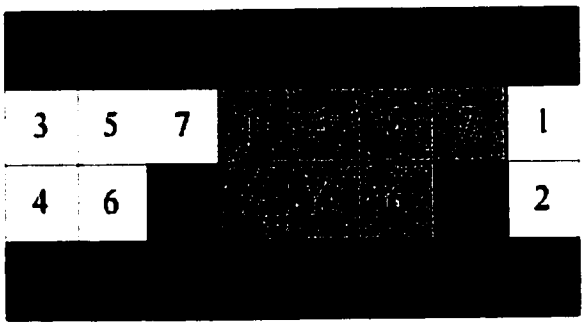

Figure B.1: The system studied in this section: an infinitely long tube of width 2 containing periodically distributed obstacles. Here the obstacles are separated by a distance $L = 4$. Therefore, there are only 7 independent sites in the system (the primitive cell is shown in gray).

Following the procedure described in Chapter 3 to obtain the reduced diffusion coefficient of a 1×1 plaquette, we need to solve the following equation:

$$A_I|n_\varepsilon\rangle = -A_\varepsilon|n_I\rangle$$

In this case, the matrix equation reads

$$\frac{1}{4}\begin{pmatrix} -3 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & -2 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & -3 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & -3 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -3 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & -2 & 0 \\ 4 & 4 & 4 & 4 & 4 & 4 & 4 \end{pmatrix} \begin{pmatrix} n_\varepsilon(1) \\ n_\varepsilon(2) \\ n_\varepsilon(3) \\ n_\varepsilon(4) \\ n_\varepsilon(5) \\ n_\varepsilon(6) \\ n_\varepsilon(7) \end{pmatrix} = \begin{pmatrix} 0 \\ -1/2 \\ 0 \\ 0 \\ 0 \\ 1/2 \\ 0 \end{pmatrix}, \tag{B.2}$$

where the last line is the normalization condition.

B.1.1 Symmetry considerations

The effect of a field is to increase the probability for the particle to be on the sites blocking the migration in the field direction and to decrease the probability for the

particle to be on the sites which are shadowed by the obstacle. In our example (the field is in the $+x$ ($\rightarrow$) direction), this means that the field will increase the probability for the particle to be on sites #5 and 6, and decrease the probability for the particle to be on sites #1 and 2. Because of the symmetry of the system, the field correction $n_\varepsilon(i)$ for a site situated to the right of the obstacle will be equal, but of opposite sign, to the field correction for a site situated at the same distance to the left. The symmetry also imposes that there is no field correction term for the sites which are half way between two obstacles or between the obstacle and the wall across the tube width above or below an obstacle. Thus for our example we have:

$$\begin{aligned} n_\varepsilon(1) &= -n_\varepsilon(5), \\ n_\varepsilon(2) &= -n_\varepsilon(6), \\ n_\varepsilon(3) &= 0, \\ n_\varepsilon(4) &= 0, \\ n_\varepsilon(7) &= 0. \end{aligned} \tag{B.3}$$

Using those simplifications, the system shown in Fig. B.1 has only two unknowns, namely $n_\varepsilon(5)$ and $n_\varepsilon(6)$. Substituting Eq.(B.3) into Eq.(B.2), we get the much simpler system:

$$\frac{1}{4}\begin{pmatrix} -3 & 1 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} n_\varepsilon(3) \\ n_\varepsilon(4) \end{pmatrix} = \begin{pmatrix} 0 \\ -1/2 \end{pmatrix}. \tag{B.4}$$

which defines the transition matrix A_4 ($A_4|n\rangle = |b\rangle$).

B.1.2 Iterative solutions

Redoing the same calculation for a system of size $L = 6$, we find:

$$A_6 = \frac{1}{4}\begin{pmatrix} -3 & 1 & 1 & 0 \\ 1 & -3 & 0 & 1 \\ 1 & 0 & -3 & 1 \\ 0 & 1 & 1 & -2 \end{pmatrix}. \tag{B.5}$$

For $L = 8$, we have:

$$A_8 = \frac{1}{4}\begin{pmatrix} -3 & 1 & 1 & 0 & 0 & 0 \\ 1 & -3 & 0 & 1 & 0 & 0 \\ 1 & 0 & -3 & 1 & 1 & 0 \\ 0 & 1 & 1 & -3 & 0 & 1 \\ 0 & 0 & 1 & 0 & -3 & 1 \\ 0 & 0 & 0 & 1 & 1 & -2 \end{pmatrix}, \tag{B.6}$$

and for $L = 10$:

$$A_{10} = \frac{1}{4}\begin{pmatrix} -3 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -3 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -3 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & -3 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -3 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & -3 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & -2 \end{pmatrix}. \tag{B.7}$$

The matrix dimension increases by two at each iteration (this is due to the fact that we add two new independent sites each time). Except for its last line, matrix A_{j-1} is entirely contained in matrix A_j. Thus when computing the diffusion coefficient for a larger system, it is possible to use the solutions found for smaller systems. Note also that the transition matrix A is band diagonal and that the lines are all very similar. In fact, the matrix has only two "types" of line, one that has 1 0 left to the diagonal term and one that has 1 1.

The last step to find the mobility μ^*, or diffusion coefficient D^*, is to multiply the probability field correction $|n_\epsilon\rangle$ by the field independent velocity $|v_I\rangle$:

$$\mu^* = \frac{\langle v_I|n_\epsilon\rangle + \langle v_\epsilon|n_I\rangle}{\mu_0}. \tag{B.8}$$

However the components of $\langle v_I|$ will all be zero except for the sites next to an obstacle. In a tube system, the only such site is the last independent one (site #6 in the tube $L = 4$). Therefore we don't need to completely solve the linear system of equations; once A_j has been reduced to a triangular matrix, the problem is solved. In the triangular form, the matrix can be reduced to:

$$A = \begin{pmatrix} \ddots & d_{i-1} & 1 & 0 & 0 & 0 \\ 0 & a_i & b_i & 1 & 0 & 0 \\ 0 & 0 & c_i & d_i & 1 & 0 \\ 0 & 0 & 0 & a_{i+1} & b_{i+1} & 1 \\ 0 & 0 & 0 & 0 & c_{i+1} & d_{i+1} \\ 0 & 0 & 0 & 0 & 0 & \ddots \end{pmatrix}. \tag{B.9}$$

We can then introduce four parameters (a_n, b_n, c_n, d_n), which completely define the matrix. It is easily shown that:

$$\begin{aligned} a_{i+1} &= -3 - \frac{1}{a_i} + \frac{d_i b_i}{a_i c_i}, \\ b_{i+1} &= 1 + \frac{b_i}{a_i c_i}, \\ c_{i+1} &= -3 - \frac{1}{c_i} + \frac{b_{i+1}(c_i - d_i)}{a_{i+1} c_i}, \\ d_{i+1} &= 1 + \frac{d_i - c_i}{c_i a_{i+1}} \end{aligned} \tag{B.10}$$

with initial conditions given by:

$$a_1 = -3,$$

$$
\begin{aligned}
b_1 &= 1, \\
c_1 &= -3 + 1/3, \\
d_1 &= 1/3.
\end{aligned}
\tag{B.11}
$$

With these parameters, we have an iterative solution providing the diffusion coefficient (or the mobility) for an increasing system size, with:

$$
\begin{aligned}
L_i &= 2(i+1)+2, \\
\mu_i &= \frac{2L_{i+1} + \frac{2}{1+c_{i+1}} - 1}{2L_{i+1} - 1},
\end{aligned}
\tag{B.12}
$$

where L_i is the length of the tube.

Note that this derivation was done for even system sizes ($L = 4, 6, 8, ...$) but the results are similar for system sizes which are odd.

B.1.3 Series expansion

To find the mobility (or diffusion coefficient) for an infinitely long tube with only one obstacle, we simply find the fixed point of the map (the iterative solution). To do so, we solve Eq.(B.10) for $a_{i+1} = a_i = a$, $b_{i+1} = b_i = b$, $c_{i+1} = c_i = c$ and $d_{i+1} = d_i = d$:

$$
\begin{aligned}
a &= -3 - \frac{1}{a} + \frac{db}{ac}, \\
b &= 1 + \frac{b}{ac}, \\
c &= -3 - \frac{1}{c} + \frac{b(c-d)}{ac}, \\
d &= 1 + \frac{d-c}{ac}
\end{aligned}
\tag{B.13}
$$

and substitute the results into the asymptotic relation

$$
\mu(L \to \infty) = 1 - \frac{2}{(2L-1)(1+c)}.
\tag{B.14}
$$

Solving Eqs.(B.13), we find:

$$\mu(L \to \infty) = 1 - \frac{(1+2\sqrt{3})C}{1-C} \cong 1 - (1+2\sqrt{3})C \cong 1 - (4.4641015...)C, \quad \text{(B.15)}$$

where $C = 1/(2L)$ is the obstacle concentration. This results is fully compatible with the numerical value obtained previously by Josée Labrie, $a_1 = -4.464102...$ [Honours project, University of Ottawa, 1998].

B.2 Wider tubes

In principle, we could use the same procedure for larger diameter tubes. The iterative solution would contain more parameters, and its complexity would increase rapidly (for instance, the number of parameters is $\propto \Phi^2$, where Φ is the tube diameter).

Until now, it has been impossible to find the fixed point algebraically, using Maple or Mathematica, for systems larger than $\Phi = 2$. We should also mention that this approach only gives the first order coefficient (a_1).

Bibliography

1. R.K. Pathria, *Statistical Mechanics 2nd ed.* (Butterworth Heinemann, Oxford, 1996).

2. C. J. O. R. Morris, *Protides of the Biological Fluids* 14th Colloquium , 543 (1967).

3. J. C. Giddings and R. Boyack, *Analytical Chemistry* 7, 1229 (1964) .

4. D. Rodbard and A. Chrambach, *Proc. Nat. Acad. Sci. USA.* 65, 970 (1970).

5. A. G. Ogston, *Trans. Faraday Soc.* 54, 1754 (1958).

6. K. A. Ferguson, *Metabolism* 13, 985 (1964).

7. D. Tietz, A. Chrambach, M. J. Dunn and B. J. Radola, *Advances in Electrophoresis* 2, 109 (1988).

8. G. W. Slater and H.L. Guo, *Electrophoresis* 16, 11 (1995).

9. G. W. Slater and H. L. Guo, *Electrophoresis* 17, 977 (1996).

10. G. W. Slater and H.L. Guo, *Electrophoresis* 17, 1407 (1996).

11. G. W. Slater and J. R. Treurniet, *J. of Chromatogr. A.* 772, 39 (1997).

12. G. W. Slater, J. Rousseau and J. Noolandi, *Biopolymers* 26, 863 (1987).

13. D. W. Heermann, *Computer simulation methods, 2nd ed.* (Springer-Verlag, 1990).

14. M. J. Saxton, *Biophys. J.* 64, 1053 (1993).

15. M. J. Saxton, *Biophys. J.* 64, 1766 (1993).

16. M. J. Saxton, *Biophys. J.* 66, 394 (1994).

17. M. J. Saxton, *Biophys. J.* 69, 389 (1995).

18. M. J. Saxton, *Biophys. J.* 70, 1250 (1996).

19. M. J. Saxton, *Biophys. J.* 72, 1744 (1997).

20. J. F. Mercier, G. W. Slater and Guo H. L. *J. Chem. Phys.* 110, 6057 (1999).

21. J. F. Mercier and G. W. Slater *J. Chem. Phys.* 110, 6050 (1999).

22. W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. T. Flannery, *Numerical recipes in Fortran 77: the art of scientific computing 2nd ed.* (Cambridge university press, Cambridge, 1992).

23. Th. M. Nieuwenhuizen, P. F. J. Van Velthoven and M. H. Ernst, *Phys. Rev. Lett.* 57, 2477 (1986).

24. G. D. L. Phillies and A. Quinlan, *Macromolecules* 25, 3110 (1992).

25. D. Stauffer and A. Aharony, *Introduction to percolation theory 2nd ed.* (Taylor & Francis, London, 1992).

26. J. F. Mercier and G. W. Slater *Electrophoresis* 19, 1560 (1998)

27. P. Serwer, *Electrophoresis* 4, 375 (1983).

28. S. Havlin and D. Ben-Avraham, *Adv. Phys.* 36, 695 (1987).

29. S. Havlin, S., G. H. Weiss, J. E. Kiefer and M. Dishon, *J. Phys. A.* 17, L347 (1984).

30. D. J. Jacobs and H. Nakanishi, *Phys. Rev. A.* 41, 706 (1990).

31. N. H. Fuchs and H. Nakanishi, *Phys. Rev. A.* 43, 1721 (1991).

32. H. Nakanishi, S. Mukherjee and N. H. Fuchs, *Phys. Rev. E.* 47, R1463 (1993).