THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

On boundary value problems for intracellular subdiffusion and signaling pathways, and for geometric flows

Tobias Gebäck

CHALMERS | GÖTEBORG UNIVERSITY

Department of Mathematical Sciences
Division of Mathematics
Chalmers University of Technology and Göteborg University
Göteborg, Sweden, 2006

On boundary value problems for intracellular subdiffusion
and signaling pathways, and for geometric flows
Tobias Gebäck
ISBN 978-91-7291-883-2

Doktorsavhandlingar vid
Chalmers Tekniska Högskola
Ny serie Nr 2564
ISSN 0346-718X

Department of Mathematical Sciences
Division of Mathematics
Chalmers University of Technology and Göteborg University
SE-412 96 Göteborg
Sweden
Telephone +46 (0)31 772 1000

Printed at the Department of Mathematical Sciences
Göteborg, Sweden, 2006

On boundary value problems for intracellular subdiffusion and signaling pathways, and for geometric flows

Tobias Gebäck

Department of Mathematical Sciences
Chalmers University of Technology and Göteborg University

Abstract

The main part of this thesis concerns mathematical models for diffusion of proteins inside cells, including reactions between the proteins. Initially, such models are applied to describe signaling pathways in yeast cells, and the properties of the model are studied, especially in relation to models that do not include diffusion. The results show that it is sometimes necessary to include diffusion in the model to capture important aspects of the biological system.

The thesis also contains work on the numerical methods used to compute solutions to the reaction-diffusion equations inside domains. Specifically, the Immersed Interface Method, which allows efficient numerical solution inside arbitrary domains using uniform rectangular grids, is applied on Boolean grids, which give the same accuracy as uniform grids while using fewer grid points.

The third part of the thesis also concerns models for protein diffusion inside cells, but now describing the phenomenon subdiffusion (or anomalous diffusion), which has been observed inside cells and manifests itself as a qualitatively different, and slower, diffusion behaviour of proteins. The cause of this phenomenon is the crowdedness of the interior of the cell, where other proteins and larger structures interfere with the motion of the proteins. In the thesis, a new mathematical model for anomalous diffusion in the form of a parabolic pseudo-differential equation is proposed, and a proof of existence of solutions for boundary value problems representing anomalous diffusion inside a cell is given. Experiments using Fluorescence Correlation Spectroscopy which support the model have also been performed.

Finally, the thesis contains a convergence result for a computational scheme for approximation of mean curvature flows inside a domain, that is the description of the motion of surfaces which move at each point with a velocity depending on the mean curvature at that point. The scheme allows a quite general dependence on the curvature and concerns the case when the moving surface is inside a domain and intersects the domain boundary at a right angle.

Keywords: boundary value problems, diffusion-reaction equations, signaling pathways, anomalous diffusion, pseudo-differential equations, fluorescence correlation spectroscopy, immersed interface method, Boolean grids, mean curvature flow

Contents

This thesis contains the following parts and papers:

Acknowledgements

First of all, I wish to thank my supervisor Alexei Heintz for his guidance during my years as a Ph.D. student. His ideas and support have been invaluable during the work on this thesis.

I also thank my co-supervisor Per Sunnerhagen for sharing his time and knowledge on the biological parts of this thesis, and Eva Asp for the good company during the warm spring days we spent in a dark cellar with the FCS microscope and for all the work that made Part III of this thesis possible. Thanks also to Claes Molin who was my collaborator during the first years.

I am very grateful to Peter Kumlin and Mohammad Asadzadeh for reading through different parts of the thesis and giving valuable comments on the contents and the presentation.

I am also grateful to the *National research school in genomics and bioinformatics*, and in particular its coordinator Anders Blomberg, for providing funding for my Ph.D. position and for giving a valuable education in cell biology and bioinformatics.

I also wish to thank the people working at the Department of Mathematical Sciences, and in particular the kinetic theory group, for making my time here enjoyable. A special thanks to my fellow Ph.D. students, and in particular the floorball and football enthusiasts who have helped keeping me in shape while working here.

Further thanks go to my friends in the guitar ensemble I Poetens Trädgård and in Haga Motettkör for giving me the opportunity to relax while indulging in music.

Finally, thanks to Carin and my family for your constant support and encouragement.

Göteborg, December 2006
Tobias Gebäck

Preface

This thesis contains four parts, as listed on the contents page. The first three parts come in a logical order describing the progress of my Ph.D. project, which was financed by the *National research school in genomics and bioinformatics* under the title *Spatial dynamic modeling of an intracellular signaling pathway.*

This is also the topic for Part I, where systems of reaction-diffusion partial differential equations are studied as models for proteins involved in signaling pathways. These equations show some interesting spatial effects which are not present in the ordinary differential equation models that are used more often in this context.

Part II describes work done on the numerical method used to solve the equations in Part I, namely the Immersed Interface Method (IIM). In this part, the IIM is implemented on Boolean grids, grids that allow us to use fewer grid nodes while retaining the same accuracy in the computations, with the result that we can make faster computations, using less memory.

In Part III, we investigate the phenomenon called subdiffusion, both experimentally using the Fluorescence Correlation Spectroscopy (FCS) method, and theoretically and numerically using a parabolic pseudo-differential equation as a model for subdiffusion. We show that this model fits the experimental data, and in the attached paper (Paper I) we prove existence and uniqueness of solutions to boundary value problems for the equation, corresponding to subdiffusion for proteins inside a cell.

The work presented in Part IV was initiated in my Master's thesis, and concerns approximation of mean curvature flows. In the paper (Paper II) contained in this part, I have proved convergence for a computational scheme for approximating mean curvature flows, for the case when the flow takes place inside a domain, with right-angle boundary conditions on the domain boundary, and the velocity is a non-linear function of the mean curvature of the moving surface.

Each of the parts contains an abstract and an introduction which describes the work in more detail.

Since this thesis may attract readers with different background knowledge, let me indicate what kind of background is needed to read the different parts. Part I and the first half of Part III are oriented towards biology, and much of it should be understood by anyone with some basic knowledge in cell biology and mathematics. Part II is purely mathematical, but not very difficult, while the two papers in Part III and Part IV require a proper background in mathematics to be fully understood.

Spatial modeling of MAPK signaling pathways

Tobias Gebäck

Abstract

MAPK signaling pathways in eukaryotic cells are sequences of enzymatic reactions that convey a signal from the cellular membrane to the nucleus, in response to some stimulus. They constitute an often vital signaling system which enables the cell to react to changes in the environment and to survive such changes.

We investigate the effects of spatial models for signaling pathways. First, we add a diffusion term to the Kholodenko model for general MAPK signaling pathways. This has a large effect on the amplitude of the oscillations that the model predicts, indicating that such oscillations have a smaller effect in a model that takes into account the spatial distribution of proteins, compared to the original space-independent model.

Second, we investigate a simple spatial model for the HOG pathway in the yeast *Saccharomyces cerevisiae*. We are able to reproduce the nuclear relocation of the Hog1 protein and also see that diffusion in the model is so fast that differences in protein concentration throughout the cell are small, even though reactions are localized only at the membrane.

The calculations are performed in three space dimensions using finite differences and the Immersed Interface Method, which is described in part II of this thesis.

Contents

Glossary

amino acid the 20 different building blocks that make up proteins

differential equation an equation involving derivatives, whose solution is a function of one or more variables

enzyme protein that speeds up specific reactions in the cell

eukaryotic cell a cell containing a nucleus, as opposed to bacteria

gene a DNA sequence coding for a protein

genome the collection of all the genes of an organism, coded for by DNA

HOG High Osmolarity Glycerol

in vitro experiment performed in an artificial environment, outside the organism

in vivo experiment performed inside a living organism

kinase enzyme whose function is to phosphorylate other enzymes

MAPK Mitogen Activated Protein Kinase

ODE Ordinary Differential Equation; a differential equation in one variable, often time

osmosis the process that strives to even out the concentration of solutes across a membrane

PDE Partial Differential Equation; a differential equation in several variables, such as space and time

phosphorylation the addition of a phosphate group (PO_3) to a protein

1. Introduction

The ordinary bakers' yeast *Saccharomyces cerevisiae*, which is a unicellular fungus, is a very widely studied organism among cell biologists. The reasons for this are many. One is that for a long time there has been a commercial interest for brewers and bakers to understand the organism in order to maximize its output of alcohol and carbon dioxide. Another reason is that it is a relatively simple unicellular eukaryotic organism that is easy to handle in the lab and can be used as a model organism for higher eukaryotic organisms, such as plant and mammal cells. Nowadays, another reason for studying yeast is that it is already very well studied, which means that more extensive studies can be performed, trying to understand more complex processes in the cell. For example, the complete yeast genome has been sequenced and it contains approximately 6.300 genes (cf. human genome approx. 30.000 genes) [1]. Furthermore a complete library of gene deletions has been set up, i.e. for (almost) every single gene, there is a yeast strain available that has this particular gene deleted from its genome, enabling biologists to easily study the effects of removing a gene from a cell under different conditions, thus hopefully learning more about the function of that gene. Also, there are many research groups continually working on different aspects of the yeast cell, making the amount of data available comparatively large.

Having said that, however, it should be noted that even the simple yeast cell is not at all understood by the biologists. There are many genes coding for proteins with unknown function and even if the gene codes for a protein that has a known function, this function may depend on other proteins and substances, so that the overall behavior is not very well understood anyway. The cell as a whole is a very complicated system, where proteins, DNA, RNA, lipids and other molecules work together to define the behavior of the cell. And although there has been a tremendous increase in knowledge about the cell during the last decades, only small parts of the complete system are well understood. For example, a single protein coded for by a single gene may be studied to determine its amino acid sequence, its three-dimensional structure, its active sites, where it may bind to other proteins, etc. This gives very valuable information about the protein, but does not tell the whole story, since questions like "When is it expressed from the DNA?", "What activates/deactivates the protein?", "Where is it located?" and so on, must also be answered to give a complete picture. The answers to this kind of questions do not depend solely on the protein itself, but also on other proteins and molecules in the cell, as well as outside stimuli and the overall "state" of the cell.

The complexity of these questions is the basis for *systems biology*, which is the research area that tries to look at larger systems of proteins and cellular functions, often using mathematical modeling in order to understand the behavior of that particular system. There is no single definition of systems biology and it is not very fruitful to try to come up with one, since these "systems" may be very different in character and the methods applied to study them may also vary accordingly. The philosophy of systems biology is not without controversies, since there is no long tradition of using mathematical modeling in cell biology, but there are a few examples where modeling has been successful as a complement to the experimental data in order to understand the behavior of a cellular

system (see [8], [12]).

1.1. SIGNALING PATHWAYS

One type of cellular system that is suited for mathematical modeling is signaling pathways. We will be concerned here with MAPK[1] pathways. Specifically, we have the High Osmolarity Glycerol (HOG) pathway in yeast in mind, but most of what is said here applies to other pathways as well. See [1, chapter 15] for general information about signaling pathways and [4] for a review of the HOG pathway. The signaling in a MAPK pathway starts at the cellular membrane, where it is activated by some stimulus, such as the presence of a specific substance (e.g. pheromones) in the environment or a more general environmental change, such as change in osmotic pressure or oxidative properties of the environment.

This stimulus is sensed in one way or another by receptors or other mechanisms at the membrane. These sensing mechanisms then convey a signal to another protein by phosphorylation (i.e. adding a phosphate group to one of the amino acids of the target). This starts a chain of phosphorylation events, which convey the signal through two or three steps, where each step consists of the phosphorylation of a target kinase, that is an enzyme which, once activated by phosphorylation, may phosphorylate other target proteins. So, as is seen in figure 2.1, the signaling cascade moves from the MAP-kinase-kinase-kinase (MKKK) which activates the MAP-kinase-kinase (MKK), which activates the MAP-kinase (MAPK), which in turn moves to the nucleus where it may activate or deactivate transcription factors that control gene expression. The MAPK may also have other functions by controlling the activity of enzymes throughout the cytoplasm and nucleus. Activation of kinases in the chain may require a double phosphorylation of two amino acids in the protein, which are both performed by the higher-level kinase.

The effect of the signaling pathway is that the cell is able to sense changes in the environment and convey the information of this change to the nucleus or other inner parts. There the cell can produce the appropriate response to the stimulus, which is often vital for the survival of the cell. With the multiple steps in the chain, the cell is able to amplify the signal and may also increase the steepness of the response, creating a switch-like response so that the signal is more or less either "on" or "off" [5]. In addition to the activating kinases, there are also deactivating phosphatases, which remove the phosphate from the enzymes, thereby deactivating them (a dephosphorylation need not be deactivating, but in this case it is). Thus, when the stimulus disappears, or the cell has adapted to the new environment, the signaling pathway switches back off through the action of the phosphatases.

It should also be mentioned that (as always in biology) things are more complicated than they seem. For example, the osmosensing mechanism is often not very well known and may include many proteins; the phosphorylation events may take place when the kinases are organized in scaffolds or large protein complexes; and the cell is full of other enzymes and molecules which may influence the signaling pathway, producing different results depending on which state the cell is in. The models we discuss here focus on

[1] Mitogen Activated Protein Kinase

the phosphorylation cascade and the movement of the phosphorylated MAPK from the membrane to the nucleus. They are of course great simplifications, but may hopefully provide some insight into the reality.

1.2. Outline

In the following sections, we study two spatial models of MAPK signaling pathways. In section 2, we investigate the effects of diffusion on the oscillations predicted by a model of a MAPK cascade, including a negative feedback loop. In section 3, we study a model of the nuclear relocalization of the yeast MAPK Hog1. The two sections are almost independent and contain separate results and discussion parts.

2. The Kholodenko model with diffusion

2.1. Introduction

In the year 2000, Boris N. Kholodenko [6] published a model of general mitogen-activated protein kinase (MAPK) signaling cascades, which was shown to give rise to oscillatory behavior for a range of parameter values. An essential feature of the model is a negative feedback loop, meaning that the end product of the pathway inhibits the activating reaction (see figure 2.1). The model received some attention, since the appearance of oscillations was a rather unexpected effect which could have interesting implications for the biology of the cells. However, no oscillations have actually been observed for signaling pathways in real cells. There could be numerous reasons for this, for example that the abundance of proteins is measured as the total protein content in a large number of cells, which means that unsynchronized oscillations will not show up in measurements. But it is perhaps more probable that the model does not agree well enough with reality, meaning that the oscillations in MAPK-pathways are artifacts which do not occur in nature. Specifically, the feedback loop in the model may be a too simple model of the feedback that is known to be present (since the signaling is turned off after a while). It is also a fact that the signaling does not take place at one point in space but involves movement of proteins through the cell. Some of the reactions take place only at the cell membrane, while some may take place wherever the participating proteins encounter each other. This should have an effect on the oscillatory behavior, since as proteins at the end of the reaction chain move away from the membrane, they cannot take part in the feedback, which should then be attenuated. The purpose of the following sections is to incorporate diffusion of proteins into the model and study the effect that this has on the oscillations and the behavior of the pathway. Although the Kholodenko model may not be an accurate model, it is quite convenient to study an oscillating system, since the oscillations are easily seen and effects on them are easily detected.

2.2. The Kholodenko model

The structure of the Kholodenko model for MAPK signaling pathways is shown in figure 2.1. The pathway is activated by a stimulus of some kind, which causes the MKKK to be phosphorylated. This activates the MKKK, which then in turn is able to phosphorylate the MKK at two different sites. The active double-phosphorylated MKK (MKKPP) may then activate the MAPK by phosphorylation, again at two sites. In a real cell, the MAPKPP continues to perform some action, which sooner or later will turn the signaling cascade off, by some means which are to a large extent unknown (and may differ between different MAPK pathways). In the Kholodenko model, this is modeled by that the MAPKPP inhibits the phosphorylation of the MKKK. This is the negative feedback loop, which is needed in order to create oscillatory behavior.

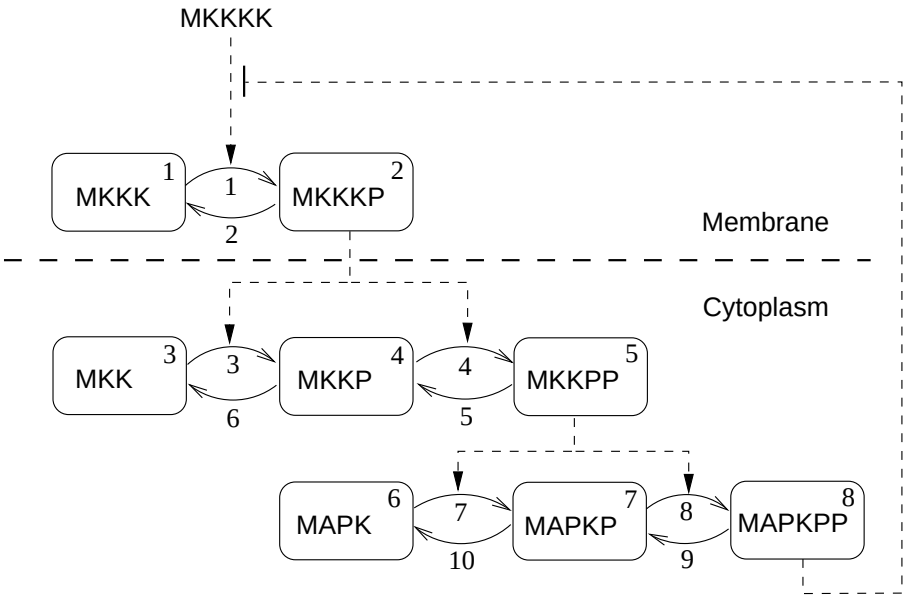

Figure 2.1: The Kholodenko model. The signaling pathway is activated by some stimulus which leads to phosphorylation of the MKKK, which in turn phosphorylates the MKK in two steps. The double-phosphorylated form of MKK then phosphorylates the MAPK, also in two steps. The end product (MAPKPP) then has the effect of inhibiting the stimulus, thus forming a negative feedback loop. Here, we also add the additional assumption that the MKKK is fixed at the cell membrane, while the other proteins are free to diffuse through the cytoplasm.

The corresponding ordinary differential equations may be written as

$$\begin{aligned}
du_1/dt &= v_2 - v_1 \\
du_2/dt &= v_1 - v_2 \\
du_3/dt &= v_6 - v_3 \\
du_4/dt &= v_3 + v_5 - v_4 - v_6 \\
du_5/dt &= v_4 - v_5 \\
du_6/dt &= v_{10} - v_7 \\
du_7/dt &= v_7 + v_9 - v_8 - v_{10} \\
du_8/dt &= v_8 - v_9
\end{aligned}$$

where the concentrations of MKKK through MAPKPP are denoted $u_j,\ j = 1, \ldots, 8$ with numbers as in the boxes in figure 2.1, and the fluxes of the reactions are denoted $v_i,\ i = 1, \ldots, 10$ and also numbered as in figure 2.1. The expressions for the fluxes are given in table 2.1.

Flux	Rate equation	Parameter values
v_1	$V_1 u_1/((1+(u_8/K_I)^n)(K_1+u_1))$	$V_1 = 2.5;\ n = 1;\ K_I = 9;\ K_1 = 10;$
v_2	$V_2 u_2/(K_2+u_2)$	$V_2 = 0.25;\ K_2 = 8;$
v_3	$k_3 u_2 u_3/(K_3+u_3)$	$k_3 = 0.025;\ K_3 = 15;$
v_4	$k_4 u_2 u_4/(K_4+u_4)$	$k_4 = 0.025;\ K_4 = 15;$
v_5	$V_5 u_5/(K_5+u_5)$	$V_5 = 0.75;\ K_5 = 15;$
v_6	$V_6 u_4/(K_6+u_4)$	$V_6 = 0.75;\ K_6 = 15;$
v_7	$k_7 u_5 u_6/(K_7+u_7)$	$k_7 = 0.025;\ K_7 = 15;$
v_8	$k_8 u_5 u_7/(K_8+u_8)$	$k_8 = 0.025;\ K_8 = 15;$
v_9	$V_9 u_8/(K_9+u_8)$	$V_9 = 0.5;\ K_9 = 15;$
v_{10}	$V_{10} u_7/(K_{10}+u_7)$	$V_{10} = 0.5;\ K_{10} = 15;$

Table 2.1: Fluxes and parameter values in the Kholodenko model. The values are the ones given in the original article [6].

The output of the model with these parameter values is shown in figure 2.2. We see the concentrations of MAPK and MAPKPP and note that they oscillate heavily, and that the oscillations are sustained. The oscillations are present for a range of parameter values, although the frequency and amplitude may vary.

2.3. ADDING DIFFUSION

We now wish to add spatial movements of proteins to the Kholodenko model. We assume that the MKKK is fixed at the membrane and that MKK and MAPK may diffuse freely through the cell cytoplasm. This means that the reactions 1, 2, 3 and 4 take place only near the membrane, where MKKK is present, while the other reactions take place all over the cell.

The geometry that we use is a near spherical cell of diameter 9 μm in three dimensions with no inner structure. This is of course a great simplification, but still it is a more advanced model than the original one, which does not include the space dimension at all. The size of the cell is the approximate size of a yeast cell, which is a rather small cell compared to other eukaryotic cells. A diffusion term is added to all the equations for the MKKs and MAPKs, yielding the eight equations

$$\begin{aligned}
du_1/dt &= v_2 - v_1 \\
du_2/dt &= v_1 - v_2 \\
du_3/dt &= d_3\Delta u_3 + v_6 - v_3 \\
du_4/dt &= d_4\Delta u_4 + v_3 + v_5 - v_4 - v_6 \\
du_5/dt &= d_5\Delta u_5 + v_4 - v_5 \\
du_6/dt &= d_6\Delta u_6 + v_{10} - v_7 \\
du_7/dt &= d_7\Delta u_7 + v_7 + v_9 - v_8 - v_{10} \\
du_8/dt &= d_8\Delta u_8 + v_8 - v_9
\end{aligned}$$

with $u_j = u_j(x,y,z,t),\ j = 1,\ldots,8$, and $\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ denoting the

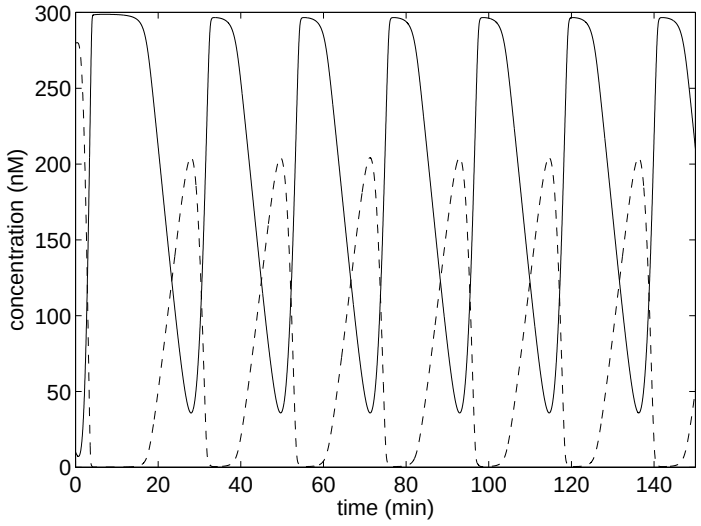

Figure 2.2: Typical behavior of the Kholodenko model. The graph shows the concentration of MAPK (solid) and MAPKPP (dashed) for times up to 150 minutes. We see the sustained oscillations in concentration.

Laplacian. The boundary conditions are

$$\hat{n} \cdot \nabla u_j = 0, \quad j = 3, \ldots, 8,$$

where $\hat{n}$ is the outward normal, meaning that there is no flux of molecules across the cell membrane. The initial values are the same as in the space-independent model, and constant throughout the cell for $u_3, \ldots, u_8$, while for the MAPKKKs u_1 and u_2 those values are only used near the boundary where the reactions 1 through 4 take place, with initial concentration zero elsewhere. The parameters d_j are the diffusion coefficients, which are now additional parameters in the model, describing how fast the diffusion of proteins is. These are known (see e.g. [10]) to be much lower inside the cell than in water (up to 10 times), because the cell is full of obstacles such as the cytoskeleton and other proteins. For simplicity, we assume that all the diffusion coefficients are equal, i.e. $d_3 = \ldots = d_8 = d$. For globular (near-spherical) proteins in water, the diffusion coefficient may be estimated from the relation

$$d \approx c \cdot W^{-1/3}, \tag{2.1}$$

where W is the molecular weight (see [9]). Fitting of measurement values for medium-sized proteins tabulated in [2, chapter 7] gives $c \approx 2.7 \cdot 10^3$ if W is measured in Daltons and d in $\mu m^2/s$. For the MAPK Hog1 in yeast, which has a molecular weight of 48.8 kDa, this gives $d = 74$ $\mu m^2/s$ for free diffusion in water, which we use as a reference value. The equations for the fluxes are the same as before, i.e. the ones given in table 2.1.

The equations were solved using the Immersed Interface Method and finite differences on uniform grids with $54 \times 54 \times 54$ nodes. The method is described in Part II of this thesis.

2.4. RESULTS

We study the results of the model for two different values of the diffusion coefficient, d, namely $d_H = d_0/10$ and $d_L = d_0/1000$, with $d_0 = 74$ $\mu m^2/s$ being the approximate diffusion coefficient in water for Hog1.

The higher value d_H for the diffusion coefficient is so large that the molecules have time to move around the entire cell faster than the reactions produce any significant changes in concentration. This means that the proteins will be evenly spread across the cell at all times. This is seen in figure 2.8 for two of the components in the pathway (MKK and MAPKPP). The figure shows concentrations for times between 0 and 50 min on a line through the center of the cell. One sees that the concentration is the same in the center of the cell as at the edge. One can also see that the oscillations are not as large as for the original model. This is shown more clearly in figures 2.4 and 2.5, which should be compared to figure 2.3 for the original model. Here concentrations for all components in the model are shown. Samples are taken at the membrane (figure 2.4) and at the center of the cell (figure 2.5). It is clear that the oscillations have a much smaller amplitude now that we have added diffusion. Furthermore, the oscillations seem to be damped, so that the amplitude decreases with time. This has been confirmed by running longer simulations, where the oscillations slowly fade away. The diffusion acts as a damper for the system.

A similar behavior is seen with the lower value $d = d_L$ for the diffusion coefficient (figures 2.6 and 2.7). The oscillations again have a lower amplitude and are again damped. Here, however, the proteins do not have time to diffuse through the cell before the reactions produce significant changes in phosphorylation levels. This is seen clearly in figure 2.9, where again concentrations of MKK and MAPKPP are shown along a line through the cell center for different times. It is clear that the concentrations are different in the center and at the membrane. However, the oscillations are present and approximately equal in period and amplitude in the center and at the membrane, and also approximately equal to the previous case (with $d = d_H$).

A final example is shown in figure 2.10. There, we show the result of the same model with slow diffusion ($d = d_L$), but in a larger cell with a diameter of about 90 μm. Now the phosphorylated proteins do not have time to move very far from the membrane before they are dephosphorylated. Therefore the oscillations occur only near the membrane, while in the center all the kinases are in their inactive (unphosphorylated) state. This of course makes the signaling pathway useless, since its main purpose is to convey the signal of phosphorylations to the cell nucleus, which it fails to do when distances are large, diffusion slow and dephosphorylation reactions comparatively fast.

It is clear that what determines the spatial behavior of the system is a combination of the length scale, the diffusion coefficient and the reaction rate. In order to see this more clearly, we may make the equations non-dimensional. To illustrate this, we take the equation for u_5 (MKKPP), but include only reaction term — the dephosphorylation

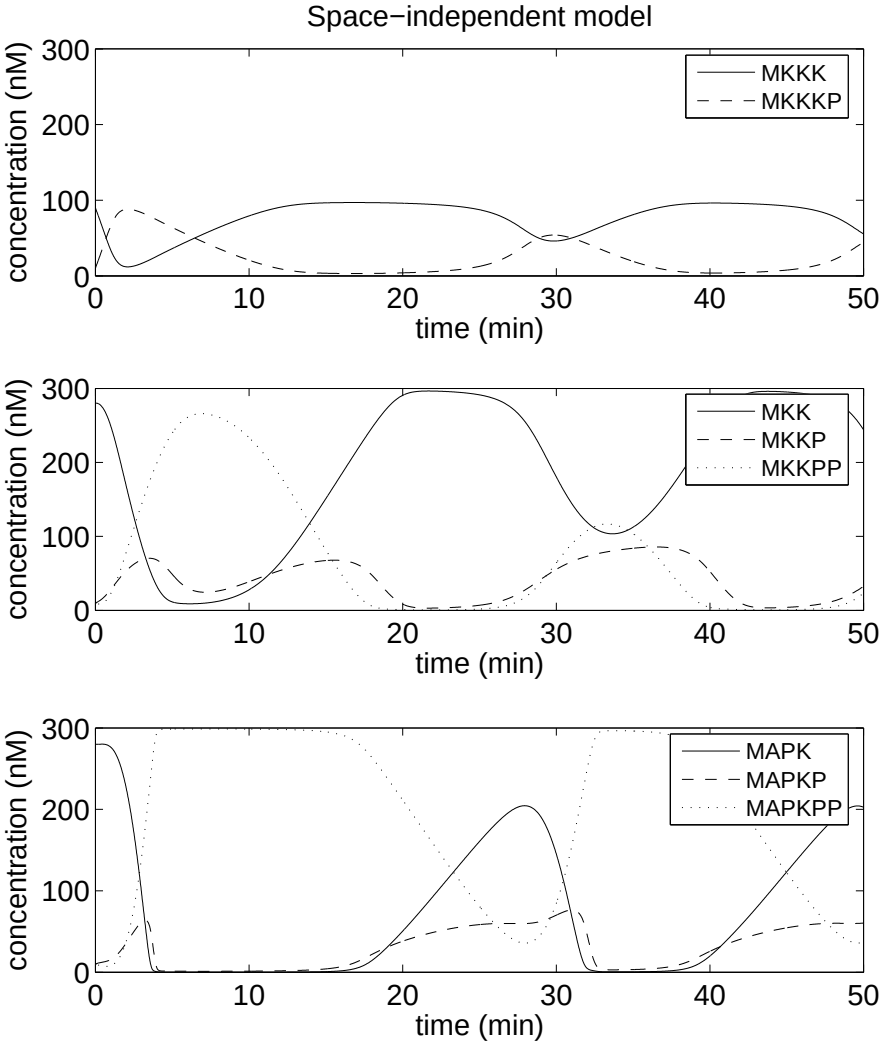

Figure 2.3: The concentrations of the eight components of the original (space-independent) model.

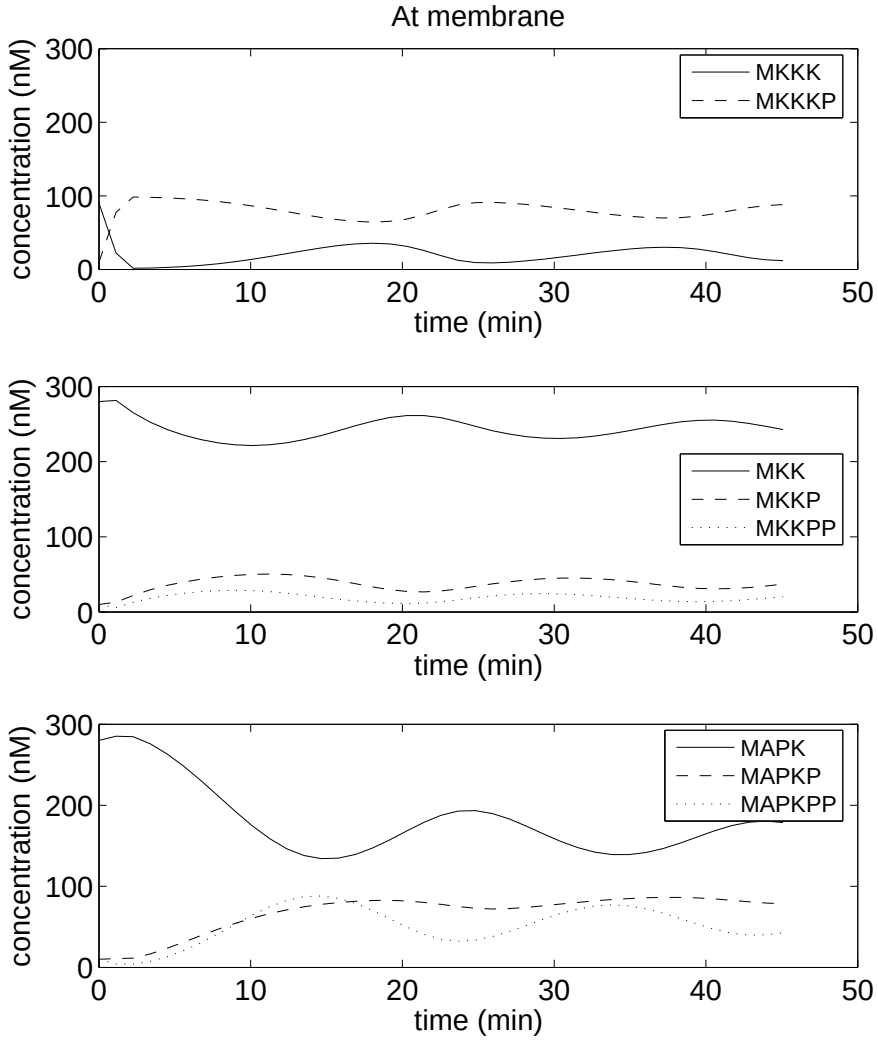

Figure 2.4: Concentrations for the eight components in the space-dependent model with fast diffusion $d = d_H$. The values are taken at the membrane.

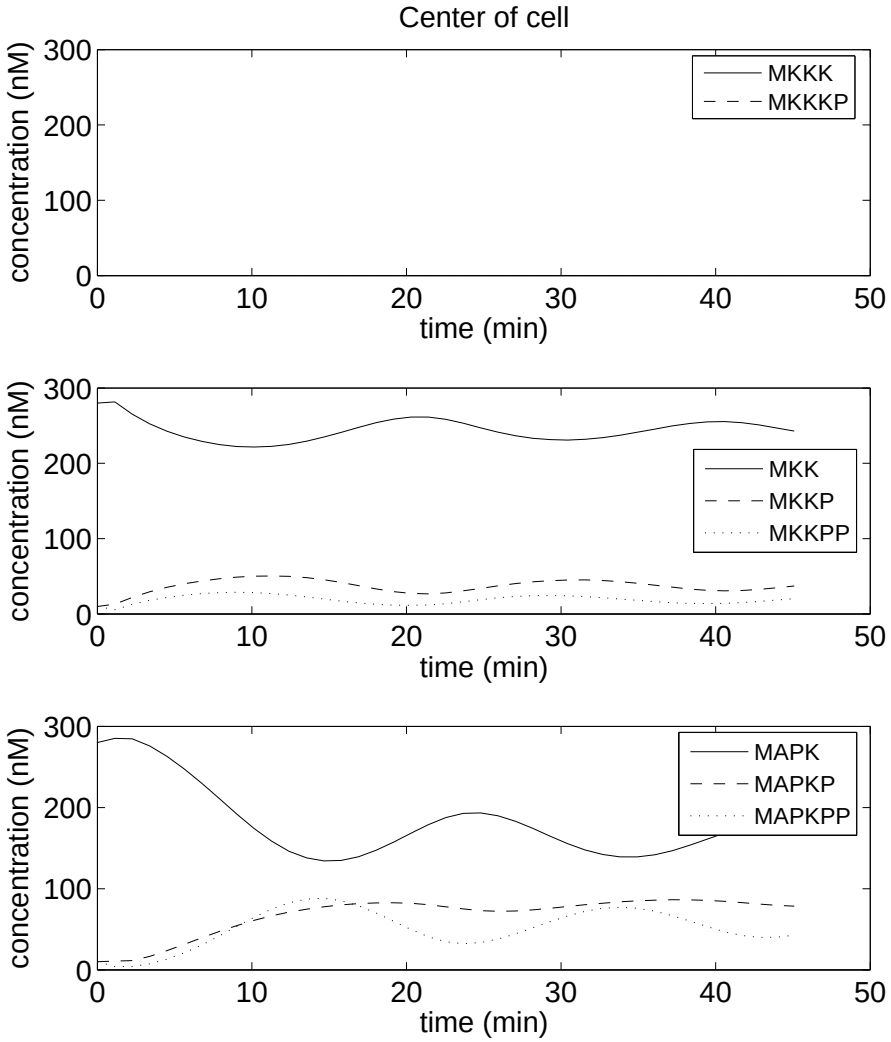

Figure 2.5: Concentrations for the eight components in the space-dependent model with fast diffusion $d = d_H$. The values are taken at the center of the cell.

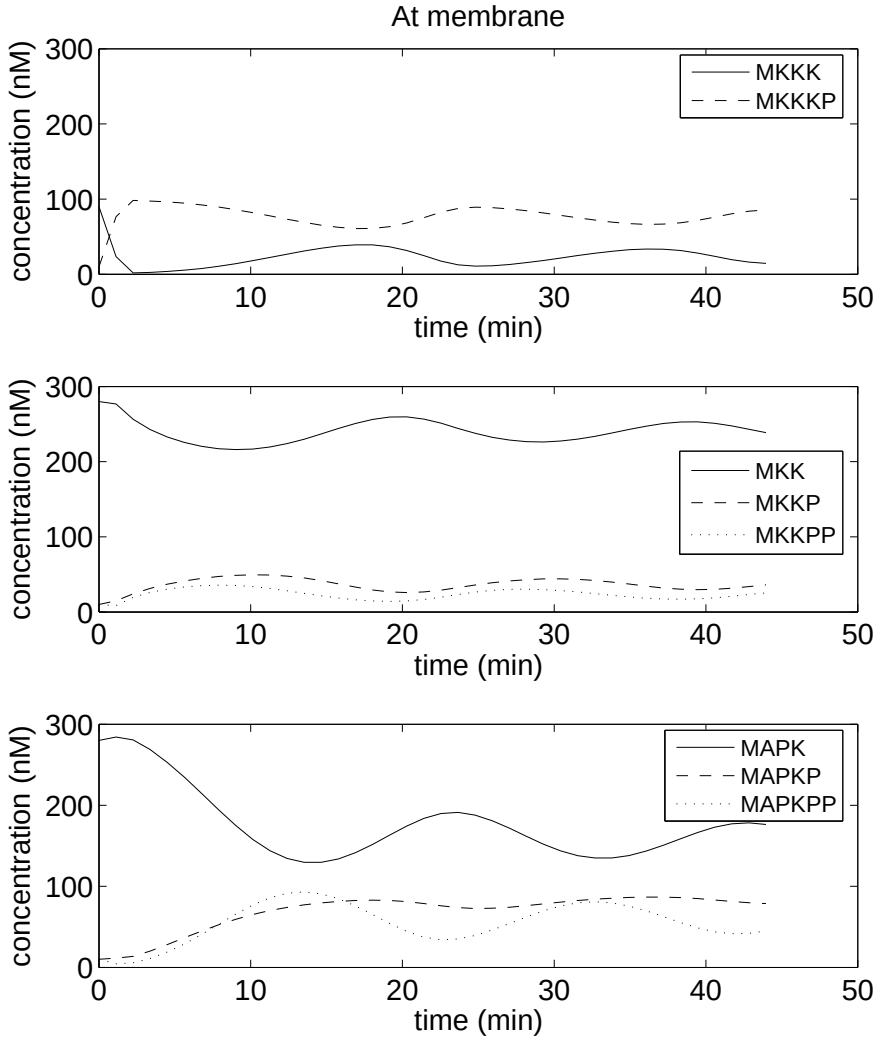

Figure 2.6: Concentrations for the eight components in the space-dependent model with slow diffusion $d = d_L$. The values are taken at the membrane.

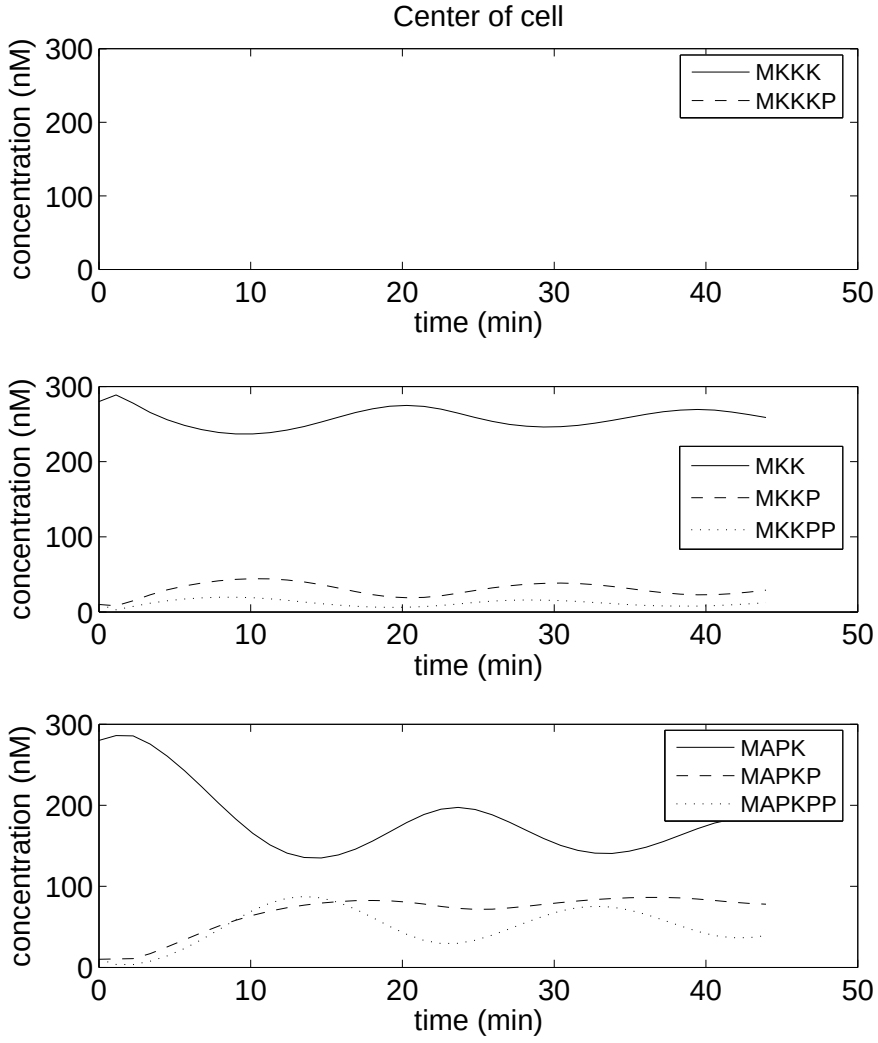

Figure 2.7: Concentrations for the eight components in the space-dependent model with slow diffusion $d = d_L$. The values are taken at the center of the cell.

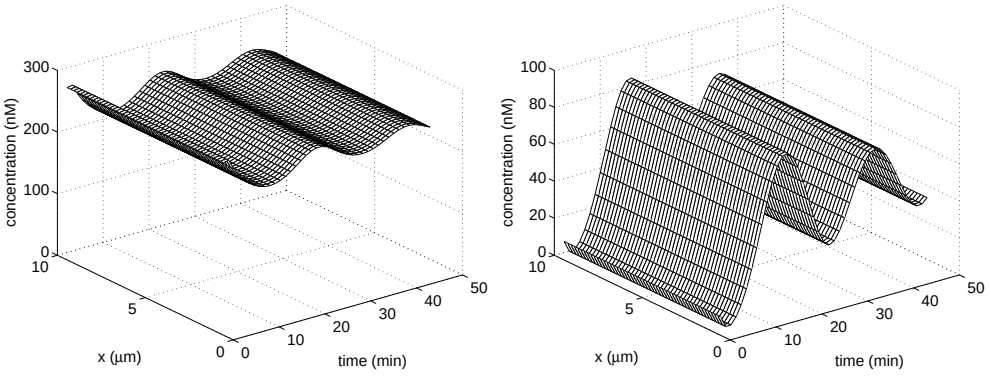

Figure 2.8: Concentrations of MKK (left) and MAPKPP (right) in the model with fast diffusion $d = d_H$. The values are taken on a line through the center of the cell for times up to about 50 minutes. The concentrations are almost equal throughout the cell.

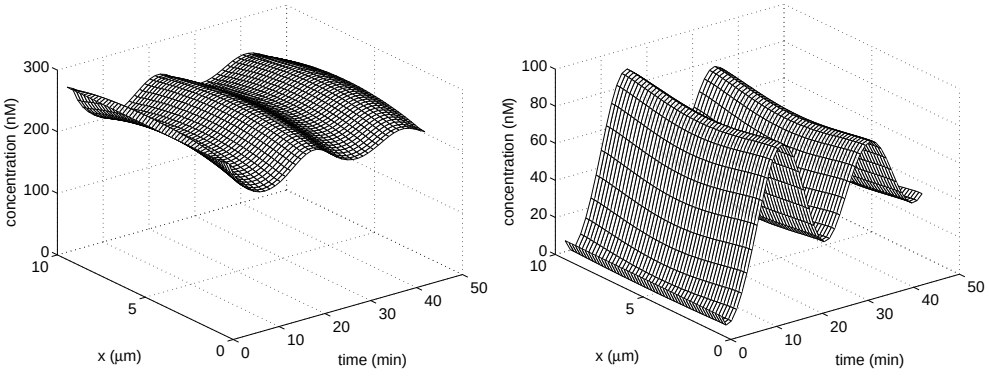

Figure 2.9: Concentrations of MKK (left) and MAPKPP (right) in the model with slow diffusion $d = d_L$. The values are taken on a line through the center of the cell for times up to about 50 minutes. The concentration gradient is clearly visible.

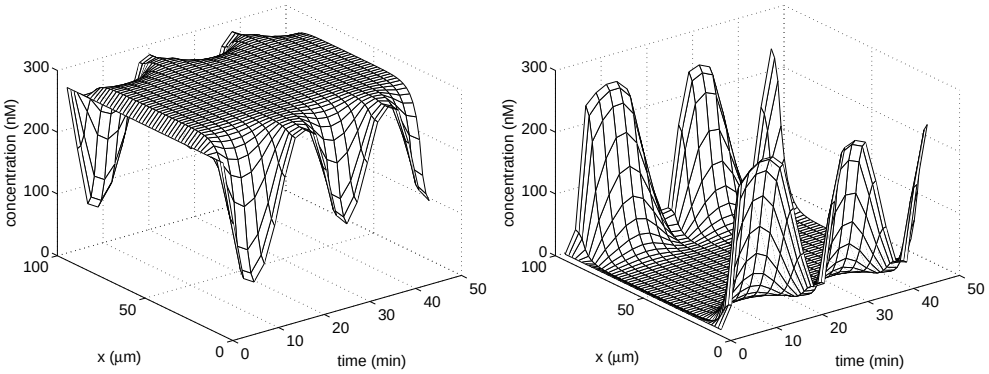

Figure 2.10: Again, concentrations of MKK and MAPKPP with slow diffusion $d = d_L$, but this time in a cell with diameter 90 μm. The values are taken on a line through the center. Here, the phosphorylated proteins do not have time to diffuse far before the phosphate is removed, so the whole oscillation takes place at the membrane.

reaction with flow v_5. The equation then becomes

$$\frac{\partial u_5}{\partial t} = d\Delta u - \frac{V_5 u_5}{K_5 + u_5}.$$

If we introduce the non-dimensional space variables $(\xi, \eta, \zeta) = h^{-1}(x, y, z)$, the non-dimensional time $\tau = d/h^2 \cdot t$ and the non-dimensional concentration $\nu = u_5/u_0$, with scaling coefficients h, k and u_0 with dimensions $[h] = \text{m}$, $[k] = \text{s}$ and $[u_0] = \text{M}$, we get

$$\frac{\partial \nu}{\partial \tau} = \Delta_\xi \nu - \frac{h^2 V_5}{d\, u_0} \frac{\nu}{K_5/u_0 + \nu},$$

with Δ_ξ denoting the Laplacian in the (ξ, η, ζ)-variables. The typical size of the diffusion term is now 1, while the typical size of the reaction term is $\lambda = h^2 V_5/(d\, u_0)$. This is the quantity that determines the influence of the reaction compared to the influence of diffusion. The values of λ for the three examples discussed here are summarized in table 2.2. We see that in the first case $\lambda \ll 1$, indicating that reactions are slow compared

Example	h (μm)	d (μm^2/s)	V_5 (nM/s)	u_0 (nM)	λ
1	10	7.39	0.75	300	$3.38 \cdot 10^{-2}$
2	10	0.0739	0.75	300	3.38
3	100	0.0739	0.75	300	$3.38 \cdot 10^{2}$

Table 2.2: The ratio λ for the three examples.

to diffusion, so that molecules have time to travel across the cell before changing their phosphorylation state. This leads to homogeneous mixing of the diffusing components and

no concentration gradients are observed. This is exactly what we see in figure 2.8 for the first example. For the second example, we have $\lambda \approx 1$ and diffusion and reaction are comparable, meaning that mixing of components is not complete. Thus we expect to see some concentration gradients, but also some diffusion of reaction products through the cell, which is what figure 2.9 shows. Finally, in the third example, $\lambda \gg 1$, meaning that reactions are much faster than diffusion, leading to highly localized reactions with small exchange of material. This is seen in figure 2.10.

2.5. Discussion

Much could be said about the Kholodenko model and whether oscillations really do occur in signaling pathways in real cells. One of the weakest points of the model is perhaps that the stimulus that activates the pathway is assumed to be "on" all the time, only being inhibited temporarily when the level of MAPKPP is high. This assumes that the cell never really responds to the signal, which is pretty absurd, since the purpose of the signal is to cause the cell to respond and adapt to new conditions. Once the cell has adapted, the signal must cease, or the cell will probably die or at least spend all its energy on useless tasks. So, it is maybe not so probable that we will observe oscillations in signaling pathways in real cells.

However, this numerical experiment has shown a number of other things of more general interest, which are easy to observe because oscillations is an effect that is easily studied. First, when we include the spatial distribution of proteins, we see that the oscillations are severely damped compared to the original model, almost independent of the diffusion coefficient as long as it is not too small. This shows that by neglecting diffusion and spatial distribution when modeling, one may overlook important aspects and draw false conclusions about the behavior of the system. Of course, the principal behavior is in large determined by the space-independent reaction terms, but our simulations show that the amplitude of the effect may be diminished significantly by the addition of diffusion. Furthermore, in our examples, the oscillations are not sustained, but seem to diminish with time, indicating that the system is damped by the diffusion. This is a different type of behavior than the sustained oscillations and is also an important thing to keep in mind when doing space-independent modeling of biochemical processes. The reason for this dependence on diffusion is that the problem is directly space-dependent, since we know that the first reactions take place only at the membrane, while the proteins at the end of the chain may move about freely. Thus, purely time-dependent modeling can be expected to produce errors for all signaling pathways, since these are space-dependent by nature, while for example space-independent metabolic models may be assumed to be more correct, since the space-dependence is not so obvious in that case.

There is of course a reason for not using full spatial modeling, in particular three-dimensional modeling, since the solving of systems of PDEs take so much more time than solving a system of ODEs. The images shown here are results of simulations that took 10 minutes or more. This should be compared to fractions of a second for solving the space-independent model. The amount of time needed to solve the space-dependent equations

makes it impossible to use for example parameter fitting algorithms, since such algorithms require a large number of simulations with different parameter settings. Therefore, space-independent ODE models are of great importance, but one should be aware that one is neglecting something and that it may be worthwhile to see what happens in a space-dependent model.

3. Spatial modeling of the HOG pathway in yeast

3.1. The HOG pathway

One rather well-studied MAPK signaling pathway is the Hog1-pathway in the yeast *Saccharomyces cerevisiae*, where HOG stands for High Osmolarity Glycerol and the MAP kinase Hog1 is the last enzyme in the signaling chain, whose purpose is to sense that the osmotic pressure on the cell membrane increases and produce the appropriate response (see [4]). That is, if solutes (e.g. salt) are added to the solution outside the cell, the process called osmosis will strive to level out the difference in solute concentration over the membrane by forcing water to flow out of the cell. This is potentially harmful for the cell, since it then starts to shrink and if that goes on, the cell can not function anymore and eventually it will die.

To avoid this sad fate, the cell has to respond in some way to this new environment. It does so by starting to produce glycerol and accumulating it inside the cell, which evens out the solute concentrations and thereby prevents water from flowing out of the cell. And the link between the sensing of osmotic pressure and the response in the form of glycerol production is the HOG pathway. It is not really well known how the actual sensing of the change in osmotic pressure takes place, but there are enzymes at the cell membrane that are somehow activated, which leads to the activation of Hog1 through a few intermediate kinases. The pathway is shown in figure 3.1 with some of its surrounding components. When Hog1 has been activated, it enters the nucleus and once there it affects transcription of several genes through the transcription factors shown at the bottom of the figure.

A central feature of the signaling pathway is that Hog1 enters the nucleus. This can be viewed in the microscope by genetically adding a Green Fluorescent Protein (GFP) tag to the Hog1 protein, which makes the molecules shine bright green when viewed under ultraviolet light. This is a powerful method to view the localization of proteins in the cell and the results look like figure 3.2. One should be aware, however, that the GFP is a protein of about the same size as Hog1, so that the Hog1-GFP fusion is a much larger protein than wild-type Hog1, which could affect the kinetics and function of the protein. Still, the Hog1-GFP fusion is functional in the sense that it is able to carry out its function in the signaling pathway.

3.2. The model

Figure 3.3 shows the model that we are studying here. It is a model of the center part of the Hog1-pathway, where Hog1 itself is involved. The focus is on the transport of activated Hog1 into the nucleus. For simplicity, we assume that the MAPKK Pbs2 sits at the membrane in its phosphorylated form, ready to phosphorylate Hog1-molecules that come close to the membrane. This assumes that the upper part of the pathway that is not included in the model has reached an equilibrium and that phosphorylation of Hog1 can only take place at the membrane, which is believed to be true. There is also a small amount of spontaneous phosphorylation of Hog1 both in the nucleus and in the cytoplasm.

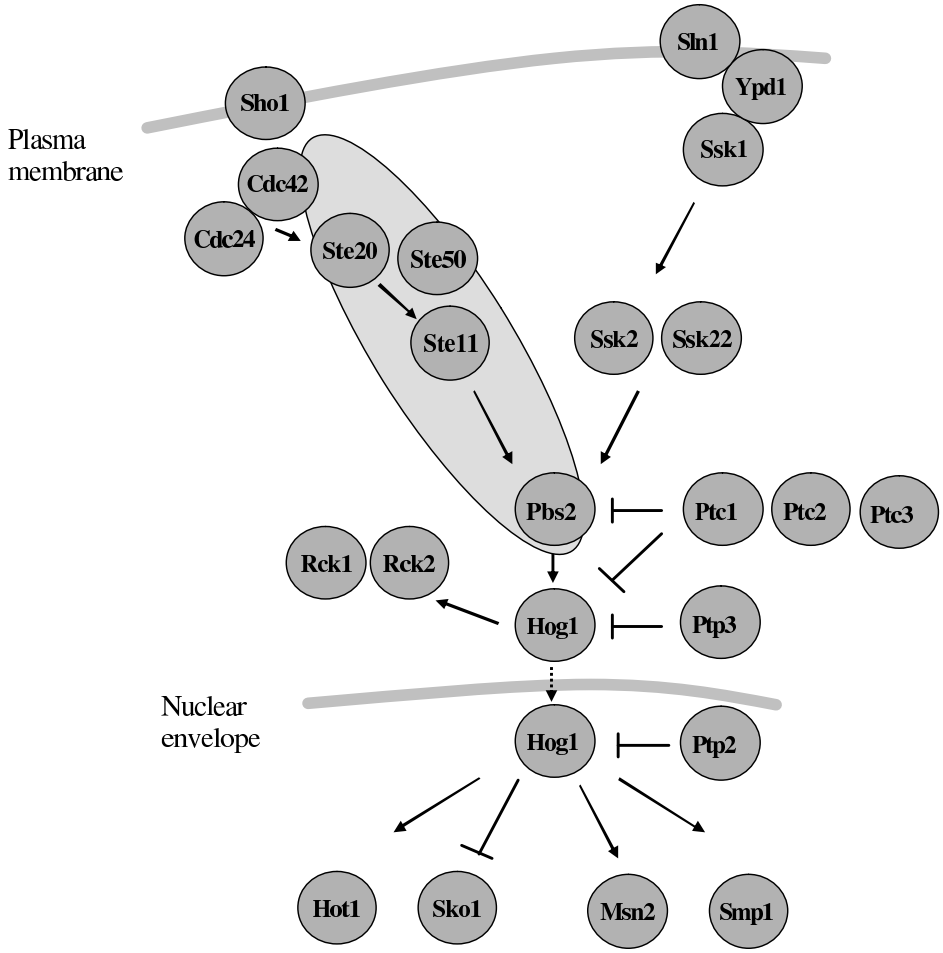

Figure 3.1: A schematic diagram showing the essentials of the HOG pathway, including the two different osmosensing mechanisms at the top, phosphatases on the right, cytoplasmic targets on the left and nuclear targets (transcription factors) at the bottom. Hog1 is seen at the center of it all and it is indicated that upon osmotic stress, it moves from the cytoplasm to the nucleus.

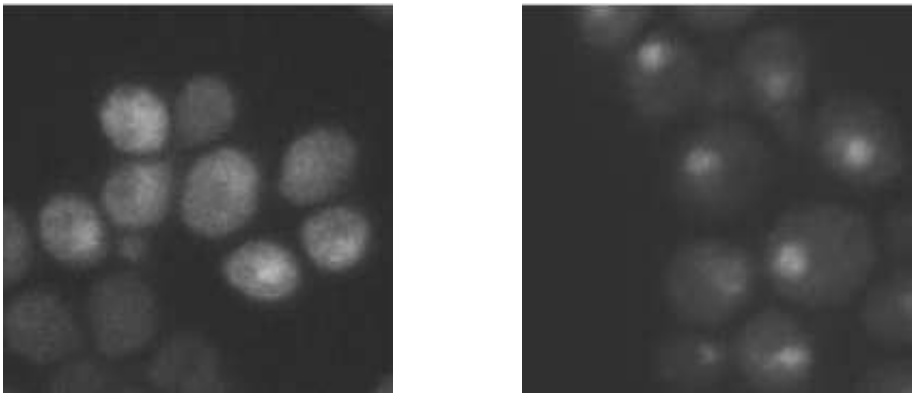

Figure 3.2: Hog1-GFP in wild-type yeast cells, under normal osmotic conditions on the left and after addition of NaCl on the right. The nuclei can be seen as bright spots in the right image, indicating that Hog1 has entered the nucleus. The positions of the nuclei may be verified by staining with a special dye (DAPI) (not shown). The images are generated by Claes Molin.

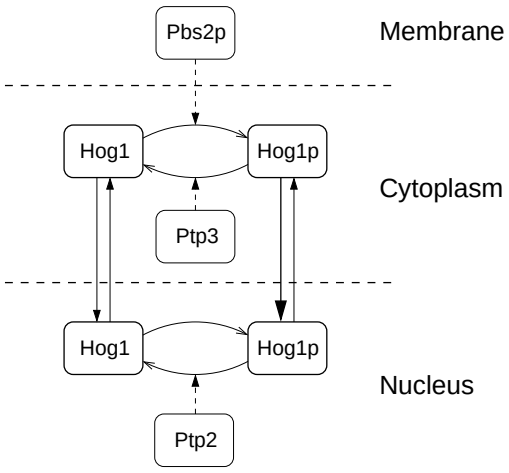

Figure 3.3: The model for the Hog1-pathway. The MAPKK Pbs2 is assumed fixed at the membrane in its phosphorylated form, meaning that the pathway is constantly active. The phosphatases Ptp2 and Ptp3 are distributed evenly throughout the nucleus and cytoplasm respectively. Hog1 is free to diffuse through the cell, but at the nuclear membrane, the transport is regulated. Unphosphorylated Hog1 is transported in and out of the nucleus at equal rates, while phosphorylated Hog1 (Hog1p) is transported into the nucleus at a higher rate than it is transported out.

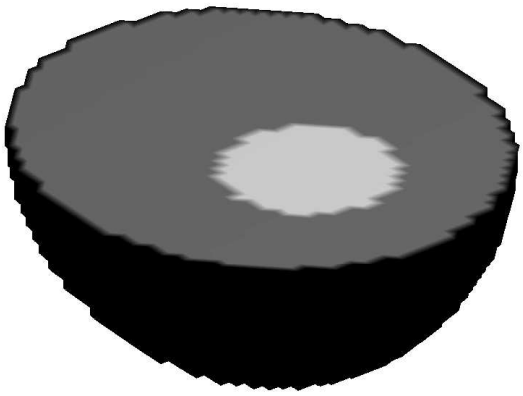

Figure 3.4: A cross section of the cell used in the model. The nucleus is shown as the bright spot.

The two phosphatases in the model, Ptp2 and Ptp3 are assumed to be evenly distributed in their respective domains, the nucleus and the cytoplasm. They are responsible for the dephosphorylation of Hog1. The only diffusing components of the model are thus Hog1 and Hog1p, which are free to move in the cytoplasm and in the nucleus, but not between the two compartments. The transport through the nuclear membrane is regulated so that unphosphorylated Hog1 is transported in and out of the nucleus at equal rates. Phosphorylated Hog1 (Hog1p) is transported out of the nucleus at the same rate, but transported into the nucleus at a much higher rate. The geometry used for the model is a simplified three-dimensional cell with an off-center spherical nucleus and a slightly ellipsoidal cell membrane. No other internal cell structure is included. A cross section of the model cell is shown in figure 3.4.

The model is of course a very simplified one. We disregard the fact that the phosphorylation of Hog1 takes place in two steps, so that the active form is double-phosphorylated. We also ignore the upper part of the pathway, the osmosensing mechanism and the phosphorylation of Pbs2. Also, the targets of Hog1 are not included in the model and there is no feedback loop to turn off the signaling pathway. The transport of proteins through the nuclear membrane is also simplified. Here, we use the assumption that the flux through the membrane is proportional to the concentration of proteins, which is of course a simplification of the rather complex transport system that shuttles the proteins in and out of the nucleus. And finally, the inner structure of the cell is not included other than in a reduction of the diffusion constant because of the obstacles.

We denote the concentrations of unphosphorylated and phosphorylated Hog1 by u_H and u_P, respectively, the nucleus by Ω_N, the cytoplasm by Ω_C and the nuclear and plasma membranes Γ_N and Γ_C. The model then leads to the following partial differential equations

$$\begin{aligned}
\partial u_H/\partial t &= d_H \Delta u_H - f_C(u_H, u_P), && \text{in } \Omega_C,\\
\partial u_P/\partial t &= d_P \Delta u_P + f_C(u_H, u_P), && \text{in } \Omega_C,\\
\partial u_H/\partial t &= d_H \Delta u_H - f_N(u_H, u_P), && \text{in } \Omega_N,\\
\partial u_P/\partial t &= d_P \Delta u_P + f_N(u_H, u_P), && \text{in } \Omega_N,\\
\partial u_H/\partial n &= 0, && \text{on } \Gamma_C\\
\partial u_P/\partial n &= 0, && \text{on } \Gamma_C\\
d_H \partial u_H/\partial n &= a_H u_H^{(N)} - b_H u_H^{(C)}, && \text{on } \Gamma_N\\
d_P \partial u_P/\partial n &= a_P u_P^{(N)} - b_P u_P^{(C)}, && \text{on } \Gamma_N
\end{aligned}$$

with

$$\begin{aligned}
f_C(u_H, u_P) &= \frac{V_{Pbs2} u_H}{K_{Pbs2} + u_H} - \frac{V_{Ptp3} u_P}{K_{Ptp3} + u_P} + k_{spC} u_H\\
f_N(u_H, u_P) &= -\frac{V_{Ptp2} u_P}{K_{Ptp2} + u_P} + k_{spN} u_H.
\end{aligned}$$

Here, d_H and d_P are diffusion constants, which we assume to be equal. The four first equations are diffusion-reaction equations with one diffusion term and one reaction term, the two first describe the cytoplasm and the two last the nucleus. The reaction terms make use of Michaelis-Menten kinetics (see e.g. [2, chapter 4]) for the enzyme reactions, which is a standard way of modeling such reactions. Next follows boundary conditions at the plasma membrane, which say that the flux through the membrane is zero, i.e. that no protein molecules may leave the cell. The two last rows are boundary conditions at the nuclear membrane, both saying that the flux out through the membrane equals a constant times the nuclear concentration at the membrane minus a constant times the cytoplasmic concentration on the outside of the membrane. We also need initial conditions that describe the concentrations at time $t = 0$. These are shown in figure 3.5 and are chosen so that the system is near its equilibrium. Most of the Hog1 molecules are in the unphosphorylated state and are evenly distributed between nucleus and cytoplasm. A small fraction of the molecules are phosphorylated and these have a higher concentration in the nucleus than in the cytoplasm.

Table 3.1 shows the parameter values used for the simulations. The diffusion coefficients are taken to be $d_0/1000$, where d_0 is the diffusion coefficient of Hog1 in water, estimated from equation (2.1). The reaction coefficients are chosen in the same range as for the Kholodenko model, but their exact values are quite arbitrarily chosen to get a result which resembles the *in vivo* behavior. The parameters a and b describing the efficiency of the nuclear transport are also rather arbitrarily chosen to get a reasonable result and are not based on experiments. For comparison, one simulation was also done with faster diffusion ($d_H = d_P = d_0/10$).

d_H	$0.074\mu m^2/s$	d_P	$0.074\mu m^2/s$
a_H	$1\ d_H$ μm/s	a_P	$1\ d_P$ μm/s
b_H	$1\ d_H$ μm/s	b_P	$3\ d_P$ μm/s
V_{Pbs2}	1.0 nM/s	K_{Pbs2}	50.0 nM
V_{Ptp3}	0.2 nM/s	K_{Ptp3}	15.0 nM
V_{Ptp2}	0.2 nM/s	K_{Ptp2}	15.0 nM
k_{spC}	0.0002 s^{-1}	k_{spN}	0.0002 s^{-1}

Table 3.1: Parameter values

The equations were solved using the Immersed Interface Method and finite differences on uniform three-dimensional grids with $54 \times 54 \times 54$ nodes. The method is described in Part II of this thesis. The final simulations shown here, involving some 200 time steps, took up to 30 minutes to complete on a standard computer.

3.3. Results

Figure 3.5 shows the initial conditions used in the 3D simulations. The images only show the concentrations in a slice through the center of the cell, but the concentrations are assumed to be uniform throughout the nucleus and the cytoplasm. The values chosen are near the equilibrium of the system when the pathway is not activated. The exact values used are 100 nM for Hog1 and 1.5 nM and 3.2 nM for Hog1p in cytoplasm and nucleus respectively, but changing these values slightly does not alter the behavior of the system much, so the exact values are not very important.

Then at time $t = 0$, the signaling is turned on, in this model by activating Pbs2, that is setting V_{Pbs2} to a non-zero value. In figure 3.6 we see the simulated response as it would appear in the fluorescence microscope, namely the total concentration of Hog1 ([Hog1] + [Hog1p]) at the start and the end of the simulation. The simulation ends at 22 minutes when the concentrations are approaching the steady-state levels with the signal turned on. In a real cell the level of phosphorylation would again turn down when the cell started to adapt, but since this adaptation is not included in the model, the concentrations just approach a steady state with high concentration in the nucleus, mostly containing phosphorylated Hog1 and lower concentrations outside.

To see the time evolution of the system, we show in figure 3.7 concentrations along a line through the center of the cell for all times and for Hog1 and Hog1p separately. We see that the levels of Hog1p are gradually increasing and that Hog1p is transported into the nucleus, while unphosphorylated Hog1 is moved out of the nucleus as the levels become lower in the cytoplasm when Hog1 is phosphorylated. One also notes that there are concentration gradients in the cytoplasm, arising from the fact that the diffusion is not fast enough to even out the concentrations at the same rate that the reactions change them. The dimensionless parameter λ defined in section 2.4 is here

$$\lambda = \frac{h^2 V_{Pbs2}}{d\, u_0} \approx \frac{(10^{-5})^2 \cdot 1.0}{7.4 \cdot 10^{-14} \cdot 100} \approx 13.5,$$

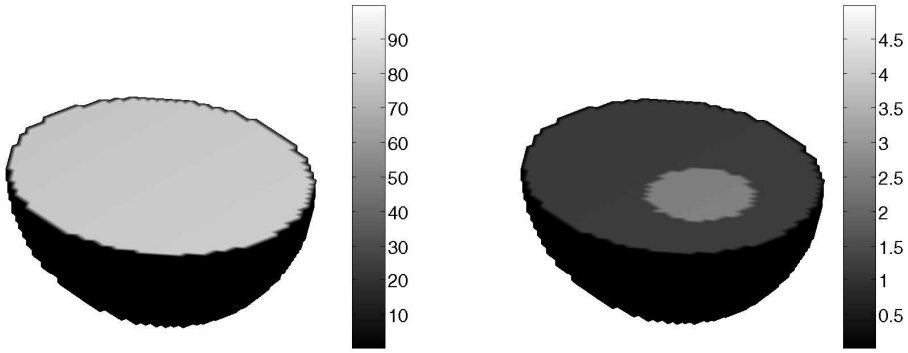

Figure 3.5: Initial conditions for the simulation in 3D, with unphosphorylated Hog1 on the left and phosphorylated Hog1 (Hog1p) on the right. The images show a slice through the center of the cell, with the nucleus visible in the right image. Note the different scales on the two images.

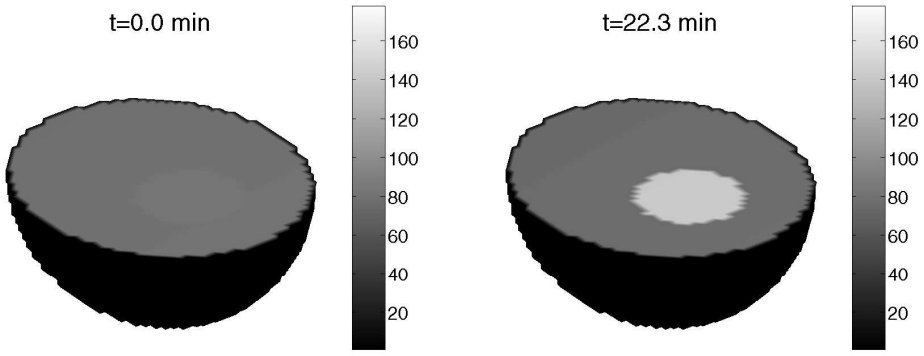

Figure 3.6: The total Hog1 concentration ([Hog1] + [Hog1p]) at the start (left) and end (right) of simulation. These images are comparable to the ones seen in the fluorescence microscope. The ragged edges are artifacts arising from the space discretization.

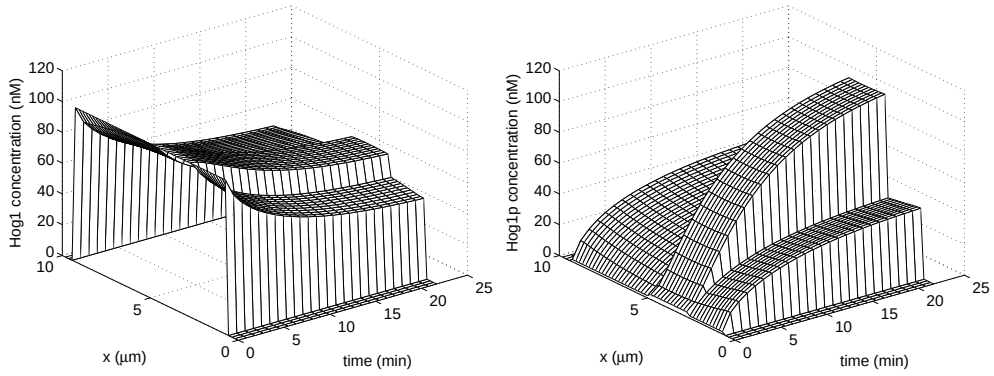

Figure 3.7: Concentrations of Hog1 (left) and Hog1p (right) for slow diffusion. The values are taken along a line through the center of the cell and the nucleus. The nuclear relocation of Hog1p is clearly seen. There is a slight concentration gradient in the cytoplasm.

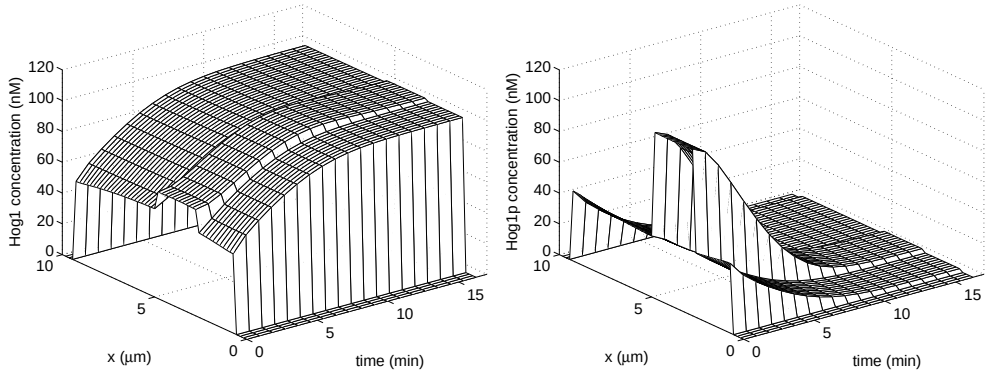

Figure 3.8: Concentrations of Hog1 (left) and Hog1p (right) for slow diffusion, when the signal is turned off after Hog1p has accumulated in the nucleus. The concentrations return to their original state.

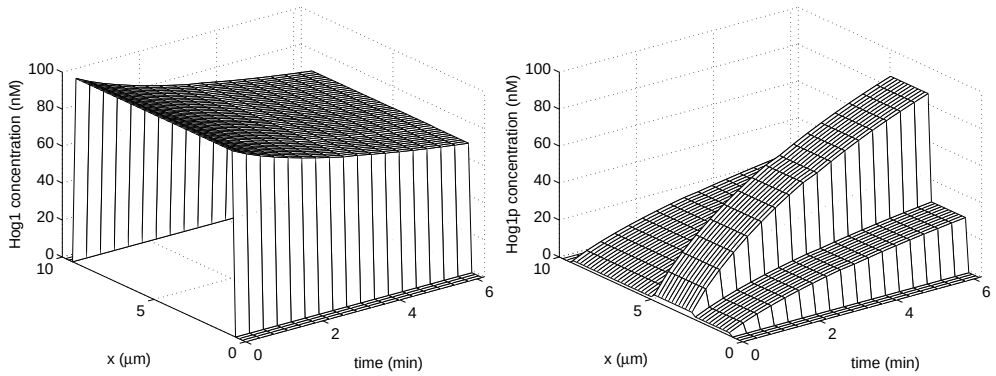

Figure 3.9: Concentrations of Hog1 (left) and Hog1p (right) for fast diffusion. Here the concentration gradient is no longer present and the response is slightly faster than in the example with slow diffusion.

indicating that reactions are slightly faster than diffusion, which should produce gradients in the concentration levels.

In figure 3.8, we also show the return to the original state when the signaling pathway is turned off. Starting at the final levels of the previous simulation, and setting $V_{Pbs2} = 0$, we see that the system returns to its original state with low levels of phosphorylated Hog1 and almost equal concentrations in nucleus and cytoplasm.

Finally, we also show a simulation with a larger diffusion coefficient ($d_H = d_P = d_0/10 \approx 7.4\mu\text{m}^2/\text{s}$). The results are shown in figure 3.9. Now, diffusion is fast enough to level out the concentrations, so that they are uniform in nucleus and cytoplasm, respectively. The response in the nucleus is also slightly faster than in the slow diffusion simulation.

3.4. Discussion

The results shown above show that it is possible to reproduce in simulations the nuclear relocation of Hog1 that is observed when yeast cells are exposed to osmotic stress. The model used is a very simple one, including only two diffusing components and only the final steps of the phosphorylation chain. The results should therefore not be used to draw far-reaching conclusions about the actual behavior of the cell.

What we see, however, is that only with the small diffusion coefficient $d_0/1000$ do we observe concentration gradients and interesting spatial effects. This value for the diffusion coefficient is much smaller than the values observed for similar proteins inside living cells (up to 10-fold reduction compared to diffusion in water [10]). For the larger diffusion coefficient $d_0/10$, which is more in agreement with observed values, concentrations are practically constant in each of the compartments. This indicates that diffusion is fast compared to the reactions involved, so that full spatial modeling with PDEs might not be of crucial importance for the simple model with the parameter values used here. However,

the behavior of the system is very much dependent on the enzyme rates, especially for the phosphatases Ptp2 and Ptp3. The values for the reaction rates used here are not based on direct measurements, but are taken to be near the values in the Kholodenko model, which are based on *in vitro* measurements for MAP kinases in mammalian cells. Those values may not be correct for the reactions in our model. So, we can not determine from our simple model whether the explicit modeling of spatial effects shown here is of crucial importance. In any case, one must remember that cell signaling is a spatial phenomenon, so that spatial features should be taken into account, in one way or another. What we have developed is a method to simulate full spatial models that could be used to compare the output with results from simpler, less computationally intense methods.

In the end, what decides whether a biological model is good or bad, is if it compares well with measurements of the real phenomenon in live cells. So, to be certain about what type of models are needed, one should compare the results to precise measurements. This, however, is not an easy task. The types of measurements available at present to determine spatial distributions of proteins inside the cell are fluorescence microscopy experiments, yielding images such as those in figure 3.2. In order to compare these to the results of the computations made here, one would like to follow a single cell and take images at intervals of seconds or less, which is rather hard to do. Furthermore, one would ideally like to be able to view phosphorylated and unphosphorylated Hog1 separately, which is not possible at present.

One may also discuss if a diffusion PDE is an appropriate model for the transport of proteins inside the cell. First of all, the number of molecules is relatively small. The estimated number of Hog1 molecules in the yeast cell is about 7000 [3], giving a concentration of around 50 nM. Since the diffusion equation is obtained as the limit when the number of molecules tend to infinity, it may not be an accurate model. The random fluctuations arising from the small number of molecules may also influence the reactions. However, as long as there is no data to compare to, it is hard to decide between models. But it would be interesting to investigate the differences between a PDE model and a stochastic model for the signaling pathway.

Another issue related to diffusion is that the cell is full of obstacles, consisting of organelles, cytoskeleton networks, large protein complexes etc. This makes the assumptions underlying the diffusion model invalid. It has been shown (see [11], [10], [7]) that these obstacles do not only cause a reduction in the diffusion coefficient, but may give rise to the phenomenon called anomalous (sub)diffusion. For anomalous diffusion, the mean square displacement of the molecules is no longer proportional to time ($\langle x(t)^2 \rangle \propto t$), but to some power $\alpha < 1$ of time ($\langle x(t)^2 \rangle \propto t^\alpha$). This may give rise to new interesting phenomena and is also a subject that would be interesting to study.

REFERENCES

[1] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Molecular Biology of the Cell.* Garland Science, 4th edition, 2002.

[2] C. P. Fall, E. S. Marland, J. M. Wagner, and J. J. Tyson, editors. *Computational Cell Biology.* Springer, 2002.

[3] S. Ghaemmaghami, W.-K. Huh, K. Bower, R. W. Howson, A. Belle, N. Dephoure, E. K. O'Shea, and J. S. Weissman. Global analysis of protein expression in yeast. *Nature*, 425:737 – 741, Oct 2003.

[4] S. Hohmann. Osmotic stress signaling and osmoadaptation in yeasts. *Microbiol Mol Biol Rev*, 66(2):300–372, 2002.

[5] C.-Y. F. Huang and J. E. Ferrell, Jr. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc. Natl. Acad. Sci. USA*, 93:10078–10083, sept 1996.

[6] B. N. Kholodenko. Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. *European Journal of Biochemistry*, 267:1583–1588, 2000.

[7] M. J. Saxton. Anomalous diffusion due to obstacles: a Monte Carlo study. *Biophys J.*, 66(2:1):394–401, 1994.

[8] I. Swameye, T. G. Müller, J. Timmer, O. Sandra, and U. Klingmüller. Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling. *PNAS*, 100(3):1028–33, 2003.

[9] G. A. Truskey, F. Yuan, and D. F. Katz. *Transport Phenomena in Biological Systems: A Textbook for Biomedical Engineers.* Prentice Hall, 2003.

[10] M. Wachsmuth, W. Waldeck, and J. Langowski. Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy. *J. Mol. Biol.*, 298:677–689, 2000.

[11] M. Weiss, H. Hashimoto, and T. Nilsson. Anomalous protein diffusion in living cells as seen by fluorescence correlation spectroscopy. *Biophys J.*, 84(6):4043–52, 2003.

[12] H. S. Wiley, S. Y. Shvartsman, and D. A. Lauffenburger. Computational modeling of the EGF-receptor system: a paradigm for systems biology. *Trends Cell Biol.*, 13(1):43–50, 2003.

The Immersed Interface Method on uniform and Boolean grids

Tobias Gebäck

Abstract

The Immersed Interface Method (IIM) is a method which allows the use of finite differences in non-rectangular domains by immersing interfaces into a rectangular domain which is discretized by a uniform grid. The finite differences near the interface are then corrected using the size and position of the jumps in the solution and its derivatives across the interface.

The method presented here is the Explicit Jump IIM developed by Wiegmann and Bube, but we also present some additional details on how to apply the method in three dimensions, using Robin boundary conditions and to time-dependent problems.

We then apply the IIM to Boolean grids. These are grids that use combinations of a number of Cartesian grids to achieve greater accuracy of approximations, while using fewer grid nodes. The use of IIM on such grids requires some new development of the estimation of jumps on the boundary. We show numerically that the resulting Boolean IIM gives second order error convergence with respect to the smallest step size in the grid, meaning that the required number of nodes needed for a given maximal error is considerably smaller than on uniform Cartesian grids.

Contents

1. The Immersed Interface Method

1.1. Introduction

The use of finite difference methods for solving partial differential equations has a few advantages, mainly that they are easy to implement and that they may be easily and quickly solved using Fast Fourier Transform (FFT) methods. However, using the FFT requires a uniform grid on a rectangular domain, so it can not be immediately applied to problems in non-rectangular domains, resulting in longer solution times for such problems.

The immersed interface method overcomes this problem, in that it allows the solution and its derivatives to be discontinuous along interfaces. This makes it possible to immerse a boundary into a rectangular grid, let the solution be zero outside the boundary and apply boundary conditions to the solution along the boundary. The finite differences near the boundary are then corrected using the jumps in the solution at the boundary, so that the differences remain valid even though the solution is discontinuous. Therefore, the method makes it unnecessary to spend time on making grids adapted to the geometry and also allows the use of FFT-based methods even though the domain is not rectangular.

The ideas behind the immersed interface method were first conceived by Peskin [13] and used in computations on heart flows, with moving boundaries. The method was further developed by LeVeque and Li [11], [12] and used with finite differences on Cartesian grids. Finally, Wiegmann and Bube [16], [15] gave the method a clearer formulation and extended it to more general problems. They call their method the "explicit jump immersed interface method" or EJIIM, and that is the method we will be concerned with here.

We will first describe the general idea behind the EJIIM and then proceed to the problem of implementing the method. This will include treatment of three-dimensional problems and time-dependent problems, which is not included in the original article. We will also discuss some additional boundary conditions which were not discussed by Wiegmann and Bube.

This presentation contains no proofs of convergence or error estimates. Wiegmann and Bube provide proofs of convergence for the method in one dimension and also in two dimensions for the special case when the jumps at the boundary are known beforehand. Other proofs or error estimates are not known and here we only confirm numerical convergence.

1.2. EJIIM theory

The basic idea of the IIM is that standard finite differences, such as

$$u_{xx}(x_i) = \frac{u(x_i + h) - 2u(x_i) + u(x_i - h)}{h^2} + O(h^2) \tag{1.1}$$

are not valid for non-smooth functions, since they are based on Taylor-expansions. However, they may be corrected using the size and position of the discontinuities in u and its derivatives. Let us denote the jump in the m:th derivative in $u : \mathbb{R} \to \mathbb{R}$ at a point $\alpha \in \mathbb{R}$

by

$$[u^{(m)}]_\alpha = \lim_{x\to\alpha^+} u^{(m)}(x) - \lim_{x\to\alpha^-} u^{(m)}(x),$$

where, of course, $u^{(0)} = u$.

Following [16], we give two lemmas that contain the essence of the EJIIM. The proofs are essentially exercises in the use of Taylor expansions and we refer the reader to the original article [16].

Lemma 1.1 (cf. [16, lemma 1]) *Let $h > 0$ and assume $u^- \in C^{l+1}([\alpha - h, \alpha])$ and $u^+ \in C^{l+1}([\alpha, \alpha + h])$. Then let*

$$u(x) = \begin{cases} u^-(x) & \text{for } x \le \alpha, \\ u^+(x) & \text{for } x > \alpha. \end{cases}$$

For $x \in (\alpha - h, \alpha)$, we then have

$$u(x+h) = \sum_{k=0}^{l} \frac{h^k}{k!} u^{(k)}(x) + \sum_{k=0}^{l} \frac{(x+h-\alpha)^k}{k!} [u^{(k)}]_\alpha + O(h^{l+1}),$$

and

$$u(x) = \sum_{k=0}^{l} \frac{(-h)^k}{k!} u^{(k)}(x+h) - \sum_{k=0}^{l} \frac{(x-\alpha)^k}{k!} [u^{(k)}]_\alpha + O(h^{l+1}),$$

Using this lemma in the expressions for finite differences on a uniform 1d-grid $\{x_j\}$ with grid spacing h, we get the next lemma.

Lemma 1.2 (cf. [16, lemma 3]) *Let $x_j \le \alpha < x_{j+1}$ and suppose $u \in C^4([x_j - h, \alpha)) \cap C^4((\alpha, x_j + h])$, with derivatives extending continuously to α. Then the following approximations hold:*

$$u_x(x_j) = \frac{u(x_{j+1}) - u(x_{j-1})}{2h} - \frac{1}{2h} \sum_{m=0}^{2} \frac{(x_{j+1}-\alpha)^m}{m!} [u^{(m)}]_\alpha + O(h^2), \tag{1.2}$$

$$u_x(x_{j+1}) = \frac{u(x_{j+2}) - u(x_j)}{2h} - \frac{1}{2h} \sum_{m=0}^{2} \frac{(x_j-\alpha)^m}{m!} [u^{(m)}]_\alpha + O(h^2), \tag{1.3}$$

$$u_{xx}(x_j) = \frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1})}{h^2} - \frac{1}{h^2} \sum_{m=0}^{3} \frac{(x_{j+1}-\alpha)^m}{m!} [u^{(m)}]_\alpha + O(h^2), \tag{1.4}$$

$$u_{xx}(x_{j+1}) = \frac{u(x_{j+2}) - 2u(x_{j+1}) + u(x_j)}{h^2} + \frac{1}{h^2} \sum_{m=0}^{3} \frac{(x_j-\alpha)^m}{m!} [u^{(m)}]_\alpha + O(h^2). \tag{1.5}$$

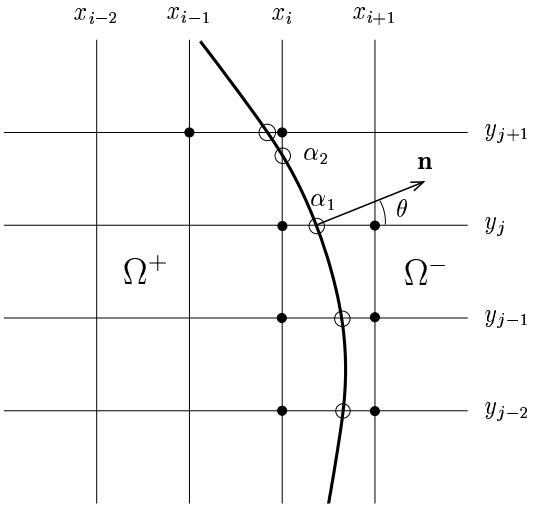

Figure 1.1: Immersed boundary in 2D. The immersed boundary intersects the grid at points $\alpha_1 = (x_{\alpha_1}, y_{\alpha_1})$ and $\alpha_2 = (x_{\alpha_2}, y_{\alpha_2})$. Interface intersection points (IIPs) are marked by '$\circ$' and anchor points are marked by '$\bullet$'.

The application of these results to higher dimensions is now straightforward. In the situation in figure 1.1, for example, the Laplacian at (x_i, y_j) would be approximated as

$$\begin{aligned}\Delta u(x_i, y_j) = &\frac{u(x_{i+1}, y_j) + u(x_{i-1}, y_j) + u(x_i, y_{j+1}) + u(x_i, y_{j-1}) - 4u(x_i, y_j)}{h^2} \\ &+ \frac{1}{h^2}\sum_{m=0}^{3} \frac{(x_{i+1} - x_{\alpha_1})^m}{m!}[u_x^{(m)}]_{\alpha_1} + \frac{1}{h^2}\sum_{m=0}^{3} \frac{(y_{j+1} - y_{\alpha_2})^m}{m!}[u_y^{(m)}]_{\alpha_2} + O(h^2).\end{aligned} \tag{1.6}$$

Here,

$$[u^{(m)}]_{\alpha_k} = u^{(m),+}(\alpha_k) - u^{(m),-}(\alpha_k) = \lim_{\substack{(x,y)\to\alpha_k \\ (x,y)\in\Omega^+}} u^{(m)}(x,y) - \lim_{\substack{(x,y)\to\alpha_k \\ (x,y)\in\Omega^-}} u^{(m)}(x,y),$$

with $\alpha_k = (x_{\alpha_k}, y_{\alpha_k})$, $k = 1, 2$, which in this case gives a different sign for the jump compared to the previous definition (for the one-dimensional case), accounting for the plus signs for the corrections in (1.6), while there is a minus sign in (1.4). This new definition makes it unnecessary to keep track of coordinate directions at the interface when adding corrections.

This means that on a uniform grid with an immersed boundary, the discrete Laplacian may be calculated as

$$\Delta_h U + \Psi C, \tag{1.7}$$

where Δ_h is the standard $N \times N$ finite difference matrix for the Laplacian, U is the vector of function values at the grid points, Ψ contains the coefficients for the correction terms from (1.6) and $C = ([u]_{\alpha_1}, [u^{(1)}]_{\alpha_1}, [u^{(2)}]_{\alpha_1}, \ldots)^T$ is a vector containing the jumps of u and its derivatives at the points where the interface intersects the grid. So (1.6) corresponds to one row of (1.7).

We will now consider the problem of solving the Poisson equation

$$\Delta u(x) = f(x) \qquad x \in \Omega^+, \tag{1.8}$$

where Ω^+ is a domain in $\mathbb{R}^2$ or $\mathbb{R}^3$. We wish to apply boundary conditions of the following types

$$\begin{aligned} [u](x) &= g(x), \\ \left[\frac{\partial u}{\partial n}\right](x) &= h(x), \\ \left[\frac{\partial u}{\partial n}\right](x) &= au^+(x) - bu^-(x), \end{aligned} \tag{1.9}$$

for $x \in \partial\Omega^+$, that is Dirichlet, Neumann or Robin boundary conditions.

In order to apply the EJIIM, we let the domain Ω^+ be embedded in a rectangular area discretized by a uniform grid and let Ω^- denote the domain outside the interface $\partial\Omega^+$. Setting $f(x) = 0$ in Ω^- and thus $u^- = 0$, the above boundary conditions become the regular Dirichlet, Neumann and Robin boundary conditions for u^+ in Ω^+. The corresponding linear system of equations becomes

$$\Delta_h U + \Psi C = F.$$

The problem here is that the jumps C are unknown, so we need to find an additional relation that determines these jumps. Some of the jumps are determined by the boundary conditions at $\Gamma = \partial\Omega^+$, while others must be determined in another way. This is done by creating an interpolation matrix D^T, which, given the function values, estimates the jumps at the boundary. Thus the entire equation system may be written

$$\begin{aligned} \Delta_h U + \Psi C &= F_1, \\ C &= D^T U + F_2, \end{aligned} \tag{1.10}$$

where F_1 contains function values of the right hand side f and F_2 contains the known jumps derived from boundary conditions.

1.3. ESTIMATING JUMPS AT BOUNDARIES

The jumps at the boundary are estimated by Lagrange interpolation, meaning that for each interface intersection point (IIP), i.e. for each point where the interface intersects the grid (see figure 1.1), we select a number of grid nodes on one side of the boundary and calculate the interpolating polynomial of degree d, say. This polynomial and its derivatives

are then evaluated at the boundary. If we do the same from the other side of the interface, we may take the difference and get an estimate of the jump at the IIP, given function values at the grid nodes. The number of grid nodes needed to construct a polynomial of degree d in n dimensions is $\binom{n+d}{d}$.

As an example, let us consider a second order polynomial in two dimensions (i.e. $d = n = 2$). We select $\binom{4}{2} = 6$ grid nodes $p_i = (x_i, y_i),\ i = 1, \ldots, 6$ with corresponding function values $u_1, \ldots, u_6$ and set up the equation system

$$\begin{bmatrix} 1 & x_1 & y_1 & x_1^2 & x_1y_1 & y_1^2 \\ 1 & x_2 & y_2 & x_2^2 & x_2y_2 & y_2^2 \\ 1 & x_3 & y_3 & x_3^2 & x_3y_3 & y_3^2 \\ 1 & x_4 & y_4 & x_4^2 & x_4y_4 & y_4^2 \\ 1 & x_5 & y_5 & x_5^2 & x_5y_5 & y_5^2 \\ 1 & x_6 & y_6 & x_6^2 & x_6y_6 & y_6^2 \end{bmatrix} \begin{bmatrix} a_{00} \\ a_{10} \\ a_{01} \\ a_{20} \\ a_{11} \\ a_{02} \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix} \tag{1.11}$$

in order to construct the polynomial

$$\mathcal{P}_2(x, y) = a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2.$$

We wish to solve this system for the coefficients a_{jk}, which leads to the requirement that the determinant of the 6×6-matrix in (1.11) (the so called Vandermonde determinant) must be non-zero. This in turn leads to restrictions on the choice of the grid nodes p_i. These restrictions were studied by Chui and Lai [2] for arbitrary n and d. They give a sufficient criterion for selection of nodes so that the Vandermonde determinant is non-zero, which they call *Node Configuration A*.

Definition 1.1 (Node Configuration A in $\mathbb{R}$) *Any set of distinct points in $\mathbb{R}$ satisfies Node Configuration A in $\mathbb{R}$.*

Definition 1.2 (Node Configuration A in $\mathbb{R}^n$) *Let $\hat{X}_d^n = \{x_1, \ldots, x_{N_d^n}\}$ be a set of $N_d^n = \binom{n+d}{d}$ distinct points in $\mathbb{R}^n$. $\hat{X}_d^n$ satisfies Node Configuration A in $\mathbb{R}^n$ if there exist $d+1$ hyperplanes $K_i^n, i = 0, \ldots, d$ with*

$$x_{N_{d-1}^n+1}, \ldots, x_{N_d^n} \in K_d^n$$

and

$$x_{N_{j-1}^n+1}, \ldots, x_{N_j^n} \in K_j^n \setminus (K_{j+1}^n \cup \ldots \cup K_d^n)$$

for $j = 0, \ldots, d-1$, and such that each set of points

$$\hat{X}_j^{n-1} = \{x_{N_{j-1}^n+1}, \ldots, x_{N_j^n}\}, \qquad 0 \le j \le d,$$

viewed as points in $\mathbb{R}^{n-1}$ satisfies Node Configuration A in $\mathbb{R}^{n-1}$.

Theorem 1.3 (cf. [2, theorem 4]) *If $\hat{X}_d^n$, with $n \ge 1$ and $d \ge 0$, satisfies Node Configuration A, then the corresponding Vandermonde determinant is non-zero.*

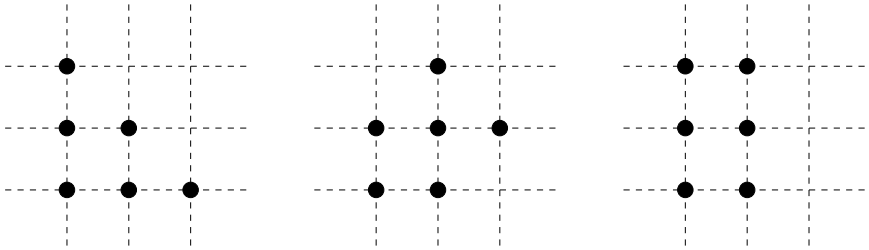

Figure 1.2: Examples of node configurations in two dimensions. The two on the left satisfy *Node Configuration A*, while the one on the right does not.

In the definition we use the convention $N^n_{-1} = 0$ for all n. Also note that $(N^n_j - N^n_{j-1}) = N^{n-1}_j$, so that the $(n-1)$-dimensional hyperplane K^n_j contains N^{n-1}_j points, which by an orthogonal transformation may be regarded as points in $\mathbb{R}^{n-1}$ for which the node configuration problem could again be posed. Therefore *Node Configuration A* is well defined.

The recursive definition of the node configuration may seem hard to interpret in all its generality, so let us see what it means in our above example with $n = 2$ and $d = 2$. In that case, the hyperplanes become lines, and the definition says that we should be able to select $d+1 = 3$ lines, such that $N^1_2 = \binom{1+2}{2} = 3$ points lie on the first one, $N^1_1 = \binom{1+1}{1} = 2$ on the second (but not on the first) and $N^1_0 = \binom{1+0}{0} = 1$ on the third. Furthermore the set of points on each of these lines should satisfy *Node Configuration A*, which they do automatically, since they are on one-dimensional sets. Two valid node configurations and one invalid are shown in figure 1.2.

In three dimensions, the hyperplanes become planes, and for a second order polynomial, we need to be able to choose 3 planes, containing 6, 3 and 1 nodes respectively, making $\binom{3+2}{2} = 10$ nodes in total. And on each of the planes, the nodes should satisfy *Node Configuration A* in two dimensions.

1.4. APPLYING BOUNDARY CONDITIONS

Having seen how to create interpolating polynomials, we return to the problem of estimating the jumps in the vector C in (1.10), i.e. we need to determine the matrix D^T and the vector F_2 in the equation

$$C = D^T U + F_2.$$

Here, we do this in three dimensions, since the two-dimensional problem is presented in [16], while three-dimensional problems are not treated. In order to determine D^T, we first find all the interface intersection points (IIPs). For each of these points, there is a coordinate direction $X_j = x, y$ or z, which is the direction of the grid line on which the IIP is located. Along this grid line, there are two anchor points, one on each side of the

interface, which are the grid nodes closest to the IIP and whose finite difference stencils are affected by the discontinuity at the IIP (see figure 1.1). We denote these p_+ and p_-.

Now, for each IIP α_j, we see from lemma 1.2 that we need to estimate the jumps

$$[u]_j = u^+(\alpha_j) - u^-(\alpha_j), \quad [u_{X_j}]_j = u^+_{X_j}(\alpha_j) - u^-_{X_j}(\alpha_j), \quad [u_{X_jX_j}]_j = u^+_{X_jX_j}(\alpha_j) - u^-_{X_jX_j}(\alpha_j)$$

in order to get accurate finite difference approximations of order $O(h)$. We will see from numerical results that it is in fact enough to have $O(h)$ approximations at the boundary in order to get overall $O(h^2)$-convergence. This behavior is confirmed for all IIM methods and may be loosely motivated by the fact that the boundary is a lower-dimensional set so that in three dimensions, for example, the number of points with $O(h)$ errors is $O(h^{-2})$ rather than $O(h^{-3})$. One may also use $O(h^2)$ estimates at the boundary. This still gives overall $O(h^2)$ errors for the solution, although the actual errors may be smaller (see [16, example 1]).

In order to approximate the jumps, we select stencils of grid nodes around each of the anchor points according to *Node Configuration A* in the previous section and define matrices $\mathcal{P}_{j+}$ and $\mathcal{P}_{j-}$ in analogy with the matrix in (1.11), using coordinates relative to p_+ and p_-, respectively. We also define restriction matrices $\mathcal{R}_{j+}$ and $\mathcal{R}_{j-}$, consisting of only 1's and 0's and serving only to select and reorder the grid nodes used in the stencils. In this way, we see that $\mathcal{P}^{-1}_{j+}\mathcal{R}_{j+}U$ is a vector containing the coefficients of an interpolating polynomial near α_j on the $+$-side of the interface. Using this we may estimate

$$\begin{bmatrix} u^+(\alpha_j) \\ u^+_x(\alpha_j) \\ u^+_{xx}(\alpha_j) \end{bmatrix} \approx \begin{bmatrix} 1 & h^+_x & h^+_y & h^+_z & (h^+_x)^2 & (h^+_y)^2 & (h^+_z)^2 & h^+_x h^+_y & h^+_x h^+_z & h^+_y h^+_z \\ 0 & 1 & 0 & 0 & 2h^+_x & 0 & 0 & h^+_y & h^+_z & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \mathcal{P}^{-1}_{j+}\mathcal{R}_{j+}U,$$

where we have set $h^+_x = x_{\alpha_j} - x_{p_+}$, etc, and with similar expressions for derivatives in other coordinate directions. Naming the matrix of coefficients on the left $\mathcal{Q}_{j+}$, we may now write

$$\begin{bmatrix} u^+(\alpha_j) \\ u^+_x(\alpha_j) \\ \vdots \\ u^-_z(\alpha_j) \\ u^-_{zz}(\alpha_j) \end{bmatrix} \approx \begin{bmatrix} \mathcal{Q}_{j+} & 0 \\ 0 & \mathcal{Q}_{j-} \end{bmatrix} \begin{bmatrix} \mathcal{P}^{-1}_{j+} & 0 \\ 0 & \mathcal{P}^{-1}_{j-} \end{bmatrix} \begin{bmatrix} \mathcal{R}_{j+} \\ \mathcal{R}_{j-} \end{bmatrix} U \equiv \mathcal{Q}_j \mathcal{P}_j^{-1} \mathcal{R}_j U. \tag{1.12}$$

It is now time to incorporate the boundary conditions (1.9) into the equation for C. This is done by creating a matrix $\mathcal{L}_j$ and a vector $F_{2,j}$. The vector $F_{2,j}$ contains the known jump information derived from the boundary conditions, while the matrix $\mathcal{L}_j$ defines linear combinations of estimated function values at α_j giving the contributions to the jumps that have to be estimated from the solution U.

In the case where Ω^+ is immersed into a rectangular domain and $u = 0$ outside Ω^+, then of course $u^- = 0$ everywhere along with its derivatives, so it is not necessary to estimate $u^-(\alpha_j)$. However, by describing the method in this more general context, it may also be applied when the interface is an interface between two domains where u is nonzero.

In the case of Dirichlet boundary conditions, this is very simple. Suppose we have a boundary condition that says $[u] = g(x)$ on the interface. Then, if for example the grid intersects the interface in the x-direction at α_j, we write

$$\begin{bmatrix} [u]_j \\ [u_x]_j \\ [u_{xx}]_j \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & -1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & -1 & \cdots & 0 \end{bmatrix} \mathcal{Q}_j \mathcal{P}_j^{-1} \mathcal{R}_j U + \begin{bmatrix} g(\alpha_j) \\ 0 \\ 0 \end{bmatrix}, \tag{1.13}$$

where $F_{2,j}$ is the vector on the right and $\mathcal{L}_j$ is the leftmost matrix, so that we get

$$\begin{bmatrix} [u]_j \\ [u_x]_j \\ [u_{xx}]_j \end{bmatrix} = \mathcal{L}_j \mathcal{Q}_j \mathcal{P}_j^{-1} \mathcal{R}_j U + F_{2,j} = \begin{bmatrix} g(\alpha_j) \\ u_x^+(\alpha_j) - u_x^-(\alpha_j) \\ u_{xx}^+(\alpha_j) - u_{xx}^-(\alpha_j) \end{bmatrix}.$$

Now setting $D_j^T = \mathcal{L}_j \mathcal{Q}_j \mathcal{P}_j^{-1} \mathcal{R}_j$, we have

$$C_j = D_j^T U + F_{2,j}. \tag{1.14}$$

By just stacking the contributions from all IIPs, we arrive at (1.10).

For other boundary conditions, things are more complicated. Consider for example Neumann conditions, $\partial u/\partial n = h(x)$ on the interface. Then we are interested in derivatives in the coordinate directions, while the boundary condition is given in the normal direction of the interface. We define the normal direction through the two angles (θ, φ), $0 \le \theta \le \pi$, $0 \le \varphi < 2\pi$, as in spherical coordinates and introduce a new Cartesian coordinate system (ξ, η, ζ) with origin at the IIP, which is oriented so that the ξ-axis points in the normal direction, the η-axis is perpendicular to the ξ-axis and lies in the (x, y)-plane, and the ζ-axis is perpendicular to the other two and has positive z-component. This coordinate transformation yields

$$\begin{bmatrix} [u_x]_j \\ [u_y]_j \\ [u_z]_j \end{bmatrix} = \begin{bmatrix} \sin\theta\,\cos\varphi & -\sin\varphi & -\cos\theta\,\cos\varphi \\ \sin\theta\,\sin\varphi & \cos\varphi & -\cos\theta\,\sin\varphi \\ \cos\theta & 0 & \sin\theta \end{bmatrix} \begin{bmatrix} [u_\xi]_j \\ [u_\eta]_j \\ [u_\zeta]_j \end{bmatrix} \tag{1.15}$$

$$\begin{bmatrix} [u_\xi]_j \\ [u_\eta]_j \\ [u_\zeta]_j \end{bmatrix} = \begin{bmatrix} \sin\theta\,\cos\varphi & \sin\theta\,\sin\varphi & \cos\theta \\ -\sin\varphi & \cos\varphi & 0 \\ -\cos\theta\,\cos\varphi & -\cos\theta\,\sin\varphi & \sin\theta \end{bmatrix} \begin{bmatrix} [u_x]_j \\ [u_y]_j \\ [u_z]_j \end{bmatrix} \tag{1.16}$$

By the boundary condition, we know that $[u_\xi]_j = h(\alpha_j)$, so using (1.15) we get $[u_x]_j$, $[u_y]_j$ and $[u_z]_j$ expressed in $h(\alpha_j)$, $[u_\eta]_j$ and $[u_\zeta]_j$, which, using (1.16) may be expressed in the quantities $u_x^+(\alpha_j) - u_x^-(\alpha_j)$, $u_y^+(\alpha_j) - u_y^-(\alpha_j)$ and $u_z^+(\alpha_j) - u_z^-(\alpha_j)$, which we get from (1.12). Thus, we get for example

$$\begin{aligned} [u_x]_j = {} & \sin\theta\,\cos\varphi\; h(\alpha_j) + (\sin^2\varphi + \cos^2\theta\,\cos^2\varphi)(u_x^+(\alpha_j) - u_x^-(\alpha_j)) \\ & + \sin\varphi\,\cos\varphi\,(\cos^2\theta - 1)(u_y^+(\alpha_j) - u_y^-(\alpha_j)) - \cos\theta\,\sin\theta\,\cos\varphi\,(u_z^+(\alpha_j) - u_z^-(\alpha_j)), \end{aligned}$$

where the first term is known and therefore included in $F_{2,j}$, while the coefficients of the second term are included in $\mathcal{L}_j$, so that we get (1.14) once more. The expressions for $[u_y]_j$ and $[u_z]_j$ are derived similarly, while $[u]_j = u^+(\alpha_j) - u^-(\alpha_j)$ and $[u_{X_j X_j}]_j = u^+_{X_j X_j}(\alpha_j) - u^-_{X_j X_j}(\alpha_j)$ for $i = 1, 2, 3$.

Finally, we consider Robin boundary conditions, i.e. $\partial u/\partial n = au^+ - bu^-$. Using the above notation, this leads to the expression

$$u^+_\xi = u^-_\xi = au^+ - bu^- = a[u] - (b-a)u^-,$$

which gives us

$$[u]_j = \frac{1}{a} u^+_\xi(\alpha_j) + (\frac{b}{a} - 1)u^-(\alpha_j).$$

Furthermore, we need to impose the implicit condition that $u^+_\xi = u^-_\xi$, which we impose by setting $h(\alpha_j) = 0$ in the Neumann case above. In summary, this leads to the above expression for $[u]_j$, expressions similar to the Neumann case (but with $F_{2,j} = 0$) for $[u_{X_j}]_j$ and the same expressions as before for $[u_{X_j X_j}]_j$.

This way, we construct the sparse matrix D^T and the vector F_2 by stacking the contributions from each IIP. If we have multiple interfaces with different boundary conditions, we just include the IIPs from all the interfaces and use the corresponding boundary condition for each IIP according to the above treatment.

Finally, we need to construct the sparse matrix Ψ in (1.10). This matrix contains the coefficients of the jumps in (1.2) to (1.5), placed at appropriate positions, so that the corrections affect the corresponding anchor points. Since the corrections are additive, the Ψ-matrix is easy to construct.

We also remark that it is possible to apply other types of boundary conditions, such as in composite material problems illustrated in [16], by the same method as shown here.

1.5. SOLVING THE LINEAR SYSTEM

Through the previous sections, we have arrived at the system of equations

$$\begin{aligned} \Delta_h U + \Psi C &= F_1, \\ C &= D^T U + F_2, \end{aligned}$$

which is (1.10). Solving for U in the first equation and inserting this into the second, we get

$$\begin{aligned} U &= \Delta_h^{-1}(-\Psi C + F_1), \\ (I + D^T \Delta_h^{-1} \Psi) C &= D^T \Delta_h^{-1} F_1 + F_2. \end{aligned}$$

Here, the second equation is a 'small' system of equations for C with $3N_{IIP}$ rows (where N_{IIP} is the number of IIPs). It may be solved using iterative methods to avoid forming the matrix on the left explicitly, which would be unfeasibly memory-consuming for three-dimensional problems, since the matrix is not sparse. We use the BiCGStab method (see

e.g. [10], [9]) to solve the system and since we then only need to be able to form the matrix-vector product $(I + D^T\Delta_h^{-1}\Psi)C$, our only remaining problem is to apply Δ_h^{-1}. Since our domain is embedded in a rectangular uniform grid, we may use the Fast Fourier Transform to achieve this efficiently (see e.g. [14, chapter 19]). If N is the total number of grid nodes, the FFT is applied in $O(N \log N)$ operations and this is the most time-consuming part of each iteration in the BiCGStab algorithm. We have used the library FFTW [5] to get a fast implementation of the FFT in C.

Having solved for C, all that remains is to compute U from the first equation by applying Δ_h^{-1} once more.

1.6. Other equations

The immersed interface method may also be applied to other types of equations. First of all, it is straightforward to apply it to the heat equation

$$u_t(x,t) - d\Delta u(x,t) = f(x,t), \qquad x \in \Omega^+,\ t \geq 0,$$

with suitable boundary conditions, using for example an implicit Euler scheme in time. In that case we get

$$\begin{aligned}(I/\Delta t - d\Delta_h)U_{n+1} - d\Psi C_{n+1} &= U_n/\Delta t + F_1,\\ C_{n+1} &= D^T U_{n+1} + F_2,\end{aligned}$$

and

$$\begin{aligned}U_{n+1} &= (I/\Delta t - d\Delta_h)^{-1}(d\Psi C_{n+1} + U_n/\Delta t + F_1),\\ (I - D^T(I/\Delta t - d\Delta_h)^{-1}d\Psi)C_{n+1} &= D^T(I/\Delta t - d\Delta_h)^{-1}(F_1 + U_n/\Delta t) + F_2,\end{aligned}$$

where the operator $(I/\Delta t - d\Delta_h)^{-1}$ may be applied using FFT and the second equation solved using BiCGStab for each time step.

It is also possible to solve equations in multiple domains with boundary conditions between the domains or with different coefficients in the equation in different domains. The technique is exactly as described in the previous sections. See section 1.7 and [16] for examples.

Other combinations of spatial derivatives than the Laplacian may also be considered, although they are not discussed here. It should be clear, however, from the discussion above how the immersed interface method should be applied to these cases.

It is also in principle possible to use the method for moving boundary problems, where one would like to avoid costly grid generation. It would then, however, be necessary to compute the matrices D^T and Ψ at every time step.

1.7. Examples

Our first example is solving the Poisson equation inside an ellipsoid E with half-axes 0.44, 0.3 and 0.3, centered at $(0.5, 0.5, 0.5)$. The equation we solve is given in spherical

coordinates as

$$\begin{aligned} \Delta u(r,\theta,\varphi) &= f(r,\theta,\varphi), & (r,\theta,\varphi) &\in E, \\ u(r,\theta,\varphi) &= g(r,\theta,\varphi), & (r,\theta,\varphi) &\in \partial E, \end{aligned}$$

where $f(r,\theta,\varphi) = -5\sin(3\varphi)$ and $g(r,\theta,\varphi) = r^2\sin^2\theta\sin(3\varphi)$, which gives the solution $u(r,\theta,\varphi) = r^2\sin^2\theta\sin(3\varphi)$ in E. The solution and the errors are shown in figure 1.3. We see that this non-trivial problem can be solved with reasonable accuracy, even on the rather coarse $40\times40\times40$ grid used here. The solution took less than a second to compute on a standard computer. The errors are of order $O(h^2)$, which is shown more clearly for another example in section 2.4.3.

The second example shows that more complicated equations and boundary conditions may be solved using the immersed interface method. We solve the heat equation in two dimensions. The domain consists of two concentric circular discs, C_1 and C_2, with C_1 inside C_2. We apply Neumann boundary conditions at ∂C_2 and Robin boundary conditions at ∂C_1. The problem may be written as

$$\begin{aligned} \frac{\partial u}{\partial t}(x,y,t) - \Delta u(x,y,t) &= 0, & (x,y) &\in C_1 \cup C_2,\ t \geq 0, \\ \frac{\partial u}{\partial n}(x,y,t) &= 0, & (x,y) &\in \partial C_2,\ t \geq 0, \\ \frac{\partial u}{\partial n}(x,y,t) &= 10u_2(x,y) - 3u_1(x,y), & (x,y) &\in \partial C_1,\ t \geq 0, \\ u(x,y,0) &= u_0(x,y), & (x,y) &\in C_1 \cup C_2. \end{aligned}$$

Here, $u_0(x,y)$ is a bell-shaped function inside C_1 shown at the top of figure 1.4. u_2 and u_1 are the values of u at ∂C_1, on the outer and inner side of the boundary.

The method used is the one described in section 1.6 with implicit Euler time stepping. The boundary conditions are implemented as described earlier.

Obviously, the exact solution is not known for this problem, so we have nothing to compare to. However, the example shows that it is possible to solve multiple boundary, time-dependent problems with a range of boundary conditions using the immersed interface method. Here, for simplicity, we have used only circular and elliptical domains, but in the method itself there are no limitations on the shape of the domain. Further examples are given by Wiegmann and Bube [16] and in part I of this thesis.

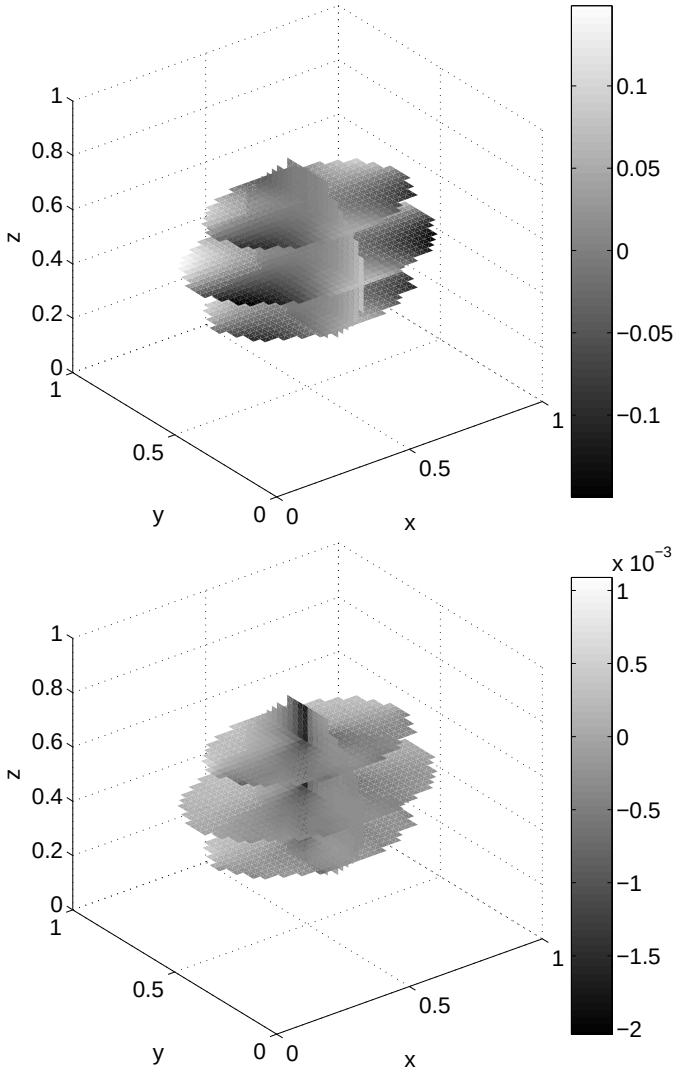

Figure 1.3: EJIIM for the Poisson equation inside an ellipsoid in 3D. The solution was computed on a $40 \times 40 \times 40$ grid and is shown at the top. Errors compared to the exact solution are shown at the bottom. Values are shown on selected slices through the domain.

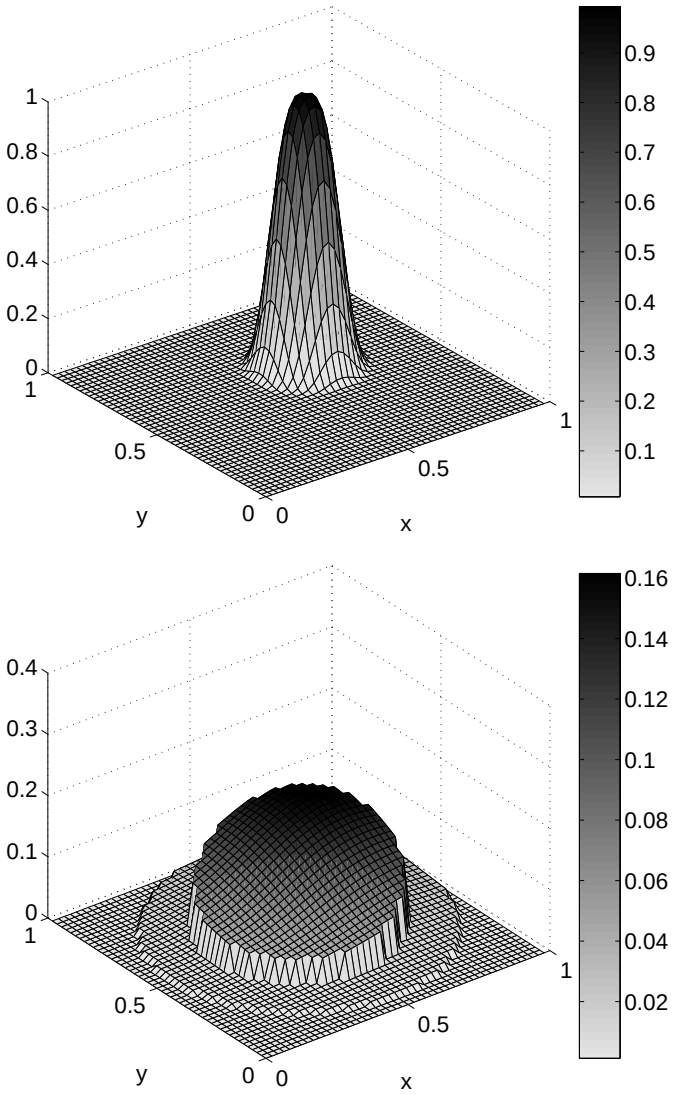

Figure 1.4: EJIIM for the heat equation inside two concentric discs. The initial data is shown at the top and the solution at $t = 0.03$ below. The domain boundaries are clearly visible in the second figure.

2. The Immersed Interface Method on Boolean grids

2.1. Introduction

There is a constant desire to make computations as fast and efficient as possible, in order to be able to solve larger problems with higher accuracy, or just to minimize the time spent waiting for the computer to carry out the calculations. The Boolean grids presented here provide one method to make computations faster by using less data, while still obtaining the same accuracy in the calculations.

The Boolean methods were originally developed in the 1960's in order to represent surfaces used in computer-aided design (CAD). They were first used by Coons [3] to create interpolatory surfaces, coinciding with prescribed values on the boundary of the unit square. The theory was then developed in a series of articles by Gordon [7], [6], resulting in an abstract theory of commutative projectors which is presented here in section 2.2.1. There is also a book by Delvos [4], where the methods are presented in some detail.

As we will see, it is straightforward to use finite differences on Boolean grids and one may use FFT-based methods to solve PDEs on such grids, gaining several orders of magnitude in the number of points needed for a given accuracy. However, just as for uniform grids, these methods can only be applied to rectangular regions. Therefore it is interesting to apply the Immersed Interface Method from the previous section to Boolean grids in order to get a similar decrease in computational time even for problems in irregularly shaped domains.

The outline of our presentation is as follows. In section 2.2, we present the abstract theory of Boolean interpolation and construct Boolean grids on which we apply finite differences and the immersed interface method. In section 2.3, we discuss how to extend the immersed interface method to work on the Boolean grids and finally, in section 2.4, we give some numerical examples of the use of Boolean grids for interpolation and equation solving.

As for the IIM on uniform grids, we do not have any proofs of convergence or error estimates for the IIM on Boolean grids. We only confirm numerical convergence and superiority to the uniform IIM.

2.2. Boolean interpolation

2.2.1. The algebraic theory

The following presentation is taken from Gordon [6] and is an abstract algebraic approach to approximation theory, giving a motivation for the use of Boolean approximations.

We consider a linear space $\mathcal{F}$ consisting of functions. On this space, we define M projectors P_j, $j = 1, \ldots, M$, meaning that $P_j : \mathcal{F} \to \Phi_j$ is a linear transformation with the property $P_j P_j = P_j$. Here Φ_j is a subspace of $\mathcal{F}$ for $j = 1, \ldots, M$.

A function $\pi = P_j f \in \Phi_j$ is called the approximation of $f \in \mathcal{F}$, and the function $f - \pi$ is called the remainder.

We define multiplication and addition of projectors in the natural way and note that the associative and distributive rules hold for projectors defined on the same domain $\mathcal{F}$. Furthermore, we assume that multiplication is commutative, i.e.

$$P_j P_k = P_k P_j \quad \text{for all } j,k = 1,\ldots,M.$$

It is obvious that the product of two commutative projectors, $A = P_j P_k$, is again a projector, since $AA = A$. This is not true, however, for the sum of two projectors, since

$$(P_j + P_k)(P_j + P_k) = P_j P_j + P_j P_k + P_k P_j + P_k P_k = P_j + P_k + 2P_j P_k \neq P_j + P_k.$$

Therefore, we introduce the *Boolean addition*, denoted by $\oplus$ and defined by

$$P_j \oplus P_k = P_j + P_k - P_j P_k. \tag{2.1}$$

It is easy to check that $P_j \oplus P_k$ is again a projector.

We also need to compare projectors in order to decide which are better than others. To this end, we introduce the ordering relation $\leq$ defined by

$$P_j \leq P_k \Leftrightarrow P_j P_k = P_j, \tag{2.2}$$

that is, if $P_j \leq P_k$, P_j removes at least as much of the function f as P_k does, or $\Phi_j \subset \Phi_k$.

Now, we may define the space Ψ as the set of all projectors which can be built up as combinations of the P_j, $j = 1,\ldots,M$, under the operations of multiplication and Boolean addition. The set Ψ is now a *distributive lattice* under the partial ordering $\leq$. This means that for all projectors $A, B, C \in \Psi$, the following properties hold (and are easily checked):

i)	reflexivity	$A \leq A$
ii)	anti-symmetry	$A \leq B$ and $B \leq A \Rightarrow A = B$
iii)	transitivity	$A \leq B$ and $B \leq C \Rightarrow A \leq C$
iv)	idempotence	$A \oplus A = A$ and $AA = A$
v)	commutativity	$A \oplus B = B \oplus A$ and $AB = BA$
vi)	associativity	$A(BC) = (AB)C$ and $A \oplus (B \oplus C) = (A \oplus B) \oplus C$
vii)	distributivity	$A(B \oplus C) = AB \oplus AC$ and $A \oplus (BC) = (A \oplus B)(A \oplus C)$
viii)	consistency	$A \leq B \Leftrightarrow AB = A \Leftrightarrow A \oplus B = B$

(2.3)

It is a property of every lattice that any pair $\{A, B\}$ of elements has both a *least upper bound* (denoted sup), that is the least element C such that $C \geq A$ and $C \geq B$, and a *greatest lower bound* (denoted inf), that is the largest element D such that $D \leq A$ and $D \leq B$. These are given explicitly by

$$\sup\{A, B\} = A \oplus B,$$
$$\inf\{A, B\} = AB.$$

It follows that every finite lattice has a unique *maximal element* $\mathcal{M} \in \Psi$, that is an element satisfying $A \leq \mathcal{M}$ for all $A \in \Psi$, and a unique *minimal element* $\mathcal{L}$, which satisfies $\mathcal{L} \leq A$ for all $A \in \Psi$. It is easily seen that

$$\mathcal{M} = \sup\{P_j\}_{j=1}^{M} = P_1 \oplus P_2 \oplus \cdots \oplus P_M,$$
$$\mathcal{L} = \inf\{P_j\}_{j=1}^{M} = P_1 P_2 \cdots P_M.$$

It is also interesting to study the range of the projectors in Ψ. It is clear that the range of $P_j \oplus P_k$ is $\Phi_j \oplus \Phi_k$ and that the range of $P_j P_k$ is $\Phi_j \cap \Phi_k$. Thus, the range of $\mathcal{M}$ is the largest space formed from the Φ_j:s, namely $\Phi_1 \oplus \cdots \oplus \Phi_M$ and the range of $\mathcal{L}$ is the smallest space, $\Phi_1 \cap \cdots \cap \Phi_M$.

Finally, we introduce the *remainder operator* or the complement of a projector A, namely

$$A' \equiv I - A,$$

where I is the identity operator. It is clear that A' is a projector, and that

$$AA' = A'A = 0, \qquad A \oplus A' = I.$$

Now for each P_j, we set $R_j \equiv P_j' = I - P_j$ and note that although in general $R_j \notin \Psi$, the set of combinations of these remainder operators also form a distributive lattice, denoted Ψ'. Also, if we allow the three operators multiplication, Boolean addition and complement to work on the projectors, we generate a Boolean algebra, where additionally de Morgan's laws hold

$$(A \oplus B)' = A'B', \qquad (AB)' = A' \oplus B'.$$

Now, the final statement is that given a commutative set of projectors $\{P_j\}_{j=1}^{M}$, the identity operator, I has a *maximal decomposition*

$$I = \mathcal{M} \oplus \mathcal{M}' = \mathcal{M} + \mathcal{M}' = P_1 \oplus \cdots \oplus P_M + (R_1 R_2 \cdots R_M), \tag{2.4}$$

and a *minimal decomposition*

$$I = \mathcal{L} \oplus \mathcal{L}' = \mathcal{L} + \mathcal{L}' = P_1 P_2 \cdots P_M + (R_1 \oplus \cdots \oplus R_M). \tag{2.5}$$

Here, $R_1 R_2 \cdots R_M = \inf\{R_j\}_{j=1}^{M} \in \Psi'$ and $R_1 \oplus \cdots \oplus R_M = \sup\{R_j\}_{j=1}^{M} \in \Psi'$ as before. This means that by choosing the algebraically maximal projector $\mathcal{M}$, we minimize the remainder and vice versa.

2.2.2. BOOLEAN GRIDS

We now wish to apply the abstract results from the previous section to the problem of creating a grid on which we wish to apply finite difference solvers. This could be done in many ways, depending on how one defines the projectors $P_1, \ldots, P_M$. Here, we only discuss the grids that we actually use and the reason for using those will be clear later on.

First, assume a two-dimensional problem, where we want to approximate the smooth function $f(x,y)$ on a rectangle $\Omega = [a_x, b_x] \times [a_y, b_y]$. We introduce step sizes $h_{1,x} = (b_x - a_x)/(N_{1,x} - 1)$ and $h_{1,y} = (b_y - a_y)/(N_{1,y} - 1)$, interpolation points $\{x_i\}_{i=1}^{N_{1,x}}$, $\{y_i\}_{j=1}^{N_{1,y}}$, with $x_i = a_x + h_{1,x}(i-1)$, $y_j = a_y + h_{1,y}(j-1)$, and define the interpolating projectors P_x^1 and P_y^1 by

$$[P_x^1 f](x,y) = \sum_{i=1}^{N_{1,x}} f(x_i, y)\varphi_i(x),$$
$$[P_y^1 f](x,y) = \sum_{j=1}^{N_{1,y}} f(x, y_j)\psi_j(y),$$

where $\varphi_i(x)$ and $\psi_j(y)$ are piecewise linear hat functions, satisfying $\varphi_i(x_k) = \delta_{ik}$ and $\psi_j(y_k) = \delta_{jk}$. Using these two projectors, we may define the algebraically maximal and minimal projectors as

$$\mathcal{M}_{xy} = P_x^1 \oplus P_y^1 = P_x^1 + P_y^1 - P_x^1 P_y^1,$$
$$\mathcal{L}_{xy} = P_x^1 P_y^1.$$

We see that the minimal projector $\mathcal{L}_{xy}$ is the regular interpolation operator, that interpolates $f(x,y)$ only on the nodes (x_i, y_j). The maximal projector, however, is different. It interpolates $f(x,y)$ along the lines $x = x_i$ and $y = y_j$ and since the values at the nodes (x_i, y_j) are included both by P_x^1 and P_y^1, one occurrence is removed by subtracting $P_x^1 P_y^1$.

Since $P_x^1 f$ is piecewise linear in x, we see that $P_x^1 f$ approximates $f(x,y)$ to an accuracy of $O(h_{1,x}^2)$, that is

$$R_x^1 f = (I - P_x^1)f = O(h_{1,x}^2).$$

This obviously holds for P_y^1 too, and we may therefore deduce from (2.4) that the remainder of the maximal projector $\mathcal{M}_{xy}$ is given by

$$\mathcal{M}'_{xy} f = (I - \mathcal{M}_{xy})f = R_x^1 R_y^1 f = O(h_{1,x}^2 h_{1,y}^2),$$

while for the minimal projector the remainder is

$$\mathcal{L}'_{xy} f = (I - \mathcal{L}_{xy})f = (R_x^1 \oplus R_y^1)f = (R_x^1 + R_y^1 - R_x^1 R_y^1)f = O(h_{1,x}^2 + h_{1,y}^2).$$

We see that the maximal projector has a fourth order error, while the standard interpolation operator, which is minimal, has a second order error. This can also be easily checked by using Taylor expansions of f. One then sees that for the maximal operator, the second order terms are eliminated by subtracting the term $P_x^1 P_y^1 f$, so that only fourth and higher order terms remain.

There is, however, a difference between $\mathcal{M}_{xy}$ and $\mathcal{L}_{xy}$ in that $\mathcal{M}_{xy} f$ is not fully discretized, since it still has a continuous variable along the lines $x = x_i$ and $y = y_j$. But along these lines, we can afford a finer discretization. We introduce new (smaller) step

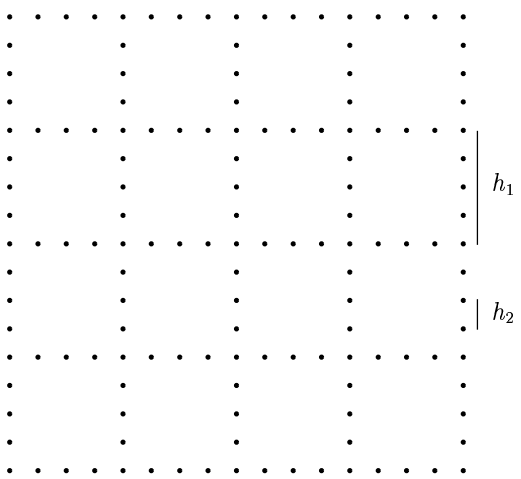

Figure 2.1: The two-dimensional Boolean grid with step sizes h_1 and h_2. The grid consists of three parts, the horizontal lines (G_{21}), the vertical lines (G_{12}) and the coarse grid (G_{11}) consisting of the intersection nodes of these lines. These three are combined as in (2.6) below to create the Boolean approximation.

sizes $h_{2,x}$ and $h_{2,y}$ and projectors P_x^2 and P_y^2, identical to P^1 except for the step size. Since the error in $P_x^2 f$ is of order $h_{2,x}^2$, we see that to preserve the $O(h_{1,x}^2 h_{1,y}^2)$ error estimate, we should set

$$h_{2,x} = h_{2,y} = h_{1,x} h_{1,y},$$

or $h_2 = h_1^2$ if the step sizes are equal in the two directions. This results in a grid like the one in figure 2.1. It is important to ensure that the nodes for P^1 form a subset of the nodes for P^2, so that P^1 and P^2 commute and that $P^1 \leq P^2$.

The resulting projector may be expressed in several ways, which are all equal, as can be shown by direct calculation using the rules (2.3) and the fact that $P_x^1 \leq P_x^2$ and $P_y^1 \leq P_y^2$. Denoting the resulting projector by $\mathcal{P}_{xy}$, we have

$$\mathcal{P}_{xy} = P_x^1 P_y^2 \oplus P_x^2 P_y^1 = P_x^2 P_y^2 (P_x^1 \oplus P_y^1) = P_x^2 P_y^1 + P_x^1 P_y^2 - P_x^1 P_y^1. \tag{2.6}$$

The last form shows the three uniform grids which we must combine in order to get the Boolean approximation, namely the grids G_{21} with steps $h_{2,x}$ and $h_{1,y}$, G_{12} with steps $h_{1,x}$ and $h_{2,y}$ and the coarse grid G_{11} with steps $h_{1,x}$ and $h_{1,y}$. Using the second form of $\mathcal{P}_{xy}$ and de Morgan's laws, we may also express the remainder as

$$\mathcal{R}_{xy} = R_x^2 \oplus R_y^2 \oplus R_x^1 R_y^1 = R_x^2 + R_y^2 + R_x^1 R_y^1 - R_x^2 R_y^1 - R_x^1 R_y^2,$$

where we again see that the resulting error will be of order $O(h_2^2 + h_1^4)$, so that a choice of $h_2 = h_1^2$ is optimal.

There is another way to view the error cancellation property of the Boolean grids, as is shown by Bungartz et al. [1]. This view is based on an error splitting for the approximation of the form

$$u^{h_x,h_y} - u = e_x(h_x) + e_y(h_y) + R(h_x, h_y),$$

where e_x only depends on h_x, x and y, e_y depends only on h_y, x and y and $|R(h_x, h_y)| \leq c(h_x h_y)^\nu$ for some constants c and ν independent of x and y. We can now study combinations as the ones above, namely

$$\hat{u}^{h_x,h_y} = u^{\alpha h_x,h_y} + u^{h_x,\alpha h_y} - u^{h_x,h_y}$$

for some $\alpha > 0$. Then, the resulting error is

$$\begin{aligned}\hat{u}^{h_x,h_y} - u &= e_x(\alpha h_x) + e_y(h_y) + R(\alpha h_x, h_y) + e_x(h_x) + e_y(\alpha h_y) + R(h_x, \alpha h_y) \\ &\quad - e_x(h_x) - e_y(h_y) - R(h_x, h_y) \\ &= e_x(\alpha h_x) + e_y(\alpha h_y) + \widehat{R}(h_x, h_y).\end{aligned}$$

Now, if e_x and e_y are the dominating terms and α is small, then we have reduced the error. For example, if $e_x(h_x) = O(h_x^2)$, $e_x(h_y) = O(h_y^2)$ and $\nu = 2$ as is the case for the projections above, and if $\alpha = h_x = h_y$, then $\hat{u}^{h,h} - u = e_x(h^2) + e_y(h^2) + \widehat{R}(h, h) = O(h^4)$, while $u^{h,h} - u = e_x(h) + e_y(h) + \widehat{R}(h, h) = O(h^2)$ exactly as above. This idea may be pursued further, taking combinations of a series of grids to eliminate errors while using as few points as possible. This *combination technique* was introduced by Griebel et al. [8] and is summarized in [1].

Assuming $N_{1,x} = N_{1,y} = N_1$ and $N_{2,x} = N_{2,y} = N_2$, the number of nodes used in the Boolean approximation (2.6) is

$$N_B = 2N_1N_2 + N_1^2 = O(N_1^3),$$

if we choose N_2 proportional to N_1^2. This should be compared to the case of a standard grid, where the number of points needed to achieve the same order of accuracy is

$$N_S = N^2 = N_2^2 = O(N_1^4),$$

since we must choose N of the same magnitude as N_2 in order to get an error of order h_2^2. We therefore see that by using a Boolean grid, we save one order of magnitude in N_1 while retaining the same accuracy of approximation.

So far for two-dimensional problems. In higher dimensions, one would expect even greater gains and we will now derive an approximation scheme and show that it is indeed so. We consider the same problem in three dimensions. As before, given a smooth function f defined on a box $\Omega = [a_x, b_x] \times [a_y, b_y] \times [a_z, b_z] \in \mathbb{R}^3$, we define the projector P_x^1 by

$$[P_x^1 f](x, y, z) = \sum_{i=1}^{N_x} f(x_i, y, z)\varphi_i(x),$$

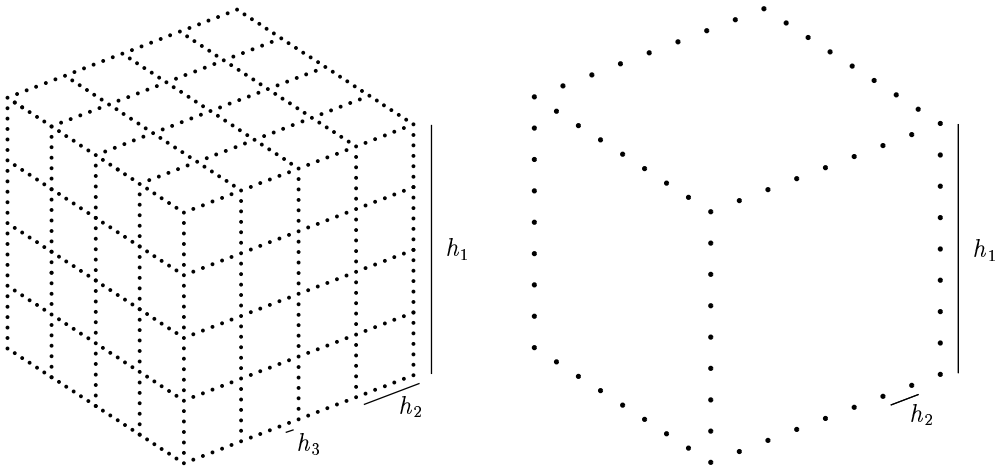

Figure 2.2: The building blocks of Boolean grids in 3D. The maximal grid corresponding to $\mathcal{P}^{\mathcal{M}}_{xyz}$ is shown on the left and the intermediate grid of $\mathcal{P}_{xyz}$ on the right. The maximal grid is created by using a two-dimensional Boolean grid on each of the sides of the cube and is a combination of 13 different grids. The intermediate grid is created by discretizing lines in each coordinate direction, and consists of only four different grids.

where φ_i are hat-functions as before, $x_i = a_x + (i-1)h_{1,x}$ and $h_{1,x} = (b_x - a_x)/(N_x - 1)$. We similarly define P^1_y, P^1_z, $h_{1,y}$, $h_{1,z}$, N_y and N_z. It is clear that $P^1_x f$ interpolates f on the planes $x = x_i$ and that the error introduced by this approximation is $O(h_x^2)$. Now, to make a full Boolean approximation, we would make an approximation on each of the planes $x = x_i$ using the two-dimensional Boolean grid introduced above, resulting in a projection for this direction given by

$$\mathcal{P}^{\mathcal{M}}_x = (P^2_y P^3_z \oplus P^2_z P^3_y) P^1_x,$$

where we have introduced the projectors P^2_α and P^3_α for $\alpha = y, z$, corresponding to the step sizes $h_{2,\alpha}$ and $h_{3,\alpha}$, respectively. Defining $\mathcal{P}^{\mathcal{M}}_y$ and $\mathcal{P}^{\mathcal{M}}_z$ similarly and taking the Boolean sum, we get the final Boolean projector in three dimensions,

$$\begin{aligned}
\mathcal{P}^{\mathcal{M}}_{xyz} &= \mathcal{P}^{\mathcal{M}}_x \oplus \mathcal{P}^{\mathcal{M}}_y \oplus \mathcal{P}^{\mathcal{M}}_z \\
&= (P^2_y P^3_z \oplus P^2_z P^3_y) P^1_x \oplus (P^2_z P^3_x \oplus P^2_x P^3_z) P^1_y \oplus (P^2_x P^3_y \oplus P^2_y P^3_x) P^1_z \\
&= P^{123}_{xyz} \oplus P^{132}_{xyz} \oplus P^{213}_{xyz} \oplus P^{231}_{xyz} \oplus P^{312}_{xyz} \oplus P^{321}_{xyz},
\end{aligned} \tag{2.7}$$

where we have used the notation $P^{ijk}_{xyz} = P^i_x P^j_y P^k_z$. This would result in a grid made up of cubes like the one on the left of figure 2.2. In order to estimate the error of this projection, we study the remainder operator

$$\mathcal{R}^{\mathcal{M}}_{xyz} = I - \mathcal{P}^{\mathcal{M}}_{xyz} = R^{123}_{xyz} R^{132}_{xyz} R^{213}_{xyz} R^{231}_{xyz} R^{312}_{xyz} R^{321}_{xyz},$$

with

$$R^{123}_{xyz} = R^1_x \oplus R^2_y \oplus R^3_z = R^1_x + R^2_y + R^3_z - R^1_x R^2_y - R^1_x R^3_z - R^2_y R^3_z + R^1_x R^2_y R^3_z$$

and so on. Using this in the expression for $\mathcal{R}^{\mathcal{M}}_{xyz}$, and remembering that $R^i_\alpha \geq R^j_\alpha$ if $i \leq j$ for $\alpha = x, y, z$, one could in principle write out the full expression for the remainder using only regular '+' and '−' operations. This expression would include terms of the type $R^1_x R^1_y R^1_z$, $R^2_x R^2_y$ and R^3_x, as well as terms which lead to higher order errors. Thus, assuming for simplicity that the step sizes are equal in all directions (i.e $h_{j,x} = h_{j,y} = h_{j,z} = h_j$ for $j = 1, 2, 3$), and using the fact that $R^j_\alpha f = O(h^2_j)$, we see that

$$\mathcal{R}^{\mathcal{M}}_{xyz} f = O(h^6_1 + h^4_2 + h^2_3),$$

from which we deduce that it is optimal if $h^6_1 = h^4_2 = h^2_3$, i.e. $h_3 = h^2_2 = h^3_1$. In that case, we will get an error of order h^6_1, while the number of points used is

$$N_B = 3N_1(2N_2N_3 - N_2^2) - 3N_1^2N_3 + N_1^3 = O(N_1^6),$$

while in order to have the same accuracy on a regular grid, we need $N_S = N_3^3 = O(N_1^9)$ grid nodes, i.e. a difference of 3 in the exponent.

This seems very good indeed, but there is a problem with this approach, which is seen if we expand (2.7) into an expression containing only regular '+' and '−' operators. Then we get 13 distinct terms, which is far too many for our purposes, since it requires a lot of overhead calculations as will be seen in the next section. It may also be a problem to actually create and use such Boolean grids, since even a fairly large h_1 will create a very small h_3 and make the number of points needed very large. Therefore, we take a middle way, reducing the number of terms but paying for this by getting a reduction in accuracy, although we still maintain a clear advantage over the regular grid.

So, using the same definition of P^j_α as above, we create the projector

$$\mathcal{P}_{xyz} = P^2_x P^1_y P^1_z \oplus P^2_y P^1_z P^1_x \oplus P^2_z P^1_x P^1_y = P^{211}_{xyz} \oplus P^{121}_{xyz} \oplus P^{112}_{xyz}, \tag{2.8}$$

which corresponds to first projecting the function onto lines $\{x = x_i,\ y = y_j\}$, which are spaced with the step size h_1, then discretizing these lines with step size h_2 and finally take the Boolean sum of the three directions. This gives a grid like the one on the right of figure 2.2. Expanding $\mathcal{P}_{xyz}$ using the algebraic rules, we see that

$$\mathcal{P}_{xyz} = P^{211}_{xyz} + P^{121}_{xyz} + P^{112}_{xyz} - 2P^{111}_{xyz}, \tag{2.9}$$

giving us just four terms. The remainder becomes

$$\mathcal{R}_{xyz} = R^{211}_{xyz} R^{121}_{xyz} R^{112}_{xyz},$$

with

$$R^{211}_{xyz} = R^2_x \oplus R^1_y \oplus R^1_z = R^2_x + R^1_y + R^1_z - R^2_x R^1_y - R^2_x R^1_z - R^1_y R^1_z + R^2_x R^1_y R^1_z,$$

which means that $\mathcal{R}_{xyz}$ includes terms of the type R_x^2 and $R_x^1 R_y^1$ as well as higher order terms. Therefore,

$$\mathcal{R}_{xyz} f = O(h_1^4 + h_2^2),$$

and we see that a choice of $h_2 = h_1^2$ gives the optimal $O(h_1^4)$ error. The number of points used is now

$$N_B = 3N_1^2 N_2 - 2N_1^3 = O(N_1^4),$$

while for a regular grid $N_S = N_2^3 = O(N_1^6)$ for the same accuracy.

2.2.3. Application to finite difference solvers

It is clear from the previous section that if we can create some approximation method with second order errors in each direction, we can use the Boolean grids defined there to improve accuracy, while using fewer grid nodes than in a uniform grid. Specifically, we want to use these grids for finite difference approximations of derivatives, in particular the Laplacian in two and three dimensions.

Starting in two dimensions, we therefore define the standard finite difference approximation of the second derivative in x by

$$D_{x,h}^2 u(x_i, y) = \frac{u(x_{i+1}, y) - 2u(x_i, y) + u(x_{i-1}, y)}{h^2},$$

which is second order accurate, that is, the error in the approximation is $O(h^2)$. $D_{y,h}^2$ is defined similarly. The discrete Laplacian is then given by

$$\Delta_{h_x, h_y} u(x_i, y_j) = D_{x,h_x}^2 u(x_i, y_j) + D_{y,h_y}^2 u(x_i, y_j).$$

Assuming for simplicity that the step sizes are equal in the two directions ($h_{1,x} = h_{1,y} = h_1$ and $h_{2,x} = h_{2,y} = h_2$), we may then investigate the Boolean approximation of the Laplacian, which according to (2.6) should be defined as

$$\Delta_{h_1,h_2}^B u(x, y) = \Delta_{h_2,h_1} u(x, y) + \Delta_{h_1,h_2} u(x, y) - \Delta_{h_1,h_1} u(x, y).$$

The accuracy of this approximation is then $O(h_1^4 + h_2^2)$, i.e. $O(h_1^4)$ if we choose $h_2 = h_1^2$. It is again easy to see why the errors cancel out. The two first approximations have one fine direction and one coarse. The derivatives in the coarse direction are canceled by subtraction of the last term, which is coarse in both directions. It should be emphasized that the $O(h_1^4)$-accuracy holds not only on the Boolean grid in figure 2.1, but at any point inside the domain, as long as we use the Boolean combination of the three components. That is, having calculated the three contributions $\Delta_{h_2,h_1} u$, $\Delta_{h_1,h_2} u$ and $\Delta_{h_1,h_1} u$ on the three grids, we may linearly interpolate on each of these and combine these interpolations by the Boolean scheme to get the value at any point. These interpolations introduce new error terms, and the error is in general larger, but the order of the error is the same as before.

In three dimensions, things are very much the same. We use the intermediate scheme defined by (2.8) on the grid on the right of figure 2.2, which consists of four grids. Three

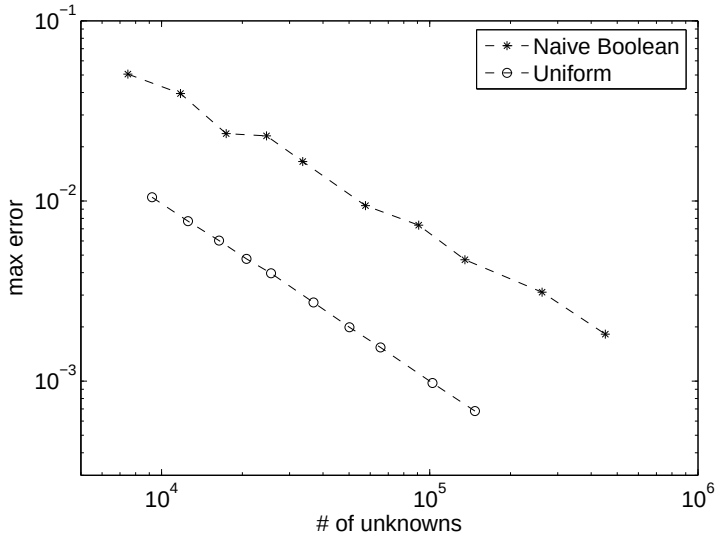

Figure 2.3: Comparison of different types of immersed interface methods applied to the Poisson equation inside a circle in $\mathbb{R}^2$, where the solution is a sixth degree polynomial. The errors for the naive Boolean IIM, where three complete solutions by IIM are combined by the Boolean scheme, are denoted by '$*$'. The errors for IIM on uniform grids are denoted by '$\circ$'. It is seen clearly that in the naive method, the errors are not eliminated in the way we want.

grids are fine in one direction and coarse in the other two, while the last grid is coarse in all directions. Again, using the appropriate Boolean scheme given by (2.9), we set

$$\begin{aligned}\Delta^B_{h_1,h_2} u(x,y,z) = \Delta_{h_2,h_1,h_1} u(x,y,z) + \Delta_{h_1,h_2,h_1} u(x,y,z) \\ + \Delta_{h_1,h_1,h_2} u(x,y,z) - 2\Delta_{h_1,h_1,h_1} u(x,y,z),\end{aligned}$$

to get an $O(h_1^4 + h_2^2)$-approximation.

Solving for example the Poisson equation $\Delta u = f$ using finite differences on Boolean grids is now straightforward. One simply solves the problem on each of the component grids and then combines the solutions using the Boolean scheme. This is shown in example 2 in section 2.4.2.

2.3. Applying IIM to Boolean grids

Having successfully applied finite differences on the Boolean grids, we now try to implement the Immersed Interface Method as described in section 1 on these grids in order to cope with domains that are not rectangular.

From the previous section, it might be suggested that one could use the Immersed Interface Method straight away on the component grids and then combine them using the Boolean scheme to get improved results. This does not work, however, since when we introduce jumps in the finite differences, the errors no longer behave in the correct way to be eliminated by the Boolean scheme. This is seen in figure 2.3.

Therefore, we need a slightly more elaborate method. The reason for not getting the desired elimination of errors is clearly the estimation of the jumps at the boundaries, since the results in section 2.4.2 show that finite differences work fine on Boolean grids. So, instead of approximating the jumps on each grid individually, we use the Boolean interpolation to get the jumps from the solution. That is, we must create a matrix D_B^T implementing this Boolean interpolation so that

$$C = D_B^T U + F_2,$$

with $U = [U_{11}^T \; U_{21}^T \; U_{12}^T]^T$ containing the solution on the three grids (supposing we are solving a two-dimensional problem). Here U_{ij} is the solution on the grid created by the projection P_{xy}^{ij}. As before, if we intend to solve the Poisson equation, we also have the equation

$$\Delta_h U + \Psi C = F_1.$$

Here Δ_h works on the three grids separately and Ψ is divided into three blocks corresponding to the grids. As in the uniform case, we also solve for C in the equation

$$(I + D_B^T \Delta_h^{-1} \Psi) C = D_B^T \Delta_h^{-1} F_1 + F_2$$

and then finally compute the solution U as

$$U^* = B\Delta_h^{-1}(-\Psi C + F_1).$$

Here B is a Boolean interpolation matrix, combining the three solutions U_{11}, U_{21} and U_{12} into the final solution U^*, which could be defined on any grid.

The procedure outlined here is our Boolean IIM, resulting in the errors shown in section 2.4.3. The only difference to the uniform IIM described in section 1 is the jump approximation matrix D_B^T and the interpolation matrix B. The interpolation matrix B uses the Boolean schemes (2.6) and (2.9) in two and three dimensions respectively, together with linear interpolation between grid nodes, in order to get the solution U^* at the desired points. The matrix D_B^T is slightly more complicated and we will now discuss how to form it.

2.3.1. BOOLEAN APPROXIMATION OF JUMPS

Just as in the IIM on uniform grids, we need to approximate the jumps of the solution and its derivatives on the boundary, using the values of the solution on the grid nodes. The boundary conditions are implemented exactly as for the uniform grids (see section 1.4), so what we need to do here is to create equivalents to matrices $\mathcal{Q}_j$, $\mathcal{P}_j$ and $\mathcal{R}_j$ in (1.12)

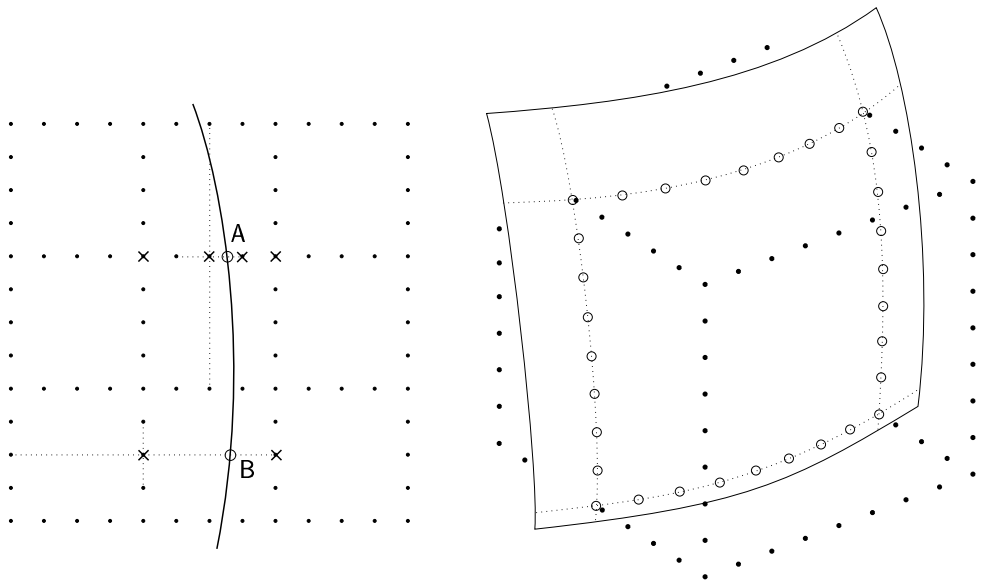

Figure 2.4: Some IIPs on the Boolean grids. The IIPs are marked with circles and the anchor points with crosses. In two dimensions, the IIPs are evenly distributed with spacing h_2. At 'A', the IIP interferes with finite differences both on the fine grid (with step size h_2) and on the coarse grid (with step size h_1). Two of the affected finite difference stencils are indicated by dotted lines. In three dimensions the IIPs are distributed along the edge of the grid cube, giving rise to a two-dimensional Boolean grid pattern.

in order to estimate the values of u and its derivatives at the interface intersection points (IIPs) from both sides of the interface.

So, first of all, we need to find all the IIPs, that is, all the points where the interface intersects the grid lines and thus interferes with the finite differences. Suppose that we want to calculate the Laplacian in two dimensions. Then the IIPs are all the points where the interface intersects the fine grid, as is shown in figure 2.4. That is, we will have as many IIPs as in the case of a uniform grid. In three dimensions, however, the IIPs are distributed on the surface in a two-dimensional Boolean grid, meaning that we have much fewer IIPs using a Boolean grid than a uniform one, namely $O(N_1N_2)$ IIPs instead of $O(N_2^2)$. This also implies that the vector C will be smaller and thus the entire system of equations will be smaller.

Having found the IIPs, we assign to each of them two (or more) anchor nodes, on both sides of the surface. These are the nodes where the finite difference is affected by the jump in the solution at the IIP. The particular difference taken may be either of size h_1 or size h_2, as is shown in figure 2.4. From (1.4), we see that this gives rise to corrections to the regular second order difference $D_h^2 u$ given by

$$\frac{d^2u}{dx^2}(x_i) = D_h^2 u(x_i) + \frac{1}{h^2}\sum_{m=0}^{3} \frac{(x_i-\alpha)^m}{m!}[u^{(m)}]_\alpha + O(h^2). \tag{2.10}$$

Here, α is the location of the IIP, so that $x_i - \alpha \le h$, and h may be either h_1 or h_2. The question now is how well we need to approximate the jumps and how many derivatives we need. We saw earlier that in the uniform case it is enough to get an approximation of order $O(h)$ on the boundary in order to get overall $O(h^2)$-convergence, and therefore that we may ignore the jumps in the third derivative. In the Boolean case, however, it is not quite clear what we need to do. We want overall $O(h_2^2)$-convergence, but if we ignore the jumps in third derivative in (2.10), we will at some points get errors of order $O(h_1)$, which seems to be too large. Fortunately, our numerical results show that it is in fact sufficient to use jumps in u, u' and u'' and use Boolean interpolation to approximate these to order $O(h_2^3)$, $O(h_2^2)$ and $O(h_2)$, respectively. From (2.10), we then see that we will get truncation errors of order $O(h_1)$ at some nodes (where $h = h_1$) and $O(h_2)$ at others, while the overall error in the solution will be of order $O(h_2^2)$. The explicit reason that this approach works is not known.

How this is done is most easily seen from figure 2.5. Suppose we want to approximate the function values and the two first derivatives in the x-direction at the IIP α_j with coordinates $(x_{\alpha_j}, y_{\alpha_j})$. We use regular Lagrange interpolation on each of the grids, that is, given grid nodes with x-coordinates $x_1, \ldots, x_L$, we set

$$\tilde{u}(x) = \sum_{l=1}^{L} u(x_l)\phi_l(x),$$

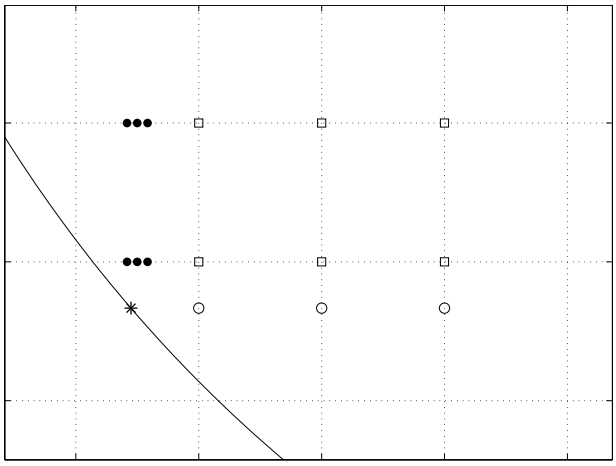

Figure 2.5: Interpolation stencil for jumps at the IIP marked by '$*$'. The nodes in the G_{21} grid are marked by '$\bullet$', in the G_{12} grid by '$\circ$' and in the G_{11} grid by '$\square$'.

with the basis functions

$$\phi_l(x) = \prod_{\substack{k=1 \\ k \neq l}}^{L} \frac{x - x_k}{x_l - x_k}.$$

To compute derivatives, we just differentiate the basis functions, getting

$$\phi_l'(x) = \sum_{\substack{i=1 \\ i \neq l}}^{L} \frac{1}{x_l - x_i} \prod_{\substack{k=1 \\ k \neq l,i}}^{L} \frac{x - x_k}{x_l - x_k},$$

and so on. In this way, we may write for each of the three grids

$$\tilde{u}(x,y) = \sum_{i=1}^{L_x} \sum_{l=1}^{L_y} u(x_i, y_l)\phi_i(x)\phi_l(y).$$

Here, $\tilde{u}$ is of course the projection of u onto the space spanned by the basis functions, so from the theory of section 2.2 we expect that if we combine the three grids according to the Boolean sum, we would eliminate the largest errors. Therefore, we choose grid nodes according to figure 2.5, with (at least) three nodes in the x-direction in order to approximate second derivatives, and (at least) two nodes in the y-direction in order to extrapolate the function to the IIP in that direction.

One could also understand the choice of nodes in the following way. What we really want is a fine (h_2) approximation in the x-direction using the grid G_{21}, which is fine in

that direction. But since we must extrapolate in the y-direction, we introduce large (h_1) errors in that direction. These are eliminated by subtracting the approximation on the coarse grid G_{11}, but this in turn adds large (h_1) errors in the x-direction. These are finally eliminated by adding the approximation on the G_{12} grid, which does not introduce new errors in the y-direction.

These principles may be used to approximate any derivative to any order of accuracy, if one uses enough grid nodes for the approximation. For our purpose it is best to use as few grid nodes as possible, since the more nodes we use, the more coupled the resulting system of equations becomes, which makes it harder to solve. Already, the minimal number of 15 nodes shown in figure 2.5 is much more than the six nodes needed on uniform grids. One should also note, however, that if the IIP lies on one of the coarse grid lines (as point A in figure 2.4), it is sufficient to use the three nodes on the fine grid G_{21}, and the Boolean combination is then not needed. But, if one would like to solve the Neumann problem, one would also need y-derivatives, which requires additional nodes in a configuration similar to the one we have studied here.

Anyway, the final scheme may be expressed in matrix form as

$$\begin{bmatrix} u^+(\alpha_j) \\ u_x^+(\alpha_j) \\ u_{xx}^+(\alpha_j) \end{bmatrix} = \mathcal{B} \begin{bmatrix} \mathcal{P}_{11}^{j,+} & 0 & 0 \\ 0 & \mathcal{P}_{21}^{j,+} & 0 \\ 0 & 0 & \mathcal{P}_{12}^{j,+} \end{bmatrix} \mathcal{R}_j^+ U,$$

where $\mathcal{B} = [-I \; I \; I]$ performs the Boolean combination, $\mathcal{R}_j^+$ consists only of 1's and 0's and just reorders and renumbers nodes, while $\mathcal{P}_{ik}^{j,+}$ performs the Lagrange interpolation to evaluate $\tilde{u}(\alpha_j)$ and its derivatives using the grid G_{ik}.

If we apply the same procedure on the other side of the interface (if necessary) to create $\mathcal{P}_{ik}^{j,-}$ and $\mathcal{R}_j^-$, and then create the matrix $\mathcal{L}_j$ exactly as in section 1.4, we may estimate jumps at the boundary as

$$\begin{bmatrix} [u]_j \\ [u_x]_j \\ [u_{xx}]_j \end{bmatrix} = \mathcal{L}_j \mathcal{B} \mathcal{P}_j \mathcal{R}_j U,$$

which should be compared to (1.13). We set $D_{B,j}^T = \mathcal{L}_j \mathcal{B} \mathcal{P}_j \mathcal{R}_j$ and stack the contributions from all the IIPs to get the matrix D_B^T and the system of equations

$$C = D_B^T U + F_2.$$

Being thus able to approximate the jumps, we may use the method outlined earlier to apply the Immersed Interface Method on Boolean grids. For problems in three dimensions, the same procedure is used. There is no need to use a three-dimensional approximation of the jumps, so we can use the stencil shown in figure 2.5 for three-dimensional problems as well.

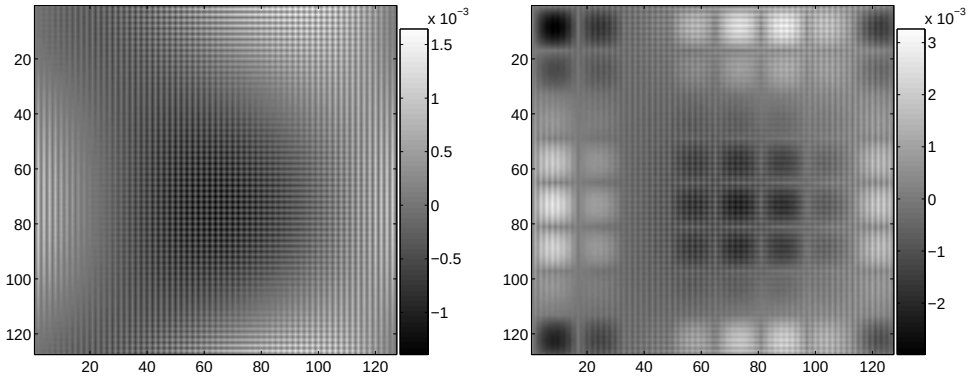

Figure 2.6: Interpolation errors using a uniform grid (left) and a Boolean grid (right). The errors are of about the same size, but the uniform grid has $64 \times 64 = 4096$ nodes, while the Boolean grid uses $2 \times 8^2 \times 8 + 8^2 = 1088$ nodes. Note the patterns of the errors, indicating the grids on which the original sampling was done.

2.4. Examples

2.4.1. Example 1 – Boolean approximation

In this example, we illustrate simple Boolean interpolation. A sixth degree polynomial in two dimensions is sampled on a uniform grid with 64×64 nodes and on a Boolean grid with $N_1 = 8$ and $N_2 = 64$. Both of these approximations are then linearly interpolated onto a uniform grid with 127×127 nodes. On the Boolean grid, we use the Boolean combination of the three different grids, while in the uniform case a regular linear interpolation is used.

Figure 2.6 shows the resulting errors. In this case, the errors on the Boolean grid are slightly larger, but of approximately the same size as on the uniform grid. However, the number of nodes used in the Boolean grid is $2 \times 8^2 \times 8 + 8^2 = 1088$, while the uniform grid uses $64 \times 64 = 4096$ nodes.

2.4.2. Example 2 – Finite differences on Boolean grids

This example illustrates the use of finite differences on Boolean grids, as explained in section 2.2.3. The Laplacian is applied to the function $f(x) = \sin(2\pi x)\sin(2\pi y)$ on the unit square using finite differences. The exact result is $\Delta f(x) = -8\pi^2 f(x)$. Figure 2.7 shows the result on a Boolean grid with $N_1 = 12$ and $N_2 = 12 \times 8 = 96$. The values on the Boolean grid are interpolated onto a uniform grid with 96×96 nodes and the errors are calculated on this grid. Note that the structure of the Boolean grid is reflected in the error image on the right of figure 2.7. This indicates that the errors are smaller on the Boolean grid than in between grid lines, so that interpolation errors are greater than errors resulting from finite differences on the grid itself.

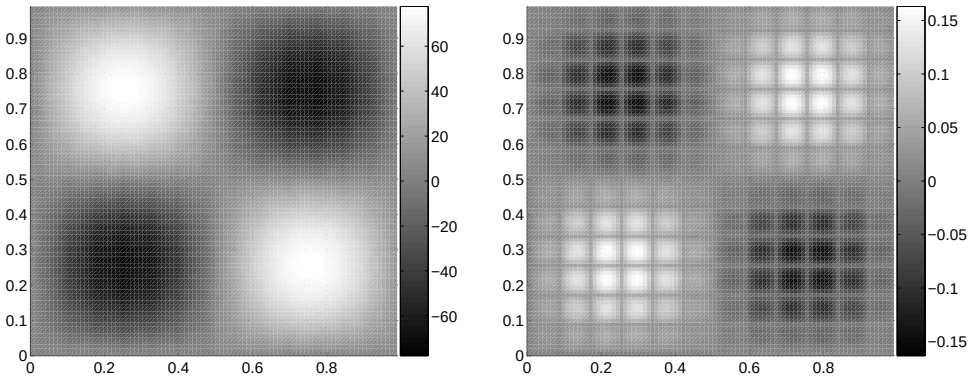

Figure 2.7: Finite differences on a Boolean grid with $N_1 = 12$ and $N_2 = 12 \times 8$. The discrete Laplacian was applied to the function $f(x) = \sin(2\pi x)\sin(2\pi y)$ on a Boolean grid and then interpolated to a full grid. The result is shown on the left with errors on the right. The error image shows that the errors are smaller on the Boolean grid than between grid lines.

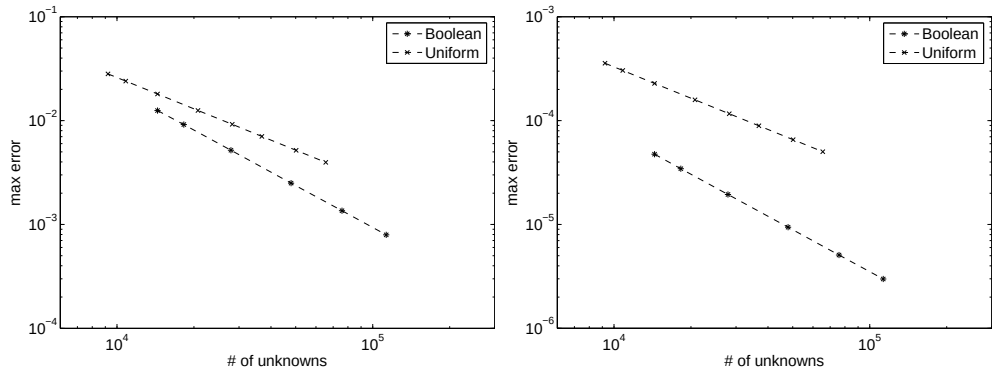

Figure 2.8: Asymptotics for finite difference approximations on two-dimensional uniform and Boolean grids for the forward Laplacian on the left and the inverse Laplacian on the right. The graphs are log-log-plots of errors as function of the number of grid nodes. In both graphs, the slopes of the lines are -1.00 for the uniform grids and -1.34 for the Boolean grids. The Boolean grids use $N_2 = N_1^2/2$ and values of N_1 range from 24 to 48 nodes.

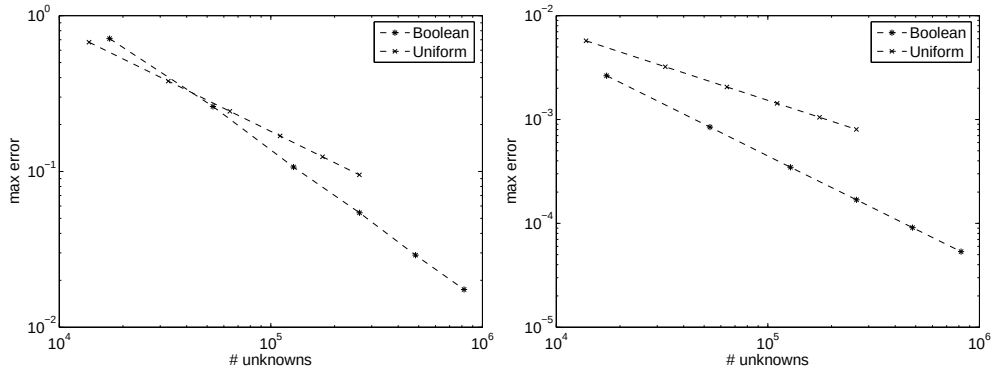

Figure 2.9: Error asymptotics for the discrete Laplacian (left) and its inverse (right) applied using finite differences on uniform and Boolean grids in 3D. The plots are log-log-plots of the number of grid nodes versus the maximum error and the slopes of the lines are near the theoretical values of $-2/3$ for the uniform grids and -1 for the Boolean grids.

Figure 2.8 shows the asymptotics of the same approximation compared to approximations on a uniform grid. On the left we see a log-log plot of the errors for the discrete Laplacian applied to $f(x)$ as above, while on the right we see errors for the inverse Laplacian applied to $-8\pi^2 f(x)$ (so that the exact solution is $f(x)$), using zero boundary conditions at the left edge ($x = 0$) and periodic boundary conditions at the other edges. The Boolean grids used have values for N_1 ranging between 24 and 48 and with $N_2 = N_1^2/2$. In the left figure, the slopes of the lines are -1.000 for the uniform grid and -1.339 for the Boolean grid, while in the right figure, the slopes are -1.000 and -1.344, respectively. This agrees well with the theoretical values, which state that on a uniform grid, the approximation should be of order $O(h^2) = O(N^{-2})$, while the number of unknowns is $O(N^2)$, giving the slope $-2/2 = -1$ in a log-log plot. On the Boolean grid, the number of unknowns is $2N_1N_2 + N_1^2 = O(N_1^3)$ and the accuracy is $O(h_2^2) = O(N_2^{-2}) = O(N_1^{-4})$, since N_2 is proportional to N_1^2. This should give a slope of $-4/3$ in the log-log plot, which agrees very well with the numerical results.

In the last figure, figure 2.9, the asymptotics for the same problem in the unit cube in 3D is shown. Here $f(x) = \sin(2\pi x)\sin(2\pi y)\sin(2\pi z)$ with $\Delta f(x) = -12\pi^2 f(x)$. The figure shows results for the Laplacian on the left and for the inverse Laplacian on the right. The Boolean grid is the intermediate grid with four components. N_1 ranges from 12 to 32 and $N_2 = N_1^2/4$. The slopes of the lines for the uniform grid are both -0.667, which agrees perfectly with the theoretical value derived from $O(h^2) = O(N^{-2})$ accuracy and $O(N^3)$ unknowns. For the Boolean grid, the slope in the figure is -0.965 for the Laplacian and -1.011 for its inverse, which agrees with the theoretical value of -1, arising from the fact that the accuracy is $O(h_2^2) = O(N_1^4)$ and the number of unknowns is $3N_1^2N_2 + N_1^3 = O(N_1^4)$.

The conclusion is that the theory works for standard finite differences and that by

using Boolean grids, the number of nodes may be reduced while still retaining the same accuracy. It should be noted that the choices of N_2 here are not necessarily optimal. N_2 is proportional to N_1^2, which it must be, in order to achieve the correct asymptotics. But given a specific N_1, it is not clear beforehand which is the optimal N_2. There is a limit where there is no use increasing N_2 further, because the dominating error terms depend only on N_1. This limit is problem dependent, since the sizes of the error terms depend on the magnitude of the derivatives of the solution f. Choosing optimal values of N_2 for all N_1 will not change the slope of the lines, only the position, that is, it will not change the asymptotics even if it changes the error values.

2.4.3. EXAMPLE 3 – IIM ON BOOLEAN GRIDS

Our final example shows the full immersed interface method on Boolean grids. We solve the Poisson equation with Dirichlet boundary conditions inside an ellipse E with center at $(0.5, 0.5)$ and half-axes 0.44 and 0.38, i.e.

$$\begin{aligned} \Delta u(x,y) &= f(x,y), \qquad (x,y) \in E, \\ u(x,y) &= g(x,y), \qquad (x,y) \in \partial E, \end{aligned}$$

with $f(x,y)$ and $g(x,y)$ chosen so that the solution $u(x,y)$ is a given sixth degree polynomial (the same as in example 1). The solution and the distribution of errors on one of the component grids (G_{12}) are shown in figure 2.10. The results for different grid sizes are listed in table 1. $\|E_N\|_\infty$ denotes the maximum error compared to the exact solution on the Boolean grid, while $\|T_N\|_\infty$ denotes the maximum truncation error, i.e. $T_N = \Delta_h u_0 + \Psi C_0 - F_1$, where u_0 and C_0 are the node values and jumps for the exact solution, respectively. The table also lists approximate computational times.

N_1	n_2	N_{tot}	time (s)	$\|E_N\|_\infty$	$\|T_N\|_\infty$
24	6	7488	< 0.1	5.37e-3	2.51
28	7	11760	0.1	3.35e-3	1.49
32	8	17408	0.2	1.76e-3	1.37
36	9	24624	0.3	1.35e-3	1.19
40	10	33600	0.5	6.90e-4	1.34
48	12	57600	0.9	3.94e-4	1.23
56	14	90944	1.4	2.68e-4	1.02
64	16	135168	2.5	1.31e-4	0.86

Table 1: Results for Boolean IIM in two dimensions. $n_2 = N_2/N_1$ and $N_{tot} = N_1^2(2n_2+1)$. See text for details.

The asymptotics of the errors are shown in figure 2.11. There, asymptotics for a corresponding three-dimensional problem is also shown, that is the Poisson equation inside an ellipsoid with Dirichlet boundary conditions, whose solution is a sixth degree polynomial. The corresponding data is shown in table 2. The numerical asymptotics for the errors

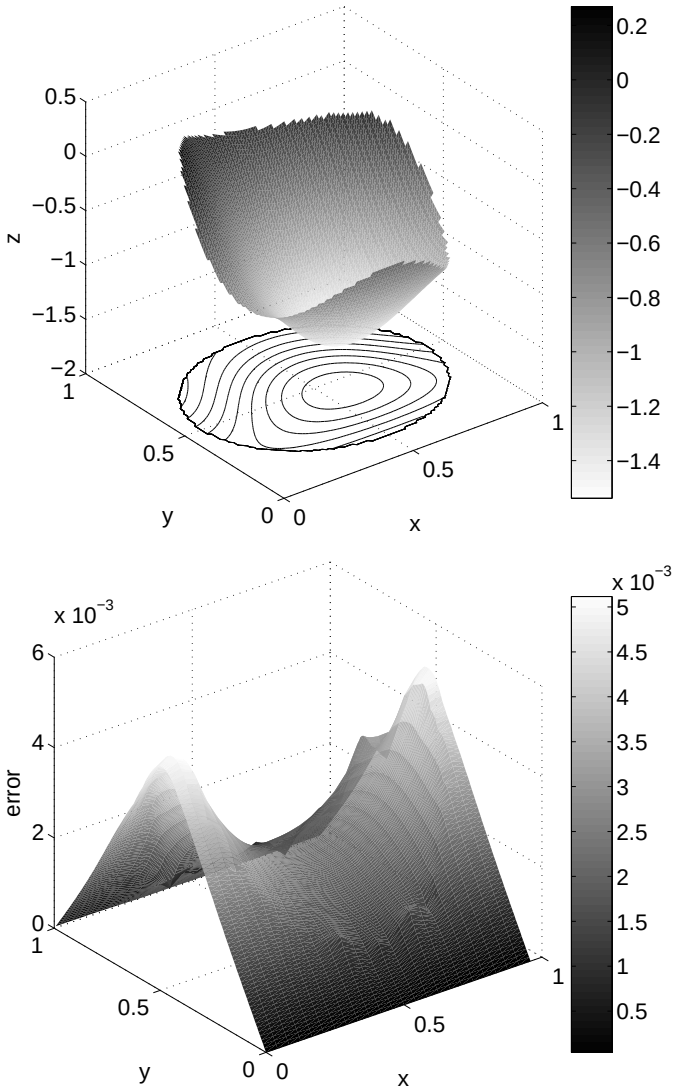

Figure 2.10: IIM on a two-dimensional Boolean grid. At the top we see the numerical solution $u(x, y)$ to the Poisson problem in an ellipse computed on a Boolean grid with $N_1 = 24$ and $N_2 = 24 \times 6 = 144$. Level curves are drawn below the surface, showing clearly the ellipse E. The solution is zero outside the domain. At the bottom, we see the errors compared to the exact solution on the G_{12} grid with small steps in the y-direction.

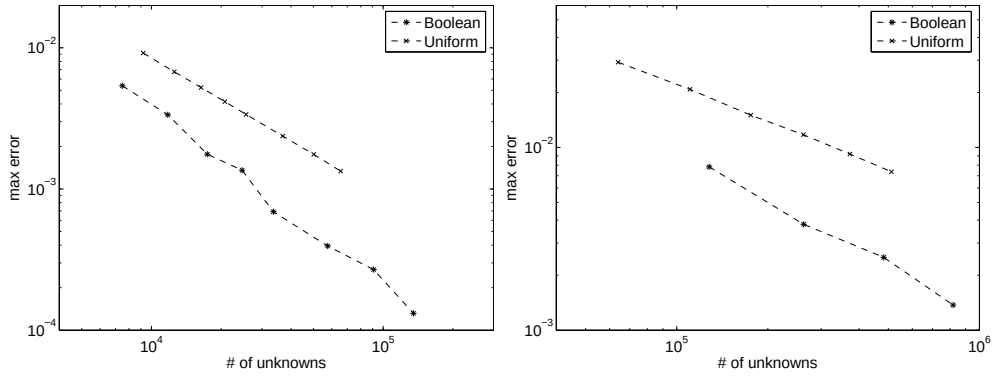

Figure 2.11: Asymptotics of errors for IIM on Boolean grids. Errors for the solution to the Poisson equation inside an ellipse in 2D on the left and inside an ellipsoid in 3D on the right. The results on Boolean grids are compared to results for regular IIM on uniform grids. On the Boolean grids, values of $N_2 = N_1^2/4$ are used throughout. The slopes of the lines are -1.27 for Boolean grids and -0.97 for uniform grids in 2D, while for the 3D problem they are -0.91 and -0.66, respectively. This should be compared to the theoretical values, which for $O(h^2)$-convergence are $-4/3$, -1, -1 and $-2/3$ in the given order (see example 2).

N_1	n_2	N_{tot}	time (s)	$\|E_N\|_\infty$	$\|T_N\|_\infty$
20	5	1.28e5	3.1	7.74e-3	2.78
24	6	2.63e5	7.1	3.76e-3	2.49
28	7	4.83e5	13.2	2.50e-3	2.19
32	8	8.19e5	26.6	1.37e-3	1.91

Table 2: Results for Boolean IIM in three dimensions. $n_2 = N_2/N_1$ and $N_{tot} = N_1^3(3n_2+1)$. See text for details.

$\|E_N\|$ agree nicely with the theoretical values for $O(h_2^2)$-behavior given in the previous example. As expected, the truncation errors exhibit $O(h_1)$-behavior in both two and three dimensions, but even though these errors are large, the final errors in the solution are small. The computational times listed indicate that the time needed to solve the problem grows slightly faster than the number of unknowns. This is because the number of iterations needed in the BiCGStab algorithm to solve the linear system of equations is not constant, but tend to increase slightly with the number of unknowns. (e.g. from 17 to 21 in the 3D problem).

We conclude that we are able to achieve $O(h_2^2)$ error asymptotics for both two- and three-dimensional problems on the Boolean grids. We also see from the error plots that the number of grid nodes needed to get a specific error is much lower for the Boolean grids than for the uniform ones.

References

[1] H. Bungartz, M. Griebel, and U. Rüde. Extrapolation, combination, and sparse grid techniques for elliptic boundary value problems. *Comput. Methods Appl. Mech. Engrg.*, 116(1-4):243–252, 1994. ICOSAHOM'92 (Montpellier, 1992).

[2] C. K. Chui and H.-C. Lai. Vandermonde determinant and Lagrange interpolation in $\mathbf{R}^s$. In *Nonlinear and convex analysis (Santa Barbara, Calif., 1985)*, volume 107 of *Lecture Notes in Pure and Appl. Math.*, pages 23–35. Dekker, New York, 1987.

[3] S. A. Coons. Surfaces for computer aided design of space forms. Technical report, Project MAC, Dept. of Mech. Engineering, MIT, 1964. Revised to MAC-TR-41, 1967.

[4] F.-J. Delvos and W. Schempp. *Boolean methods in interpolation and approximation*, volume 230 of *Pitman Research Notes in Mathematics Series*. Longman Scientific & Technical, Harlow, 1989.

[5] FFTW, `http://www.fftw.org`.

[6] W. J. Gordon. Distributive lattices and the approximation of multivariate functions. In *Approximations with Special Emphasis on Spline Functions (Proc. Sympos. Univ. of Wisconsin, Madison, Wis., 1969)*, pages 223–277. Academic Press, New York, 1969.

[7] W. J. Gordon. Blending-function methods of bivariate and multivariate interpolation and approximation. *SIAM J. Numer. Anal.*, 8:158–177, 1971.

[8] M. Griebel, M. Schneider, and C. Zenger. A combination technique for the solution of sparse grid problems. In *Iterative methods in linear algebra (Brussels, 1991)*, pages 263–281. North-Holland, Amsterdam, 1992.

[9] IML++ (Iterative Methods Library), `http://math.nist.gov/iml++/`.

[10] C. T. Kelley. *Iterative methods for linear and nonlinear equations*, volume 16 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1995. With separately available software.

[11] R. J. LeVeque and Z. L. Li. The immersed interface method for elliptic equations with discontinuous coefficients and singular sources. *SIAM J. Numer. Anal.*, 31(4):1019–1044, 1994.

[12] Z. Li. *The immersed interface method: a numerical approcah to partial differential equations with interfaces*. PhD thesis, Dept. of Applied Mathematics, University of Washington, Seattle, 1994.

[13] C. S. Peskin. Lectures on mathematical aspects of physiology. In *Mathematical aspects of physiology (Proc. Summer Sem., Univ. Utah, Salt Lake City, Utah, 1980)*, volume 19 of *Lectures in Appl. Math.*, pages 1–107. Amer. Math. Soc., Providence, RI, 1981.

[14] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in C.* Cambridge University Press, Cambridge, second edition, 1992. The art of scientific computing.

[15] A. Wiegmann. *The explicit jump immersed interface method and interface problems for differential equations.* PhD thesis, Department of Mathematics, University of Washington, Seattle, 1998.

[16] A. Wiegmann and K. P. Bube. The explicit-jump immersed interface method: Finite difference methods for PDEs with piecewise smooth solutions. *SIAM J. Numer. Anal.*, 37(3):827–862, 2000.

A model for intracellular subdiffusion

Tobias Gebäck

Abstract

This part of the thesis contains Paper I, in which a parabolic pseudo-differential equation describing subdiffusion (or anomalous diffusion) inside a cell is introduced and studied. We prove existence of unique solutions to boundary value problems both with Dirichlet conditions and conditions specifying the flux at the boundary, corresponding to Neumann problems for the ordinary diffusion equation. We also show some numerical solutions to our equation and compare them to the ordinary diffusion case.

The introductory material before the paper provides the motivation for studying this equation. In particular, we have performed fluorescence correlation spectroscopy (FCS) experiments in fission yeast, and observed anomalous diffusion inside the cells and showed that our model fits the data. The fitting of a model to FCS data requires derivation of autocorrelation functions from the model, and such functions have been derived for two different cases.

Contents

1. Introduction

In this part of the thesis, we will discuss the background and motivation for studying the space-anomalous diffusion equation, which is the topic for the paper that follows. The main motivation is the observation of so called anomalous diffusion, or subdiffusion, for proteins inside yeast cells. One useful method for measuring the properties of diffusion inside cells is Fluorescence Correlation Spectroscopy (FCS), which will be described in the next section.

After the description of the method, we will go into a description of anomalous diffusion, which is a type of motion for the molecules in the cell where the classical model for diffusion is observed to be no longer valid, and the data from FCS experiments call for another diffusion model. Possible models for anomalous diffusion will also be discussed, in particular the space-anomalous diffusion equation, which is the one we have investigated further. Then some necessary tools for analyzing the FCS data will be derived, namely theoretical autocorrelation functions derived from the diffusion models. Finally, some experimental results from measurements inside yeast cells of the species *Schizosaccharomyces pombe* supporting our model will be shown and discussed.

The FCS experiments were conducted in collaboration with Eva Asp and Per Sunnerhagen at the Department of Cell and Molecular Biology, Göteborg University, using the microscope facilities at the SweGene centre for cellular imaging [19].

2. Fluorescence Correlation Spectroscopy

2.1. Introduction

Fluorescence Correlation Spectroscopy (FCS) is a method which enables us to look into the smallest building blocks of the cell, into the world of proteins and large molecules, and study their behaviour. It enables us to investigate the motion of these molecules and draw conclusions about the environment in which they move, by means of mathematical models and analysis.

As the name indicates, the method is based on fluorescence, and fluorescent molecules are excited by a highly focused laser beam, which means that the measurements are conducted within a very small volume, where the fluctuations in the fluorescence signal are recorded and analyzed, as will be described in the next section.

The method was invented already in 1974, by Magde, Elson and Webb [2, 6], but it is not until recent years that the equipment has been sufficiently enhanced to enable measurements with high precision on small concentrations and in a small enough volume, to be able to conduct measurements inside cells. These developments were mainly due to Rudolf Rigler and coworkers at Karolinska Institute in the 1990's, where they combined the FCS technique with the confocal microscope technique.

Even though the method has been vastly improved and simplified through the design of special purpose microscopes, namely the Zeiss ConfoCor microscopes [5, 25], it is still a

rather complicated method, and in our own experience, some mathematical understanding is useful for interpreting the results, making the technique ideal for collaboration between experimentalists and theoreticians.

In the following sections, we will first go more into details of the FCS method, describing the measurements, the data, and the analysis and comparison to models for the motion of molecules. The purpose of this description is just to give enough understanding of the method so that the reader can appreciate the results in later sections. For a more detailed introduction, see the compendium by Schwille and Haustein [18], and for a more detailed and extensive description of the possibilities and limitations of FCS, see the book [14], edited by Rigler and Elson.

2.2. Basic principles and experimental setup

In order to measure the motion of molecules, we need to 'see' the molecules, which we are able to do through fluorescence, i.e. the emission of light from the molecules after excitation by light of higher energy. Therefore, the proteins we wish to study must be tagged with a fluorescent tag, a part of another protein which is fluorescent by nature (the most common is GFP — Green Fluorescent Protein, but there are tags with different colours available). This is achieved by taking the DNA encoding for the protein we are interested in, adding the DNA encoding for the fluorescent tag right next to it, and inserting the whole DNA sequence into the yeast cells, where the new DNA is incorporated into the cells' own DNA, and the new, fluorescently tagged, proteins are expressed from the DNA using the cells' machinery for protein construction. The result is that the cells create variants of the protein that we are interested in which we can actually see, if not by naked eye, so at least using microscopes and lasers to assist our eyes. Furthermore, the DNA encoding for the fluorescent proteins is inherited to the next generation of yeast cells, so we will quickly get a large amount of cells having this new, visible protein. What this may look like in a microscope is shown in figure 1, where we see fission yeast cells where the MAPKAP-kinase Mkp1 has been labeled with GFP. These methods and manipulations are common tools for the modern cell biologist, but it is fascinating that it works.

Now then, when we can see the molecules, we wish to conduct measurements. The basic microscope setup for FCS is shown in figure 2. The basic components are first of all a laser and a system of lenses, which focus the laser beam to a very small volume (< 1 fl) in the sample. In this volume, the laser light excites the fluorescent molecules, which then emit light (of longer wavelength). This light is collected by a photon counter with very high time resolution (12.5 ns), which counts the individual photons coming from the measurement volume. A typical photon count curve might look like the one in figure 3. The curve looks like pure noise, but the data really contains a lot of information, including information about how molecules move in and out of the measurement volume, and about the excitation of molecules and their fluorescence properties.

In order to extract this information in a more understandable way, the autocorrelation function of the signal is computed, that is, if we denote the photon count signal by $F(t)$,

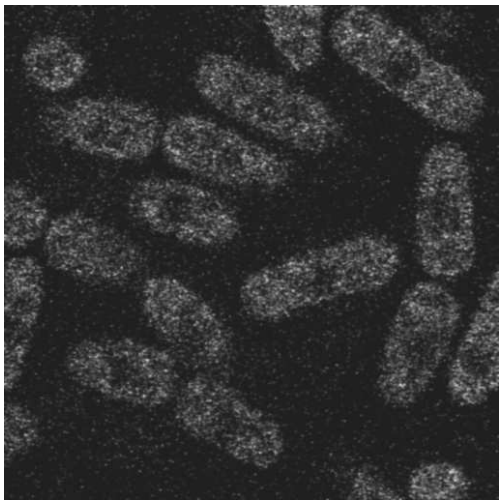

Figure 1: Yeast cells of the species *S. pombe*, where the protein Mkp1 has been labeled with GFP and can be seen as shining green in the microscope. Some cells seem shorter because they are at an angle with the focal plane. The darker spots in the center of the cells are the nuclei, where the concentration of Mkp1 is lower than in the cytoplasm.

we compute

$$G(\tau) = \frac{\langle F(t)F(t+\tau)\rangle - \langle F(t)\rangle^2}{\langle F(t)\rangle^2} = \frac{\langle \delta F(t)\delta F(t+\tau)\rangle}{\langle F(t)\rangle^2}, \tag{2.1}$$

where $\delta F(t) = F(t) - \langle F(t)\rangle$ and $\langle\cdot\rangle$ denotes the mean over $t \in [0, T]$, where T is the measurement time, and τ is a time delay which is varied between approximately 1 μs $-$ 1 s. An autocorrelation function computed from an experiment may look like the one in figure 4, although this is data from a calibration experiment using the fluorescent dye Rhodamine Green in aqueous solution, so it is a near perfect autocorrelation curve.

The intuition behind the appearance of the autocorrelation function is that for small time delays τ, the molecules will not have time to move far between times t and $t+\tau$, and so the fluorescence signal will be almost the same at these two time points, and therefore the autocorrelation will be high for small τ. For larger τ, though, the molecules will have time to diffuse out of the measurement volume, and new molecules will have time to come into the volume, and so the fluorescence at times t and $t+\tau$ will be essentially uncorrelated, making the autocorrelation function decrease towards zero when τ gets large. Thus, the value of τ for which the slope is maximal gives us a measure of the diffusion speed of the molecules. For very small τ, we may see other effects, which are more related to the excitation and fluorescence process itself.

If we wish to get something more useful, such as diffusion constants, out of the autocorrelation curve, we need to derive theoretical autocorrelation curves from models of

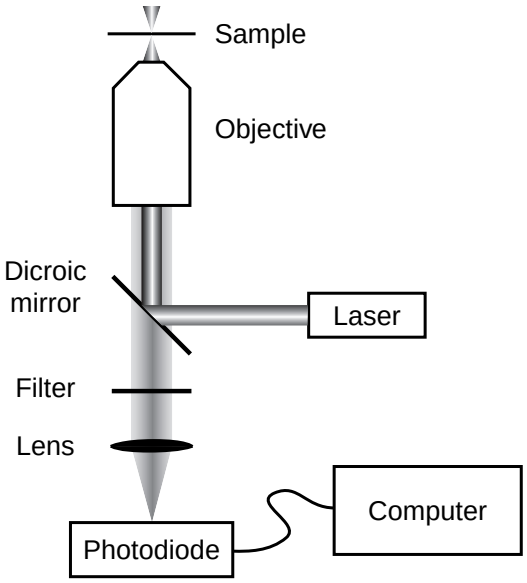

Figure 2: Schematic microscope setup for FCS measurements.

diffusion and molecular behaviour and fit them to the data. Such curves will be derived in the next section. Let us first note, however, that there are other ways to analyze the raw photon count data to extract more information than what one gets from the autocorrelation function. Among these are high-order autocorrelation analysis, see [20, 11, 13], and photon count histogram, see [10], but since we have not used them we will not discuss them much further.

2.3. Theoretical autocorrelation curves

We will now derive the theoretical autocorrelation function for diffusing molecules, which will show which parameters that influence the appearance of the function, as well as give us the necessary framework for deriving autocorrelation functions for other models, which we will do in section 4. The following derivation is essentially taken from [18], although somewhat simplified.

As before, we denote the fluorescence signal by $F(t)$, and define

$$\delta F(t) = F(t) - \langle F(t) \rangle , \qquad \langle F(t) \rangle = \frac{1}{T} \int_0^T F(t)\, dt.$$

Assuming that the fluctuations in fluorescence arise only from fluctuations in concentration and fluorescence properties of molecules inside the measurement volume V, essentially

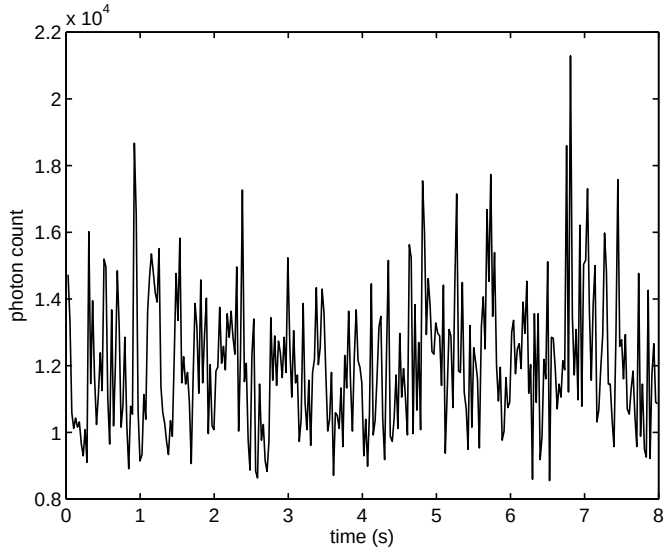

Figure 3: Photon counts in a typical FCS measurement. The counts are binned in time intervals of length 20 ms, and plotted for each such interval. Note that the scale on the vertical axis does not start at zero.

determined by the size of the focal spot of the laser beam, we may write

$$\delta F(t) = \kappa \int_V I(\mathbf{x}) S(\mathbf{x}) \delta\left(\sigma q C(\mathbf{x}, t)\right) d\mathbf{x} \tag{2.2}$$

where the following parameters are used:

κ	overall detection efficiency
$I(\mathbf{x})$	spatial distribution of the excitation energy
$S(\mathbf{x})$	optical transfer function, which describes the spatial collection efficiency for the emitted photons
σ	molecular absorption cross-section
q	quantum yield
$C(\mathbf{x}, t)$	the local particle concentration

In order to simplify this expression, we set $\eta = I_0 \kappa \sigma q$ and $W(\mathbf{x}) = I(\mathbf{x})/I_0 \cdot S(\mathbf{x})$, where I_0 is the maximum excitation energy (i.e. the maximum of $I(\mathbf{x})$). Now $W(\mathbf{x})$ is a dimensionless function which describes the spatial distribution of the emitted light, and η is a measure of the photon count rate per detected molecule per second.

With this new notation, (2.2) becomes

$$\delta F(t) = \int_V W(\mathbf{x}) \delta\left(\eta C(\mathbf{x}, t)\right) d\mathbf{x},$$

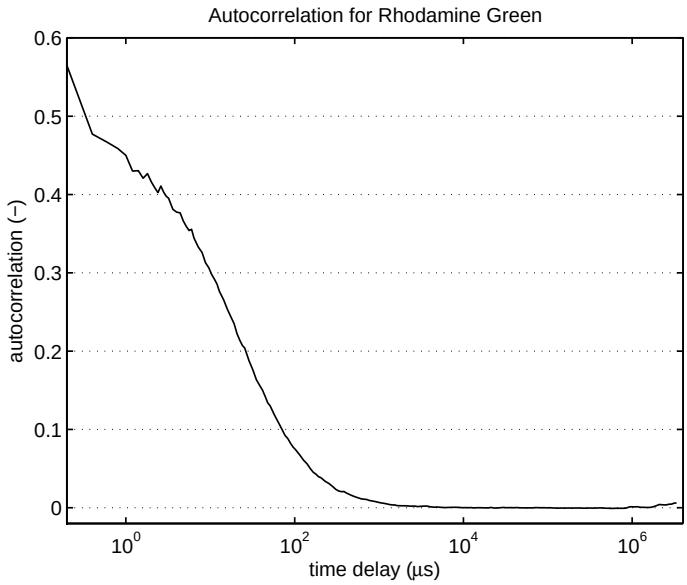

Figure 4: Autocorrelation curve for Rhodamine Green in aqueous solution. Note the logarithmic scale on the horizontal axis.

and if we insert this expression into the definition of the autocorrelation function (2.1), we get

$$G(\tau) = \frac{\int_V \int_V W(\mathbf{x}) W(\mathbf{x}') \big\langle \delta(\eta C(\mathbf{x},t))\, \delta(\eta C(\mathbf{x}', t+\tau)) \big\rangle \, d\mathbf{x}\, d\mathbf{x}'}{\big(\int_V W(\mathbf{x}) \big\langle \delta\left(\eta C(\mathbf{x},t)\right) \big\rangle \, d\mathbf{x}\big)^2}.$$

If we assume that the fluorescence properties do not change within the measurement time, that is η is constant, then this expression may be further simplified to

$$G(\tau) = \frac{\int_V \int_V W(\mathbf{x}) W(\mathbf{x}') \big\langle \delta C(\mathbf{x},0) \delta C(\mathbf{x}', \tau) \big\rangle \, d\mathbf{x}\, d\mathbf{x}'}{\big(\langle C \rangle \int_V W(\mathbf{x})\, d\mathbf{x}\big)^2}.$$

Now, the spatial distribution $W(\mathbf{x})$ is often approximated by a three-dimensional Gaussian

$$W(\mathbf{x}) = e^{-2\frac{x^2+y^2}{r_0^2}} e^{-2\frac{z^2}{z_0^2}}, \tag{2.3}$$

where z is in the axial direction, in which the focal point is broader than in the lateral x- and y-directions (so $z_0 > r_0$). And furthermore, the autocorrelation term

$$\langle \delta C(\mathbf{x}, 0) \delta C(\mathbf{x}', \tau) \rangle$$

may be interpreted as the probability that a molecule starting at $\mathbf{x}$ at time 0 will be at $\mathbf{x}'$ at time τ, times the concentration. This quantity is given by the fundamental solution ϕ of the equation that describes the motion of the particles, that is

$$\langle \delta C(\mathbf{x},0)\delta C(\mathbf{x}',\tau)\rangle = \langle C\rangle\, \phi(\mathbf{x},\mathbf{x}',\tau).$$

Thus, in general, the autocorrelation function is described by the expression

$$G(\tau) = \frac{1}{\langle C\rangle}\frac{\iint W(\mathbf{x})W(\mathbf{x}')\phi(\mathbf{x},\mathbf{x}',\tau)\, d\mathbf{x}\, d\mathbf{x}'}{\left(\int W(\mathbf{x})\, d\mathbf{x}\right)^2}, \tag{2.4}$$

where we may as well integrate over the whole space, since W decreases quickly anyway.

Assume now that we are studying ordinary diffusion in three dimensions with a diffusion constant D, described by the diffusion equation

$$\frac{\partial u}{\partial t}(\mathbf{x},t) - D\Delta u(\mathbf{x},t) = 0, \qquad \mathbf{x}\in\mathbb{R}^3,\ t\geq 0.$$

The fundamental solution is then given by

$$\phi(\mathbf{x},\mathbf{x}',\tau) = \frac{1}{(4\pi D\tau)^{3/2}} e^{-\frac{|\mathbf{x}-\mathbf{x}'|^2}{4D\tau}}, \qquad \mathbf{x},\mathbf{x}'\in\mathbb{R}^3.$$

If we insert this into (2.4), and define the effective focal volume V_{eff} to be

$$V_{\text{eff}} = \frac{\left(\int W(\mathbf{x})\, d\mathbf{x}\right)^2}{\int (W(\mathbf{x}))^2\, d\mathbf{x}} = \pi^{3/2} r_0^2 z_0,$$

we can compute the integrals and get the autocorrelation function for ordinary three-dimensional diffusion

$$G(\tau) = \frac{1}{V_{\text{eff}}\,\langle C\rangle}\frac{1}{1+\frac{\tau}{\tau_D}}\frac{1}{\sqrt{1+\frac{1}{S^2}\frac{\tau}{\tau_D}}}, \tag{2.5}$$

where $S = z_0/r_0$ and the diffusion time τ_D is given by

$$\tau_D = \frac{r_0^2}{4D}.$$

In figure 5 the autocorrelation function (2.5) has been plotted for a few different τ_D. Changing τ_D only shifts the curve left or right. We note also that for $\tau = 0$, only the first factor remains, and $V_{\text{eff}}\,\langle C\rangle$ is obviously the average number of particles in the effective volume, so $G(0)$ gives a measure of the concentration of the molecule we are measuring and the average number of molecules that we conduct our measurements on. When fitting the curve to experimental data, and in figure 5, we denote the quantity $G(0) = (V_{\text{eff}}\,\langle C\rangle)^{-1}$ by C_0. It may be noted that if the number of molecules is very large, the amplitude of the autocorrelation function is very small, and the function is hidden in noise. Therefore, the main importance of the recent development of the FCS method is that the measurement volume has been reduced so that the average number of particles is small, thereby increasing the signal-to-noise ratio.

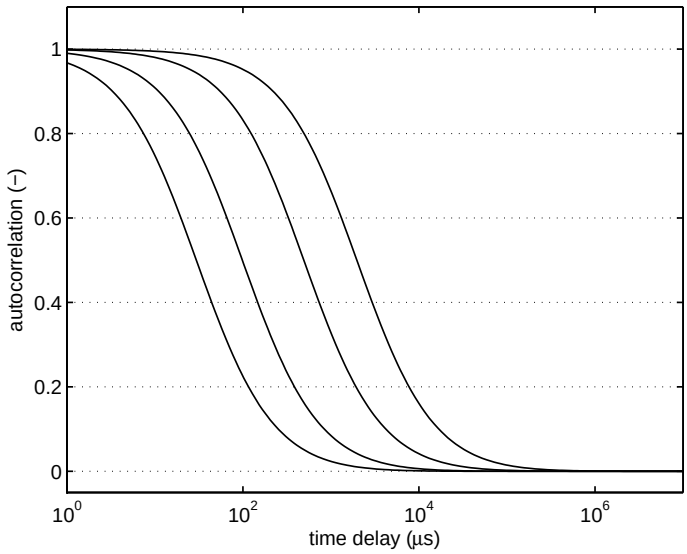

Figure 5: The autocorrelation for ordinary three-dimensional diffusion (2.5) plotted for $\tau_D = 30, 100, 500, 2000$. Note that increasing τ_D only shifts the curve to the right and that the value of τ_D corresponds to the value of τ where the slope of the curve is maximal. The other parameter values are $C_0 = 1$ and $S = 8$.

2.4. Other phenomena

2.4.1. Multiple species

When conducting an FCS experiment inside a cell, we can not be sure that the molecule that is fluorescently tagged is only diffusing by itself. It may bind to other molecules and form complexes with a distinctly different diffusion behaviour than the molecule by itself. Therefore we need a theoretical autocorrelation function where we can take this into account and fit multiple diffusion parameters to the experimental curve. This is done through the autocorrelation function

$$G(\tau) = \frac{1}{V_{\text{eff}}} \frac{\sum_i \eta_i^2 \langle C_i \rangle M_i(\tau)}{\left(\sum_i \eta_i \langle C_i \rangle\right)^2}$$

where the sum is over all fluorescent species, and $M_i(\tau)$ are individual motility terms such as the one in (2.5) (that is, if molecules are freely diffusing in three dimensions, $M_i(\tau)$ is the two last factors in this expression). We assume here that the motions of the different species are uncorrelated.

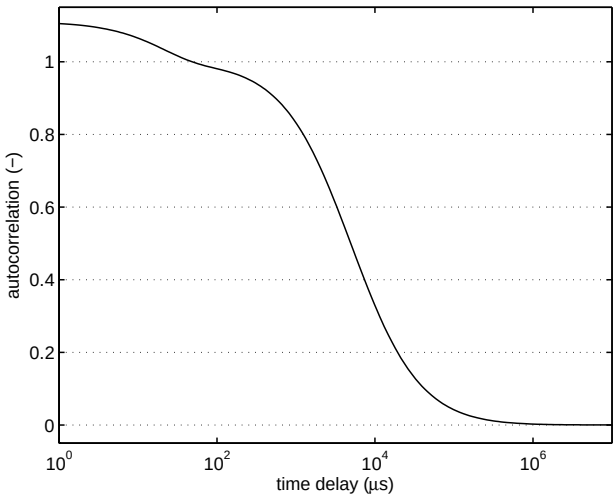

Figure 6: An autocorrelation curve for molecules with a triplet state. The parameters used are $C_0 = 1$, $\tau_D = 5000$ µs, $F_{\text{tr}} = 0.1$, $\tau_{\text{tr}} = 20$ µs, and $S = 8$. A large value of τ_D was chosen to show the effect of the triplet state correction. See also figure 12, where the same function has been fitted to experimental data.

2.4.2. Triplet states

In the previous derivation, we assumed that the fluorescence properties were constant, i.e. η was constant. This is not quite true. In the excitation process, the molecules may end up in a so called triplet state, the transition into which is forbidden by quantum mechanics and therefore results in a relatively long relaxation time, before the molecule is back in its ground state again. During this relaxation time, the molecule does not emit any light, and therefore this phenomenon shows up in the autocorrelation function as an additional hump for small time delays (around $1 - 20$ µs), see figure 6.

In order to correct for this behaviour, but avoid complicated computations, one usually assumes that the triplet state formation occurs on a much faster time-scale than the diffusional motion, and that the triplet state does not affect the diffusion of the molecules. In that case, that effect may be approximated by a multiplicative factor $X(\tau)$, so that

$$G(\tau) = X(\tau) G_{\text{diff}}(\tau)$$

where $G_{\text{diff}}(\tau)$ is given by (2.4) and $X(\tau)$ may be approximated by

$$X(\tau) = \frac{1 - F_{\text{tr}}(1 - e^{-\tau/\tau_{\text{tr}}})}{1 - F_{\text{tr}}}.$$

This correction is sometimes necessary to achieve a good fit with the experimental curve.

2.4.3. Photobleaching

Another phenomenon that has to be considered when conducting FCS experiments is photobleaching. This effect is caused by the fact that the fluorescent molecules have a limited lifetime, so that they stop emitting light after a period of repeated excitations. This has two important effects on the autocorrelation curve. First, if a molecule stops emitting light while it traverses the measurement volume, the effect is the same as if it had left the volume. Therefore, the autocorrelation for large τ decreases, leading to an over-estimation of the diffusion speed. Second, if we conduct measurements inside a cell, where the number of fluorescent molecules is very limited, photobleaching will lead to a notable decrease in the number of fluorescent molecules available for measurement. This is seen as an overall decrease in fluorescence signal, and also affects the autocorrelation function, which will not approach zero for large τ, but decrease even further, since there is a negative correlation for large τ.

In order to minimize these effects, it is important to use as low laser power as possible, as this reduces the rate of photobleaching, while ensuring that the signal is still strong enough to give reliable measurements. One may also try to compensate for the bleaching by correcting the autocorrelation function, but this is rather difficult and unreliable.

3. Anomalous diffusion

Anomalous diffusion is a phenomenon that has been observed in a large number of environments in recent years, when measurement techniques, including FCS, have reached a level of accuracy where the effect may be observed. Here, we focus on biological systems, and in particular diffusion inside cells, but the phenomenon has also been observed for for example charged carriers in amorphous semiconductors, diffusion in percolative and porous systems, and diffusion in polymeric networks. A list of references may be found in [9]. In biological cells, anomalous diffusion has been observed by for example Wachsmuth et al. [21] and Weiss, Nilsson et al. [24, 23, 22], and we will also see examples later where our own FCS measurements indicate anomalous diffusion.

There are at least two ways to characterize anomalous diffusion. One is to study moving particles and observe the mean squared displacement $\langle x^2(t)\rangle$. For Brownian motion, the stochastic model describing ordinary diffusion, one observes that

$$\left\langle x^2(t)\right\rangle \sim Dt$$

where D is the diffusion coefficient. But in the situations mentioned above one has observed instead

$$\left\langle x^2(t)\right\rangle \sim Kt^\beta$$

for some $\beta > 0$, and it is this behaviour that is called anomalous diffusion. For $\beta < 1$, one gets subdiffusion (or slow diffusion), while for $\beta > 1$ one gets superdiffusion (or fast diffusion). Both types of anomalous diffusion have been observed, but here we are only interested in subdiffusion, since it is the type that may occur inside biological cells.

Another way to describe anomalous diffusion is on the level of differential equations. The diffusion equation

$$\frac{\partial u}{\partial t}(x,t) - D\Delta u(x,t) = 0 \tag{3.1}$$

for a molecule concentration u is derived from two relations. The first is the conservation law

$$\frac{\partial u}{\partial t}(x,t) + \nabla \cdot \Phi(x,t) = 0, \tag{3.2}$$

where Φ is the flux of molecules, which is derived using Gauss' divergence theorem, and expresses the fact that the increase in the amount of molecules inside a test volume is equal to the net amount of material that flows into the volume (assuming, as we have done here, that no production takes place). The second relation is Fick's second law, which states that

$$\Phi(x,t) = -D\nabla u(x,t).$$

This is an experimental result which holds for systems exhibiting Brownian motion, but it does not hold in the case of anomalous diffusion, and in that case a different relation has to be derived from experimental data.

Before discussing the resulting equations, we will briefly discuss the possible causes of subdiffusive behaviour. It is clear that since subdiffusion is slower than ordinary diffusion, there is some kind of obstruction that causes the subdiffusion. In a series of articles, Saxton [16, 17] used Monte-Carlo simulations to investigate the effects of physical obstacles and of molecular binding on the diffusion behaviour. He found that it was possible to get subdiffusive behaviour both in the case where the diffusing molecules were obstructed by obstacles of certain configurations, and in the case when diffusing molecules were temporarily bound at predefined sites with varying relaxation time, although not all such configurations lead to subdiffusion. Another hint is that subdiffusive behaviour is found when studying diffusion on fractals (see [1] and [9]), which indicates that highly obstructed diffusion may lead to subdiffusion. It is also a fact that the interior of a cell is full of obstacles for a diffusing molecule the size of a protein. These obstacles may consist of both larger structures (like organelles) and smaller ones (like molecular complexes), compact ones (like ribosomes) and web-like ones (like the cytoskeleton), and the diffusing protein may also bind to other large molecules and structures, in order to perform its tasks in the cellular machinery. It is therefore not surprising that a protein diffusing in a cell may exhibit subdiffusive behaviour, even if the exact reasons for it remain unclear.

We will now discuss two types of pseudo-differential equations that may be used to model subdiffusion. Let us first remark, though, that there is also a number of stochastic models which we do not discuss here. For a more complete discussion and a list of references on this topic, see [9].

3.1. THE TIME-ANOMALOUS DIFFUSION EQUATION

The time-anomalous diffusion equation (or (time-)fractional diffusion equation) may be derived using a continuous time random walk (CTRW) model (see [9]). In that model,

one studies a particle performing jumps on a lattice, with a given waiting time probability distribution $w(t)$ for the time between two successive jumps, and a jump length probability distribution $\lambda(x)$ for the length of the jumps. If $w(t)$ has a finite mean (or expected waiting time), and $\lambda(x)$ has finite variance, one obtains the ordinary diffusion equation (3.1) as a lowest order approximation. But if the expected waiting time is infinite, that is the waiting time distribution $w(t)$ has a broad tail with high probability for long waiting times between jumps, chosen to give the desired behaviour $\langle x^2(t)\rangle \sim K_\beta t^\beta$, then the same arguments (see [9]) give the time-anomalous diffusion equation

$$\frac{\partial u}{\partial t}(\mathbf{x},t) = {}_0D_t^{1-\beta} K_\beta \Delta u(\mathbf{x},t) \tag{3.3}$$

with initial condition $u(\mathbf{x},0) = u_0(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^3$ and $0 < \beta < 1$, and where the so called Riemann-Liouville operator ${}_0D_t^{1-\beta}$ is defined through

$${}_0D_t^{1-\beta} u(\mathbf{x},t) = \frac{\partial}{\partial t} {}_0D_t^{-\beta} u(\mathbf{x},t) = \frac{1}{\Gamma(\beta)} \frac{\partial}{\partial t} \int_0^t \frac{u(\mathbf{x},t')}{(t-t')^{1-\beta}}\, dt'. \tag{3.4}$$

This operator is called a fractional derivative of order $1-\beta$, a name motivated by the relation

$${}_0D_t^{1-\beta} t^p = \frac{\Gamma(1+p)}{\Gamma(p+\beta)} t^{p+\beta-1}$$

for any $p \in \mathbb{R}$, which is a generalization of integer order derivatives. This implies that when this operator is applied to a constant, it does not give zero as integer order differentiation does, but instead (with $p = 0$)

$${}_0D_t^{1-\beta} 1 = \frac{1}{\Gamma(\beta)} t^{\beta-1}.$$

A more thorough discussion of the Riemann-Liouville operator and the time-anomalous diffusion equation (3.3) may be found in [9] and [12]. Here, we just remark that the integral in (3.4) expresses the fact that the diffusion depends on the concentration at previous times, or in other words that the system has a memory. This is thought to model that molecules bind to structures in a cell, and that this is the cause for subdiffusion. But exactly this property also makes the model somewhat problematic, since it puts time $t = 0$ in a very special position, in that the change in concentration $u(\mathbf{x}, t_1)$ at time t_1 is influenced by the solution for all times from $t = 0$ to $t = t_1$, but not by $u(\mathbf{x}, t)$ for $t < 0$. This means that the solution to the initial value problem (3.3) depends on what we choose as time $t = 0$, and this seems a bit unsatisfactory. The fact that u depends on previous states may be expressed as the model not having the Markov property, and this is a consequence of the CTRW model used, and the fact that we want to have $\langle x^2(t)\rangle \sim K_\beta t^\beta$ with $\beta < 1$.

3.2. THE SPACE-ANOMALOUS DIFFUSION EQUATION

Motivated by the above discussion, we seek to find an equation which has the Markov property, that is has no memory, but still is capable of modeling subdiffusion. For this, we

introduce the space-anomalous diffusion equation

$$\begin{cases} \dfrac{\partial u}{\partial t}(\mathbf{x},t) + K_\alpha(-\Delta)^\alpha u(\mathbf{x},t) = 0, & \mathbf{x}\in\mathbb{R}^3,\ t>0, \\ u(\mathbf{x},0) = u_0(\mathbf{x}), & \mathbf{x}\in\mathbb{R}^3 \end{cases} \tag{3.5}$$

with $\alpha > 1$ and diffusion parameter K_α with dimension $\mathrm{m}^{2\alpha}\mathrm{s}^{-1}$. The pseudo-differential operator $(-\Delta)^\alpha$ is defined through the Fourier transform as

$$(-\Delta)^\alpha u(\mathbf{x},t) = \mathscr{F}^{-1}\left(|\xi|^{2\alpha}\widehat{u}(\xi,t)\right).$$

This equation satisfies the conservation law (3.2) with a flux

$$\Phi(\mathbf{x},t) = K_\alpha(-\Delta)^{\alpha-1}\nabla u(\mathbf{x},t)$$

which is non-local in space, meaning that not just the local behaviour, but the behaviour of u far away from $\mathbf{x}$, influences the concentration u at $\mathbf{x}$. How to interpret this property is not quite clear, but it may be noted that when $1 < \alpha < 2$, we are somewhere between the diffusion equation ($\alpha = 1$) and elasticity equations ($\alpha = 2$), which indicates that equation (3.5) describes a situation where molecules interact with each other, so to say pulling at each other, making the motion slower and different in character compared to ordinary diffusion.

The space-anomalous diffusion equation (3.5) also has some questionable properties. To begin with, we may compute the mean squared displacement $\langle x^2(t)\rangle$ using the fundamental solution $\phi(\mathbf{x},t)$ to (3.5) (with $u_0(\mathbf{x}) = \delta_0(\mathbf{x})$, the Dirac delta-function), which is found by applying the Fourier transform in the $\mathbf{x}$-variable to (3.5). When we do that, we get

$$\begin{cases} \dfrac{\partial\widehat{\phi}}{\partial t}(\boldsymbol{\xi},t) + K_\alpha|\boldsymbol{\xi}|^2\widehat{\phi}(\boldsymbol{\xi},t) = 0, & \boldsymbol{\xi}\in\mathbb{R}^3,\ t>0, \\ \widehat{\phi}(\boldsymbol{\xi},0) = \widehat{\delta_0}(\boldsymbol{\xi}) = 1, \end{cases}$$

which is an ODE in t with solution

$$\widehat{\phi}(\boldsymbol{\xi},t) = e^{-tK_\alpha|\boldsymbol{\xi}|^{2\alpha}}. \tag{3.6}$$

There is a representation of $\phi(\mathbf{x},t)$ as an infinite sum, but we do not need it here.

Using $\widehat{\phi}(\boldsymbol{\xi},t)$, we can compute

$$\left\langle x^2(t)\right\rangle = \sum_{i=1}^3 \int_{\mathbb{R}^3} x_i^2\phi(\mathbf{x},t)\,d\mathbf{x} = \sum_{i=1}^3 \left\langle -(\delta_0)''_{ii}, \widehat{\phi}(\cdot,t)\right\rangle = -\sum_{i=1}^3 \frac{\partial^2}{\partial{\xi_i}^2}\widehat{\phi}(\boldsymbol{\xi},t)\bigg|_{\boldsymbol{\xi}=0} = 0$$

if $\alpha > 1$. From this we see that not only do we not get the relation $\langle x^2(t)\rangle \sim Kt^\beta$ for some β, but also that $\phi(\mathbf{x},t)$ must be negative somewhere, since otherwise we could not get zero, and therefore solutions to (3.5) may have negative values, which is of course a problem if we want the solution $u(\mathbf{x},t)$ to be a concentration. However, when we study

the equation inside bounded domains, as we will in the paper that follows, we need not get non-positive solutions since the solution may be everywhere far from zero, and tends to smooth out with time. And even if we do not get the same time dependence of $\langle x^2(t)\rangle$ as for the time-anomalous diffusion equation, we will see further on that the solutions to the space-anomalous diffusion equation (3.5) fit well to experimental data — just as well as solutions to the time-anomalous diffusion equation.

It should also be noted here that the operator $(-\Delta)^\alpha$ in (3.5) is not the only possible choice for a space-anomalous diffusion equation, although it is probably the simplest, which is why we use it here and in the paper that follows. In fact, the symbol $|\boldsymbol{\xi}|^{2\alpha}$ in (3.6) could be exchanged for anything that has desirable properties, and it would be interesting to investigate if it is possible to find a symbol that gives solutions that fit the experimental data, while for example giving a positive fundamental solution.

4. New autocorrelation curves

4.1. Autocorrelation curve for space-anomalous diffusion

In this section, we derive the autocorrelation curve for space-anomalous diffusion, that is the type of diffusion that is described by equation (3.5).

According to (2.4), we need to compute the integral

$$\iint W(\mathbf{x})W(\mathbf{x}')\phi(\mathbf{x}-\mathbf{x}',\tau)\,d\mathbf{x}'\,d\mathbf{x} \tag{4.1}$$

in order to find the autocorrelation function $G(\tau)$. Here $\phi(\mathbf{x},\tau)$ is the fundamental solution (3.6) to (3.5).

With the use of Parseval's formula, we may write the integral (4.1) as

$$\int_{\mathbb{R}^3} W(\mathbf{x})\,[W*\phi]\,(\mathbf{x},\tau)\,d\mathbf{x} = \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3}\widehat{W}(\boldsymbol{\xi})\widehat{W}(\boldsymbol{\xi})\widehat{\phi}(\boldsymbol{\xi},\tau)\,d\boldsymbol{\xi}. \tag{4.2}$$

As before,

$$W(\mathbf{x}) = \exp\left(-2\frac{x^2+y^2}{r_0^2} - 2\frac{z^2}{z_0^2}\right),$$

and we may easily compute

$$\widehat{W}(\boldsymbol{\xi}) = \frac{\pi^{3/2}r_0^2 z_0}{2\sqrt{2}}\exp\left(-\frac{1}{8}(z_0^2\zeta^2 + r_0^2(\xi^2+\eta^2))\right),$$

where $\boldsymbol{\xi} = (\xi,\eta,\zeta)$. If we insert this into (4.2), set $S = z_0/r_0$, $\tau_\alpha = \frac{r_0^{2\alpha}}{4^\alpha K_\alpha}$, and make the

change of variables $\boldsymbol{\xi}' = \frac{r_0}{2}\boldsymbol{\xi}$, we get

$$\begin{aligned}
&\frac{\pi^3 r_0^4 z_0^2}{8(2\pi)^3} \iiint_{\mathbb{R}^3} \exp\left(-\frac{z_0^2\zeta^2}{4} - \frac{r_0^2}{4}(\xi^2+\eta^2) - |\boldsymbol{\xi}|^{2\alpha} K_\alpha \tau\right) d\xi\, d\eta\, d\zeta \\
&\quad = \frac{r_0^2 z_0 S}{8} \iiint_{\mathbb{R}^3} \exp\left(-S^2\zeta'^2 - \xi'^2 - \eta'^2 - \frac{\tau}{\tau_\alpha}|\boldsymbol{\xi}'|^{2\alpha}\right) d\xi'\, d\eta'\, d\zeta' \\
&\quad = \frac{\pi r_0^2 z_0 S}{4} \int_{\mathbb{R}} e^{-S^2\zeta'^2} \int_0^\infty r \exp\left(-r^2 - \frac{\tau}{\tau_\alpha}(r^2+\zeta'^2)^\alpha\right) dr\, d\zeta' \\
&\quad \equiv \frac{\pi r_0^2 z_0 S}{4} M(\tau)
\end{aligned}$$

where we changed to polar coordinates in (ξ', η') in the last step. The double integral $M(\tau)$ will have to be computed numerically, but this can be done quite efficiently as the integrand is smooth and decays quickly to zero. Finally, to get the autocorrelation function, we also need to compute the denominator in (2.4), but

$$\left(\int_{\mathbb{R}^3} W(\mathbf{x})\, d\mathbf{x}\right)^2 = \left(\frac{\pi^{3/2} r_0^2 z_0}{2\sqrt{2}}\right)^2$$

so the autocorrelation function for space-anomalous diffusion becomes

$$G(\tau) = \frac{1}{\langle C\rangle}\frac{2S}{\pi^2 r_0^2 z_0} M(\tau) = \frac{1}{V_{\text{eff}}\langle C\rangle}\frac{2S}{\sqrt{\pi}} \int_{\mathbb{R}} e^{-S^2\zeta^2} \int_0^\infty r e^{-r^2-(r^2+\zeta^2)^\alpha \tau/\tau_\alpha}\, dr\, d\zeta. \tag{4.3}$$

It may be noted that for $\tau = 0$, the integrals may be computed explicitly and we get

$$G(0) = \frac{1}{V_{\text{eff}}\langle C\rangle}$$

just as for the ordinary diffusion case. Also, if we set $\alpha = 1$, (4.3) reduces to (2.5) as expected.

In figure 7, we see the autocorrelation function (4.3) plotted for a number of different α. Most notably, changing α changes the slope of the autocorrelation function, while changing the anomalous diffusion time τ_α only shifts the curve left or right.

4.2. AUTOCORRELATION CURVE FOR TIME-ANOMALOUS DIFFUSION

As we have seen, time-anomalous diffusion is described by the equation

$$\begin{cases} \dfrac{\partial u}{\partial t}(\mathbf{x},t) - {}_0D_t^{1-\beta} K_\beta \Delta u(\mathbf{x},t) = 0 & \mathbf{x}\in\mathbb{R}^3,\ t>0 \\ u(\mathbf{x},0) = u_0(\mathbf{x}) & \mathbf{x}\in\mathbb{R}^3 \end{cases} \tag{4.4}$$

with $0 < \beta < 1$ and the Riemann-Liouville operator ${}_0D_t^{1-\beta}$ defined in (3.4), and K_β a diffusion parameter of dimension $\text{m}^2\text{s}^{-\beta}$. The fundamental solution of (4.4) may be

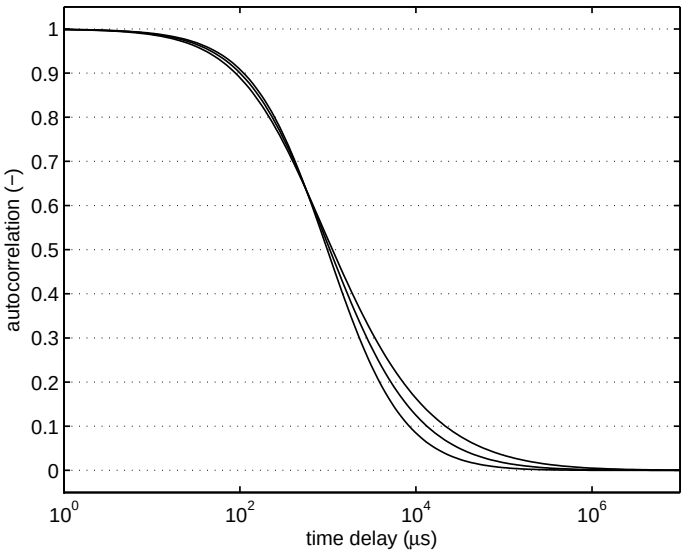

Figure 7: Autocorrelation functions for space-anomalous diffusion, for $\alpha = 1.0, 1.25, 1.5$, with $\tau_\alpha = 1000$, $C_0 = 1$ and $S = 8$. Increasing α makes the curve less steep, and for $\alpha = 1.0$ we get the ordinary diffusion curve.

described as a so called Fox function [9, 8, 7], which has a rather complicated definition and can not be computed really efficiently. It is therefore not very practical to insert this fundamental solution into (2.4) and derive the autocorrelation function to use it for fitting to experimental curves, since this would be computationally infeasible.

Instead, a more practical approach is usually employed. As mentioned in section 3.1, solutions to (4.4) have the property that

$$\left\langle x^2(t) \right\rangle \sim K_\beta t^\beta.$$

For this reason, it is argued that we may replace τ/τ_D in the autocorrelation function for ordinary diffusion (2.5) with $(\tau/\tau_\beta)^\beta$, thus getting

$$G(\tau) = \frac{1}{V_{\text{eff}} \langle C \rangle} \frac{1}{1 + \left(\frac{\tau}{\tau_\beta}\right)^\beta} \frac{1}{\sqrt{1 + \frac{1}{S^2}\left(\frac{\tau}{\tau_\beta}\right)^\beta}} \tag{4.5}$$

as the autocorrelation for time-anomalous diffusion. Here $(\tau_\beta)^\beta = r_0^2/\Gamma_\beta$, where Γ_β is a constant related to the diffusion parameter K_β.

In [22], Weiss et al claim that they have numerically computed the correct autocorrelation function from the fundamental solution in terms of Fox functions, and compared this

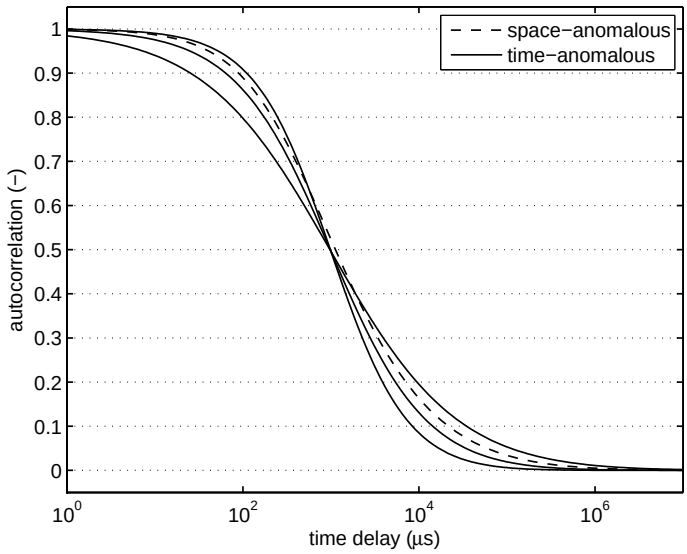

Figure 8: Autocorrelation functions for the time-anomalous diffusion equation, with $\beta = 0.6, 0.8, 1.0$ (solid lines). Lower β gives a less steep curve, and $\beta = 1$ gives the ordinary diffusion curve. For reference, an autocorrelation function for the space-anomalous case (with $\alpha = 1.5$) has been included (dashed line), and we see that the time-anomalous curves are not equivalent to the space-anomalous curves. The other parameter values used are $C_0 = 1$, $\tau_\beta = \tau_\alpha = 1000$, $S = 8$.

function to the expression for $G(\tau)$ in (4.5) with good agreement, namely that the when fitting one of the autocorrelation functions to the other, the exponents β differed by less than 10%, which is within the accuracy of experimental data. So, it seems like (4.5) is correct enough to be used for fitting with data. In figure 8, we see the autocorrelation function (4.5) plotted for a few different values of β, and compared to the space-anomalous autocorrelation function (4.3).

4.3. Autocorrelation curve for diffusion inside a closed domain

Since we make FCS measurements inside a cell which has a volume comparable to the measurement volume (the radius of the measurement volume is typically $r_0 \approx 0.15$ µm, and the yeast cells are about $5 - 10$ µm in diameter), we would expect a slightly different appearance of the autocorrelation function. To derive the autocorrelation function for this situation, we obviously need to take the boundary conditions into account. We note that

the numerator in the autocorrelation function (2.4) is

$$\left\langle W, \phi * W(\cdot, \tau) \right\rangle_2 = \left\langle W, u(\cdot, \tau) \right\rangle_2,$$

where ϕ is the fundamental solution for the equation that models the diffusion, and $u(\mathbf{x}, t)$ consequently is the solution to the diffusion equation with initial value $u_0 = W$. So what we want to compute for our situation is $\left\langle u_0, u \right\rangle_2$ when u is the solution to a diffusion equation inside a domain with impenetrable boundary, and with the initial value u_0 being a three-dimensional Gaussian as in (2.3).

In order to simplify the calculations, we make everything spherically symmetric. We let the domain Ω be a ball of radius R_0 in $\mathbb{R}^3$, we let $S = 1$ in (2.3) (although this is not true for the experimental setup), and we use a Neumann condition on the boundary, to model that the molecules may not move across the membrane.

The resulting boundary value problem is

$$\begin{cases} u_t(\mathbf{x}, t) - D\Delta u(\mathbf{x}, t) = 0, & |\mathbf{x}| < R_0, t > 0 \\ \frac{\partial u}{\partial n}(\mathbf{x}, t) = 0 & |\mathbf{x}| = R_0, t > 0 \\ u(\mathbf{x}, 0) = u_0(\mathbf{x}) = \langle C \rangle \, e^{-2\frac{|\mathbf{x}|^2}{r_0^2}} & |\mathbf{x}| \leq R_0 \end{cases}, \tag{4.6}$$

for $\mathbf{x} \in \mathbb{R}^3$, where we assume that $r_0 < R_0$.

This problem is solved by separation of variables and expansion in eigenfunctions. The solution to similar problems can be found for example in Folland [3, chapter 6], and what follows is just an application of this standard method to our problem. We set $r = |x|$ and make the ansatz $u(r, t) = T(t)R(r)$. Inserting this into the equation, using the expression for the Laplacian in spherical coordinates, we get the two equations

$$\begin{aligned} T'(t) &= -\lambda^2 T(t) \\ R'(r) &= r^2 R''(r) + 2rR'(r) + \lambda^2 r^2 R(r) = 0 \end{aligned}$$

where λ is an eigenvalue to be determined. The equation for T obviously has the solution $T(t) = T(0)\exp(-\lambda^2 t)$, and the equation for R can be transformed into Bessel's equation, which gives the solution

$$R(r) = \frac{1}{\sqrt{r}} J_{1/2}(\lambda r),$$

where $J_{1/2}$ is the Bessel function of order $1/2$, which may be written

$$J_{1/2}(s) = \sqrt{\frac{2}{\pi s}} \sin s.$$

Now, we need to determine the eigenvalues λ. To do that, we use the boundary condition, which may now be written $R'(R_0) = 0$, or

$$0 = \frac{d}{dr}\left(\sqrt{\frac{2}{\pi}} \frac{\sin(\lambda r)}{\sqrt{\lambda} r} \right)\Bigg|_{r=R_0} = \sqrt{\frac{2}{\pi}} \frac{\lambda R_0 \cos(\lambda R_0) - \sin(\lambda R_0)}{\sqrt{\lambda} R_0^2},$$

which amounts to

$$\lambda R_0 = \tan(\lambda R_0),$$

where we only consider positive solutions λ, giving us $\{\lambda_k\}_{k=1}^{\infty} = \{z_k/R_0\}_{k=1}^{\infty}$ where z_k are the positive solutions to $\tan z = z$. The first solution is really $\lambda_0 = 0$, but for this value of λ, the equation for R reduces to

$$r^2 R''(r) + 2rR'(r) = 0,$$

with the solution $R(r) = A_1 + A_2 r^{-2}$. But since we want R to be bounded, we must have $A_2 = 0$, and we also choose $A_1 = 1$, since the solution will be normalized in the end anyway.

The full solution to (4.6) is therefore

$$u(r,t) = c_0 + \sum_{k=1}^{\infty} c_k e^{-\lambda_k^2 t} \frac{\sin \lambda_k r}{\sqrt{\lambda_k} r},$$

where the coefficients c_k are given by

$$c_k = \frac{\langle u_0, F_k \rangle_2}{\|F_k\|_2^2}, \quad k = 0, 1, 2, \ldots$$

with

$$F_0(r) = 1$$
$$F_k(r) = \sqrt{\frac{2}{\pi}} \frac{\sin \lambda_k r}{\sqrt{\lambda_k} r}, \quad k = 1, 2, \ldots$$

With u_0 as in (4.6), we compute

$$c_0 = \left(\frac{4\pi R_0^3}{3}\right)^{-1} 4\pi \int_0^{R_0} e^{-2r^2/r_0^2} r^2 \, dr = -\frac{3r_0^2}{4R_0} e^{-2R_0^2/r_0^2} + \frac{3\sqrt{\pi} r_0^3}{8\sqrt{2} R_0^3} \mathrm{Erf}(\sqrt{2}\tfrac{R_0}{r_0})$$

and, for $k \geq 1$,

$$\begin{aligned} c_k &= \frac{R_0 \lambda_k}{4R_0^2 \sin^2 R_0 \lambda_k} 4\pi \sqrt{\frac{2}{\pi}} \int_0^{R_0} e^{-2r^2/r_0^2} \frac{\sin \lambda_k r}{\sqrt{\lambda_k} r} r^2 \, dr \\ &= \frac{\pi R_0^{5/2}}{\sqrt{2}} \frac{r_0^3}{8R_0^3} e^{-r_0^2 z_k^2 / 8R_0^2} \Psi(\tfrac{r_0}{R_0}, z_k) - \sqrt{2\pi R_0} r_0^2 e^{-2R_0^2/r_0^2} \frac{\sin z_k}{z_k} \end{aligned}$$

where $\Psi(\frac{r_0}{R_0}, z_k) = \mathrm{Erf}\big(\sqrt{2}\frac{R_0}{r_0} - i\frac{r_0 z_k}{4R_0}\big) + \mathrm{Erf}\big(\sqrt{2}\frac{R_0}{r_0} + i\frac{r_0 z_k}{4R_0}\big)$ and $\mathrm{Erf}(t) = \int_0^t e^{-\tau^2} \, d\tau$ is the error function. Now, if R_0/r_0 is large ($\gtrsim 10$ or so), $\Psi(\frac{r_0}{R_0}, z_k) \approx 2$ for small z_k, and for large z_k the exponential factor in front will dominate, so we may set $\Psi(\frac{r_0}{R_0}, z_k) = 2$. Furthermore, $e^{-2R_0^2/r_0^2}$ is small compared to the other terms, so we approximate it by zero; and finally, $\mathrm{Erf}(\sqrt{2}R_0/r_0) \approx 1$ for large R_0/r_0, yielding

$$c_0 \approx -\frac{3r_0^2}{4R_0} e^{-2R_0^2/r_0^2} \qquad c_k \approx \frac{\pi r_0^3}{4\sqrt{2R_0}} e^{-r_0^2 z_k^2/8R_0^2}, \quad k = 1, 2, \ldots$$

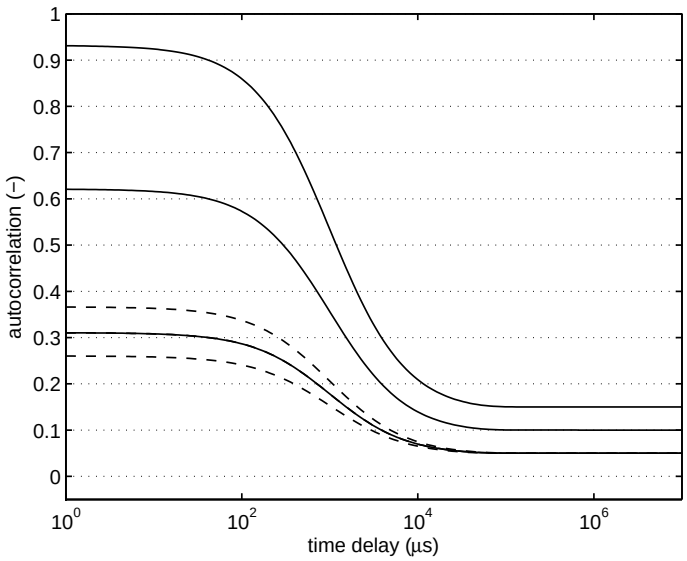

Figure 9: The autocorrelation curve for ordinary diffusion in a closed domain. The function (4.7) is plotted for $R_0/r_0 = 10$ and $(|\Omega|\,\langle C\rangle)^{-1} = 0.05,\ 0.10,\ 0.15$ (solid lines, from bottom up), and with $(|\Omega|\,\langle C\rangle)^{-1} = 0.05$ and $R_0/r_0 = 9,\ 11$ (dashed lines, from bottom up). The autocorrelation curve does not approach zero as τ gets large.

But what we really want to compute is the correlation function $G(\tau)$ given by (2.4), which in the current setting amounts to

$$G(\tau) = \frac{\big\langle u(r,\tau), u_0(r)\big\rangle_2}{\langle C\rangle\,\big(\int_\Omega u_0(\mathbf{x})\,d\mathbf{x}\big)^2} = \frac{8}{\langle C\rangle\,\pi^3 r_0^6}\sum_{k=0}^{\infty}\frac{|\big\langle u_0, F_k\big\rangle_2|^2}{\|F_k\|_2^2}\exp\big(-\big(\tfrac{z_k}{R_0}\big)^2 D\tau\big).$$

With the approximate c_k above and since

$$\frac{|\big\langle u_0, F_k\big\rangle_2|^2}{\|F_k\|_2^2} = \|F_k\|_2^2\,c_k^2, \qquad k = 0, 1, 2, \ldots$$

we get

$$G(\tau) \approx \frac{1}{|\Omega|\,\langle C\rangle}\left(1 + \frac{1}{12}\sum_{k=1}^{\infty}\Big(z_k + \frac{1}{z_k}\Big)\exp\Big(-\tfrac{z_k^2}{4}\big(\tfrac{r_0}{R_0}\big)^2\big(1 + \tfrac{\tau}{\tau_D}\big)\Big)\right) \tag{4.7}$$

with $|\Omega| = \frac{4}{3}\pi R_0^3$ and $\tau_D = r_0^2/4D$. Note that $z_k + z_k^{-1} = z_k/\sin^2 z_k$, since $\tan z_k = z_k$. The resulting autocorrelation curve is shown in figure 4.3.

The most notable fact about this autocorrelation function is that it does not decrease to zero when τ is large. Instead, we get

$$G(\tau) \to \frac{1}{|\Omega| \langle C \rangle} \approx \frac{1}{N_{\text{tot}}}, \quad \tau \to \infty,$$

where N_{tot} is the total number of molecules in the cell. This reflects the fact that in a small domain, the molecules have the possibility to move out from the measurement volume, 'bounce' at the domain boundary, and move back into the measurement volume, thereby giving a contribution to the autocorrelation. In practice, however, such correlations on long time scales will not be observed, since other effects that affect the molecules, such as photobleaching and protein degradation, will come into play and decrease the correlation. But this does not mean that we can not observe any effects due to the small domain, just that it will not be as easy to distinguish them from other effects.

5. FCS IN S. POMBE

This section describes briefly our FCS measurements done in the yeast species *Schizosaccharomyces pombe*, or fission yeast. What the cells look like is shown in figure 1. They are approximately 10 µm long and 3 µm wide. We have conducted measurements on three proteins involved in a signaling cascade, these are Yap1, Mkp1 and Mkp2, and they are all kinases of moderate size (40–70 kDa[†]). All three show similar results, but we only show data for Mkp1 here.

The main purpose here is to show that we actually get data that fits the anomalous diffusion model, and in particular that the space-anomalous diffusion model fits the data just as well as the time-anomalous model.

5.1. EXPERIMENTAL CONSIDERATIONS

There were some issues to take into consideration when performing FCS measurements inside the small yeast cells. First of all, the number of molecules inside each cell is limited, so photobleaching is a problem, as has already been discussed in section 2.4.3. Furthermore, the yeast cells have a tendency to move around in their medium, and therefore the cells had to be mixed into an agarose gel and kept at a temperature at which the gel was solid enough so that the cells could not move during the measurement time. And finally, we had to try to measure at a good place inside the cell, which is not so easy since we do not see any internal structure inside the cell (except perhaps the nucleus, as in figure 1). Sometimes we seemed to hit some larger structure in the cell, and that showed up in the measurement. Therefore, not all measurements gave nice autocorrelation curves, and what we show below is one of the curves that fit nicely to the anomalous diffusion model. It should be noted, though, that even if the anomalous diffusion model does not always fit

[†]1 Da (Dalton) = 1 u = $\frac{1}{12}$ of the mass of a ^{12}C-atom. Proteins vary in size between approximately 2 kDa to 500 kDa

the data, it does not mean that the ordinary diffusion model fits the data better. It just means that there are phenomena influencing the correlation curve that are not included in the models we use, which is not surprising since there is very much going on inside the cell which we do not know about and can not include in our models, and we have to choose those measurements where those unwanted phenomena are kept to a minimum.

5.2. Results

Figure 10 shows an autocorrelation curve computed from one of our FCS experiments on the protein Mkp1 tagged with GFP in *S. pombe*. The measurements lasted for 60 seconds and the laser beam was focused in an area well outside the nucleus, to avoid the large structures in and near the nucleus. After the measurement it was checked that the cell had not moved, and that the correlation curve looked reasonable, otherwise the measurement was discarded for reasons discussed in the previous section. But many curves gave similar results as the one shown in figure 10, though more experiments would have to be performed in order to draw proper conclusions about parameter values, which are seen to vary between individual measurements, and also between the three proteins that we conducted experiments on.

A standard calibration experiment using the dye Rhodamine Green in water solution was conducted in order to determine the parameters r_0 and S in (2.3). Fitting the ordinary diffusion correlation curve (2.5) to this experimental curve yielded $S \approx 8.3$ and $\tau_D \approx 20$ µs, which, using the known diffusion coefficient for Rhodamine Green $D \approx 2.8 \cdot 10^{-10}$ m^2/s [15], yields the approximate radius of the measurement volume $r_0 \approx 0.15$ µm.

The curve fits in figure 10 are made using theoretical autocorrelation functions for space-anomalous, time-anomalous and single species ordinary diffusion (see sections 4.1, 4.2, 2.3), with an additional factor taking care of the contribution from triplet states, as in section 2.4.2, and using the value of S from the calibration experiment. The parameter values achieved in the fitting are shown in table 1.

	Space-anomalous	Time-anomalous	Ordinary
C_0	0.22	0.23	0.20
τ_α, τ_β, τ_D (µs)	731	726	1100
α, β	1.38	0.78	N/A
F_{tr}	0.11	0.08	0.18
τ_{tr} (µs)	15	9.3	40
K_α, Γ_β, D	$3.3 \cdot 10^{-23}$	$2.4 \cdot 10^{-21}$	$5.2 \cdot 10^{-12}$

Table 1: Parameter values for the curve fits in figure 10. K_α has units m$^{2\alpha}$s^{-1}, Γ_β has units m^{2}s$^{-\beta}$ and D has units m^{2}s^{-1}

The parameter C_0 equals $G(0) = (V_{\text{eff}} \langle C \rangle)^{-1} = (\langle N \rangle)^{-1}$, where $\langle N \rangle$ is the average number of particles inside the measurement volume during the measurement, which from our data is approximately 4–5. For a cylinder-shaped cell with base diameter 3 µm and length 10 µm, this gives an estimated total number of molecules inside the cell of about

1200, which is in the same order of magnitude as estimates for similar proteins in the closely related species *S. cerevisiae* [4].

In figure 10 we see that the autocorrelation functions for space-anomalous and time-anomalous diffusion fit the data rather well, except for small τ where there is much noise, and for large τ where the experimental curve does not approach zero quite as fast as the theoretical ones. The ordinary diffusion curve, however, does not fit the data at all well in the region that matters, namely where the autocorrelation curve decreases quickly due to diffusion of molecules. The misfit is even more clear in figure 11, where we see the differences between theoretical and experimental curves. It is also clear that the space-anomalous and time-anomalous curves are almost identical, and we can not say that one is better than the other. One may also note that the triplet time τ_{tr} is larger for the ordinary diffusion model. This is just another indication of the misfit, since as the diffusion part of the correlation function can not be fitted, the curve fitting algorithm tries to change τ_{tr} to achieve a better fit for the diffusion part using the triplet part of the function. In fact there was a limit of 40 µs set for τ_{tr}, since otherwise it would have been increased even more to achieve a better fit, using unrealistic parameter values.

For comparison, we also made a measurement on the same proteins, where we took a number of cells and destroyed their membranes so that the contents of the cells was mixed in a solution. The autocorrelation curve from this measurement on Mkp1 is shown in figure 12 and fits very well with an ordinary diffusion autocorrelation function. Here, we get parameters

$$C_0 = 0.42, \quad \tau_D = 215 \text{ µs}, \quad F_{\text{tr}} = 0.16, \quad \tau_{\text{tr}} = 7.8 \text{ µs},$$

giving a diffusion coefficient $D \approx 2.6 \cdot 10^{-11}$ m^2/s, which is nearer the expected diffusion coefficient for the protein in water, suggesting that in this case the protein is not hindered much by the obstacles in the solution, and there is no trace of anomalous diffusion.

5.3. Discussion

The results in the previous section indicate a number of properties of the diffusion inside our cells. First of all, it is clear that inside the cell the ordinary diffusion model does not fit the data. We need to have a model that changes the slope of the autocorrelation curve, and both the anomalous diffusion models do this and fit the data well, although not perfectly. There might be other models that could be made to fit the data, for example one could use a multi-component ordinary diffusion curve as in section 2.4.1, and this would surely yield a better fit, since we have more parameters to vary (in fact, this has been tried and the fit becomes almost as good as for the anomalous diffusion models), but such a model would imply that the fluorescent protein (Mkp1-GFP) is bound to other proteins so that it would diffuse both in free form and as part of a larger complex. But the autocorrelation curve in figure 12 when the cells have been destroyed show no traces of such complexes, and although this is no proof that they do not exist inside the cell, it speaks in favor of an anomalous diffusion model compared to a multi-component model. It is evident that the

protein does bind to other molecules now and then to perform the tasks it is there for (and this might even be the cause of the anomalous diffusion), but it seems reasonable to believe that these bindings and complex formations occur on a much shorter time scale, meaning that they are not stable enough to show up as a second component in the autocorrelation curve. Therefore, our results indicate (as do the results in [23], [21], and others) that anomalous diffusion models are relevant models of the motion of proteins inside cells.

It is also clear from figures 10 and 11 that we can not distinguish between the two anomalous diffusion models and say that one is better than the other, judging only from the fitting of autocorrelation curves. Other experiments or data analysis is necessary to distinguish between the two; and the two models are not quite equivalent, not even in terms of autocorrelation functions, so there is a possibility that one could use FCS experiments to decide which model to prefer based on experimental data. Furthermore, using only the autocorrelation curve computed from fluorescence fluctuations, we use only a small part of the information in the data, and other analysis methods might give complementary information that could be useful. In fact, as mentioned earlier, there are at least two attempts at such analysis methods, namely the high-order correlation analysis [20, 11, 13], where correlation functions are computed for powers of the fluorescence values, that is $\langle F(t)^k F(t+\tau)^m \rangle$ with m and k integers larger than 1, and the photon count histogram (PCH) analysis [10], where histograms over the number of photons counted in a given (small) time interval are analyzed. The effect of anomalous diffusion on these quantities has not yet been investigated.

It is also interesting to see that the experimental autocorrelation curve in figure 10 does not decrease to zero as fast as it should according to the anomalous diffusion models. This might be an effect due to the fact that we make measurements of diffusion inside a small domain, which should have similar effects on the autocorrelation curve as we saw in section 4.3.

In conclusion, the experimental data presented above justifies the study of mathematical models such as the space-anomalous diffusion equation (3.5). In the paper that follows, we will see how we can define and solve boundary value problems for that equation, corresponding to anomalous diffusion inside a cell.

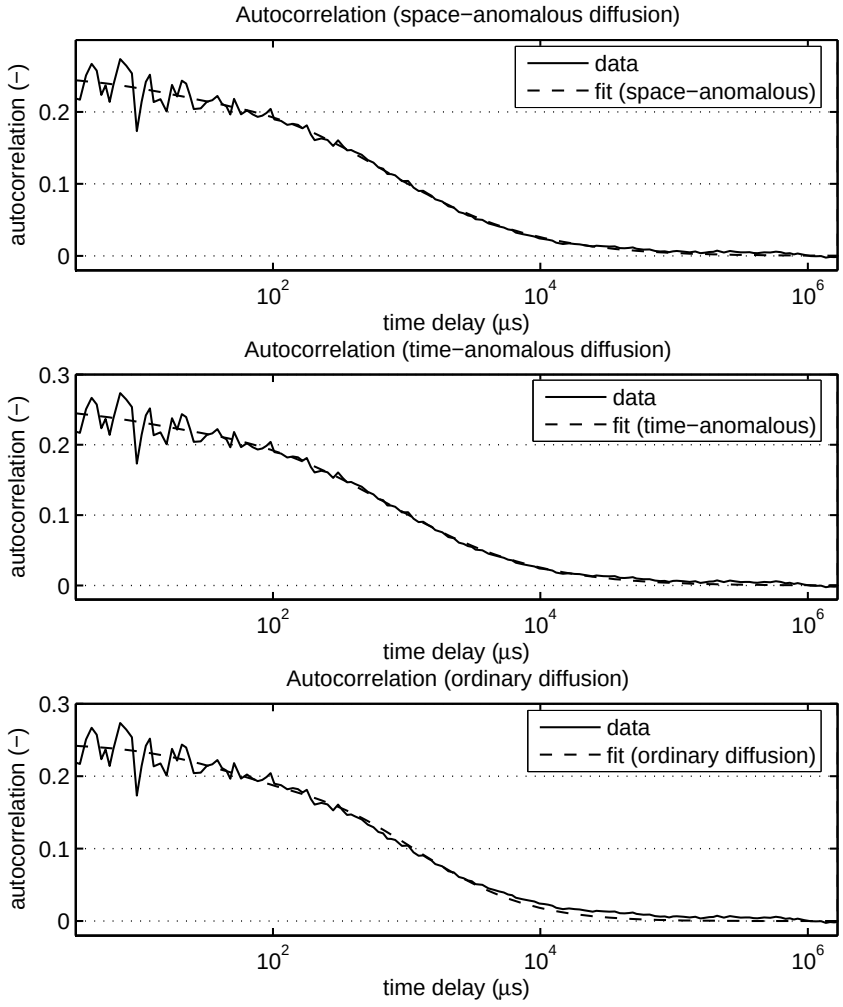

Figure 10: Experimental data with fitted theoretical autocorrelation curves for space-anomalous diffusion (top), time-anomalous diffusion (middle) and ordinary diffusion (bottom), all including a triplet state correction. The two anomalous diffusion models fit the data well in the central region where the slope that characterizes the diffusion is. The ordinary diffusion model, however, does not fit in this region, although the misfit might not appear very large at first glance.

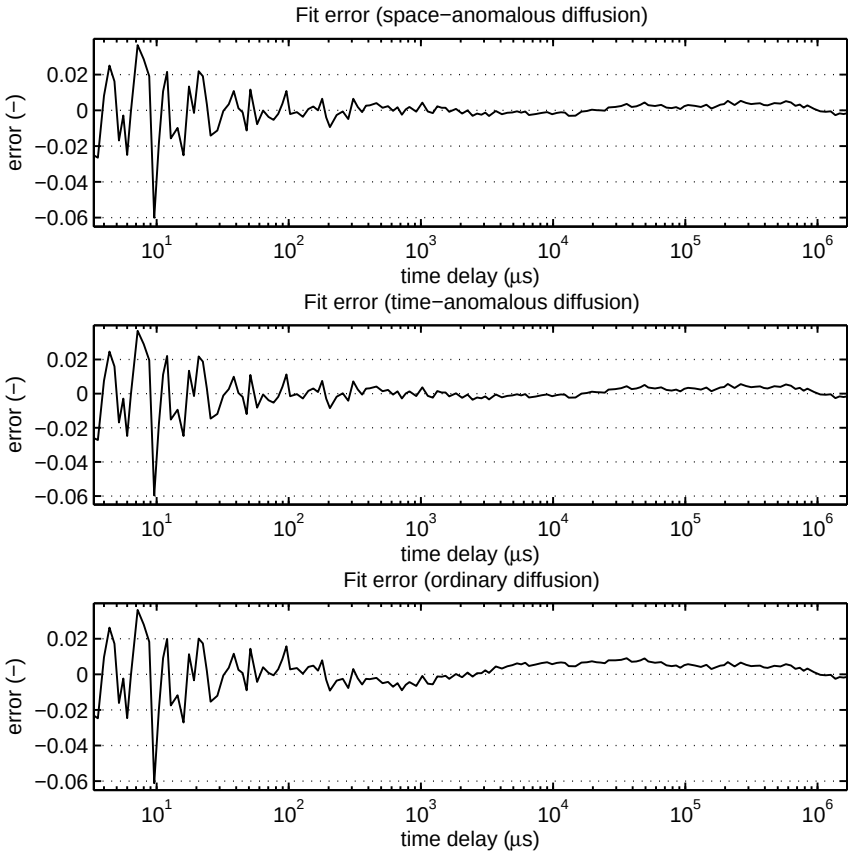

Figure 11: Errors plots for the curve fits in figure 10. Here it is seen more clearly that the ordinary diffusion model does not fit as well as the other two in the important region ($\sim 10^2 - 10^5$ μs).

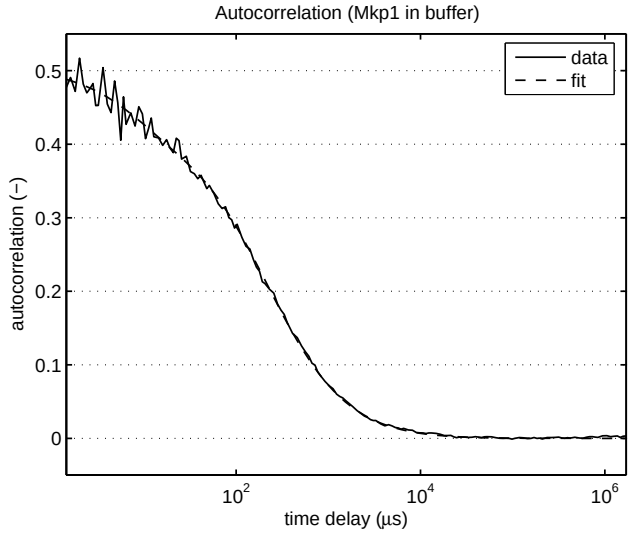

Figure 12: Autocorrelation function for Mkp1-GFP in buffer.

References

[1] D. ben-Avraham and S. Havlin. *Diffusion and Reactions in Fractals and Disordered Systems*. Cambridge University Press, 2000.

[2] E. Elson and D. Magde. Fluorescence correlation spectroscopy, I: Conceptual basis and theory. *Biopolymers*, 13(1):1–27, 1974.

[3] G. B. Folland. *Fourier Analysis and Its Applications*. Brooks/Cole Publishing Company, 1992.

[4] S. Ghaemmaghami, W.-K. Huh, K. Bower, R. W. Howson, A. Belle, N. Dephoure, E. K. O'Shea, and J. S. Weissman. Global analysis of protein expression in yeast. *Nature*, 425:737 – 741, Oct 2003.

[5] T. Jankowski and R. Janka. ConfoCor 2 – the second generation of fluorescence correlation microscopes. In Rigler and Elson [14], chapter 15, pages 331–345.

[6] D. Magde, E. Elson, and W. W. Webb. Fluorescence correlation spectroscopy. II: An experimental realization. *Biopolymers*, 13(1):29–61, 1974.

[7] F. Mainardi, Y. Luchko, and G. Pagnini. The fundamental solution of the space-time fractional diffusion equation. *Fract. Calc. Appl. Anal.*, 4(2):153–192, 2001.

[8] F. Mainardi, G. Pagnini, and R. K. Saxena. Fox h functions in fractional diffusion. *J. Comp. Appl. Math.*, 178:321–331, 2005.

[9] R. Metzler and J. Klafter. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.*, 339:1–77, 2000.

[10] J. D. Müller, Y. Chen, and E. Gratton. Photon counting histogram statistics. In Rigler and Elson [14], chapter 20, pages 410–437.

[11] A. G. Palmer III and N. L. Thompson. High-order fluorescence fluctuation analysis of model protein clusters. *Proc. Natl. Acad. Sci. USA*, 86(16):6148–52, 1989.

[12] I. Podlubny. *Fractional differential equations*, volume 198 of *Mathematics in Science and Engineering*. Academic Press Inc., San Diego, CA, 1999.

[13] H. Qian and E. L. Elson. Distribution of molecular aggregation by analysis of fluctuation moments. *Proc. Natl. Acac. Sci. USA*, 87:5479–83, 1990.

[14] R. Rigler and E. S. Elson, editors. *Fluorescence Correlation Spectroscopy : theory and applications*, volume 65 of *Springer series in chemical physics*. Springer Verlag, 2001.

[15] R. Rigler, Ülo Mets, J. Widengren, and P. Kask. Fluorescence correlation spectroscopy with high count rate and low background: analysis of translational diffusion. *Eur. Biophys. J.*, 22(3):169–175, 1993.

[16] M. J. Saxton. Anomalous diffusion due to obstacles: a Monte Carlo study. *Biophys J.*, 66(2:1):394–401, 1994.

[17] M. J. Saxton. Anomalous diffusion due to binding: a Monte Carlo study. *Biophys. J.*, 70:1250–62, 1996.

[18] P. Schwille and E. Haustein. Fluorescence correlation spectroscopy - an introduction to its concepts and applications. *Biophysics Textbook Online*, 2002.

[19] SweGene, CCI, `http://www.swegene.org/cellular_imaging`.

[20] N. L. Thompson and J. L. Mitchell. High order autocorrelation in fluorescence correlation spectroscopy. In Rigler and Elson [14], chapter 21, pages 438–458.

[21] M. Wachsmuth, W. Waldeck, and J. Langowski. Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy. *J. Mol. Biol.*, 298:677–689, 2000.

[22] M. Weiss, M. Elsner, F. Kartberg, and T. Nilsson. Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells. *Biophys. J.*, 87:3518–24, 2004.

[23] M. Weiss, H. Hashimoto, and T. Nilsson. Anomalous protein diffusion in living cells as seen by fluorescence correlation spectroscopy. *Biophys J.*, 84(6):4043–52, 2003.

[24] M. Weiss and T. Nilsson. In a mirror dimly: tracing the movements of molecules in living cells. *Trends Cell Biol.*, 14(5):267–273, 2004.

[25] Zeiss website, `http://www.zeiss.com/micro`.

BOUNDARY VALUE PROBLEMS FOR A SPACE-ANOMALOUS DIFFUSION EQUATION

TOBIAS GEBÄCK* AND ALEXEI HEINTZ*

Abstract. We propose a new equation describing subdiffusion inside a bounded domain in $\mathbb{R}^n$. The equation is a parabolic pseudo-differential equation based on the conservation of mass. We pose boundary value problems for the equation, both Dirichlet problems and problems specifying the flux in the normal direction at the boundary, similar to Neumann problems for the ordinary diffusion equation. We then prove existence and uniqueness results for these problems using Hilbert space methods and semigroups, and show some numerical results.

1. Introduction. In recent years, experiments have shown that the ordinary diffusion model does not suffice to describe the motion of particles and molecules in crowded environments. Such an environment is for example inside a living cell, where the medium in which the molecules move is so full of other molecules that interact with the moving molecule that it makes the diffusion behaviour of large molecules (e.g. proteins) qualitatively different from diffusion in water or solvent. This phenomenon is called anomalous diffusion or subdiffusion, indicating that it is slower than ordinary diffusion. It has been observed using Fluorescence Correlation Spectroscopy (FCS) inside cells from different organisms, see e.g. [18, 19]. This technique measures fluctuations in the fluorescence signal from proteins with a fluorescent tag, and from the autocorrelation of these fluctuations, information about the diffusion of proteins may be obtained (see [13] and [11] for an introduction to FCS).

In order to analyze the data acquired using FCS, it is necessary to have a model giving a theoretical autocorrelation curve that can be fitted to the experimental curve. There are at least two models that may be fitted to subdiffusion data. The first, and more widely used, is the time-anomalous diffusion equation,

$$\frac{\partial u}{\partial t}(x,t) = {}_0D_t^{1-\beta} K_\beta \Delta u(x,t) \tag{1.1}$$

with $\beta < 1$, K_β a constant, and the Riemann-Liouville operator ${}_0D_t^{1-\beta}$ defined by

$${}_0D_t^{1-\beta} f(t) = \frac{\partial}{\partial t} {}_0D_t^{-\beta} f(t) = \frac{1}{\Gamma(\beta)} \frac{\partial}{\partial t} \int_0^t \frac{f(t')}{(t-t')^{1-\beta}}\, dt'. \tag{1.2}$$

This diffusion equation can be derived using Continuous Time Random Walk (CTRW, see [9] and [6]), and contains the non-local operator ${}_0D_t^{1-\beta}$, which models memory properties of the physical system studied. A lot of work has been done studying this equation and operators like (1.2), including work by Mainairdi, Gorenflo, Podlubny, Lubich and many others, see e.g. [6, 5, 10, 8]. However, these studies do not include boundary value problems of the type we are interested in. Also, we wish to describe subdiffusion with a model without memory, since the operator (1.2) gives a special significance to the time $t = 0$, and we are faced with the problem of choosing a time $t = 0$ before which there is no memory, which seems like an unnatural question for the situation we wish to model.

*Mathematical Sciences, Chalmers University of Technology and Göteborg University, SE-412 96 Göteborg, Sweden, (tobiasg@math.chalmers.se, heintz@math.chalmers.se)

Therefore, in this paper we discuss the space-anomalous diffusion equation

$$\frac{\partial u}{\partial t} + K_\alpha(-\Delta)^\alpha u = 0, \tag{1.3}$$

where K_α is a constant and α is a positive real number, implying a non-local pseudo-differential operator in the space variable, with symbol $|\xi|^{2\alpha}$ in the Fourier domain. When $\alpha = 1$ we get the ordinary diffusion equation, when $\alpha < 1$ we get superdiffusion (faster than ordinary diffusion) corresponding to so called Lévy flights (see [9]), and when $\alpha > 1$ we get subdiffusion. Our own experiments with FCS in fission yeast (see [4]) show that the autocorrelation function derived from this model may be fitted to data just as well as the time-anomalous autocorrelation function related to (1.1), and it would require a different analysis of experimental data, and perhaps more accurate data, to decide between the two models. In our experiments, the observed values of α were between 1 and $\frac{3}{2}$. Equation (1.3), with $\alpha > 1$, can not be derived from a CTRW model and the motivation for (1.3) is instead through its fundamental solution and the fitting to experimental FCS data. It may be noted, though, that when $1 < \alpha < 2$ we are somewhere between ordinary diffusion ($\alpha = 1$) and elasticity ($\alpha = 2$), which may suggest an interpretation where molecules attach to each other to a certain degree, producing elastic effects. It should also be noted that one can imagine using other pseudo-differential operators than $(-\Delta)^\alpha$ in (1.3), which perhaps fit the experimental data even better and have other desirable properties. The approach used in this paper can be generalized to a larger family of pseudo-differential operators.

The purpose of this paper is to define boundary value problems for the space-anomalous diffusion equation (1.3), prove existence of solutions to those problems, and to show some numerical results for them. The problem we have in mind is to describe anomalous diffusion inside a cell, bounded by membranes, which are impermeable to the proteins inside, or where the passage of proteins through the membrane is highly regulated. For ordinary diffusion, this situation is described by Neumann-type problems

$$-\hat{n} \cdot \nabla u(x) = G(x, u), \quad x \in \partial\Omega, \tag{1.4}$$

in some domain Ω with boundary $\partial\Omega$ and outward unit normal $\hat{n}$. The left hand side of (1.4) is the flux of material perpendicular to the boundary $\partial\Omega$, and G is typically linear in u, or $G \equiv 0$ if the membrane is impermeable. We therefore wish to define boundary value problems for equation (1.3), with $\alpha > 1$, corresponding to this situation, and we do this in the next section. Existence and uniqueness of solutions will be proved for $1 < \alpha < 1 + n/2$, where n is the number of space dimensions.

There has been at least two previous attempts to formulate boundary value problems for pseudo-differential equations such as (1.3). These are the approaches by Vishik and Eskin [1, 17], and the one initiated by Seeley [14, 15]. The main reason for not using those approaches is that in our approach, we get a natural definition of the flux, and therefore natural boundary conditions corresponding to (1.4). Furthermore, the operator we define will give zero when applied to constants, which is also a desirable property, and which circumvents some of the problems in the Vishik-Eskin approach. It is possible that there is a formulation equivalent to ours within the Vishik-Eskin framework, where potentials are added to the domain boundary to

compensate for the singularities that may appear, but our approach is more straightforward and uses simpler Hilbert-space theory and semigroups to define solutions. In the Seeley-approach, powers are taken of operators with boundary conditions, and this theory does not allow for the fractional Neumann-type boundary conditions that we will derive in the next section, so that is not a satisfactory approach for our problem.

It could also be noted that there have been attempts to define boundary value problems to (1.3) with $\alpha < 1$ (the Lévy-diffusion case) using variants of the Riemann-Liouville operator (1.2) for the non-local operator in the space-variable (see [3, 12]). These approaches are similar to ours in the sense that they modify Fick's law in a similar way, at least in one space dimension, but only Dirichlet problems have been considered, and the extension to higher dimensions is very different.

In the next section, we start out by deriving our variant of equation (1.3) which allows for weak solutions. This new variant is equivalent to (1.3) if studied in the whole space, but for the problem inside a domain there are differences. After discussing some properties of the operator we derive, we prove the existence of unique solutions to the stationary problem, and then to the time-dependent problem with different boundary conditions, including the Neumann-type condition. The method used is standard Hilbert space methods and semi-groups, which are described thoroughly in Showalter [16], and we will refer to this book throughout this paper although the same results may be found in other places too. Finally, we discuss the fact that solutions to (1.3) need not be positive for positive initial conditions, and also show some results of numerical calculations, illustrating what solutions may look like, including differences compared to the ordinary diffusion equation.

2. Deriving the equation. To derive our equation, we begin with the conservation of mass, which may be expressed as

$$\frac{\partial u}{\partial t}(x,t) + \nabla \cdot \Phi(x,t) = F(x,t), \tag{2.1}$$

where $u(x,t)$ is the concentration of a substance at a point $x \in \mathbb{R}^n$ at time t, Φ is the flux of the substance, and F is the production of substance (which may be negative). This is a fundamental relationship, which is derived mathematically through Gauss' theorem, and which should hold for any model describing the movement of a substance.

For ordinary diffusion, the flux Φ is determined by Fick's law,

$$\Phi(x,t) = -D\nabla u(x,t), \tag{2.2}$$

with diffusion coefficient D, which is an experimentally motivated law, and can thus only be assumed to hold in the conditions under which it has been tested. As mentioned in the introduction, there are situations in which Fick's law, and therefore the ordinary diffusion equation, is not valid, but instead the space-anomalous diffusion equation (1.3) fits the experimental data. We will therefore modify Fick's law (2.2) to get an equation that resembles (1.3).

Based on this, we propose that the flux in (2.1) should be defined as

$$\Phi(x,t) \equiv -K_\alpha(-\Delta)^{\alpha-1}\nabla u(x,t), \tag{2.3}$$

with anomalous diffusion constant $K_\alpha > 0$ (with units $m^{2\alpha}s^{-1}$), which gives the following new anomalous diffusion equation:

$$u_t(x,t) - K_\alpha \nabla \cdot (-\Delta)^{\alpha-1} \nabla u(x,t) = 0. \tag{2.4}$$

The operator $(-\Delta)^\beta$, with non-integer β, is a non-local pseudo-differential operator, which is defined in $\mathbb{R}^n$ through the Fourier transform as

$$(-\Delta)^\beta u = \mathscr{F}^{-1}\left(|\xi|^{2\beta}\widehat{u}(\xi)\right),$$

for suitable u, and so equation (2.4) contains a number of subtleties which will be discussed below. The definition of the Fourier transform and its inverse used throughout this paper is

$$[\mathscr{F}(u)](\xi) = \int_{\mathbb{R}^n} e^{-i\xi\cdot x} u(x)\, dx$$
$$[\mathscr{F}^{-1}(\widehat{u})](x) = \int_{\mathbb{R}^n} e^{i\xi\cdot x} \widehat{u}(\xi)\, d\xi,$$

with the notation $\widehat{u} = \mathscr{F}(u)$. The variable ξ will be used throughout as the variable in Fourier space.

It should be noted that if $u \in L^2(\mathbb{R}^n)$ and is smooth enough for the operations to make sense, then $-\nabla \cdot (-\Delta)^{\alpha-1}\nabla u = (-\Delta)^\alpha u$, since

$$-\nabla \cdot (-\Delta)^{\alpha-1}\nabla u = \mathscr{F}^{-1}\Big(-i\xi \cdot (|\xi|^{2(\alpha-1)} i\xi)\widehat{u}\Big) = \mathscr{F}^{-1}\left(|\xi|^{2\alpha}\widehat{u}\right) = (-\Delta)^\alpha u.$$

But when we study the equation inside a domain, and not in the whole space, there will be a difference.

The formulation where the spatial derivative on u is decomposed into three parts, $\nabla\cdot$, $(-\Delta)^\alpha$ and ∇, has some implications. First of all, by first taking the gradient of u, before applying the pseudo-differential operator, we ensure that the whole operator applied to a constant gives zero, which is not the case otherwise. This idea is similar to the so called 'Caputo derivative' for fractional derivatives in one dimension, defined by switching the order of integration and differentiation in (1.2), see e.g. [6]. Second, from the conservation of mass, we have a natural definition of the flux (namely (2.3)), and thus we may formulate boundary value problems where the boundary conditions specify the flux in the normal direction at the boundary, similar to the Neumann problem for the ordinary diffusion equation. Thus, the equations we are mainly interested in take the form

$$\begin{cases} u_t(x,t) - K_\alpha \nabla \cdot (-\Delta)^{\alpha-1}\nabla u(x,t) = F(x,t), & (x,t) \in \Omega \times (0,T), \\ -\hat{n} \cdot K_\alpha (-\Delta)^{\alpha-1}\nabla u(x,t) = G(u,x), & (x,t) \in \partial\Omega, \times (0,T) \\ u(x,0) = u_0(x), & x \in \Omega. \end{cases} \tag{2.5}$$

Here, Ω is a bounded domain in $\mathbb{R}^n$ with smooth boundary, G is linear in u, and $\alpha > 1$. Further restrictions will be discussed below. For simplicity, we will set $K_\alpha = 1$ in the following.

As the equation stands here, there are a number of questions that may be asked, especially regarding how the operator we have introduced acts, and how to define it properly. In the next section, we will make this clear.

3. The operator A. Motivated by the discussion in the previous section, we introduce the operator A, which we write as

$$Au = -\nabla_\Omega \cdot (-\Delta)^\varepsilon \nabla_\Omega u \tag{3.1}$$

with $\varepsilon = \alpha - 1 > 0$, and with Ω an open, bounded domain in $\mathbb{R}^n$. We will assume $u \in L^2(\Omega)$ and discuss further restrictions on u below.

First of all, we have written ∇_Ω instead of ∇, since as the operator $(-\Delta)^\varepsilon$ is defined through the Fourier transform, the function it is operating on must be defined in the whole space, while the gradient should only operate inside Ω. Thus, ∇_Ω may be thought of as $\chi_\Omega \nabla$, where χ_Ω is the characteristic function of Ω, meaning that the function is first differentiated inside Ω, and then extended to be zero outside Ω, ignoring anything that may happen on the boundary. We can also make a proper definition in the sense of distributions as follows. We let $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_n)$, with $\varphi_j \in C^\infty(\mathbb{R}^n)$, $j = 1, \dots, n$, which we will write as $\boldsymbol{\varphi} \in C^\infty(\mathbb{R}^n)$. Then, for $v \in L^2(\Omega)$, ∇_Ω is defined by

$$\langle \nabla_\Omega v, \boldsymbol{\varphi} \rangle = -\langle v, \nabla \cdot \boldsymbol{\varphi} \rangle_{L^2(\Omega)}, \quad \text{for all } \boldsymbol{\varphi} \in C^\infty(\mathbb{R}^n) \tag{3.2}$$

which defines $\nabla_\Omega v$ as a distribution on $\mathbb{R}^n$ with compact support (in Ω), i.e. $\nabla_\Omega v \in \mathcal{E}'(\mathbb{R}^n)$. We will use the notation $\langle \cdot, \cdot \rangle$ to denote a distribution acting on a test function, while scalar product in $L^2(\Omega)$ and $L^2(\mathbb{R}^n)$ will be denoted by the same symbol, but with corresponding index, as in the right hand side of (3.2). Of course the distributional gradient ∇ is also defined by the relation (3.2), but for all $\boldsymbol{\varphi} \in C_0^\infty(\Omega)$, which defines $\nabla v \in \mathcal{D}'(\Omega)$, a distribution on Ω.

We remark that if $v \in C^1(\overline{\Omega})$ and $\boldsymbol{\varphi} \in C^\infty(\mathbb{R}^n)$, then

$$\langle \nabla_\Omega v, \boldsymbol{\varphi} \rangle = -\langle v, \nabla \cdot \boldsymbol{\varphi} \rangle_{L^2(\Omega)} = -\int_\Omega v \, \nabla \cdot \boldsymbol{\varphi} \, dx = -\int_{\partial\Omega} v \, \boldsymbol{\varphi} \cdot \hat{n} \, dS + \int_\Omega \nabla v \cdot \boldsymbol{\varphi} \, dx.$$

In this sense the difference between ∇_Ω and ∇ is a measure on the boundary of Ω. In particular, if $v \in C_0^1(\overline{\Omega})$, then $\nabla_\Omega v = \nabla v$ in $\mathcal{E}'(\mathbb{R}^n)$ (if ∇v is extended by zero outside Ω and interpreted as a distribution on $\mathbb{R}^n$).

Now, the next step is the definition of $(-\Delta)^\varepsilon$. As mentioned earlier, this operator is defined through the Fourier transform by

$$(-\Delta)^\varepsilon w = \mathscr{F}^{-1}\left(|\xi|^{2\varepsilon} \widehat{w}(\xi) \right),$$

whenever this makes sense, which it clearly does for example if $\varepsilon > -n/2$ and w belongs to the Schwartz class $\mathcal{S}(\mathbb{R}^n)$ of rapidly decreasing infinitely differentiable functions, since then $\widehat{w} \in \mathcal{S}(\mathbb{R}^n)$ and $|\xi|^{2\varepsilon} \widehat{w}(\xi) \in L^1(\mathbb{R}^n)$. But it may also be interpreted in the sense of distributions, i.e. if $|\xi|^{2\varepsilon} \widehat{w}(\xi) \in \mathcal{S}'(\mathbb{R}^n)$, since the Fourier transform is defined for the tempered distributions $\mathcal{S}'(\mathbb{R}^n)$. In the definition of the operator A, we apply $(-\Delta)^\varepsilon$ to each of the components of $\nabla_\Omega u$.

Finally, we define the divergence as usual, in the sense of distributions, noting that we are only interested in the result inside Ω, although $(-\Delta)^\varepsilon \nabla_\Omega u$ is defined in the whole of $\mathbb{R}^n$, so the operator of restriction to Ω is implicit in the divergence, which is indicated by using the symbol ∇_Ω.

By combining these three operators, we get the following definition of the operator A in the sense of distributions.

DEFINITION 3.1. *Given $\varepsilon > -n/2$, the operator $A : L^2(\Omega) \to \mathcal{D}'(\Omega)$ is defined for $u \in L^2(\Omega)$ by*

$$\langle Au, \varphi \rangle = -\langle u, \nabla \cdot (-\Delta)^{\varepsilon} \nabla \varphi \rangle_{L^2(\Omega)}, \qquad \text{for all } \varphi \in C_0^\infty(\Omega) \tag{3.3}$$

One of the reasons for factorizing the operator A into three parts, is that it is now possible to integrate by parts, and if u and v are smooth enough, say $u, v \in C^\infty(\overline{\Omega})$ with $Au \in L^2(\Omega)$, it holds that

$$\begin{aligned} \langle Au, v \rangle_{L^2(\Omega)} &= \int_\Omega (-\nabla_\Omega \cdot (-\Delta)^{\varepsilon} \nabla_\Omega u)\, v \, dx \\ &= -\int_{\partial\Omega} \hat{n} \cdot ((-\Delta)^{\varepsilon} \nabla_\Omega u)\, v \, dS + \int_\Omega (-\Delta)^{\varepsilon} \nabla_\Omega u \cdot \nabla_\Omega v \, dx. \end{aligned} \tag{3.4}$$

But using Parseval's formula, this last integral over Ω may be written as

$$\begin{aligned} \int_\Omega (-\Delta)^{\varepsilon} \nabla_\Omega u \cdot \nabla_\Omega v \, dx &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} |\xi|^{2\varepsilon} \widehat{\nabla_\Omega u} \cdot \widehat{\nabla_\Omega v} \, d\xi \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} |\xi|^{\varepsilon} \widehat{\nabla_\Omega u} \cdot |\xi|^{\varepsilon} \widehat{\nabla_\Omega v} \, d\xi \\ &= \int_{\mathbb{R}^n} (-\Delta)^{\varepsilon/2} \nabla_\Omega u \cdot (-\Delta)^{\varepsilon/2} \nabla_\Omega v \, dx. \end{aligned} \tag{3.5}$$

And, furthermore, the boundary integral in (3.4) contains the flux across the boundary as defined in (2.3), so that if u either satisfies the zero flux Neumann-type boundary condition

$$-\hat{n} \cdot ((-\Delta)^{\varepsilon} \nabla_\Omega u(x,t)) = \hat{n} \cdot \Phi(x,t) = 0$$

or the Dirichlet condition $u = 0$ on the boundary, then

$$\langle Au, u \rangle = \int_{\mathbb{R}^n} |(-\Delta)^{\varepsilon/2} \nabla_\Omega u|^2 \, dx.$$

4. The space $\mathcal{H}^\alpha(\Omega)$. Led by the above discussion, we define the space $\mathcal{H}^\alpha(\Omega)$, $\alpha > 1$, associated with the operator A, as the space of all $u \in L^2(\Omega)$ for which the norm

$$\|u\|_{\mathcal{H}^\alpha(\Omega)} \equiv \left(\|u\|^2_{L^2(\Omega)} + \|(-\Delta)^{\varepsilon/2} \nabla_\Omega u\|^2_{L^2(\mathbb{R}^n)} \right)^{1/2},$$

with $\varepsilon = \alpha - 1 > 0$, is finite. Here, we have $\mathbf{v} = (-\Delta)^{\varepsilon/2} \nabla_\Omega u$ in $\mathcal{D}'(\mathbb{R}^n)$ for $u \in L^2(\Omega)$ if

$$\langle \mathbf{v}, \boldsymbol{\varphi} \rangle = -\langle u, \nabla \cdot (-\Delta)^{\varepsilon/2} \boldsymbol{\varphi} \rangle_{L^2(\Omega)} = -\int_\Omega u \, \nabla \cdot (-\Delta)^{\varepsilon/2} \boldsymbol{\varphi} \, dx \tag{4.1}$$

for all $\boldsymbol{\varphi} \in C_0^\infty(\mathbb{R}^n)$.

The space $\mathcal{H}_0^\alpha(\Omega)$ is the closure of $C_0^\infty(\Omega)$ in this norm. We first prove that $\mathcal{H}^\alpha(\Omega)$ and $\mathcal{H}_0^\alpha(\Omega)$ are Hilbert spaces.

PROPOSITION 4.1. *$\mathcal{H}^\alpha(\Omega)$ and $\mathcal{H}_0^\alpha(\Omega)$ are Hilbert spaces.*

Proof. It is obvious that $\|\cdot\|_{\mathcal{H}^\alpha(\Omega)}$ is a norm with a corresponding scalar product. It therefore only remains to show that $\mathcal{H}^\alpha(\Omega)$ is complete. We therefore choose a Cauchy sequence $\{u_m\}$ in $\mathcal{H}^\alpha(\Omega)$, and since $L^2(\Omega)$ and $L^2(\mathbb{R}^n)$ are complete, we know that there exist $u \in L^2(\Omega)$ and $\mathbf{v} \in L^2(\mathbb{R}^n)$ such that

$$u_m \to u \text{ in } L^2(\Omega), \text{ and } (-\Delta)^{\varepsilon/2}\nabla_\Omega u_m \to \mathbf{v} \text{ in } L^2(\mathbb{R}^n).$$

We need only show that $\mathbf{v} = (-\Delta)^{\varepsilon/2}\nabla_\Omega u$ in $L^2(\mathbb{R}^n)$. We therefore choose $\boldsymbol{\varphi} \in C_0^\infty(\mathbb{R}^n)$ and compute

$$\begin{aligned}\langle \mathbf{v}, \boldsymbol{\varphi}\rangle &= \langle \mathbf{v}, \boldsymbol{\varphi}\rangle_{L^2(\mathbb{R}^n)} = \lim_{m\to\infty} \langle (-\Delta)^{\varepsilon/2}\nabla_\Omega u_m, \boldsymbol{\varphi}\rangle_{L^2(\mathbb{R}^n)} \\ &= \lim_{m\to\infty} -\langle u_m, \nabla\cdot(-\Delta)^{\varepsilon/2}\boldsymbol{\varphi}\rangle_{L^2(\Omega)} = -\langle u, \nabla\cdot(-\Delta)^{\varepsilon/2}\boldsymbol{\varphi}\rangle_{L^2(\Omega)} \\ &= \langle (-\Delta)^{\varepsilon/2}\nabla_\Omega u, \boldsymbol{\varphi}\rangle\end{aligned}$$

by (4.1), where we note that $\nabla\cdot(-\Delta)^{\varepsilon/2}\boldsymbol{\varphi}\big|_\Omega \in L^2(\Omega)$, so the limits are correctly taken in L^2, and thus $\mathbf{v} = (-\Delta)^{\varepsilon/2}\nabla_\Omega u$ in $\mathcal{D}'(\mathbb{R}^n)$, and therefore in $L^2(\mathbb{R}^n)$.

Finally, $\mathcal{H}_0^\alpha(\Omega)$ is also a Hilbert space since it is a closed subspace of $\mathcal{H}^\alpha(\Omega)$. ∎

We now prove a few useful results about the nature of $\mathcal{H}^\alpha(\Omega)$ and $\mathcal{H}_0^\alpha(\Omega)$.

PROPOSITION 4.2. *If $u \in \mathcal{H}_0^\alpha(\Omega)$ and $\varepsilon = \alpha - 1 > 0$, then*

$$\|(-\Delta)^{\varepsilon/2}\nabla_\Omega u\|_{L^2(\mathbb{R}^n)} = \|(-\Delta)^{\alpha/2}u\|_{L^2(\mathbb{R}^n)}.$$

Proof. Since $u \in \mathcal{H}_0^\alpha(\Omega)$, we have $\nabla_\Omega u = \nabla u$ as distributions on $\mathbb{R}^n$, and therefore

$$\begin{aligned}\|(-\Delta)^{\varepsilon/2}\nabla_\Omega u\|^2_{L^2(\mathbb{R}^n)} &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} ||\xi|^\varepsilon i\xi\widehat{u}(\xi)|^2\,d\xi = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} |\xi|^{2+2\varepsilon}|\widehat{u}(\xi)|^2\,d\xi \\ &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} |\left(|\xi|^2\right)^{\alpha/2}\widehat{u}(\xi)|^2\,d\xi = \|(-\Delta)^{\alpha/2}u\|^2_{L^2(\mathbb{R}^n)}\end{aligned}$$

using Parseval's formula. ∎

PROPOSITION 4.3. *If $u \in L^2(\Omega)$, extended with zero outside Ω, and $(-\Delta)^\beta u \in L^2(\mathbb{R}^n)$, where Ω is a bounded domain and $\beta > 0$, then*

$$\|u\|_{L^2(\Omega)} \le C\,\|(-\Delta)^\beta u\|_{L^2(\mathbb{R}^n)}$$

with $C = \frac{n+4\beta}{n(2\pi)^{4\beta}}\left(V_n|\Omega|\left(1+\frac{n}{4\beta}\right)\right)^{4\beta/n}$, where V_n is the volume of the n-dimensional unit ball.

Proof. We first note that $u \in L^1(\Omega)$ with $\|u\|_{L^1(\Omega)} \le |\Omega|^{1/2}\|u\|_{L^2(\Omega)}$ by the Hölder inequality, and furthermore that this implies that $\widehat{u}$ is continuous with $\sup|\widehat{u}| \le \|u\|_{L^1(\Omega)}$, if we extend u by zero outside Ω.

We now apply Parseval's formula, getting

$$\begin{aligned}\|u\|^2_{L^2(\Omega)} &= \frac{1}{(2\pi)^n}\|\widehat{u}\|^2_{L^2(\mathbb{R}^n)} = \frac{1}{(2\pi)^n}\int_{|\xi|<r}|\widehat{u}(\xi)|^2\,d\xi + \frac{1}{(2\pi)^n}\int_{|\xi|\geq r}|\widehat{u}(\xi)|^2\,d\xi \\ &\leq \frac{V_n r^n}{(2\pi)^n}(\sup|\widehat{u}|)^2 + \frac{1}{(2\pi)^n}\int_{|\xi|\geq r}\frac{|\xi|^{4\beta}}{r^{4\beta}}|\widehat{u}(\xi)|^2\,d\xi \\ &\leq \frac{V_n|\Omega|r^n}{(2\pi)^n}\|u\|^2_{L^2(\Omega)} + \frac{1}{r^{4\beta}}\|(-\Delta)^\beta u\|^2_{L^2(\mathbb{R}^n)}\end{aligned}$$

for any $r > 0$ and $\beta > 0$. This implies

$$\|u\|^2_{L^2(\Omega)} \leq \frac{r^{-4\beta}}{1 - V_n|\Omega|\left(\frac{r}{2\pi}\right)^n}\|(-\Delta)^\beta u\|^2_{L^2(\mathbb{R}^n)},$$

and taking infimum over those r that give a positive constant, we get C as above. ∎

PROPOSITION 4.4. *If Ω is bounded and $\alpha_1 \geq \alpha_2 \geq 1$, then*

$$\mathcal{H}^{\alpha_1}(\Omega) \subset \mathcal{H}^{\alpha_2}(\Omega) \subset H^1(\Omega)$$

with corresponding norm-inequalities. Here, $H^1(\Omega)$ is the standard Sobolev space of order 1.

Proof. The right inclusion follows from proposition 4.3 applied to $\nabla_\Omega u$, since if $u \in H^1(\Omega)$, then $\nabla_\Omega u \in L^2(\Omega)$.

The left inclusion follows from taking $u \in \mathcal{H}^{\alpha_1}(\Omega)$, letting $\varepsilon_2 = \alpha_2 - 1$ and $\varepsilon_1 = \alpha_1 - 1$, so that $\varepsilon_2 \leq \varepsilon_1$, and computing

$$\begin{aligned}\|(-\Delta)^{\varepsilon_2/2}\nabla_\Omega u\|^2_{L^2(\mathbb{R}^n)} &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}|\xi|^{2\varepsilon_2}|\widehat{\nabla_\Omega u}|^2\,d\xi \\ &= \frac{1}{(2\pi)^n}\int_{|\xi|<1}|\xi|^{2\varepsilon_2}|\widehat{\nabla_\Omega u}|^2\,d\xi + \frac{1}{(2\pi)^n}\int_{|\xi|\geq 1}|\xi|^{2\varepsilon_2}|\widehat{\nabla_\Omega u}|^2\,d\xi \\ &\leq \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}|\widehat{\nabla_\Omega u}|^2\,d\xi + \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}|\xi|^{2\varepsilon_1}|\widehat{\nabla_\Omega u}|^2\,d\xi \\ &= \|\nabla_\Omega u\|^2_{L^2(\Omega)} + \|(-\Delta)^{\varepsilon_1/2}\nabla_\Omega u\|^2_{L^2(\mathbb{R}^n)} \\ &\leq C\|(-\Delta)^{\varepsilon_1/2}\nabla_\Omega u\|^2_{L^2(\mathbb{R}^n)}\end{aligned}$$

by proposition 4.3. ∎

One may well ask the question what $u \in \mathcal{H}^\alpha(\Omega)$ may look like. It is obvious that if $\nabla_\Omega u \in C_0^\infty(\Omega)$, then $u \in \mathcal{H}^\alpha(\Omega)$, but are there other possibilities? In particular, since $\nabla_\Omega u$ typically has a jump across the boundary of Ω, one may ask what $(-\Delta)^{\varepsilon/2}\nabla_\Omega u$ looks like in that case. As a partial answer to this, we will see what happens to the characteristic function of a sphere in $\mathbb{R}^n$, i.e. investigate $(-\Delta)^{\varepsilon/2}\chi_{|x|\leq R}(x)$, $x \in \mathbb{R}^n$. Now, the Fourier transform of $\chi_{|x|\leq R}$ has a decay of $O(|\xi|^{-K})$ as $|\xi| \to \infty$, with $K = \min\left(\frac{n+1}{2}, 2\right)$ (which may be checked by computing the Fourier transform explicitly),

so $\mathscr{F}\big((-\Delta)^{\varepsilon/2}\chi_{|x|\le R}(x)\big) = O(|\xi|^{\varepsilon-K})$. If we consider only asymptotic decay, then for this function to belong to $L^2(\mathbb{R}^n)$ in Fourier space, we must have

$$2(\varepsilon - K) + n - 1 < -1,$$

which amounts to $\varepsilon < \min(\frac{1}{2}, 2 - \frac{n}{2})$. Thus, at least for $n \le 3$ and $\alpha < 3/2$, functions in $\mathcal{H}^\alpha(\Omega)$ need not have zero derivative at the boundary, which allows for more interesting solutions to the equations.

5. Stationary problems and Poincaré inequalities. Our aim in this section is to prove the existence of unique weak solutions in $\mathcal{H}^\alpha(\Omega)$ to the problem

$$\begin{cases} -\nabla_\Omega \cdot (-\Delta)^\varepsilon \nabla_\Omega u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{5.1}$$

where $f \in L^2(\Omega)$, and Ω is a bounded, open domain in $\mathbb{R}^n$ with C^1-boundary. Proving this will give us a number of results which are useful for proving the existence of solutions to the time-dependent equation (2.5), which we treat in the next section.

The method we use is the standard Hilbert space approach (using in particular the Lax-Milgram theorem) to define and prove existence of weak solutions. We will refer to the book by Showalter [16] for the standard theory. Weak solutions are defined as follows. Formally, we multiply the equation by a function $v \in \mathcal{H}^\alpha(\Omega)$ and integrate over Ω, then integrate by parts as in (3.4) and (3.5) and get

$$\langle Au, v\rangle = a(u,v) - \int_{\partial\Omega} \hat{n} \cdot ((-\Delta)^\varepsilon \nabla_\Omega u)\, v\, dS, \tag{5.2}$$

where the bilinear form $a(u,v)$ is defined by

$$a(u,v) = \int_\Omega (-\Delta)^{\varepsilon/2}\nabla_\Omega u \cdot (-\Delta)^{\varepsilon/2}\nabla_\Omega v\, dx \qquad u, v \in \mathcal{H}^\alpha(\Omega). \tag{5.3}$$

We see that if $v \in \mathcal{H}_0^\alpha(\Omega)$, then the boundary integral in (5.2) vanishes, and so we define $u \in \mathcal{H}_0^\alpha(\Omega)$ to be a weak solution to (5.1) if

$$a(u,v) = \langle f, v\rangle_{L^2(\Omega)}, \qquad \text{for all } v \in \mathcal{H}_0^\alpha(\Omega).$$

Now the Lax-Milgram theorem guarantees the existence of a unique solution u, provided that we prove that the bilinear form a is coercive on $\mathcal{H}_0^\alpha(\Omega)$, i.e. that

$$a(u,u) \ge C\|u\|^2_{\mathcal{H}^\alpha(\Omega)},$$

for all $u \in \mathcal{H}_0^\alpha(\Omega)$. But this result follows from the Poincaré inequality below, which in turn follows immediately from the results in the previous section.

PROPOSITION 5.1 (Poincaré inequality on $\mathcal{H}_0^\alpha(\Omega)$). *If $u \in \mathcal{H}_0^\alpha(\Omega)$, with $\alpha > 1$, then*

$$\|u\|_{L^2(\Omega)} \le C\|(-\Delta)^{\alpha/2}u\|_{L^2(\mathbb{R}^n)}.$$

Proof. By propositions 4.2 and 4.3 we have

$$\|(-\Delta)^{\varepsilon/2}\nabla_\Omega u\|_{L^2(\mathbb{R}^n)} = \|(-\Delta)^{\alpha/2}u\|_{L^2(\mathbb{R}^n)} \geq C\|u\|_{L^2(\Omega)}$$

■

We summarize the preceding discussion in the following corollary.

COROLLARY 5.2. *If $\alpha > 1$, the bilinear form a defined by (5.3) is $\mathcal{H}_0^\alpha(\Omega)$-coercive, that is*

$$|a(u,u)| \geq C\|u\|^2_{\mathcal{H}^\alpha(\Omega)}$$

for all $u \in \mathcal{H}_0^\alpha(\Omega)$, and therefore the Dirichlet problem (5.1) has a unique weak solution $u \in \mathcal{H}_0^\alpha(\Omega)$ for any $f \in L^2(\Omega)$.

For time-dependent problems we will also be interested in the Neumann-type problem (2.5), and will therefore need a similar result on $\mathcal{H}^\alpha(\Omega)$, when u need not be zero on the boundary. In that case, though, we can not expect the Poincaré inequality to be on the form above, since $u \in \mathcal{H}^\alpha(\Omega)$ can assume any value on the boundary, and therefore we can make the L^2-norm as big as we like by adding constants. Therefore, we introduce the notation

$$(u)_\Omega = \frac{1}{|\Omega|}\int_\Omega u(x)\,dx$$

for the average of $u \in L^1(\Omega)$ over Ω. We then set out to prove a Poincaré inequality for $u - (u)_\Omega$. The first step is the following theorem on compact imbedding in $L^2(\Omega)$.

PROPOSITION 5.3. *The space $\mathcal{H}^\alpha(\Omega)$, $\alpha \geq 1$, is compactly imbedded in $L^2(\Omega)$ if the boundary $\partial\Omega$ is C^1.*

Proof. Obviously, $\mathcal{H}^\alpha(\Omega) \subset L^2(\Omega)$, so we only need to show that if $\{u_m\}_{m=1}^\infty$ is a bounded sequence in $\mathcal{H}^\alpha(\Omega)$, then there is a subsequence $\{u_{m_k}\}_{k=1}^\infty$ which converges in $L^2(\Omega)$.

But $\|\cdot\|_{H^1(\Omega)} \leq C\|\cdot\|_{\mathcal{H}^\alpha(\Omega)}$ by proposition 4.4, and since $H^1(\Omega)$ is compactly embedded in $L^2(\Omega)$ (see e.g. Evans [2, thm 5.7.1]), then so is $\mathcal{H}^\alpha(\Omega)$. ■

We may now prove the Poincaré inequality on $\mathcal{H}^\alpha(\Omega)$. The proof is adopted from Evans [2, thm 5.8.1].

THEOREM 5.4 (Poincaré inequality on $\mathcal{H}^\alpha(\Omega)$). *Let Ω be a bounded and connected domain in $\mathbb{R}^n$ with C^1-boundary and let $u \in \mathcal{H}^\alpha(\Omega)$, $\alpha \in (1, 1+\frac{n}{2})$. Then there is a constant $C = C(n,\alpha,\Omega)$ such that*

$$\|u - (u)_\Omega\|_{L^2(\Omega)} \leq C\|(-\Delta)^{\varepsilon/2}\nabla_\Omega u\|_{L^2(\mathbb{R}^n)}$$

where $\varepsilon = \alpha - 1$, as before.

Proof. We assume that the estimate is false and try to derive a contradiction. In that case, for each integer $k = 1, 2, \ldots$ there is a function $u_k \in \mathcal{H}^\alpha(\Omega)$ such that

$$\|u_k - (u_k)_\Omega\|_{L^2(\Omega)} > k\|(-\Delta)^{\varepsilon/2}\nabla_\Omega u_k\|_{L^2(\mathbb{R}^n)}. \tag{5.4}$$

We set

$$v_k = \frac{u_k - (u_k)_\Omega}{\|u_k - (u_k)_\Omega\|_{L^2(\Omega)}}$$

for $k = 1, 2, \ldots$, so that $\|v_k\|_{L^2(\Omega)} = 1$ and $(v_k)_\Omega = 0$. Furthermore, (5.4) implies

$$\|(-\Delta)^{\varepsilon/2}\nabla_\Omega v_k\|_{L^2(\mathbb{R}^n)} < \frac{1}{k},$$

for $k = 1, 2, \ldots$, so that the sequence $\{v_k\}_{k=1}^\infty$ is bounded in $\mathcal{H}^\alpha(\Omega)$. Therefore, by proposition 5.3, there is a subsequence $\{v_{k_j}\}_{j=1}^\infty$ and a function $v \in L^2(\Omega)$ such that

$$v_{k_j} \to v \text{ in } L^2(\Omega) \text{ as } j \to \infty.$$

It follows that $\|v\|_{L^2(\Omega)} = 1$ and $(v)_\Omega = 0$. On the other hand, the distributional gradient $\nabla v = 0$ in $\mathcal{D}'(\Omega)$. To prove this, choose a test function $\boldsymbol{\varphi} \in C_0^\infty(\Omega)$ and set $\boldsymbol{\psi} = (-\Delta)^{-\varepsilon/2}\boldsymbol{\varphi}$, noting that $\boldsymbol{\psi} \in C^\infty(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ if $\varepsilon < n/2$. Then,

$$\begin{aligned}
\langle \nabla v, \boldsymbol{\varphi} \rangle &= -\langle v, \nabla \cdot \boldsymbol{\varphi} \rangle_{L^2(\Omega)} = -\lim_{j\to\infty} \langle v_{k_j}, \nabla \cdot \boldsymbol{\varphi} \rangle_{L^2(\Omega)} \\
&= -\lim_{j\to\infty} \langle v_{k_j}, \nabla \cdot (-\Delta)^{\varepsilon/2}\boldsymbol{\psi} \rangle_{L^2(\Omega)} = \lim_{j\to\infty} \langle (-\Delta)^{\varepsilon/2}\nabla_\Omega v_{k_j}, \boldsymbol{\psi} \rangle_{L^2(\mathbb{R}^n)} \\
&= \lim_{j\to\infty} \int_{\mathbb{R}^n} (-\Delta)^{\varepsilon/2}\nabla_\Omega v_{k_j}(x)\ \boldsymbol{\psi}(x)\, dx
\end{aligned}$$

and since

$$\begin{aligned}
\left| \int_{\mathbb{R}^n} (-\Delta)^{\varepsilon/2}\nabla_\Omega v_{k_j}(x) \cdot \boldsymbol{\psi}(x)\, dx \right| &\le \|(-\Delta)^{\varepsilon/2}\nabla_\Omega v_{k_j}\|_{L^2(\mathbb{R}^n)} \|\boldsymbol{\psi}\|_{L^2(\mathbb{R}^n)} \\
&< \frac{\|\boldsymbol{\psi}\|_{L^2(\mathbb{R}^n)}}{k_j} \to 0 \quad \text{as } j \to \infty,
\end{aligned}$$

we have $\nabla v = 0$. But since Ω is connected, then v must be constant, and since $(v)_\Omega = 0$ this constant must be zero. This contradicts $\|v\|_{L^2(\mathbb{R}^n)} = 1$ and proves the theorem. ∎

6. Time-dependent problems. We now set out to prove existence of unique solutions to time-dependent parabolic boundary value problems involving the operator A. We begin with the Dirichlet problem and then consider Neumann-type problems with boundary conditions involving the flux (2.3) across the boundary.

6.1. The Dirichlet problem. We consider the equation

$$\begin{cases} u_t(x,t) + Au(x,t) = 0, & (x,t) \in \Omega \times (0,\infty), \\ u(x,t) = 0, & (x,t) \in \partial\Omega \times (0,\infty), \\ u(x,0) = u_0(x), & x \in \Omega. \end{cases} \tag{6.1}$$

We let $a(u,v)$ be the bilinear form defined by (5.3), and study it on $\mathcal{H}_0^\alpha(\Omega)$, which corresponds to the homogeneous Dirichlet problem above. Using Riesz' theorem, we define the operator $\mathcal{A} \in L(D(\mathcal{A}), L^2(\Omega))$ by

$$a(u,v) = \langle \mathcal{A}u, v \rangle_{L^2}, \qquad u \in D(\mathcal{A}),\ v \in \mathcal{H}_0^\alpha(\Omega),$$

with

$$D(\mathcal{A}) = \{u \in \mathcal{H}_0^\alpha(\Omega) \mid |a(u,v)| \le K_u \|v\|_{L^2(\Omega)}, v \in \mathcal{H}_0^\alpha(\Omega)\}.$$

The operator $\mathcal{A}$ is of course an extension of the operator A, because of (3.4) and (3.5), where the boundary integral vanishes because v is zero at the boundary.

By corollary 5.2, we know that $a(u,v)$ is coercive on $\mathcal{H}_0^\alpha(\Omega)$ (and in fact it is $\mathcal{H}_0^\alpha(\Omega)$-elliptic, that is Re $a(u,u) \geq C\|u\|^2_{\mathcal{H}^\alpha(\Omega)}$ if $u \in \mathcal{H}_0^\alpha(\Omega)$). As is shown in [16, chapter IV], this implies that $D(\mathcal{A})$ is dense in $L^2(\Omega)$, and $-\mathcal{A}$ generates an analytic contraction semigroup $S(t)$, $t \geq 0$, which gives a unique solution $u(t) = S(t)u_0 \in C([0,\infty), L^2(\Omega)) \cap C^\infty((0,\infty), L^2(\Omega))$ to the abstract initial value problem

$$\begin{cases} u'(t) + \mathcal{A}u(t) = 0 \quad t > 0 \\ u(0) = u_0 \end{cases} \tag{6.2}$$

for each $u_0 \in L^2(\Omega)$. Furthermore, $u(t) \in D(\mathcal{A}^m)$ for $t > 0$ and every integer $m \geq 1$, meaning in particular that $u(t) \in \mathcal{H}_0^\alpha(\Omega)$ for each $t > 0$, so that the boundary conditions are fulfilled. In this sense we have therefore found a unique weak solution to the Dirichlet problem for the space-anomalous diffusion equation (6.1). We may also add a right hand side $f(t) : [0,\infty) \to L^2(\Omega)$, if it is Hölder continuous in t, but then we only get one continuous time-derivative of $u(t)$. In either case, there is a measurable representative $U \in L^2(\Omega \times [0,T])$ for any $T > 0$, with $U(\cdot,t) = u(t)$ and $\partial_t U(\cdot,t) = u'(t)$ in $L^2(\Omega)$ for $t \in [0,T]$.

The inhomogeneous Dirichlet problem $u = g$ on $\partial\Omega$ may also be solved if we find a $w \in \mathcal{H}^\alpha(\Omega)$ with $w = g$ on $\partial\Omega$ and $Aw \in L^2(\Omega)$, because then the equation may be transformed to a homogeneous Dirichlet problem with right-hand side $-Aw$ and initial condition $u_0 - w$. These results are summarized in the following theorem.

THEOREM 6.1. *Let Ω be a bounded domain in $\mathbb{R}^n$ with C^1-boundary and let $\varepsilon > 0$. Then the space-anomalous diffusion equation*

$$\begin{cases} u_t(t,x) - \nabla_\Omega \cdot (-\Delta)^\varepsilon \nabla_\Omega u(x,t) = f(x,t), & (x,t) \in \Omega \times (0,\infty) \\ u(x,t) = g(x), & (x,t) \in \partial\Omega \times (0,\infty) \\ u(0,x) = u_0(x) & x \in \Omega \end{cases}$$

has a unique solution $u \in L^2(\Omega \times [0,T])$ for any $T > 0$, if $f(\cdot,t) \in L^2(\Omega)$, $\forall t \geq 0$ and the mapping $t \mapsto f(\cdot,t)$ is Hölder continuous, $u_0 \in L^2(\Omega)$, and $g \in L^2(\partial\Omega)$ with $g = \gamma w$ for some $w \in \mathcal{H}^\alpha(\Omega)$ with $Aw \in L^2(\Omega)$, where γ is the trace operator.

6.2. The Neumann-type problem. Turning now to the Neumann-type problem

$$\begin{cases} u_t(x,t) + Au(x,t) = 0, & (x,t) \in \Omega \times (0,\infty), \\ -\hat{n} \cdot (-\Delta)^\varepsilon \nabla_\Omega u(x,t) = 0, & (x,t) \in \partial\Omega \times (0,\infty), \\ u(x,0) = u_0(x), & x \in \Omega, \end{cases} \tag{6.3}$$

we use the same approach as for the Dirichlet problem, but we need a bit more reasoning, since solutions will not be zero on the boundary and we need to use theorem 5.4 instead of proposition 5.1 to prove coercivity.

We therefore define the quotient spaces

$$W = L^2(\Omega)/\{u \in L^2(\Omega) \mid u = c\} \quad \text{and} \quad V = \mathcal{H}^\alpha(\Omega)/\{u \in \mathcal{H}^\alpha(\Omega) \mid u = c\}$$

with norms

$$\|u\|_W = \|u - (u)_\Omega\|_{L^2(\Omega)}, \qquad \|u\|_V = \|u - (u)_\Omega\|_{\mathcal{H}^\alpha(\Omega)},$$

and let V_0 be the closure of $C_0^\infty(\Omega)$ in V, or $V_0 = \mathcal{H}_0^\alpha(\Omega)/\{u = c\}$. The norms above are not the standard norms on the quotient spaces, but it is easy to verify that W, V, and V_0 are Hilbert spaces.

With the bilinear form $a(u,v)$ defined for $u, v \in V$ by (5.3), we follow the standard arguments in sections IV.7 and III.2 in Showalter [16], and define the bounded operator $A_1 : V \to V_0'$ by

$$[A_1 u](v) = a(u,v), \qquad u \in V,\ v \in V_0$$

and let $D_1 \equiv \{u \in V \mid A_1 u \in W\}$, implying by Riesz' theorem that if $u \in D_1$, then $[A_1 u](v) = \langle A_1 u, v\rangle_W$ (since $V_0 \subset W = W' \subset (V_0)'$).

Next, we investigate the difference between $a(u,v)$ and $[A_1 u](v)$ for $u \in D_1$. That is, fix $u \in D_1$ and study the functional

$$\phi(v) = a(u,v) - \langle A_1 u, v\rangle_W, \quad v \in V.$$

Comparing this expression to (5.2), we see that $\phi(v)$ should correspond to the boundary integral, and we will see that this is so. Note first that $\phi \in V'$ and that $\phi|_{V_0} = 0$, which is what characterizes the range of $q' : (V/V_0)' \to V'$, the dual of the quotient map $q : V \to V/V_0$. This implies that there is a unique $F \in (V/V_0)'$ such that $F \circ q = \phi$ in V'.

We then let $\gamma : V \to L^2(\partial\Omega)/\{u = c\}$ be the standard trace map modulo constants (which exists since $\mathcal{H}^\alpha(\Omega) \subset H^1(\Omega)$) with kernel V_0 and a range which we denote B. This induces the bijection $\widetilde{\gamma} : V/V_0 \to B$ with $\widetilde{\gamma} \circ q = \gamma$, and we can define a norm on B by $\|\cdot\|_B = \|\widetilde{\gamma}^{-1}(\cdot)\|_{V/V_0}$, which makes the dual operator $\widetilde{\gamma}' : B' \to (V/V_0)'$ a bijection. Therefore, given the functional $F \in (V/V_0)'$ above, there is a unique element $\partial \in B'$ such that $F = \widetilde{\gamma}'(\partial)$, or in other words $F = \partial \circ \widetilde{\gamma}$, which implies that $\phi = \partial \circ \widetilde{\gamma} \circ q = \partial \circ \gamma$. We conclude that there exists a linear mapping $\partial_{2\varepsilon+1} : D_1 \to B'$ such that for $u \in D_1$,

$$a(u,v) - \langle A_1 u, v\rangle_W = \partial_{2\varepsilon+1}(u)(\gamma(v)), \quad v \in V \tag{6.4}$$

which is an extension of the normal flux operator $-\hat{n} \cdot (-\Delta)^\varepsilon \nabla_\Omega|_{\partial\Omega}$.

Now, theorem 5.4 implies that $a(u,v)$ is coercive (and elliptic) on V if $\varepsilon \in (0, \frac{n}{2})$, and just as for the Dirichlet problem, if we set $D(\mathcal{A}) = \{u \in V \mid |a(u,v)| \le K_u \|v\|_W\}$, $a(u,v)$ defines an operator $\mathcal{A} : D(\mathcal{A}) \to W$ through

$$a(u,v) = \langle \mathcal{A}u, v\rangle_W, \qquad u \in D(\mathcal{A}),\ v \in V,$$

because of Riesz' theorem (since V is dense in W). The operator $-\mathcal{A}$ now generates an analytic semigroup on W, which solves (6.3) for each $u_0 \in W$, where the solution u has the same properties as for the Dirichlet problem above. In particular $u(t) \in D(\mathcal{A})$ for all $t > 0$ and it is easily seen that $D(\mathcal{A}) \subset D_1$, $\mathcal{A} = A_1|_{D(\mathcal{A})}$ since V_0 is dense in W, and thus from (6.4) that

$$\partial_{2\varepsilon+1}(u(t)) = 0 \quad \text{in } B',$$

for all $t > 0$, and in this sense u fulfills the Neumann-type boundary condition.

With this solution to the Neumann-type problem on the quotient space W in hand, which for each $t \geq 0$ may be interpreted as a function in $L^2(\Omega)$ with mean zero, it is obvious that we may add constants C to the initial value u_0 and get the corresponding solution $u + C$, since $A(u + C) = Au$. We may also add a right-hand side $g \in L^2(\partial\Omega)$ to the boundary condition and still get the same result. Furthermore, just as for the Dirichlet problem, we may add a right-hand side $f(t) : [0, \infty) \to L^2(\Omega)$ to the equation if f is Hölder continuous in t, and again we have a representative U for the solution u, with $U \in L^2(\Omega \times [0, T])$ for any $t \in [0, T]$.

It is also possible to have Robin type (or mixed) boundary conditions, that is

$$-\hat{n} \cdot ((-\Delta)^{\varepsilon} \nabla_\Omega u(x,t)) + b(x)u(x,t) = g(x), \quad x \in \partial\Omega, \ t > 0$$

with $b \in L^\infty(\partial\Omega)$ strictly positive ($b > b_0 > 0$ a.e.) and $g \in L^2(\partial\Omega)$. We then define the bilinear form as

$$a(u,v) = \left\langle (-\Delta)^{\varepsilon/2} \nabla_\Omega u, (-\Delta)^{\varepsilon/2} \nabla_\Omega v \right\rangle_\Omega + \left\langle b\gamma u, \gamma v \right\rangle_{\partial\Omega}.$$

For $\alpha > 1$, this is coercive, because then $u \in H^1(\Omega)$ and proposition III.5.C in [16] is applicable, which says that

$$\int_\Omega |u|^2 \, dx \leq 2 \operatorname{diam}(\Omega) \int_{\partial\Omega} |\gamma u|^2 \, dS + (2 \operatorname{diam}(\Omega))^2 \int_\Omega |\nabla u|^2 \, dx$$

so coercivity follows directly.

We summarize these results in the following theorem.

THEOREM 6.2. *Let Ω be a bounded, connected domain in $\mathbb{R}^n$ with C^1-boundary and $\varepsilon \in (0, \frac{n}{2})$. Then the space-anomalous diffusion equation*

$$\begin{cases} u_t(t,x) - \nabla_\Omega \cdot (-\Delta)^{\varepsilon} \nabla_\Omega u(x,t) = f(x,t), & (x,t) \in \Omega \times (0, \infty) \\ u(0,x) = u_0(x) & x \in \Omega \end{cases}$$

has a unique weak solution $u \in L^2(\Omega \times [0,T])$ for any $T > 0$, if $f(\cdot,t) \in L^2(\Omega)$, $\forall t \geq 0$ and the mapping $t \mapsto f(\cdot,t)$ is Hölder continuous, and $u_0 \in L^2(\Omega)$, for each of the boundary conditions

$$\begin{aligned} -\hat{n} \cdot (-\Delta)^{\varepsilon} \nabla_\Omega u(x) &= g(x) \\ -\hat{n} \cdot (-\Delta)^{\varepsilon} \nabla_\Omega u(x) + b(x)u(x) &= g(x), \end{aligned}$$

for $x \in \partial\Omega$, if $g \in L^2(\partial\Omega)$ and $b \in L^\infty(\Omega)$ with $b > b_0 > 0$ a.e..

7. Non-positivity of solutions. We will now briefly discuss the issue of positivity of solutions. In general, the solutions to the space-anomalous diffusion equation (1.3) in the whole of $\mathbb{R}^n$ are *not* positive if $\alpha > 1$, given a positive initial value u_0. This is seen as follows: Let $\phi(x,t)$ be the fundamental solution to (1.3), namely

$$\phi(x,t) = \mathscr{F}^{-1}\left(e^{-t|\xi|^{2\alpha}}\right),$$

which is found by taking the Fourier transform of (1.3) and solving the resulting ODE in t with initial value $\widehat{\phi}(\xi,0) = 1 = \mathscr{F}(\delta_0(x))$. Next consider the integral

$$\int_{\mathbb{R}^n} x_i^2 \phi(x,t)\,dx = \langle x_i^2, \phi(x,t)\rangle = \frac{1}{(2\pi)^n}\langle -(\delta_0)''_{ii}(\xi), \widehat{\phi}(\xi,t)\rangle = -\frac{1}{(2\pi)^n}\frac{\partial^2}{\partial {\xi_i}^2}\widehat{\phi}(\xi,t)\Big|_{\xi=0}$$

Now, the right hand side may be easily evaluated, and if $\alpha > 1$ we get zero, so that

$$\int_{\mathbb{R}^n} x_i^2 \phi(x,t)\,dx = 0,$$

which implies that $\phi(x,t)$, $t > 0$ is negative for some x, since it is not everywhere zero. Thus, the fundamental solution is non-positive, which implies that the solution to (1.3) is non-positive for many initial conditions.

It would of course be desirable to have positive solutions, since the solution u is supposed to describe a concentration, and so the space-anomalous diffusion equation has a weakness when it comes to being a model for subdiffusion. But, on the other hand, we study the equation inside a domain, where the initial condition may be sufficiently far away from zero to prevent the solution from ever becoming negative. Furthermore, the Neumann-type boundary condition implies that the total mass is constant, and the solution will tend to smooth out with time, decreasing the concentration differences between different locations. Thus, regions with low concentration will get higher concentration after a while, instead of decreasing below zero, even though it may decrease during a short time interval. This is seen in the numerical examples in the next section.

One may also argue that the model we use is not valid for small concentrations, since then the crowdedness that causes the anomalous behaviour is not present. And so perhaps a non-linear model with different behaviour when u is small would be more correct. On the other hand the anomalous behaviour might be caused by other factors than interactions between molecules of the same species and thus be present regardless of the concentration.

Finally, it should not come as a surprise that solutions may be non-positive, when one considers that for $1 < \alpha < 2$, we are somewhere between ordinary diffusion and elasticity, and that in an elastic material we may have propagating waves, so it is traces of this effect that show up in our equation.

8. Examples. In this section, we will show some numerical solutions to the space-anomalous diffusion equation inside a ball in $\mathbb{R}^3$, with Neumann-type boundary conditions, and compare the ordinary diffusion case $\alpha = 1$ to an anomalous diffusion case ($\alpha = 1.4$).

The computations were made using a finite difference approximation of the operator A on a uniform rectangular grid, keeping the scheme as simple as possible in order to get some numerical results quickly. The approximation matrix was divided into three parts, corresponding to the three parts of A, as

$$-\nabla_\Omega \cdot (-\Delta)^\varepsilon \nabla_\Omega u \approx \sum_{i=1}^{3} \widetilde{D}_i B_\varepsilon D_i U,$$

where U is a vector with function values at the grid points. D_i is a matrix with finite difference coefficients for approximating the derivative in the x_i-direction, using

central differences inside the domain, and one-sided differences near the boundary in order to avoid problems with the boundary. The matrix $\widetilde{D}_i$ also contains the central difference coefficients away from the boundary, but at the grid points next to the boundary, it instead contains the i:th component of the normal vector at that point, so that the equations at these grid points implement the boundary condition

$$-\hat{n} \cdot (-\Delta)^{\varepsilon} \nabla_{\Omega} u(x) = g(x).$$

Finally, B_{ε} is a finite difference approximation of the operator $(-\Delta)^{\varepsilon}$, derived in the same way as the so called Grünwald-Letnikov approximation [5]

$$\nabla_h^{\alpha} u(x) = h^{-\alpha} \sum_{j=0}^{\infty} (-1)^j \binom{\alpha}{j} u(x - jh) = D^{\alpha} u(x) + O(h), \tag{8.1}$$

with step size $h > 0$, where

$$\binom{\alpha}{j} = \frac{\Gamma(\alpha + 1)}{\Gamma(\alpha - j + 1)\, j!}$$

and D^{α} is a fractional derivative of Riemann-Liouville type, as in (1.2). The approximation (8.1) is based on the series expansion

$$(1 - z)^{\alpha} = \sum_{j=0}^{\infty} (-1)^j \binom{\alpha}{j} z^j, \quad z \in \mathbb{C}, \tag{8.2}$$

with z replaced by the shift operator E^{-h}, defined by $E^h u(x) = u(x + h)$. We note that $h^{-1}(I - E^{-h})u(x) = h^{-1}(u(x) - u(x - h))$ is a one-sided finite difference approximation of the derivative, which motivates that the Grünwald-Letnikov formula (8.1) approximates a fractional derivative of order α (note that in Fourier-space, the shift operator E^{-h} becomes the multiplicative factor $z = e^{-i\xi h}$). It may also be noted that if α is an integer, the sum becomes finite, and we get standard approximations of higher order derivatives.

If we instead consider the standard second order approximation of the Laplacian

$$\Delta_h u(x) = \frac{u(x + h) - 2u(x) + u(x - h)}{h^2} = h^{-2}\big(-2I + E^h + E^{-h}\big)u(x) = \Delta u(x) + O(h^2)$$

in one dimension, or in three dimensions

$$\Delta_h u(x_1, x_2, x_3) = h^{-2}\Big(-6I + \sum_{i=1}^{3} (E_{x_i}^{-h} + E_{x_i}^{h})\Big) u(x_1, x_2, x_3),$$

we can use the same series expansion (8.2) to get

$$\begin{aligned}
(-\Delta_h)^{\alpha} &= \frac{6^{\alpha}}{h^{2\alpha}} \sum_{j=0}^{\infty} (-6)^{-j} \binom{\alpha}{j} \big(E_{x_1}^{h} + E_{x_1}^{-h} + E_{x_2}^{h} + E_{x_2}^{-h} + E_{x_3}^{h} + E_{x_3}^{-h}\big)^j \\
&= \frac{6^{\alpha}}{h^{2\alpha}} \sum_{j=0}^{\infty} (-6)^{-j} \binom{\alpha}{j} \sum_{|\mathbf{k}|=j} \frac{j!}{\mathbf{k}!} E_{x_1}^{k_1 h - k_2 h} E_{x_2}^{k_3 h - k_4 h} E_{x_3}^{k_5 h - k_6 h} \\
&\equiv \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{m,n,p} E_{x_1}^{mh} E_{x_2}^{nh} E_{x_3}^{ph}.
\end{aligned}$$

Here, the coefficients $c_{m,n,p}$ are invariant under permutation and negation of the indices and

$$c_{m,n,p} = \frac{6^\alpha}{h^{2\alpha}} \sum_{\substack{j=m+n+p \\ j+=2}}^{\infty} (-6)^{-j} \frac{\Gamma(\alpha+1)}{\Gamma(\alpha-j+1)} \sum_{|\mathbf{k}|=j}^{*} \frac{1}{\mathbf{k}!} \tag{8.3}$$

where "$j+=2$" means that j is increased in steps of 2, and the starred sum is taken over $\mathbf{k} = (k_1, \ldots, k_6)$ with $\sum_{k=1}^{6} k_i = j$, $k_1 - k_2 = m$, $k_3 - k_4 = n$ and $k_5 - k_6 = p$, which may also be written as

$$\sum_{|\mathbf{k}|=j}^{*} \frac{1}{\mathbf{k}!} = \sum_{a=0}^{\frac{j-m-n-p}{2}} \sum_{b=0}^{\frac{j-m-n-p-2a}{2}} \frac{1}{d!(m+d)!a!(n+a)!b!(p+b)!}$$

with $d = \frac{1}{2}(j-m-n-p-2a-2b)$. It is this approximation that is used in the matrix B_ε, which becomes a full matrix containing the coefficients $c_{m,n,p}$. These coefficients should be computed with the infinite sum (8.3), but fortunately the terms decay rather quickly, so it is enough to take some 50–100 terms in the sum.

If we set $A_{h,\alpha} = \sum_{i=1}^{3} \widetilde{D}_i B_\varepsilon D_i$, and use a backward Euler method for the time-derivative, the resulting system of equations is

$$(\tilde{I}/\Delta t - A_{h,\alpha})U_k = \tilde{I}U_{k-1}/\Delta t, \quad k = 1, \ldots, K \tag{8.4}$$

where $\tilde{I}$ is a restriction to the interior points, meaning that the boundary condition is included in (8.4), and U_0 is a given vector with the initial condition. The linear system of equations is solved for U_k at each time-step using the iterative solver BiCGStab (see e.g. [7]), meaning that we only need to be able to multiply by $A_{h,\alpha}$, and this is done by direct matrix multiplication with the sparse matrices D_i and $\widetilde{D}_i$ and the full matrix B_ε. More efficient implementations using the Fast Fourier Transform for B_ε will be considered in the future.

We will now show results from computations for the space-anomalous diffusion equation with homogeneous Neumann-type boundary conditions, as in (6.3). In the example, Ω is a ball with radius 0.45 centered at $(x, y, z) = (0.5, 0.5, 0.5)$, and the initial condition is $u_0(x, y, z) = G(\rho)$, where $\rho = |(x - 0.7, y - 0.5, z - 0.5))|$ and $G(\rho) = 1.2$ for $\rho < 0.3$, $G(\rho) = 0.2$ for $\rho > 0.4$ and with a smooth transition between these two regions. The problem was solved for both $\alpha = 1.4$ (subdiffusion) and $\alpha = 1.0$ (ordinary diffusion) in order to see the difference. The parameter K_α was set to 1, and the diffusion coefficient D (in the $\alpha = 1$ case) was chosen so that the two solutions are comparable.

In figure 8.1, the solution after a short time ($t = 10^{-4}$) is shown for $z = 0.5$, for the two cases. The solutions are still very close to the initial condition, and give an idea what it looks like. But we can also see differences near the boundary for the two cases. The ordinary diffusion solution is flatter near the boundary, because of its Neumann boundary condition, while the subdiffusion solution with its non-local boundary condition shows a slightly different behaviour.

In figure 8.2, we see the solution for the two values of α and at two different time points. At the first time point, we see a difference between the two cases in that in

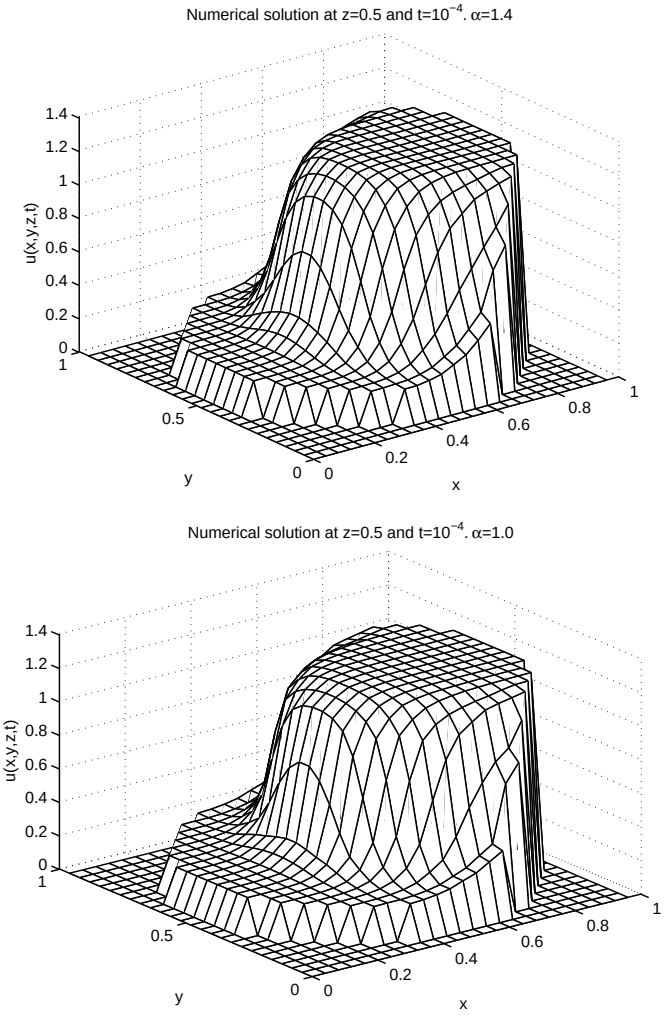

FIG. 8.1. *The numerical solution at $z = 0.5$ and $t = 10^{-4}$ for subdiffusion ($\alpha = 1.4$, top) and ordinary diffusion ($\alpha = 1.0$, bottom). Note the differences near the boundary.*

the subdiffusive case the solution remains high further from the center of the plateau than in the ordinary diffusion case, so that the solution is a bit more reluctant to smooth out. Still, at the steepest part, the subdiffusive solution is less steep than the ordinary diffusion solution. At the second time point, however, the two solutions are almost identical, and they both tend to become flatter as time passes.

The differences are rather small, but we must remember that we are only looking at the solution along a line through the center of a three-dimensional volume, so the differences are larger than they seem. Also, even very small differences could have a large effect if these diffusion equations include a reaction term in the right hand side, which is the case if we study a biological system such as in [4, Part I]. This will be subject to future research.

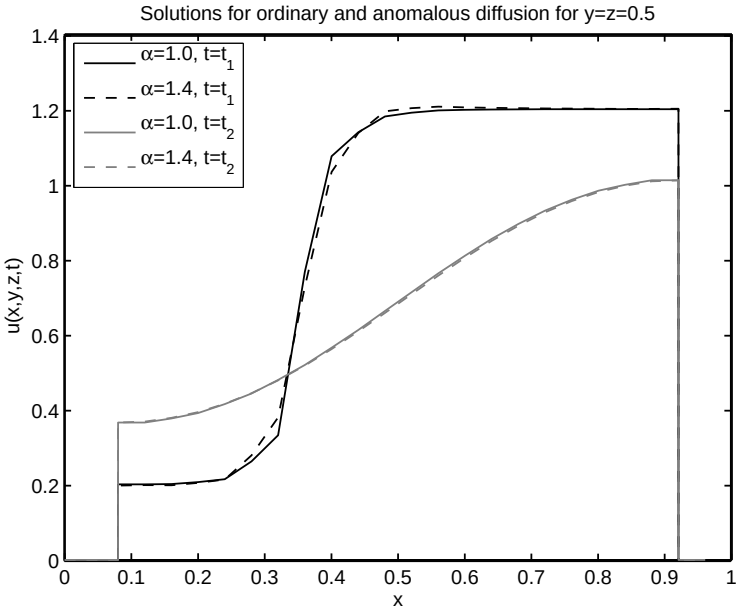

FIG. 8.2. *The numerical solution at $y = z = 0.5$ for $t = 10^{-4}$ (black) and $t = 3 \cdot 10^{-3}$ (gray), for subdiffusion ($\alpha = 1.4$, dashed) and ordinary diffusion ($\alpha = 1.0$, solid).*

REFERENCES

[1] G. I. Èskin. *Boundary value problems for elliptic pseudodifferential equations.* American Mathematical Society, 1981.

[2] L. C. Evans. *Partial Differential Equations.* American Mathematical Society, 1998.

[3] G. J. Fix and J. P. Roop. Least squares finite-element solution of a fractional order two-point boundary value problem. *Comput. Math. Appl.*, 48(7-8):1017–1033, 2004.

[4] T. Gebäck. *On boundary value problems for intracellular subdiffusion and signaling pathways, and for geometric flows.* PhD thesis, Chalmers University of Technology, 2007.

[5] R. Gorenflo. Fractional calculus: some numerical methods. In *Fractals and fractional calculus in continuum mechanics (Udine, 1996)*, volume 378 of *CISM Courses and Lectures*, pages 277–290. Springer, Vienna, 1997.

[6] R. Gorenflo, A. Vivoli, and F. Mainardi. Discrete and continuous random walk models for space-time fractional diffusion. *Nonlinear Dynam.*, 38(1-4):101–116, 2004.

[7] C. T. Kelley. *Iterative methods for linear and nonlinear equations*, volume 16 of *Frontiers in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1995. With separately available software.

[8] C. Lubich. Discretized fractional calculus. *SIAM J. Math. Anal.*, 17(3):704–719, 1986.

[9] R. Metzler and J. Klafter. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.*, 339:1–77, 2000.

[10] I. Podlubny. *Fractional differential equations*, volume 198 of *Mathematics in Science and Engineering.* Academic Press Inc., San Diego, CA, 1999.

[11] R. Rigler and E. S. Elson, editors. *Fluorescence Correlation Spectroscopy : theory and applications*, volume 65 of *Springer series in chemical physics.* Springer Verlag, 2001.

[12] J. P. Roop. Computational aspects of FEM approximation of fractional advection dispersion equations on bounded domains in $\mathbb{R}^2$. *J. Comput. Appl. Math.*, 193(1):243–268, 2006.

[13] P. Schwille and E. Haustein. Fluorescence correlation spectroscopy - an introduction to its concepts and applications. *Biophysics Textbook Online*, 2002.

[14] R. Seeley. Fractional powers of boundary problems. In *Actes du Congrès International des Mathématiciens (Nice, 1970), Tome 2*, pages 795–801. Gauthier-Villars, Paris, 1971.
[15] R. Seeley. Interpolation in L^p with boundary conditions. *Studia Math.*, 44:47–60, 1972.
[16] R. E. Showalter. *Hilbert Space Methods for Partial Differential Equations*. Pitman Publishing Ltd., 1979.
[17] M. Višik and G. I. Èskin. Parabolic convolution equations in a bounded region. *Amer. Math. Soc. Transl.*, 95(2):131–162, 1970.
[18] M. Wachsmuth, W. Waldeck, and J. Langowski. Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy. *J. Mol. Biol.*, 298:677–689, 2000.
[19] M. Weiss, H. Hashimoto, and T. Nilsson. Anomalous protein diffusion in living cells as seen by fluorescence correlation spectroscopy. *Biophys J.*, 84(6):4043–52, 2003.

Approximation of mean curvature flows

Tobias Gebäck

Abstract

This part of the thesis contains Paper II, in which we prove convergence of a computational scheme for approximation of generalized mean curvature flows with right-angle boundary conditions. A mean curvature flow is the evolution of a surface which at each point moves with a velocity in the normal direction that depends on the mean curvature of the surface at that point. We study the case when the velocity is a general non-decreasing function of the mean curvature, and when the evolution takes place inside a domain, and the surface intersects the domain boundary at a right angle whenever they intersect.

Contents

1. Introduction

This part of the thesis concerns mean curvature flows. In short, it is the study of continuously moving surfaces, where the speed with which they move depend on the mean curvature of the surface at each point. In the paper that follows this introduction, we prove convergence of a computational scheme for approximating such mean curvature flows, in the case when the surface sits inside a bounded domain and intersects the boundary of this domain at a right angle whenever they intersects. Furthermore, the speed with which the surface moves may be a rather general increasing function of the mean curvature.

In the following introductory sections, we will describe in more detail what mean curvature flows are and the basics for the approximation scheme that we use. There will also be a brief introduction to the topic of viscosity solutions, since it is this notion of solutions that is used for the mean curvature flow equations.

The paper has been submitted, but not yet accepted for publication.

2. Mean curvature flows

Consider a hypersurface Γ_0 in $\mathbb{R}^n$. At each point $x \in \Gamma_0$, assign a velocity v in the normal direction (see figure 1), so that Γ_0 moves at each point with velocity $v(x)\hat{n}(x)$, creating a new hypersurface. Continuing the process, a family $\{\Gamma_t\}_{t\geq 0}$ is created, where the hypersurfaces Γ_t evolve according to the normal velocity $v(x,t)$. If we take $v(x,t) = \kappa$, the mean curvature of Γ_t at x, we get the *mean curvature flow*. Mean curvature flows have been studied since the 1970:s, first by parametric methods from differential geometry, although it was soon clear that for $n \geq 3$ these methods ran into problems even for smooth hypersurfaces Γ_0, as the mean curvature flow could develop singularities (so that the curvature is not defined for some x), see figure 2. One method to overcome these problems, was introduced by Brakke [4], who used varifold theory to define weak notions of mean curvature flow (see also [7] for a modernization of these results). This method, however, does not give unique solutions.

Then, following ideas from Osher and Sethian [16], Evans and Spruck [9] developed a new approach to motion by mean curvature in which the hyper-surface Γ_0 is viewed as a level set of a continuous function $f : \mathbb{R}^n \to \mathbb{R}$, so that

$$\Gamma_0 = \{x \in \mathbb{R}^n \mid f(x) = \lambda\}$$

for some $\lambda \in \mathbb{R}$. The mean curvature flow is then studied through the PDE

$$\begin{cases} \dfrac{\partial u}{\partial t} = |\nabla u| \operatorname{div}\left(\dfrac{\nabla u}{|\nabla u|}\right) & \text{in } \mathbb{R}^n \times (0,\infty), \\ u = f & \text{on } \mathbb{R}^n \times \{t = 0\}, \end{cases} \tag{2.1}$$

which ensures that the level sets of u evolve according to their mean curvature, at least as long as $\nabla u \neq 0$, since $\operatorname{div}(\nabla u/|\nabla u|)$ is the divergence of the outward unit normal of

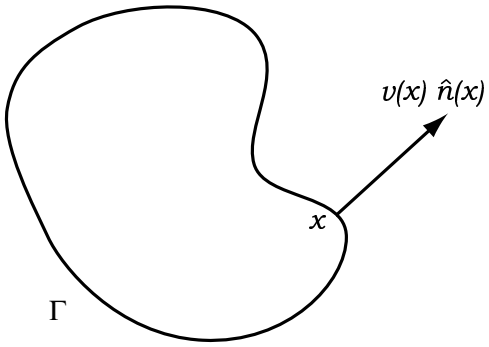

Figure 1: A surface Γ with a velocity v assigned in the normal direction at a point x.

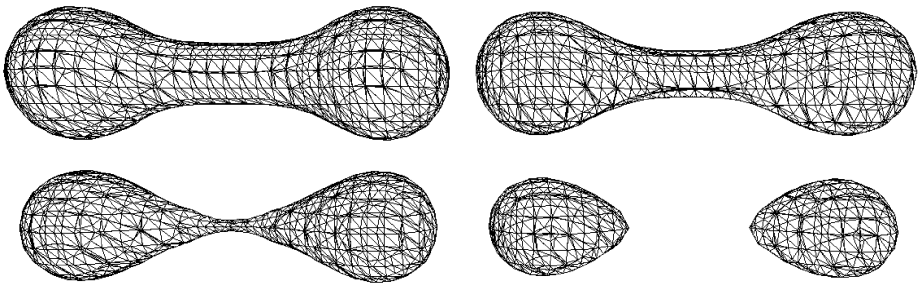

Figure 2: The mean curvature evolution of a dumbbell-shaped surface. The surface develops a singularity after a finite time and is split into two. The image was produced by R. Grzibovskis.

the level hyper-surfaces of u, which equals the mean curvature of the hyper-surfaces at each point x. The PDE (2.1) is nonlinear and degenerate parabolic and has a singularity for $\nabla u = 0$, which makes it rather hard to handle. However, the theory of *viscosity solutions* (see section 4) provides a well-suited tool to handle such equations and to prove the existence of a unique solution, which was done in [9].

A number of generalizations of this approach have appeared. Existence of a unique solution to more general level set equations has been established by Chen, Giga and Goto [5] and for very general cases, including the case we study, where $v = g(\kappa)$ and g is any non-decreasing, continuous function, by Ishii and Souganidis [15]. The Neumann problem for the mean curvature equation has also been studied. In this case, the whole evolution takes place inside a domain Ω, whose boundary is intersected at right angles by the level sets of u, i.e. the surfaces Γ. This is the situation that is treated in the article. The existence of unique solutions in this case was established by Sato [18] for convex Ω and by Giga and Sato [10] for more general Ω, but less general dependence on the curvature.

Curvature flows arise naturally in a range of situations, including fast reaction-slow diffusion problems (see [17]) and image processing (see [1]).

3. The convolution thresholding scheme

In 1992, Bence, Merriman and Osher [3] presented a scheme for approximation of mean curvature flows using the level set approach. The basic observation behind this scheme is that mean curvature flows appear for the reaction fronts in fast reaction-slow diffusion equations with a bi-stable reaction term described by a double-well potential with wells of equal depth (see [17]). That is, if a substance diffuses through space, and at the same time is subject to fast degradation and production configured so that the concentration quickly becomes either 0 or 1, then there will be a reaction front between the region with concentration 1 and the region with concentration 0, and this front will move with a normal velocity proportional to its mean curvature if the reaction is given by the above potential. But the solution to the reaction-diffusion equation may be approximated by first solving the diffusion equation for a short time, and then applying the reaction, which amounts to setting the solution to 1 if it is above a certain threshold (in our case $\frac{1}{2}$), and to 0 if it is below the threshold. And, since the solution to the diffusion equation is given by convolution with the heat kernel, we get a so called convolution-thresholding scheme, which in turn can be generalized to use a more general convolution kernel ρ, for example the characteristic function of a ball.

The resulting scheme can be described as follows (cf. [12]). First, fix a radially symmetric convolution kernel, $\rho : \mathbb{R}^n \to \mathbb{R}$, and define its contraction $\rho^{\sqrt{h}}(x) = h^{-n/2}\rho(x/\sqrt{h})$. Then, given a set $C_0 \subset \mathbb{R}^n$, choose a time-step h and compute the convolution $M^{C_0}(x,h) = (\rho^{\sqrt{h}} * \chi_{C_0})(x)$, where χ_{C_0} is the characteristic function of C_0. Set

$$C_1 = \{x \in \mathbb{R}^n \mid M^{C_0}(x,h) \geq \frac{1}{2}\int_{\mathbb{R}^n} \rho^{\sqrt{h}}(y)\, dy\}$$

and continue the process by computing $M^{C_1}(x,h)$ and defining C_2 and so on. We then end up with a sequence $\{C_k\}_0^K$ of closed sets in $\mathbb{R}^n$ and if we set

$$\Gamma_k = \partial C_k, \quad k = 0, 1, \ldots, K,$$

we get a sequence $\{\Gamma_k\}_0^K$ of surfaces which approximate the motion by mean curvature for the initial surface Γ_0. The principle is shown in figure 3.

The convergence of this scheme to the true mean curvature flow as $h \to 0$ was proved using different approaches by Evans [8], Barles and Georgelin [2] and Ishii [12]. General thresholding schemes were also studied by Ishii, Pires and Souganidis [14]. Later, two significant generalizations of the scheme have been developed. In 2002, Ishii and Ishii [13] published an algorithm for mean curvature flow with right-angle boundary conditions, and about at the same time, Grzibovskis and Heintz [11] developed a scheme for motion with normal velocity equal to a (nonlinear) non-decreasing function of the mean curvature (which is what we call *generalized* mean curvature flow). These generalizations are

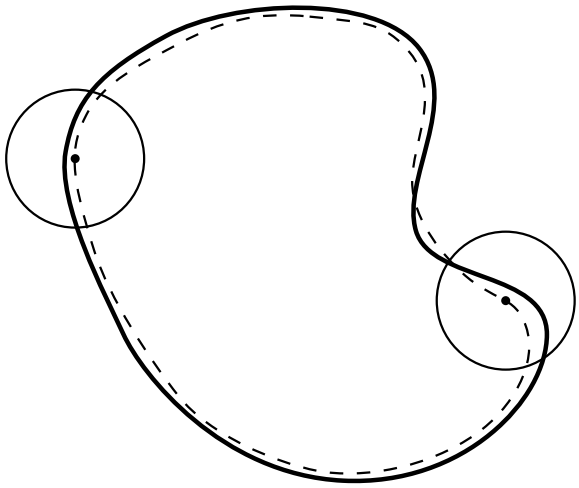

Figure 3: Illustration of the convolution-thresholding scheme. The initial set C_0 is bounded by the solid line. Convolutions are computed for the characteristic function of this set with the kernel $\rho^{\sqrt{h}}$ (with support inside the small circle), and the boundary for the new set C_1 (dashed line) is set where the value of the convolution is equal to half the mass of the kernel.

described in the paper and it is these two schemes that are combined to give a scheme for generalized mean curvature flow with right-angle boundary conditions, and the main result in the paper is the convergence of this scheme as the time step $h \to 0$.

4. Viscosity solutions

4.1. Background

The theory of viscosity solutions was developed during the 1980's by M.G. Crandall, L.C. Evans, H. Ishii, P.-L. Lions and others while seeking solutions to the Hamilton-Jacobi equations. The name viscosity solutions originates from the method of "vanishing viscosity" which was used to solve first-order equations and which was consistent with the new theory being developed. Now, however, viscosity solutions do not generally have much to do with viscosity. The theory provides very general existence and uniqueness results and allows merely continuous functions to be solutions of fully nonlinear second-order equations. An excellent account of the theory may be found in the *"User's guide to viscosity solutions"* by Crandall, Ishii and Lions [6]. Here, we give a short introduction to the theory and introduce some concepts which will be used later on.

4.2. THEORY OF VISCOSITY SOLUTIONS

The theory of viscosity solutions applies to equations of the form

$$F(x, u, \nabla u, D^2 u) = 0, \tag{4.1}$$

where $x \in \mathbb{R}^n$, $u = u(x)$ is a real-valued function, $\nabla u \in \mathbb{R}^n$ its gradient and $D^2 u \in \mathcal{S}(n)$ the matrix of second derivatives of u. $\mathcal{S}(n)$ is the set of real, symmetric $n \times n$ matrices, which is partially ordered by the relation $\leq$, where $Y \leq X$ means $\xi^t Y \xi \leq \xi^t X \xi$ for all $\xi \in \mathbb{R}^n$. We also equip $\mathcal{S}(n)$ with the norm $\|X\| = \max\{|X\xi| \mid \xi \in \mathbb{R}^n, |\xi| = 1\}$. Finally, $F : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathcal{S}(n) \to \mathbb{R}$ is a function, which can take many forms.

For the theory to apply, we require F to satisfy the monotonicity conditions

$$F(x, r, p, X) \leq F(x, s, p, X) \text{ if } r \leq s \tag{4.2}$$

and

$$F(x, r, p, X) \leq F(x, r, p, Y) \text{ if } Y \leq X, \tag{4.3}$$

where $r, s \in \mathbb{R}$, $x, p \in \mathbb{R}^n$ and $X, Y \in \mathcal{S}(n)$. If (4.3) holds, F is said to be *degenerate elliptic* and if (4.2) also holds, F is *proper*.

Now suppose that F is proper and that $u \in C^2(\mathbb{R}^n)$ is a subsolution to $F = 0$, i.e. solves

$$F(x, u(x), \nabla u(x), D^2 u(x)) \leq 0$$

for all $x \in \mathbb{R}^n$. Choose a test function φ that is also C^2, and suppose that $u - \varphi$ has a local maximum at $\hat{x}$. Then we have $\nabla(u - \varphi)(\hat{x}) = 0$ and $D^2(u - \varphi)(\hat{x}) \leq 0$, i.e. $\nabla u(\hat{x}) = \nabla \varphi(\hat{x})$ and $D^2 u(\hat{x}) \leq D^2 \varphi(\hat{x})$, and by (4.3),

$$F(\hat{x}, u(\hat{x}), \nabla \varphi(\hat{x}), D^2 \varphi(\hat{x})) \leq F(\hat{x}, u(\hat{x}), \nabla u(\hat{x}), D^2 u(\hat{x})) \leq 0.$$

We have thus replaced the derivatives of u with derivatives of test functions φ, which we choose to be well-behaved (i.e. at least C^2). One could now try to define an arbitrary function u to be a weak or generalized *subsolution* of (4.1) if

$$F(\hat{x}, u(\hat{x}), \nabla \varphi(\hat{x}), D^2 \varphi(\hat{x})) \leq 0$$

whenever φ is C^2 and $u - \varphi$ has a local maximum at $\hat{x}$. However, a slightly different definition proves more useful and we therefore note that since $u - \varphi$ has a maximum at $\hat{x}$, $u(x) \leq u(\hat{x}) - \varphi(\hat{x}) + \varphi(x)$ for x near $\hat{x}$, so a Taylor expansion of φ at $\hat{x}$ gives

$$u(x) \leq u(\hat{x}) + \langle p, x - \hat{x} \rangle + \frac{1}{2} \langle X(x - \hat{x}), x - \hat{x} \rangle + o(|x - \hat{x}|^2) \tag{4.4}$$

with $p = \nabla \varphi(\hat{x})$ and $X = D^2 \varphi(\hat{x})$. Also, if (4.4) holds for some $(p, X) \in \mathbb{R}^n \times \mathcal{S}(n)$ and $u \in C^2$, then $p = \nabla u(\hat{x})$ and $D^2 u(\hat{x}) \leq X$, so that if u solves $F \leq 0$, it follows that $F(\hat{x}, u(\hat{x}), p, X) \leq 0$ whenever (4.4) is true. Letting $\mathcal{O} \subset \mathbb{R}^n$ be locally compact, $u : \mathcal{O} \to \mathbb{R}$ and $\hat{x} \in \mathcal{O}$, we therefore define the second-order "superjet" of u at $\hat{x}$ by

$$J^{2,+}_{\mathcal{O}} u(\hat{x}) = \{(p, X) \mid (4.4) \text{ holds as } x \to \hat{x}, \text{ with } x \in \mathcal{O}\}. \tag{4.5}$$

Reversing the inequality sign in (4.4) gives us the definition of the "subjet" $J^{2,-}_{\mathcal{O}}u(\hat{x})$, or equivalently, $J^{2,-}_{\mathcal{O}}u(\hat{x}) = -J^{2,+}_{\mathcal{O}}(-u)(\hat{x})$. We also state the definitions of the sets of upper and lower semicontinuous functions

$$\begin{aligned} \mathrm{USC}(\mathcal{O}) &= \{u : \mathcal{O} \to [-\infty, \infty) \mid u^{-1}([-\infty, a)) \text{ is open in } \mathcal{O} \text{ for each } a \in \mathbb{R}\} \\ \mathrm{LSC}(\mathcal{O}) &= \{u : \mathcal{O} \to (-\infty, \infty] \mid u^{-1}((a, \infty]) \text{ is open in } \mathcal{O} \text{ for each } a \in \mathbb{R}\} \end{aligned} \tag{4.6}$$

Finally, we are able to define viscosity solutions for the equation (4.1).

Definition 4.1 *Let F be proper and $\mathcal{O} \subset \mathbb{R}^n$. Then $u \in \mathrm{USC}(\mathcal{O})$ is a* subsolution *of $F = 0$ on $\mathcal{O}$ if*

$$F(x, u(x), p, X) \leq 0 \text{ for all } x \in \mathcal{O} \text{ and } (p, X) \in J^{2,+}_{\mathcal{O}}u(x). \tag{4.7}$$

Similarly, $u \in \mathrm{LSC}(\mathcal{O})$ is a supersolution *of $F = 0$ on $\mathcal{O}$ if*

$$F(x, u(x), p, X) \geq 0 \text{ for all } x \in \mathcal{O} \text{ and } (p, X) \in J^{2,-}_{\mathcal{O}}u(x). \tag{4.8}$$

Finally, u is a (viscosity) solution *of $F = 0$ in $\mathcal{O}$ if it is both a subsolution and a supersolution.*

We note that since viscosity solutions are both upper and lower semicontinuous, they are continuous. Also, in view of the discussion above, it may be noted that

$$\begin{aligned} J^{2,+}_{\mathcal{O}}u(\hat{x}) &= \{(\nabla\varphi(\hat{x}), D^2\varphi(\hat{x})) \mid \varphi \in C^2 \text{ and } u - \varphi \text{ has a local maximum at } \hat{x}\}, \\ J^{2,-}_{\mathcal{O}}u(\hat{x}) &= \{(\nabla\varphi(\hat{x}), D^2\varphi(\hat{x})) \mid \varphi \in C^2 \text{ and } u - \varphi \text{ has a local minimum at } \hat{x}\}, \end{aligned}$$

which may be used to facilitate the use of definition 4.1. We also note that the semijets only depend on the set $\mathcal{O}$ if $\hat{x} \in \partial\mathcal{O}$, so if that is not the case, we may drop the subscript.

We also need to define the closures of the semijets for $x \in \mathcal{O}$ as

$$\begin{aligned} \overline{J}^{2,\pm}_{\mathcal{O}}u(x) = \{(p, X) \in \mathbb{R}^n \times \mathcal{S}(n) \mid \exists (x_k, p_k, X_k) \in \mathcal{O} \times \mathbb{R}^n \times \mathcal{S}(n) : \\ (p_k, X_k) \in J^{2,\pm}_{\mathcal{O}}u(x_k) \text{ and } (x_k, u(x_k), p_k, X_k) \to (x, u(x), p, X)\} \end{aligned}$$

and note that if u is a subsolution of $F = 0$ in $\mathcal{O}$, then $F(x, u(x), p, X) \leq 0$ for $x \in \mathcal{O}$ and $(p, X) \in J^{2,+}_{\mathcal{O}}u(x)$. If F is lower semicontinuous, this remains true if $(p, X) \in \overline{J}^{2,+}_{\mathcal{O}}u(x)$. Similar remarks are true for supersolutions and solutions.

4.3. Boundary conditions

Viscosity solutions also allow precise formulations of boundary conditions. Consider the boundary value problem

$$\begin{cases} F(x, u(x), \nabla u(x), D^2u(x)) = 0, & x \in \Omega \\ B(x, u(x), \nabla u(x)) = 0, & x \in \partial\Omega \end{cases} \tag{4.9}$$

in an open set $\Omega \subset \mathbb{R}^n$, where F and B are both proper functions. The correct definition of a viscosity solution of (4.9) is then

Definition 4.2 *A function* $u \in \mathrm{USC}(\overline{\Omega})$ *is a* subsolution *of (4.9) if*

$$\begin{cases} F(x,u(x),p,X) \le 0 & x \in \Omega,\ (p,X) \in \overline{J}^{2,+}_{\overline{\Omega}} u(x), \\ B(x,u(x),p) \wedge F(x,u(x),p,X) \le 0 & x \in \partial\Omega,\ (p,X) \in \overline{J}^{2,+}_{\overline{\Omega}} u(x). \end{cases} \tag{4.10}$$

$u \in \mathrm{LSC}(\overline{\Omega})$ *is a* supersolution *of (4.9) if*

$$\begin{cases} F(x,u(x),p,X) \ge 0 & x \in \Omega,\ (p,X) \in \overline{J}^{2,-}_{\overline{\Omega}} u(x), \\ B(x,u(x),p) \vee F(x,u(x),p,X) \ge 0 & x \in \partial\Omega,\ (p,X) \in \overline{J}^{2,-}_{\overline{\Omega}} u(x). \end{cases} \tag{4.11}$$

Finally, u *is a* solution *if it is both a subsolution and a supersolution.*

Here, $a \vee b = \max\{a,b\}$ and $a \wedge b = \min\{a,b\}$, so what the definition basically means is that on the boundary, either the boundary condition or the equation should hold. That we can not expect the boundary conditions to hold in a stronger sense is demonstrated by an example in [6].

4.4. SINGULAR EQUATIONS

Since the equation we are interested in has a singularity for $\nabla u = 0$, we need to introduce a third definition of viscosity solutions.

To start with, given a function $u : \Omega \to \mathbb{R}$, with $\Omega \subset \mathbb{R}^n$, we introduce the upper and lower semicontinuous relaxations

$$u^*(x) = \lim_{\varepsilon \to 0} \sup\{u(y) \mid y \in \Omega; |x-y| < \varepsilon\}, \tag{4.12}$$

$$u_*(x) = \lim_{\varepsilon \to 0} \inf\{u(y) \mid y \in \Omega; |x-y| < \varepsilon\}, \tag{4.13}$$

which are defined on $\overline{\Omega}$ and take values in $\mathbb{R} \cup \{\infty\}$ and $\mathbb{R} \cup \{-\infty\}$ respectively.

Also, supposing our function F is only defined on a dense subset W of $L(\Omega) = \Omega \times \mathbb{R} \times \mathbb{R}^n \times \mathcal{S}(n)$, we may similarly define the relaxations F^* and F_* on $\overline{W} = L(\Omega)$ as

$$F^*(x,u,p,X) = \lim_{\varepsilon \to 0} \sup\{F(y,v,q,Y) \mid (y,v,q,Y) \in W,\ \|(x,u,p,X) - (y,v,q,Y)\| < \varepsilon\},$$

$$F_*(x,u,p,X) = \lim_{\varepsilon \to 0} \inf\{F(y,v,q,Y) \mid (y,v,q,Y) \in W,\ \|(x,u,p,X) - (y,v,q,Y)\| < \varepsilon\},$$

where $\|(x,u,p,X)\|$ is just the sum of the norms for each component. Then, we can make the following definition of viscosity solutions:

Definition 4.3 *A function* $u : \Omega \to \mathbb{R}$ *is a* subsolution *of (4.1) if* $u^* < \infty$ *in* Ω *and*

$$F_*(x,u^*(x),p,X) \le 0 \text{ for all } x \in \Omega, (p,X) \in \overline{J}^{2,+}_{\Omega} u^*(x),$$

a supersolution *if* $u_* > -\infty$ *in* Ω *and*

$$F^*(x,u_*(x),p,X) \ge 0 \text{ for all } x \in \Omega, (p,X) \in \overline{J}^{2,-}_{\Omega} u_*(x),$$

and a solution *if it is both sub- and supersolution.*

This definition is adopted from [5], where it is used for proving existence of solutions of curvature flow equations and similar.

4.5. Comparison

We finally make some comments on the method generally used for proving existence of viscosity solutions to suitable equations. Such proofs in general consist of three steps. The first is to establish a comparison result, i.e that if u is a subsolution and v is a supersolution, then $u \leq v$. From this result it follows immediately that if there is a solution, it must be unique. It also follows that if u is a solution by the definition for singular equations given above, then u is continuous. Furthermore, the comparison result proves helpful to us in our proof of convergence in the article that follows.

The second step of a proof of existence is to construct a subsolution and a supersolution. The third step is to invoke Perron's method to show that in that case, there exists a solution. See Crandall, Ishii, Lions [6] for details. For applications of the method, see also the proofs of existence of solutions to level set equations for mean curvature motion, e.g [5], [10], [18].

References

[1] L. Alvarez, F. Guichard, P.-L. Lions, and J.-M. Morel. Axioms and fundamental equations of image processing. *Arch. Rational Mech. Anal.*, 123(3):199–257, 1993.

[2] G. Barles and C. Georgelin. A simple proof of convergence for an approximation scheme for computing motions by mean curvature. *SIAM J. Numer. Anal.*, 32(2):484–500, 1995.

[3] J. Bence, B. Merriman, and S. Osher. Diffusion generated motion by mean curvature. In *Computational Crystal Growers Workshop*, pages 73–83. American Mathematical Society, 1992.

[4] K. A. Brakke. *The Motion of a Surface by its Mean Curvature.* Princeton University Press, Princeton, NJ, 1978.

[5] Y.-G. Chen, Y. Giga, and S. Goto. Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations. *J. Differential Geom.*, 33:749–786, 1991.

[6] M. G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.*, 27(1):1–67, 1992.

[7] K. Ecker. *Regularity Theory for Mean Curvature Flow.* Birkhäuser Verlag, 2004.

[8] L. C. Evans. Convergence of an algorithm for mean curvature motion. *Indiana Univ. Math. J.*, 42:533–556, 1993.

[9] L. C. Evans and J. Spruck. Motion of level sets by mean curvature, I. *J. Differential Geom.*, 33:635–681, 1991.

[10] Y. Giga and M.-H. Sato. Neumann problem for singular degenerate parabolic equations. *Differential and Integral Equations*, 6(6):1217–1230, 1993.

[11] R. Grzhibovskis and A. Heintz. A convolution-thresholding approximation of generalized curvature flows. *SIAM J. Numer Anal.*, 42(6):2652–2670, 2005.

[12] H. Ishii. A generalization of the Bence, Merriman and Osher algorithm for motion by mean curvature. In *Curvature flows and related topics (Levico, 1994)*, volume 5 of *GAKUTO Internat. Ser. Math. Sci. Appl.*, pages 111–127. Gakkōtosho, Tokyo, 1995.

[13] H. Ishii and K. Ishii. An approximation scheme for motion by mean curvature with right-angle boundary condition. *SIAM J. Math. Anal.*, 33(2):369–389, 2001.

[14] H. Ishii, G. E. Pires, and P. E. Souganidis. Threshold dynamics type approximation schemes for propagating fronts. *J. Math. Soc. Japan*, 51(2):267–308, 1999.

[15] H. Ishii and P. Souganidis. Generalized motion of noncompact hypersurfaces with velocity having arbitrary growth on the curvature tensor. *Tohoku Math. J.*, 47:227–250, 1995.

[16] S. Osher and J. A. Sethian. Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.*, 79:12–49, 1988.

[17] J. Rubinstein, P. Sternberg, and J. Keller. Fast reaction, slow diffusion and curve shortening. *SIAM J. Appl. Math.*, 49:116–133, 1989.

[18] M.-H. Sato. Interface evolution with Neumann boundary condition. *Adv. Math. Sci. Appl.*, 4(1):249–264, 1994.

APPROXIMATION OF GENERALIZED MEAN CURVATURE FLOW WITH RIGHT-ANGLE BOUNDARY CONDITIONS

TOBIAS GEBÄCK*

Abstract. In this paper, we prove the convergence of a convolution-thresholding scheme for computing the evolution of surfaces, which move at each point with a velocity equal to an increasing function of the mean curvature at that point. Furthermore, the entire evolution is assumed to take place inside a convex domain and wherever the surface intersects the domain boundary, it should do so at a right angle. We show that the approximations given by the scheme converge to the viscosity solution of the corresponding level set PDE as the time step tends to zero.

The scheme presented here is a generalization of both the scheme presented by Ishii and Ishii for regular mean curvature evolution with right-angle boundary conditions, and the scheme by Grzhibovskis and Heintz for the case when the velocity equals an increasing function of the mean curvature, without boundary conditions. These algorithms are in turn based on the convolution-thresholding scheme devised by Bence, Merriman and Osher.

Key words. generalized mean curvature flow, right-angle boundary conditions, convolution-thresholding scheme, viscosity solutions

AMS subject classifications. 35K65, 35K55, 65M12, 53C44

1. Introduction. In this paper, we prove convergence of a convolution-thresholding scheme for generalized mean curvature motion with right-angle boundary conditions. By a generalized mean curvature flow, we mean a family $\{\Gamma_t\}_{t\geq 0}$ of hypersurfaces in $\mathbb{R}^n$, dependent on time t, where the surfaces move with a normal velocity equal to an increasing, continuous function of the mean curvature at each point. By right-angle boundary conditions we mean the fact that we let the surface evolution take place inside a domain, and wherever the surface intersects the domain boundary, it should do so at a right angle.

A number of methods have been suggested for describing mean curvature flows in order to overcome the problems with singularities that may appear in the flow, even for smooth initial data (see e.g. [3], [5], [7]). The approach that we will use here is the level set method, which was introduced by Evans and Spruck [7], following ideas from Osher and Sethian [13]. The idea is to let the hypersurfaces Γ_t be level sets of a continuous function $u : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ and describe the properties of this function through an evolution equation, which in the case of ordinary mean curvature flow is

$$u_t = |\nabla u| \ \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right). \tag{1.1}$$

If $u(x,t)$ solves this equation, then the level sets of u, i.e. $\Gamma_t = \{x \in \mathbb{R}^n \mid u(x,t) = \lambda\}$, move with a normal velocity equal to the mean curvature of the level set, as long as the mean curvature is well-defined. Evans and Spruck [7] proved that the evolution equation (1.1) has a unique viscosity solution (see [4]), which need only be continuous. This means that a mean curvature flow for a surface may be defined even if the function u lacks derivatives, so that the mean curvature is not defined everywhere on the level

*Mathematical Sciences, Chalmers University of Technology, S-412 96 Göteborg, Sweden, tobiasg@math.chalmers.se

set. The evolution equation that applies to the present case will be discussed in section 2.

In 1992, Bence, Merriman and Osher [2] presented a convolution-thresholding scheme for approximating mean curvature flows using the level set approach. The convergence of this scheme was proved using different approaches by Evans [6], Barles and Georgelin [1] and Ishii [10]. General thresholding schemes were later studied by Ishii, Pires and Souganidis [12]. More recently, two significant generalizations of the original scheme have been developed. In 2002, Ishii and Ishii [11] published a scheme for mean curvature flow with right-angle boundary conditions, and at about the same time, Grzhibovskis and Heintz [9] developed a scheme for approximating motion with normal velocity equal to an increasing, continuous function of the mean curvature (which is what we here call generalized mean curvature flow), but only for flows without boundary conditions.

In this paper, we introduce a scheme which extends both these results to the case of generalized mean curvature flow with right-angle boundary conditions. We use the same method as Ishii and Ishii [11] to prove that our scheme converges to the viscosity solution of the level set equation as the time step tends to zero. In order to understand our scheme, we will first briefly discuss the previous results.

1.1. The BMO-scheme. A (slightly modified) version of the Bence, Merriman and Osher-scheme (BMO-scheme) can be described as follows (cf. Ishii [10]). First, fix a radially symmetric convolution kernel, ρ, and define its contraction $\rho^{\sqrt{h}}(x) = h^{-n/2}\rho(x/\sqrt{h})$. Then, given a set $C_0 \subset \mathbb{R}^n$, choose a time-step h and compute the convolution $\widetilde{M}^{C_0}(x,h) = (\rho^{\sqrt{h}} * \chi_{C_0})(x)$, where χ_{C_0} is the characteristic function of C_0. Set

$$C_1 = \{x \in \mathbb{R}^n \mid \widetilde{M}^{C_0}(x,h) \geq \frac{1}{2}\int_{\mathbb{R}^n} \rho^{\sqrt{h}}(y)\, dy\}$$

and continue the process by computing $\widetilde{M}^{C_1}(x,h)$ and defining C_2 and so on. We then end up with a sequence $\{C_k\}_{k\in\mathbb{N}}$ of closed sets in $\mathbb{R}^n$ and we set

$$C_t^h = C_k \quad \text{if } kh \leq t < (k+1)h,\ t \geq 0.$$

Now, letting $h \to 0$, we obtain a flow of closed subsets in $\mathbb{R}^n$ whose boundaries move with a normal velocity equal to a constant times its mean curvature, where the constant depends only on n and the choice of ρ.

In the original scheme, $\widetilde{M}^{C_0}(x,t)$ was the solution to the heat equation with initial data χ_{C_0}, which corresponds to the choice of ρ as the Gauss kernel, and which leads to motion by $(n-1)$ times the mean curvature.

1.2. Right-angle boundary conditions. As was already mentioned, the above scheme was extended by Ishii and Ishii [11] to the case of right-angle boundary conditions. The extension works as follows. Given an open domain $\Omega \subset \mathbb{R}^n$ with C^2-boundary, an initial set C_0, and a convolution kernel ρ, we define

$$M^{C_0}(x,h) = \int_\Omega \rho^{\sqrt{h}}(y-x)\chi_{C_0}(y)\, dy \tag{1.2}$$

and set

$$C_1 = \{x \in \mathbb{R}^n \mid M^{C_0}(x,h) \geq \frac{1}{2}\int_\Omega \rho^{\sqrt{h}}(y-x)\,dy\},$$

which is the same as before, except that the integrals are computed over Ω instead of $\mathbb{R}^n$. Defining the sequence $\{C_k\}_{k\in\mathbb{N}}$ and C_t^h as above and letting $h \to 0$, we get a flow of sets whose boundaries not only move by a constant times mean curvature but also intersect the boundary of Ω at a right angle, at least in the sense of viscosity solutions.

1.3. Generalized mean curvature motion. The last extension of the BMO-scheme we will discuss is the scheme by Grzhibovskis and Heintz [9], that lets the boundaries of the sets move with a normal velocity $v = g(\kappa)$, where κ is the mean curvature and $g : \mathbb{R} \to \mathbb{R}$ is an increasing, continuous function.

The scheme uses two different radially symmetric convolution kernels, ρ_1 and ρ_2 and given a set C we can define

$$\begin{aligned}\widetilde{N}_i^C(x,h) &= \widetilde{M}_i^C(x,h) - \frac{1}{2}\int_{\mathbb{R}^n}\rho_i^{\sqrt{h}}(y-x)\,dy \\ &= \int_{\mathbb{R}^n}\rho_i^{\sqrt{h}}(y-x)\chi_C(y)\,dy - \frac{1}{2}\int_{\mathbb{R}^n}\rho_i^{\sqrt{h}}(y-x)\,dy\end{aligned} \tag{1.3}$$

for $i = 1, 2$. Now, a crucial part of all the proofs of convergence of these convolution-thresholding schemes is an expansion of $\widetilde{N}_i^C(x,h)$ in h of the form

$$\widetilde{N}_i^C(x,h) = a_i\sqrt{h}v(x) - b_i\sqrt{h}\kappa(x) + o(h). \tag{1.4}$$

with

$$\begin{aligned}a_i &= \int_{\mathbb{R}^{n-1}}\rho_i(y',0)\,dy', \\ b_i &= \frac{1}{2}\int_{\mathbb{R}^{n-1}} y_1^2\rho_i(y',0)\,dy'.\end{aligned}$$

Clearly, setting $\widetilde{N}_i^C(x,h) = 0$ gives $v = \frac{b_i}{a_i}\kappa + o(\sqrt{h})$, which corresponds to the original BMO-scheme.

Now, using two convolution kernels, (1.4) gives us two linear equations for v and κ. Solving these, we get

$$\begin{cases} v = \dfrac{1}{\sqrt{h}}\dfrac{b_2\widetilde{N}_1 - b_1\widetilde{N}_2}{d} + o(\sqrt{h}) \\ \kappa = \dfrac{1}{\sqrt{h}}\dfrac{a_2\widetilde{N}_1 - a_1\widetilde{N}_2}{d} + o(\sqrt{h}) \end{cases},$$

where $d = a_1b_2 - a_2b_1$ is the determinant. Thus, since we want to have $v = g(\kappa)$, we define

$$F(N_1,N_2) = v - g(\kappa) = \frac{1}{\sqrt{h}}\frac{b_2N_1 - b_1N_2}{d} - g\left(\frac{1}{\sqrt{h}}\frac{a_2N_1 - a_1N_2}{d}\right) \tag{1.5}$$

and a thresholding scheme

$$C_{k+1} = \{x \in \mathbb{R}^n \mid F(\widetilde{N}_1^{C_k}(x,h), \widetilde{N}_2^{C_k}(x,h)) \geq 0\},$$

for $k \in \mathbb{N}$.

In order for this scheme to converge to the actual generalized mean curvature motion, it turns out that F must satisfy the condition $\partial F/\partial N_i > 0,\ i = 1,2$, which leads to the restrictions

$$\begin{aligned} &d = a_1 b_2 - a_2 b_1 > 0, \\ &0 < \frac{b_1}{a_1} < g' < \frac{b_2}{a_2}, \end{aligned} \tag{1.6}$$

on g and ρ_i, saying both that g must have bounded derivative both from above and below, and that given a function g, the convolution kernels must be chosen with some care to fulfill the inequalities. If one would like to use a function g with unbounded, zero or non-existent derivative, it is possible to use uniform approximations g_ν to g and still get a scheme that converges as h and ν tend to zero.

1.4. Outline. The structure of the following sections is as follows. In section 2, our extension of the scheme is presented in detail and all assumptions stated. In section 3, a few crucial lemmas are proved, the most important one being proposition 3.3. Finally, in section 4, we prove in theorems 4.1 and 4.6 that the scheme converges to the solution of the level set equation as the time-step tends to zero.

2. The approximation scheme. We now turn to the problem of extending the scheme for approximating generalized curvature flow to the case of right-angle boundary conditions and proving its convergence as the time step tends to zero. For that purpose, let Ω be an open, bounded domain in $\mathbb{R}^n$ with C^2-boundary $\partial\Omega$. Given $u_0 \in C(\overline{\Omega})$, we consider the level set equation

$$\begin{cases} \dfrac{\partial u}{\partial t}(x,t) - |\nabla u(x,t)|\ g(\mathrm{curv}(u(x,t))) = 0, & x \in \Omega,\ t \in (0,T) \\ \dfrac{\partial u}{\partial \hat{n}}(x,t) = 0, & x \in \partial\Omega,\ t \in (0,T) \\ u(x,0) = u_0(x), & x \in \overline{\Omega} \end{cases} \tag{2.1}$$

for $T > 0$, where

$$\mathrm{curv}(u(x)) = \mathrm{div}\left(\frac{\nabla u(x)}{|\nabla u(x)|}\right) = \frac{1}{|\nabla u(x)|}\sum_{i,j=1}^{n}\left(\delta_{ij} - \frac{u_{x_i}(x)u_{x_j}(x)}{|\nabla u(x)|^2}\right)u_{x_i x_j}(x) \tag{2.2}$$

is $n-1$ times the mean curvature of the level set of u passing through the point x, $\hat{n}$ is the outward unit normal to Ω and $g : \mathbb{R} \to \mathbb{R}$ fulfills the conditions

$$g \in C(\mathbb{R}),\ g(0) = 0 \tag{2.3a}$$

$$g(x) = O(x) \text{ as } x \to \pm\infty \tag{2.3b}$$

$$g \text{ is increasing.} \tag{2.3c}$$

The PDE describes a function whose level sets $\{x \in \mathbb{R}^n \mid u(x,t) = \lambda\}$ move with normal velocity $g(\mathrm{curv}(u(x,t)))$ and intersect $\partial\Omega$ at a right angle, at least formally.

The equation is degenerate parabolic and has singularities for $\nabla u = 0$, but in spite of these difficulties, Sato [15] showed that if Ω is convex, the equation has a unique viscosity solution in $C(\overline{\Omega} \times [0,T))$ for any $T > 0$. Furthermore, if g is linear, Giga and Sato [8] proved that there is a unique viscosity solution even if Ω is not convex. Since we are interested in nonlinear functions g, we need the additional assumption that

$$\Omega \text{ is convex.}$$

It should be noted, however, that we only use this assumption through the use of the comparison principle from Sato [15]. So if a proof of existence of unique solutions is constructed for the case of non-convex Ω for nonlinear g, the proof of convergence of the scheme will be valid for this case too.

Because of (1.6), the convergence of the scheme also requires that

$$g \in C^1(\mathbb{R}) \text{ and } \exists \zeta_1, \zeta_2 > 0 : \forall x \in \mathbb{R} : g'(x) \in (\zeta_1, \zeta_2), \tag{2.3d}$$

which of course implies (2.3b) and (2.3c). But we will then also show how to get around this problem if we can find uniform approximations $g_\nu \to g$, where g fulfills (2.3a)–(2.3c), but has unbounded, zero or non-existent derivative and g_ν fulfills (2.3a)–(2.3d) for all ν.

To formulate the approximation scheme, we choose non-negative, measurable, radially symmetric weight functions ρ_1 and ρ_2, satisfying the conditions

$$\int_{\mathbb{R}^n} \rho_i(x)\, dx < \infty \tag{2.4a}$$

$$\int_{\mathbb{R}^{n-1}} \rho_i(\xi, 0)\, d\xi < \infty \tag{2.4b}$$

$$\operatorname{supp} \rho_i \text{ is compact,} \tag{2.4c}$$

as well as the conditions (1.6) depending on g, which we state again for convenience:

$$\begin{aligned} &d = a_1 b_2 - a_2 b_1 > 0, \\ &0 < \frac{b_1}{a_1} < g' < \frac{b_2}{a_2}, \end{aligned} \tag{1.6}$$

where

$$\begin{aligned} a_i &= \int_{\mathbb{R}^{n-1}} \rho_i(y', 0)\, dy', \\ b_i &= \frac{1}{2} \int_{\mathbb{R}^{n-1}} y_1^2 \rho_i(y', 0)\, dy'. \end{aligned}$$

The assumption (2.3d) ensures that this is possible. Condition (2.4c) is not really necessary and could be replaced by conditions for rapid decrease of ρ_i, but since the proofs are much simpler when we assume compact support, we use that assumption. See [11] for the proof when g is linear including the case when ρ has non-compact support.

Inspired by (1.3) and (1.2), we define

$$N_i^C(x,h) = \int_\Omega \rho_i^{\sqrt{h}}(y-x)\chi_C(y)\, dy - \frac{1}{2} \int_\Omega \rho_i^{\sqrt{h}}(y-x)\, dy$$

for $i = 1, 2$ and a mapping $\mathcal{G}_h$, that maps subsets of $\mathbb{R}^n$ to subsets of $\mathbb{R}^n$, by

$$\mathcal{G}_h(C) = \{x \in \mathbb{R}^n \mid F(N_1^C(x,h), N_2^C(x,h)) \geq 0\} \tag{2.5}$$

for $C \subset \mathbb{R}^n$ with F defined by (1.5), that is

$$F(N_1, N_2) = \frac{1}{\sqrt{h}} \frac{b_2 N_1 - b_1 N_2}{d} - g\left(\frac{1}{\sqrt{h}} \frac{a_2 N_1 - a_1 N_2}{d}\right).$$

The conditions (2.3d) and (1.6) make F an increasing function in both arguments. In order to prove that the mapping $\mathcal{G}_h$ produces a generalized mean curvature flow as $h \to 0$, we need to connect it to the PDE (2.1). For that purpose, given a function $\varphi \in C(\overline{\Omega})$ and a real number λ, we consider the super-level set $\{\varphi \geq \lambda\} \equiv \{x \in \mathbb{R}^n \mid \varphi(x) \geq \lambda\}$ and set

$$N_i(\lambda) \equiv N_i^{\{\varphi \geq \lambda\}}(x,h) = \int_\Omega \rho_i^{\sqrt{h}}(y-x)\chi_{\{\varphi \geq \lambda\}}(y)\, dy - \frac{1}{2}\int_\Omega \rho_i^{\sqrt{h}}(y-x)\, dy,$$
$$\widetilde{N}_i(\lambda) \equiv \widetilde{N}_i^{\{\varphi \geq \lambda\}}(x,h) = \int_{\mathbb{R}^n} \rho_i^{\sqrt{h}}(y-x)\chi_{\{\varphi \geq \lambda\}}(y)\, dy - \frac{1}{2}\int_{\mathbb{R}^n} \rho_i^{\sqrt{h}}(y-x)\, dy,$$

for i=1,2. This notation does not explicitly show the dependence on x, h and φ, but that will be clear from the context. Finally, we also define mappings $G_h, \widetilde{G}_h : C(\overline{\Omega}) \to C(\overline{\Omega})$, corresponding to $\mathcal{G}_h$, by

$$[G_h\varphi](x) = \sup\{\lambda \in \mathbb{R} \mid F(N_1(\lambda), N_2(\lambda)) \geq 0\} \tag{2.6}$$
$$[\widetilde{G}_h\varphi](x) = \sup\{\lambda \in \mathbb{R} \mid F(\widetilde{N}_1(\lambda), \widetilde{N}_2(\lambda)) \geq 0\} \tag{2.7}$$

for $h > 0$ and $\varphi \in C(\overline{\Omega})$. Note that all symbols with tilde ($\widetilde{G}$, $\widetilde{N}$ etc.) denote entities in the case of no boundary conditions, while the same symbols without the tilde denote the same entity in the domain Ω.

The main result, theorem 4.1, is now that the repeated application of the mappings G_h gives us an approximation of the solution to the level set equation (2.1). For clarity, we state the algorithm explicitly:

ALGORITHM 2.1.

Given Ω and g, choose functions ρ_1 and ρ_2 according to the assumptions above.
Choose an initial set C_0.
For each iteration k,
  Choose a time-step t_k,
  For each point $x \in \Omega$,
    Calculate $N_i^{C_{k-1}}(x, t_k)$, $i = 1, 2$.
    Evaluate the function $F(N_1^{C_{k-1}}(x,t_k), N_2^{C_{k-1}}(x,t_k))$.
    If $F \geq 0$, let x belong to C_k.
  End loop
End loop

For efficient implementation of this and other BMO-type schemes, see Ruuth [14] and Grzhibovskis, Heintz [9].

3. Properties of $\mathcal{G}_h$ and G_h. In this section, we prove some crucial properties of the operators $\mathcal{G}_h$ and G_h. We start with the inclusion principle for $\mathcal{G}_h$.

PROPOSITION 3.1. *Let $\mathcal{G}_h$ be defined by (2.5). Then for all $h > 0$ and all closed sets $C_1, C_2 \subset \overline{\Omega}$, we have*

$$C_1 \subset C_2 \Rightarrow \mathcal{G}_h(C_1) \subset \mathcal{G}_h(C_2).$$

Proof. Since the weight functions ρ_i are positive, $C_1 \subset C_2$ implies $N_i^{C_1}(x,h) \leq N_i^{C_2}(x,h)$, $i = 1,2$ for all $x \in \overline{\Omega}$ and $h > 0$. Since F is increasing in both arguments, we have $F(N_1^{C_1}, N_2^{C_1}) \leq F(N_1^{C_2}, N_2^{C_2})$ and therefore

$$\{F(N_1^{C_1}, N_2^{C_1}) \geq 0\} \subset \{F(N_1^{C_2}, N_2^{C_2}) \geq 0\}$$

and $\mathcal{G}_h(C_1) \subset \mathcal{G}_h(C_2)$. □

From this principle, some properties of G_h follow. Note that by the definitions of G_h and $\mathcal{G}_h$, the connection between the two is

$$[G_h\varphi](x) = \sup\{\lambda \in \mathbb{R} \mid x \in \mathcal{G}_h(\{\varphi \geq \lambda\})\}. \tag{3.1}$$

PROPOSITION 3.2. *For all $h > 0$ and $u, u_1, u_2 \in C(\overline{\Omega})$,*

(i) $G_h(u + C) = G_h u + C$, *for all* $C \in \mathbb{R}$,
(ii) $G_h(\theta \circ u) = \theta \circ (G_h u)$, *for any increasing, continuous function* $\theta : \mathbb{R} \to \mathbb{R}$,
(iii) *if* $u_1(x) \leq u_2(x)$ *for all* $x \in \overline{\Omega}$, *then* $[G_h u_1](x) \leq [G_h u_2](x)$ *for all* $x \in \overline{\Omega}$,
(iv) $\|G_h u_1 - G_h u_2\| \leq \|u_1 - u_2\|$ *in sup-norm*,
(v) $G_h u(x) = \inf\{\lambda \in \mathbb{R} \mid F(N_1(\lambda), N_2(\lambda)) \leq 0\}$.

Proof.

(i) This follows directly from the relation (3.1), since $\{u \geq \lambda\} = \{u + C \geq \lambda + C\}$.
(ii) This also follows from (3.1), since $\{u \geq \lambda\} = \{\theta \circ u \geq \theta(\lambda)\}$, and since θ is continuous, it can be moved outside the supremum.
(iii) In order to obtain a contradiction, suppose $u_1 \leq u_2$ in $\overline{\Omega}$ and that there is an $x_0 \in \overline{\Omega}$ such that $[G_h u_1](x_0) > [G_h u_2](x_0)$. Set $\lambda_1 = [G_h u_1](x_0)$ and $\lambda_2 = [G_h u_2](x_0)$, that is

$$\lambda_1 = \sup\{\lambda \mid x_0 \in \mathcal{G}_h(\{u_1 \geq \lambda\})\}, \qquad \lambda_2 = \sup\{\lambda \mid x_0 \in \mathcal{G}_h(\{u_2 \geq \lambda\})\}, \tag{3.2}$$

so that $\lambda_1 > \lambda_2$. Then, for every $\varepsilon \geq 0$, since also $u_1 \leq u_2$,

$$\{u_1 \geq \lambda_1 - \varepsilon\} \subset \{u_2 \geq \lambda_1 - \varepsilon\} \subset \{u_2 \geq \lambda_2 - \varepsilon\},$$

which by proposition 3.1 implies

$$\mathcal{G}_h(\{u_1 \geq \lambda_1 - \varepsilon\}) \subset \mathcal{G}_h(\{u_2 \geq \lambda_2 - \varepsilon\}).$$

This in turn implies that $\lambda_1 \leq \lambda_2$ by (3.2), which is a contradiction. Therefore we must have $[G_h u_1](x) \leq [G_h u_2](x)$ for all $x \in \overline{\Omega}$ and we are done.
(iv) Let $u_1, u_2 \in C(\overline{\Omega})$. It is enough to prove that

$$\|(G_h u_1 - G_h u_2)^{\pm}\|_\infty \leq \|(u_1 - u_2)^{\pm}\|_\infty,$$

with $(\cdot)^+ = \max\{0, \cdot\}$ and $(\cdot)^- = -\min\{0, \cdot\}$. The result then follows immediately. We prove only the plus-case, since the minus-case follows by the same argument.
So, aiming for a contradiction, we assume that there is an $x_0 \in \overline{\Omega}$, such that

$$[G_h u_1](x_0) - [G_h u_2](x_0) > \|(u_1 - u_2)^+\|_\infty \equiv C.$$

Then, by (i), we get

$$[G_h u_1](x_0) - [G_h(u_2 + C)](x_0) = [G_h u_1](x_0) - [G_h u_2](x_0) - C > 0$$

But since $u_1 \leq (u_2 + C)$ in $\overline{\Omega}$, this contradicts the result in (iii) with $u_1 = u_1$ and $u_2 = u_2 + C$.

(v) Fix $x \in \overline{\Omega}$, let r_i be the radius of the support of ρ_i for $i = 1, 2$ and assume $r_1 < r_2$. We are interested in values of λ for which $F(N_1(\lambda), N_2(\lambda)) = 0$. Since F is increasing in both variables and $F(0,0) = 0$, we must then have $N_1(\lambda) \geq 0$ and $N_2(\lambda) \leq 0$ or the other way around. Thus we must have

$$\{u \geq \lambda\} \cap B_n(x, r_i\sqrt{h}) \neq \varnothing \quad \text{and} \quad \{u \geq \lambda\} \cap B_n(x, r_i\sqrt{h}) \neq B_n(x, r_i\sqrt{h})$$

for $i = 1$ or 2 or both. The only other possibility would be $\{u \geq \lambda\} \cap B_n(x, r_1\sqrt{h}) = \varnothing$ and $\{u \geq \lambda\} \cap B_n(x, r_2\sqrt{h}) = B_n(x, r_2\sqrt{h})$, but this is impossible since $B_n(x, r_1\sqrt{h}) \subset B_n(x, r_2\sqrt{h})$.
But then, for any λ_1 and λ_2 near where $F = 0$, with $\lambda_1 < \lambda_2$, we have

$$\{u \geq \lambda_2\} \cap B_n(x, r_i\sqrt{h}) \subsetneq \{u \geq \lambda_1\} \cap B_n(x, r_i\sqrt{h})$$

and therefore $N_i(\lambda_2) < N_i(\lambda_1)$ with strict inequality for at least one of N_1 and N_2. We also have $N_i(\lambda_2) \leq N_i(\lambda_1)$ for the other one and therefore the function $\lambda \mapsto F(N_1(\lambda), N_2(\lambda))$ is strictly decreasing near the points where it is zero, and thus

$$\begin{aligned}[G_h u](x) &\equiv \sup\{\lambda \in \mathbb{R} \mid F(N_1(\lambda), N_2(\lambda)) \geq 0\} \\ &= \inf\{\lambda \in \mathbb{R} \mid F(N_1(\lambda), N_2(\lambda)) \leq 0\}.\end{aligned}$$

□

Finally, the following proposition, which is analogous to lemma 3.1 in [11], is a crucial part of the proof of our main theorem.

PROPOSITION 3.3. *For all $\varphi \in C^2(\overline{\Omega})$, $z \in \overline{\Omega}$ and $\varepsilon > 0$, there is a $\delta > 0$ such that*

(i) *If $z \in \Omega$ and $\nabla\varphi(z) \neq 0$, then*

$$[G_h\varphi](x) \leq \varphi(x) + h|\nabla\varphi(z)|\, g(\mathrm{curv}(\varphi(z))) + \varepsilon h, \quad x \in B_n(z, \delta), h \in (0, \delta) \quad (3.3)$$

and

$$[G_h\varphi](x) \geq \varphi(x) + h|\nabla\varphi(z)|\, g(\mathrm{curv}(\varphi(z))) - \varepsilon h, \quad x \in B_n(z, \delta), h \in (0, \delta) \quad (3.4)$$

(ii) *If $z \in \partial\Omega$ and $\partial\varphi/\partial\hat{n}(z) > 0$, then (3.3) holds for all $x \in B_n(z, \delta) \cap \overline{\Omega}$ and $h \in (0, \delta)$.*

(iii) *If* $z \in \partial\Omega$ *and* $\partial\varphi/\partial\hat{n}(z) < 0$, *then* (3.4) *holds for all* $x \in B_n(z,\delta) \cap \overline{\Omega}$ *and* $h \in (0,\delta)$.

The proof uses the same idea as in [11], namely to compare this case to the problem without boundary conditions in the whole of $\mathbb{R}^n$. The analogue of proposition 3.3 for that case can be found in a different shape in Grzhibovskis and Heintz [9, lemma 2] and is stated as lemma 3.4 below. The rest of the proof is mainly concerned with comparing $N_i(\lambda)$ and $\widetilde{N}_i(\lambda)$ (with and without boundary conditions). This requires the rather lengthy proofs of lemmas 3.5 and 3.6. Lemma 3.5 is proved exactly as in [11, lemma 3.1, case 1] and the proof is therefore not given here. The setup for that proof is identical to the one in the proof of lemma 3.6, which is given here and which we need because we have two convolution kernels instead of one.

LEMMA 3.4. (see [9, lemma 2]) *Let* $\varphi \in C^2(\mathbb{R}^n)$ *and* $z \in \mathbb{R}^n$. *If* $\nabla\varphi(z) \neq 0$, *then for each* $\varepsilon > 0$, *there is a* $\delta > 0$ *such that*

$$[\widetilde{G}_h\varphi](x) \leq \varphi(x) + h|\nabla\varphi(z)|\ g(\mathrm{curv}(\varphi(z))) + \varepsilon h,$$
$$[\widetilde{G}_h\varphi](x) \geq \varphi(x) + h|\nabla\varphi(z)|\ g(\mathrm{curv}(\varphi(z))) - \varepsilon h$$

for all $x \in B_n(z,\delta)$ *and* $h \in (0,\delta)$.

We are now ready to prove proposition 3.3.

Proof. (of proposition 3.3) As in [11], the main idea behind the proof is to compare $[G_h\varphi](x)$ to $[\widetilde{G}_h\varphi](x)$ and use lemma 3.4 to get the desired result.

Let $\varphi \in C^2(\overline{\Omega})$ and $\varepsilon > 0$. If $z \in \Omega$ and $D\varphi(z) \neq 0$, then since $\operatorname{supp}\rho_i$ is compact, there is a $\delta > 0$, such that for all $x \in B_n(z,\delta) \subset \Omega$ and $h \in (0,\delta)$, we have

$$[G_h\varphi](x) = [\widetilde{G}_h\varphi](x).$$

So proposition 3.3, part (i), follows directly from lemma 3.4. To prove the rest of the proposition, we assume $z \in \partial\Omega$ and $\partial\varphi/\partial\hat{n}(z) > 0$ and prove part (ii), noting that part (iii) may be proved similarly.

Since $\varphi \in C^2(\overline{\Omega})$ and $\partial\Omega$ is C^2, we may extend φ so that $\varphi \in C^2(B_n(z,r_0))$ for some $r_0 > 0$. Using proposition 3.2 (v), we also set

$$\lambda \equiv [G_h\varphi](x) = \inf\{\mu \in \mathbb{R} \mid F(N_1(\mu), N_2(\mu)) \leq 0\}$$

and

$$\widetilde{\lambda} \equiv [\widetilde{G}_h\varphi](x) = \inf\{\mu \in \mathbb{R} \mid F(\widetilde{N}_1(\mu), \widetilde{N}_2(\mu)) \leq 0\}.$$

Now, we wish to show that there is an $h_0 > 0$ such that $\lambda \leq \widetilde{\lambda}$ for each $x \in B_n(z,r_0) \cap \overline{\Omega}$ and each $h \in (0,h_0)$. From the definition of $\widetilde{\lambda}$, we see that $F(\widetilde{N}_1(\widetilde{\lambda}), \widetilde{N}_2(\widetilde{\lambda})) = 0$ and by the definition (1.5) of $F(N_1,N_2)$, we know $F(0,0) = 0$. Therefore, if we can show that $N_i(\widetilde{\lambda}) \leq 0$, $i = 1,2$, we will immediately have $F(N_1(\widetilde{\lambda}), N_2(\widetilde{\lambda})) \leq 0$ (since F is increasing in both variables) and thus

$$\lambda = \inf\{\mu \in \mathbb{R} \mid F(N_1(\mu), N_2(\mu)) \leq 0\} \leq \widetilde{\lambda}. \tag{3.5}$$

The same result will be obtained if we can show that $N_i(\widetilde{\lambda}) \leq \widetilde{N}_i(\widetilde{\lambda})$, $i = 1,2$.

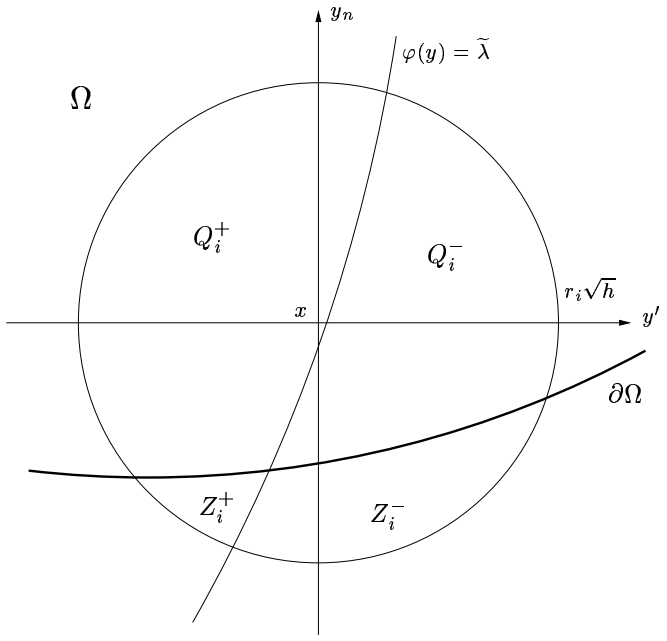

FIG. 3.1. *The sets $Q_i^{\pm}$ and $Z_i^{\pm}$ in case 1.*

To prove that $N_i(\widetilde{\lambda}) \leq 0$ or that $N_i(\widetilde{\lambda}) \leq \widetilde{N}_i(\widetilde{\lambda})$, we first define

$$Q_i^+ \equiv \{y \in B_n(x, r_i\sqrt{h}) \mid \varphi(y) < \widetilde{\lambda}\} \cap \Omega,$$
$$Q_i^- \equiv \{y \in B_n(x, r_i\sqrt{h}) \mid \varphi(y) > \widetilde{\lambda}\} \cap \Omega,$$

for $i = 1, 2$, with $r_i = \inf\{r \in \mathbb{R} \mid \operatorname{supp}\rho_i \subset B_n(0, r)\}$. See figure 3.1. We can then express $N_i(\widetilde{\lambda})$ as

$$\begin{aligned} N_i(\widetilde{\lambda}) &= \int_\Omega \rho_i^{\sqrt{h}}(y-x)\chi_{\{\varphi \geq \widetilde{\lambda}\}}(y)\, dy - \frac{1}{2}\int_\Omega \rho_i^{\sqrt{h}}(y-x)\, dy \\ &= \int_{Q_i^-} \rho_i^{\sqrt{h}}(y-x)\, dy - \frac{1}{2}\int_{Q_i^+ \cup Q_i^-} \rho_i^{\sqrt{h}}(y-x)\, dy \\ &= \frac{1}{2}\left(\int_{Q_i^-} \rho_i^{\sqrt{h}}(y-x)\, dy - \int_{Q_i^+} \rho_i^{\sqrt{h}}(y-x)\, dy\right). \end{aligned} \tag{3.6}$$

We now consider one ρ_i at a time (i.e. $i = 1$ or 2) and divide the treatment into three cases (see figures 3.1 and 3.2).

Case 1: $\{y \in B_n(x, r_i\sqrt{h}) \mid \varphi(y) = \widetilde{\lambda}, y \in \partial\Omega\} \neq \varnothing$
In this case, the computation (3.6) and lemma 3.5 below show that $N_i(\widetilde{\lambda}) \leq 0$ and lemma 3.6 gives $N_i(\widetilde{\lambda}) \leq \widetilde{N}_i(\widetilde{\lambda})$.

Case 2: $\{y \in B_n(x, r_i\sqrt{h}) \mid \varphi(y) = \widetilde{\lambda}\} \subset \Omega$
Here, the intersection between the hypersurfaces $\{\varphi = \widetilde{\lambda}\}$ and $\partial\Omega$ lies outside

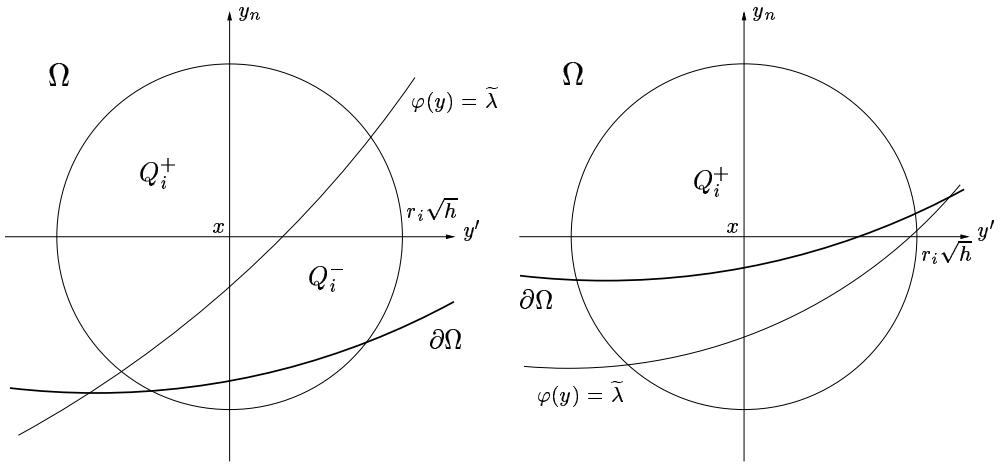

FIG. 3.2. *Case 2 on the left and case 3 on the right.*

the ball $B_n(x, r_i\sqrt{h})$. We then see that $Q_i^+ \cap \Omega^c = \varnothing$ and therefore, continuing from (3.6), we get

$$N_i(\widetilde{\lambda}) = \frac{1}{2}\left(\int_{Q_i^-} \rho_i^{\sqrt{h}}(y-x)\,dy - \int_{Q_i^+} \rho_i^{\sqrt{h}}(y-x)\,dy\right)$$
$$\leq \frac{1}{2}\left(\int_{\{\varphi \geq \widetilde{\lambda}\}} \rho_i^{\sqrt{h}}(y-x)\,dy - \int_{\{\varphi < \widetilde{\lambda}\}} \rho_i^{\sqrt{h}}(y-x)\,dy\right) = \widetilde{N}_i(\widetilde{\lambda}).$$

Case 3: $\{y \in B_n(x, r_i\sqrt{h}) \mid \varphi(y) = \widetilde{\lambda}\} \subset \Omega^c$
In this case, $Q_i^- = \varnothing$ and since ρ_i is positive, we can conclude from (3.6) that $N_i(\widetilde{\lambda}) \leq 0$.

Now, we need to see what happens when we have two convolution kernels. Set $r_i = \inf\{r \in \mathbb{R} \mid \operatorname{supp}\rho_i \subset B_n(0,r)\}$, $i = 1,2$ and assume $r_1 < r_2$. Then, if case 1 applies to ρ_1, it also applies to ρ_2, so that $N_i(\widetilde{\lambda}) \leq 0$ (and $N_i(\widetilde{\lambda}) \leq \widetilde{N}_i(\widetilde{\lambda})$) for $i = 1,2$ and therefore $\lambda \leq \widetilde{\lambda}$ by the argument leading to (3.5). If case 3 applies to ρ_1, then either case 1 or 3 applies to ρ_2, but in both cases we have again $N_i(\widetilde{\lambda}) \leq 0$ for $i = 1,2$ and $\lambda \leq \widetilde{\lambda}$. Finally, if case 2 applies to ρ_1, then either case 1 or 2 applies to ρ_2. In both cases, we have $N_i(\widetilde{\lambda}) \leq \widetilde{N}_i(\widetilde{\lambda})$ for $i = 1,2$ and thus $\lambda \leq \widetilde{\lambda}$. Therefore, once we prove lemmas 3.5 and 3.6, the proof is completed. □

All that now remains is to prove the next two lemmas. As mentioned earlier, lemma 3.5 is proved by Ishii and Ishii [11] for regular mean curvature flow and there are only very minor differences in our case, so we omit the proof. Besides, the same setup is used in the proof of lemma 3.6, which we give in detail and which is needed in our case in view of the discussion following the three cases above, because we have two convolution kernels.

LEMMA 3.5. (cf. [11, lemma 3.1, case 1]) *Let r_i, Q_i^+ and Q_i^- be defined as above, and assume that $z \in \partial\Omega$ and $\partial\varphi/\partial\hat{n}(z) > 0$. Then there is an $r_0 > 0$ and an $h_0 > 0$ such that for all $x \in B_n(z, r_0)$ and $h \in (0, h_0)$, it holds that if*

$\{y \in B_n(x, r_i\sqrt{h}) \mid \varphi(y) = \widetilde{\lambda}, y \in \partial\Omega\} \neq \varnothing$ *(i.e. case 1 applies) then*

$$\int_{Q_i^+} \rho_i^{\sqrt{h}}(y-x)\, dy \geq \int_{Q_i^-} \rho_i^{\sqrt{h}}(y-x)\, dy, \tag{3.7}$$

and thus $N_i(\widetilde{\lambda}) \leq 0$.

LEMMA 3.6. *Let* r_i *be defined as above, and assume* $z \in \partial\Omega$ *and* $\partial\varphi/\partial\hat{n}(z) > 0$. *Then there is an* $r_0 > 0$ *and an* $h_0 > 0$ *such that for all* $x \in B_n(z, r_0)$ *and* $h \in (0, h_0)$ *it holds that if* $\{y \in B_n(x, r_i\sqrt{h}) \mid \varphi(y) = \widetilde{\lambda}, y \in \partial\Omega\} \neq \varnothing$ *(i.e. case 1 applies), then*

$$N_i(\widetilde{\lambda}) \leq \widetilde{N}_i(\widetilde{\lambda}).$$

Proof. We wish to prove that $N_i(\widetilde{\lambda}) \leq \widetilde{N}_i(\widetilde{\lambda})$, which is the same as

$$\int_{Q_i^-} \rho_i^{\sqrt{h}}(y-x)\, dy - \int_{Q_i^+} \rho_i^{\sqrt{h}}(y-x)\, dy \leq \int_{\widetilde{Q}_i^-} \rho_i^{\sqrt{h}}(y-x)\, dy - \int_{\widetilde{Q}_i^+} \rho_i^{\sqrt{h}}(y-x)\, dy,$$

with $Q_i^{\pm}$ defined above and

$$\widetilde{Q}_i^+ = \{y \in B_n(x, r_i\sqrt{h}) \mid \varphi(y) < \widetilde{\lambda}\}$$
$$\widetilde{Q}_i^- = \{y \in B_n(x, r_i\sqrt{h}) \mid \varphi(y) > \widetilde{\lambda}\},$$

so that $Q_i^{\pm} = \widetilde{Q}_i^{\pm} \cap \Omega$. Rewriting this expression once more, we get

$$\int_{Z_i^-} \rho_i^{\sqrt{h}}(y-x)\, dy - \int_{Z_i^+} \rho_i^{\sqrt{h}}(y-x)\, dy \geq 0 \tag{3.8}$$

with

$$Z_i^+ = \{y \in B_n(x, r_i\sqrt{h}) \cap \Omega^c \mid \varphi(y) < \widetilde{\lambda}\},$$
$$Z_i^- = \{y \in B_n(x, r_i\sqrt{h}) \cap \Omega^c \mid \varphi(y) > \widetilde{\lambda}\}$$

as is illustrated in figure 3.1.

Given $z \in \partial\Omega$ with $\partial\varphi/\partial\hat{n}(z) > 0$, there is an $r_0 > 0$ such that φ can be extended to be in $C^2(B_n(z, r_0))$ (since $\partial\Omega$ is C^2 and $\varphi \in C^2(\overline{\Omega})$). The idea of the proof is now that for any $x \in B_n(z, r_0) \cap \overline{\Omega}$, we may approximate the hypersurfaces $\{\varphi = \widetilde{\lambda}(x)\}$ and $\partial\Omega$ by hyperplanes and show that for small h, since $\partial\varphi/\partial\hat{n} > 0$, the contribution from the set denoted W in figure 3.3 is greater than the contributions from the sets enclosed by dashed lines, within which the two hypersurfaces must lie. But first we need to introduce some notation.

We take a parameterization $\psi \in C^2(\mathbb{R}^{n-1})$ of $\partial\Omega$ and choose coordinate system so that

$$y_n - z_n = \psi(y' - z') \text{ for all } y = (y', y_n) \in B_n(z, r_0) \cap \partial\Omega,$$
$$\nabla'\psi(0) = 0,$$
$$y_n - z_n > \psi(y' - z') \text{ for all } y \in B_n(z, r_0) \cap \Omega,$$

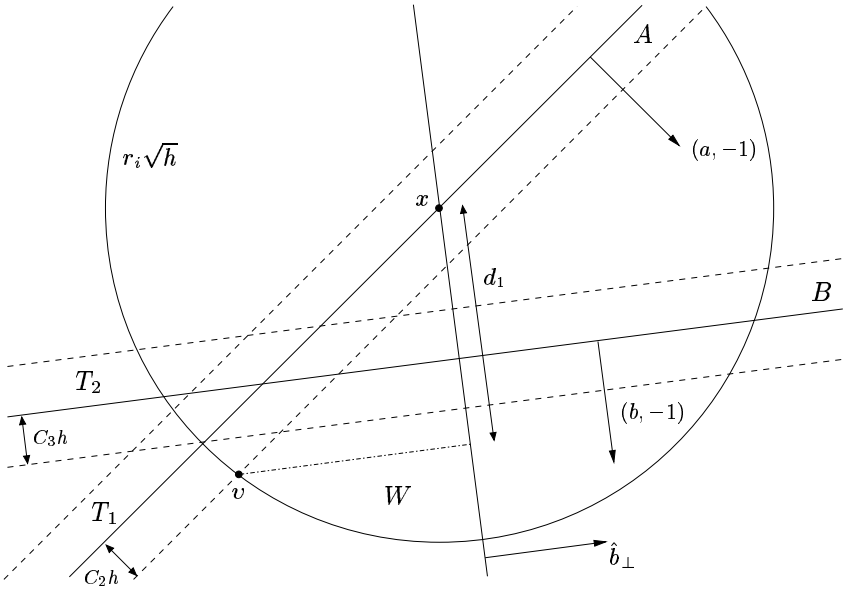

FIG. 3.3. *The plane spanned by* $(a,-1)$ *and* $(b,-1)$*, showing the hyperplanes* A *and* B *as well as the position of* v *and the distance* d_1.

where $\nabla' = (\partial/\partial x_1, \dots, \partial/\partial x_{n-1})$. Taking r_0 smaller if necessary, we may assume

$$\nabla_n\varphi(y) \le -\gamma, \quad |\nabla'\varphi(y)| \le K, \quad |\nabla'\psi(y'-z')| \le \varepsilon \text{ for all } y \in B_n(z,r_0),$$

where $\nabla_n = \partial/\partial x_n$ and $\gamma, K > 0$ are independent of $\varepsilon > 0$ and r_0.

Now fix $x \in B_n(z,r_0) \cap \overline{\Omega}$ and a small $h_0 > 0$. Let $h \in (0,h_0)$ and choose a point $\xi \in \partial\Omega \cap \{\varphi = \widetilde{\lambda}\} \cap B_n(x, r_i\sqrt{h})$. We then set

$$a(x) = \frac{\nabla'\varphi(x)}{\nabla_n\varphi(x)}, \qquad b(\xi') = \nabla'\psi(\xi')$$

and define the hyperplanes A and B by

$$A = \{y \in \mathbb{R}^n \mid \langle (a,-1), y-x\rangle = 0\},$$
$$B = \{y \in \mathbb{R}^n \mid \langle (b,-1), y-\xi\rangle = 0\} = \{y \in \mathbb{R}^n \mid \langle (b,-1), y-x\rangle = c\},$$

with $c = \langle (b,-1), \xi - x\rangle$, see figure 3.3. We also note that $|a| \le K/\gamma$ and $|b| \le \varepsilon$, so that

$$|\langle (a,-1),(b,-1)\rangle| = |\langle a,b\rangle + 1| \ge 1 - \frac{\varepsilon K}{\gamma} \ge \frac{1}{2} \tag{3.9}$$

if $\varepsilon \in (0,\varepsilon_0)$ with $\varepsilon_0 = \gamma/(2K)$, which in particular means that A and B are not perpendicular.

We now need three elementary lemmas with counterparts in [11].

LEMMA 3.7. (cf. Ishii and Ishii [11, lemma 3.4]) *There exists a* $\delta > 0$ *and a* $C_1 > 0$ *independent of* $x \in B_n(z,\delta)$ *such that if* $h \in (0,\delta)$ *and* $x \in B_n(z,\delta)$*, then* $|\widetilde{\lambda} - \varphi(x)| \le C_1 h$.

Proof. This follows directly from lemma 3.4, using the fact that $\nabla\varphi(z) \neq 0$. □

LEMMA 3.8. *There exists a $C_2 > 0$, which is independent of $x \in B_n(z, r_0)$, and an $h_1 > 0$, such that if $y \in B_n(x, r_i\sqrt{h})$ satisfies $\varphi(y) = \widetilde{\lambda}$, with $h \in (0, h_1)$, and if $a(x) = \nabla'\varphi(x)/\nabla_n\varphi(x)$, then $|y_n - x_n - \langle a(x), y' - x'\rangle| \leq C_2 h$.*

Proof. This follows easily from lemma 3.7, exactly as in [11, lemma 3.6]. □

LEMMA 3.9. ([11, lemma 3.3(1)]) *For any $r > 0$, $a, b \in \mathbb{R}^{n-1}$ and $c \in \mathbb{R}^n$, with $|a| \leq K_1$, $|b| \leq \varepsilon_1$, and $\varepsilon_1 K_1 < 1$, there is a $\theta \in (0,1)$, depending only on ε_1 and K_1, such that if*

$$\{y \in B_n(0,r) \mid y_n = \langle b, y'\rangle + c,\ y_n = \langle a, y'\rangle\} \neq \varnothing,$$

then

$$\frac{|c|}{\sqrt{1+|b|^2}} \leq \theta r.$$

(Note that the quantity on the left is the orthogonal distance from the hyperplane $\langle (b,-1), y\rangle = c$ to the origin).

Proof. This is part 1 of lemma 3.3 in [11]. The proof is given there and is a straightforward geometrical argument. □

Now, by lemma 3.8, we have

$$|\langle (a,-1), y - x\rangle| \leq C_2 h, \qquad \text{if } \varphi(y) = \widetilde{\lambda},\ y \in B_n(x, r_i\sqrt{h}), \tag{3.10}$$

and since $\psi \in C^2(\mathbb{R}^{n-1})$, we also have

$$|\langle (b,-1), y - x\rangle - c| \leq C_3 h, \qquad \text{for all } y \in \partial\Omega \cap B_n(x, r_i\sqrt{h}), \tag{3.11}$$

for some $C_3 > 0$, independent of $x \in B_n(z, r_0)$ and $h \in (0, h_0)$.

We now restrict our view to the plane spanned by $(a,-1)$ and $(b,-1)$, which is obviously perpendicular to both A and B. This is the view shown in figure 3.3.

We set

$$v = x + \beta_1(a,-1) + \beta_2(b,-1),$$

with

$$\beta_1 = \frac{1}{(1+|a|^2)^{1/2}}\left(C_2 h - \frac{1+\langle a,b\rangle}{D^{1/2}}\sqrt{r_i^2 h - C_2^2 h^2}\right),$$
$$\beta_2 = \frac{(1+|a|^2)^{1/2}}{D^{1/2}}\sqrt{r_i^2 h - C_2^2 h^2}$$

defined so that v is located as in figure 3.3, that is $\langle (a,-1), v - x\rangle/(1+|a|^2)^{1/2} = C_2 h$ and $|v - x| = r_i\sqrt{h}$. Here, $D = (1+|a|^2)(1+|b|^2) - (1+\langle a,b\rangle)^2$. Expanding this expression, it is easily seen that $D \geq |a-b|^2$, so that it is zero only when $a = b$. Assume for the moment that $a \neq b$, so that $D \neq 0$ and v is well-defined.

Now,

$$\begin{aligned} d_1 &= \frac{\langle v - x, (b,-1)\rangle}{(1+|b|^2)^{1/2}} \\ &= \frac{1+\langle a,b\rangle}{(1+|a|^2)^{1/2}(1+|b|^2)^{1/2}}\left(C_2h - \frac{1+\langle a,b\rangle}{D^{1/2}}\sqrt{r_i^2h - C_2^2h^2}\right) \\ &\quad + \frac{(1+|a|^2)^{1/2}(1+|b|^2)^{1/2}}{D^{1/2}}\sqrt{r_i^2h - C_2^2h^2} \\ &= \left(\frac{D}{(1+|a|^2)(1+|b|^2)}\right)^{1/2}\sqrt{r_i^2h - C_2^2h^2} + \frac{1+\langle a,b\rangle}{(1+|a|^2)^{1/2}(1+|b|^2)^{1/2}}C_2h. \end{aligned}$$

Investigating the first term, we see first that $0 < \sqrt{r_i^2h - C_2^2h^2} \le r_i\sqrt{h}$ if $h \in (0,h_0)$ with $h_0 < r_i^2/C_2^2$. Using (3.9), we also see that

$$\begin{aligned} \frac{D}{(1+|a|^2)(1+|b|^2)} &= \frac{(1+|a|^2)(1+|b|^2) - (1+\langle a,b\rangle)^2}{(1+|a|^2)(1+|b|^2)} = 1 - \frac{(1+\langle a,b\rangle)^2}{(1+|a|^2)(1+|b|^2)} \\ &\le 1 - \frac{1}{4(1+(K/\gamma)^2)(1+\varepsilon_0^2)} \equiv \theta_1^2 < 1. \end{aligned}$$

For the second term, we may now use (3.9) to get the estimate

$$\frac{1+\langle a,b\rangle}{(1+|a|^2)^{1/2}(1+|b|^2)^{1/2}}C_2h \le \frac{3}{2}C_2h$$

Thus,

$$d_1 \le \theta_1 r_i\sqrt{h} + \frac{3}{2}C_2h \le \frac{\theta_1+1}{2}r_i\sqrt{h} < r_i\sqrt{h} \tag{3.12}$$

if $h \in (0,h_0)$ with $h_0 \le (r_i(1-\theta_1)/(3C_2))^2$.

Also, by lemma 3.9, there is a $\theta_2 \in (0,1)$ such that $|c|/\sqrt{1+|b|^2} \le \theta_2 r_i\sqrt{h}$. We note that

$$\theta_2 r_i\sqrt{h} + C_3h \le \frac{\theta_2+1}{2}r_i\sqrt{h} < r_i\sqrt{h} \tag{3.13}$$

if $h \in (0,h_0)$ with $h_0 \le (r_i(1-\theta_2)/(2C_3))^2$. We therefore set $\theta = \max\{\theta_1+1, \theta_2+1\}/2 < 1$ and $h_0 \le \min\{(r_i(1-\theta_1)/(3C_2))^2, (r_i(1-\theta_2)/(2C_3))^2\}$.

We are now ready to estimate the integrals over Z^+ and Z^- in (3.8). Letting $\hat{b}_\perp$ denote the unit vector in the $(a,-1)$,$(b,-1)$-plane with $\langle (b,-1), \hat{b}_\perp\rangle = 0$ and $\langle (a,-1), \hat{b}_\perp\rangle > 0$, we define

$$\begin{aligned} W &= \{y \in B_n(x, r_i\sqrt{h}) \mid \langle (b,-1), y-x\rangle/\sqrt{1+|b|^2} \ge \theta r_i\sqrt{h},\ \langle \hat{b}_\perp, y-x\rangle < 0\} \\ T_1 &= \{y \in B_n(x, r_i\sqrt{h}) \mid |\langle (a,-1), y-x\rangle| \le C_2h\} \\ T_2 &= \{y \in B_n(x, r_i\sqrt{h}) \mid |\langle (b,-1), y-x\rangle - c| \le C_3h\} \end{aligned}$$

(see figure 3.3) and note that

$$\begin{aligned} Z^- &\supset (\{y \in B_n(x, r_i\sqrt{h}) \mid \langle (b,-1), y-x\rangle > c + C_3h,\ \langle \hat{b}_\perp, y-x\rangle > 0\} \setminus T_1) \cup W, \\ Z^+ &\subset (\{y \in B_n(x, r_i\sqrt{h}) \mid \langle (b,-1), y-x\rangle > c - C_3h,\ \langle \hat{b}_\perp, y-x\rangle < 0\} \cup T_1) \setminus W. \end{aligned}$$

The definition of θ ensures that $W \neq \varnothing$. From this, it is clear that

$$\begin{aligned}\int_{Z^-} \rho_i^{\sqrt{h}}(y-x)\,dy - \int_{Z^+} \rho_i^{\sqrt{h}}(y-x)\,dy \geq \\ 2\int_W \rho_i^{\sqrt{h}}(y-x)\,dy - \left(\int_{T_1} + \int_{T_2}\right)\rho_i^{\sqrt{h}}(y-x)\,dy\end{aligned} \tag{3.14}$$

Now, we see from (3.12) and (3.13) that

$$\mathcal{L}^n(W) = \alpha h^{n/2},$$

for some $\alpha > 0$ depending only on θ and r_i, where $\mathcal{L}^n$ is the n-dimensional Lebesgue measure. From this we conclude that

$$\int_W \rho_i^{\sqrt{h}}(y-x)\,dy \geq C_4,$$

for some $C_4 > 0$, independent of $x \in B_n(z, r_0)$, $\varepsilon \in (0, \varepsilon_0)$ and $h \in (0, h_0)$.

Furthermore, it is clear that $\mathcal{L}^n(T_1) \leq C_5 h^{(n+1)/2}$ for some $C_5 > 0$ depending only on C_2, r_i and n, and that $\mathcal{L}^n(T_2) \leq C_6 h^{(n+1)/2}$ for some C_6 depending only on C_3, r_i and n. Thus, by changing variables $(y-x)/\sqrt{h} \to \hat{y}$, we see

$$\begin{gathered}\left(\int_{T_1} + \int_{T_2}\right)\rho_i^{\sqrt{h}}(y-x)\,dy = \left(\int_{T_1(h,x)} + \int_{T_2(h,x)}\right)\rho_i(\hat{y})\,d\hat{y} \\ \mathcal{L}^n(T_1(h,x)) \leq C_5\sqrt{h}, \qquad \mathcal{L}^n(T_2(h,x)) \leq C_6\sqrt{h} \\ T_j(h,x) = \{(y-x)/\sqrt{h} \mid y \in T_j\},\ j = 1,2.\end{gathered}$$

Therefore, we conclude that there is an $h_0 > 0$ such that if $h \in (0, h_0)$, then

$$2\int_W \rho_i^{\sqrt{h}}(y-x)\,dy - \left(\int_{T_1} + \int_{T_2}\right)\rho_i^{\sqrt{h}}(y-x)\,dy \geq 0,$$

for all $x \in B_n(z, r_0)$, which proves $N_i(\widetilde{\lambda}) \leq \widetilde{N}_i(\widetilde{\lambda})$ by (3.14) and the discussion leading to (3.8).

Finally, we need to cover the case $a = b$. In this case, we define

$$W = \{y \in B_n(x, r_i\sqrt{h}) \mid \langle (b,-1), y-x\rangle > \max\{c + C_3 h, C_2 h\}\},$$

with $W \neq \varnothing$ if h is small enough. Then $Z^- \supset W \setminus T_1$ and $Z^+ \subset T_1 \cup T_2$, and the same argument holds. □

4. The convergence theorem. In this section, we prove the convergence of the output of algorithm 2.1 to the viscosity solution of the level set PDE (2.1) as the time step h tends to zero. The proof is based on the proofs by Ishii and Ishii [11] and Ishii [10].

We begin by defining the approximations u^m as follows. Given a function $f \in C(\overline{\Omega})$, let $u^m \in C(\overline{\Omega} \times [0,T))$ be defined for $m \in \mathbb{Z}_+$ by

$$u^m(x,t) = [G_{t-lh} \circ (G_h)^l f](x), \tag{4.1}$$

where $h = T/m$ and $l \in \mathbb{N}$ is chosen so that $lh \le t < (l+1)h$. The main convergence theorem is then the following.

THEOREM 4.1. *Choose $f \in C(\overline{\Omega})$ and let $\{u^m\}_{m=1}^{\infty}$ be defined by (4.1). Then $u^m \to u$ locally uniformly on $\overline{\Omega} \times [0,T)$ as $m \to \infty$, where u is the unique viscosity solution of the PDE*

$$\begin{cases} \dfrac{\partial u}{\partial t}(x,t) - |\nabla u(x,t)| g(\operatorname{curv}(u(x,t))) = 0 & x \in \Omega,\ t \in (0,T), \\ \dfrac{\partial u}{\partial \hat{n}}(x,t) = 0 & x \in \partial\Omega,\ t \in (0,T), \\ u(x,0) = f(x) & x \in \overline{\Omega}, \end{cases}$$

which exists by theorem 3.12 in Sato [15].

The idea of the proof is to define $\overline{u}(x,t)$ and $\underline{u}(x,t)$ by

$$\begin{aligned} \overline{u}(x,t) &= \lim_{\varepsilon \to 0} \sup\{u^m(y,s) \mid m > \varepsilon^{-1}, (y,s) \in \overline{\Omega} \times [0,T), |x-y| + |s-t| < \varepsilon\}, \\ \underline{u}(x,t) &= \lim_{\varepsilon \to 0} \inf\{u^m(y,s) \mid m > \varepsilon^{-1}, (y,s) \in \overline{\Omega} \times [0,T), |x-y| + |s-t| < \varepsilon\}, \end{aligned} \tag{4.2}$$

and prove that these are sub- and supersolution respectively of the level set PDE. It then follows from the comparison result by Sato [15] that if $\overline{u}(x,0) \le \underline{u}(x,0)$, then $\overline{u}(x,t) \le \underline{u}(x,t)$ for $(x,t) \in \overline{\Omega} \times [0,T)$ and thus (since $\overline{u} \ge \underline{u}$ by definition) that $u = \overline{u} = \underline{u}$ is a solution.

First, we state the definition of a viscosity solution to the PDE (2.1) (cf. [15] and [4]).

DEFINITION 4.2. *A function $u \in C(\overline{\Omega} \times [0,T))$ is a viscosity* subsolution *of* (2.1) *if for any $\varphi \in C^2(\overline{\Omega} \times [0,T))$ such that $u - \varphi$ has a maximum at $(x_0,t_0) \in \overline{\Omega} \times [0,T)$, then*

$$\varphi_t'(x_0,t_0) - |\nabla\varphi(x_0,t_0)| g\big(\operatorname{curv}(\varphi(x_0,t_0))\big) \le 0$$

if $x_0 \in \Omega$ and $\nabla\varphi(x_0,t_0) \ne 0$ or $x_0 \in \partial\Omega$ and $\partial\varphi/\partial\hat{n}(x_0,t_0) > 0$, and

$$\varphi_t'(x_0,t_0) \le 0$$

if $x_0 \in \Omega$, $\nabla\varphi(x_0,t_0) = 0$ and $D^2\varphi(x_0,t_0) = 0$.

Analogously, $u \in C(\overline{\Omega} \times [0,T))$ is a viscosity supersolution *of* (2.1) *if for any $\varphi \in C^2(\overline{\Omega} \times [0,T))$ such that $u - \varphi$ has a minimum at $(x_0,t_0) \in \overline{\Omega} \times [0,T)$, then*

$$\varphi_t'(x_0,t_0) - |\nabla\varphi(x_0,t_0)| g\big(\operatorname{curv}(\varphi(x_0,t_0))\big) \ge 0$$

if $x_0 \in \Omega$ and $\nabla\varphi(x_0,t_0) \ne 0$ or $x_0 \in \partial\Omega$ and $\partial\varphi/\partial\hat{n}(x_0,t_0) < 0$, and

$$\varphi_t'(x_0,t_0) \ge 0$$

if $x_0 \in \Omega$, $\nabla\varphi(x_0,t_0) = 0$ and $D^2\varphi(x_0,t_0) = 0$.

The function u is a viscosity solution *if it is both a sub- and a supersolution.*

We remark that there is no condition in the case $\nabla\varphi(x_0,t_0) = 0$, $D^2\varphi(x_0,t_0) \ne 0$, since it follows from the other cases by the argument given by Barles and Georgelin [1,

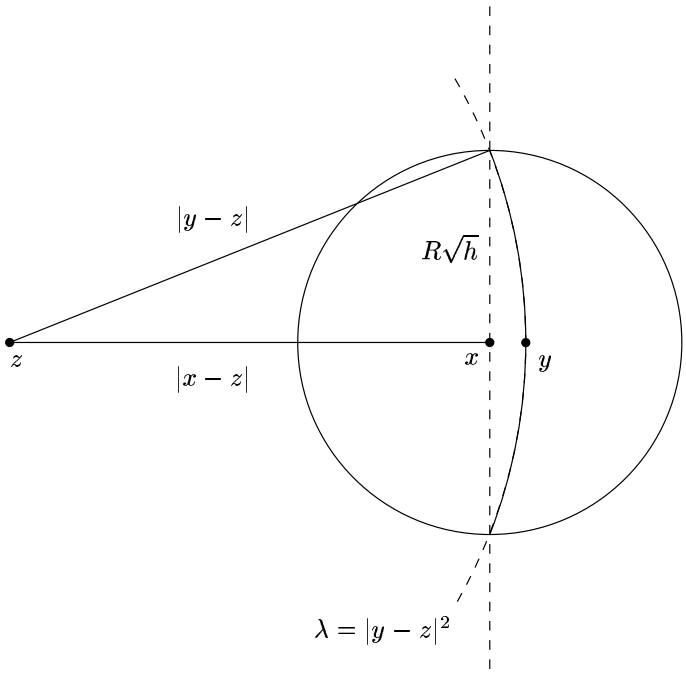

FIG. 4.1. *The setting in the proof of lemma 4.3.*

proposition 2.2], using the hypothesis (2.3b) that g must not grow faster than linearly towards infinity.

Before we begin with the actual proof of the convergence theorem, we need a few lemmas.

LEMMA 4.3. *For each $z \in \Omega$, there is a $C > 0$, a $\delta > 0$ and an $h_0 > 0$ such that*

$$[G_h(|\cdot - z|^2)](x) \leq |x-z|^2 + Ch,$$
$$[G_h(-|\cdot - z|^2)](x) \geq -|x-z|^2 - Ch$$

for all $x \in B_n(z,\delta)$ and $h \in (0, h_0)$.

Proof. Fix $z \in \Omega$. Since the convolution kernels ρ_i have compact support, there is an $R > 0$, such that $\operatorname{supp} \rho_i^{\sqrt{h}}(\cdot - x) \subset B_n(x, R\sqrt{h})$, $i = 1, 2$, for any $x \in \overline{\Omega}$ and any $h > 0$. Furthermore, there is a $\delta > 0$ and an $h_0 > 0$ such that if $h \in (0, h_0)$, then $B_n(x, R\sqrt{h}) \subset \Omega$ for any $x \in B_n(z, \delta)$.

We fix such R, h_0, δ and an $x \in B_n(z,\delta)$. In order to estimate $[G_h(|\cdot - z|^2)](x)$, we need to see which level sets of the function $\xi \mapsto |\xi - z|^2$ that may reach x. Therefore, we investigate the situation in figure 4.1, where y is on the level set $\{\xi \in \Omega \mid |\xi - z|^2 = \lambda\}$, which is chosen so that it intersects $\partial B_n(x, R\sqrt{h})$ on the hyperplane $\{\xi \in \mathbb{R}^n \mid \langle \xi - x, x - z\rangle = 0\}$.

Since $F(N_1, N_2)$ is increasing in both variables and $F(0,0) = 0$, we need to have either $N_1 \geq 0$ or $N_2 \geq 0$ in order to get $F(N_1, N_2) \geq 0$. And since in the situation in figure 4.1, no part of the super-level set $\{\xi \in \Omega \mid |\xi - z|^2 \geq \lambda\}$ is in the left half of

$B_n(x, R\sqrt{h})$, obviously $N_i(x,h) < 0$, $i = 1,2$ and thus $F(N_1, N_2) < 0$ so that

$$x \notin \mathcal{G}_h(\{\xi \in \Omega \mid |\xi - z|^2 \geq \lambda = |y - z|^2\}).$$

Therefore,

$$[G_h(|\cdot - z|^2)](x) - |x - z|^2 \leq |y - z|^2 - |x - z|^2 = (R\sqrt{h})^2 = R^2 h$$

by the Pythagorean theorem, which gives the desired result with $C = R^2$.

A similar argument gives the result for $-|x - z|^2$. □

LEMMA 4.4. (cf. [11, lemma 4.1]) *Let $f \in C^2(\overline{\Omega})$ in (4.1). If $\partial f/\partial \hat{n} > 0$ on $\partial\Omega$, then there is a constant $C > 0$ and an $M \in \mathbb{Z}_+$, such that*

$$\sup_{x \in \overline{\Omega}, m \geq M} (u^m(x,t) - f(x)) \leq Ct$$

for all $t \in [0,T)$. If instead $\partial f/\partial \hat{n} < 0$ on $\partial\Omega$, then

$$\inf_{x \in \overline{\Omega}, m \geq M} (u^m(x,t) - f(x)) \geq -Ct.$$

Proof. We assume $\partial f/\partial \hat{n} > 0$ on $\partial\Omega$ and prove the first inequality. The other case may be proved similarly.

We prove that there is a $C > 0$ and an $h_0 > 0$ such that

$$G_h f(x) \leq f(x) + Ch, \qquad \text{for all } x \in \overline{\Omega}, h \in (0, h_0), \tag{4.3}$$

which, if M is chosen so that $T/M \leq h_0$, may be iterated to give the desired result.

In order to prove (4.3), we fix $z \in \overline{\Omega}$. If $\nabla f(z) \neq 0$, it follows from proposition 3.3 with $\varepsilon = 1$ that there is a $\delta_1 > 0$ such that

$$G_h f(x) \leq f(x) + (|\nabla f(z)| g(\mathrm{curv}(f(z))) + 1)h \tag{4.4}$$

holds for all $x \in B_n(z, \delta_1) \cap \overline{\Omega}$ and $h \in (0, \delta_1)$.

Now suppose that $\nabla f(z) = 0$. Since we assume $\partial f/\partial \hat{n} > 0$ on $\partial\Omega$, we then have $z \in \Omega$. Since $f \in C^2(\overline{\Omega})$, there is a $\delta_2 > 0$, such that

$$|f(x) - f(z)| \leq \|D^2 f\|_\infty |x - z|^2 \tag{4.5}$$

for all $x \in B_n(z, \delta_2) \subset \Omega$. Taking $C_1 = 2\|f\|_\infty/\delta_2^2 + \|D^2 f\|_\infty$, we then get

$$f(x) \leq f(z) + C_1 |x - z|^2$$

for all $x \in \overline{\Omega}$. Applying G_h to both sides of this inequality with z fixed, and using in turn proposition 3.2 (i)-(iii), lemma 4.3 and (4.5), we get

$$\begin{aligned} [G_h f](x) &\leq f(z) + C_1 [G_h(|\cdot - z|^2)](x) \leq f(z) + C_1(|x - z|^2 + C_2 h) \\ &\leq f(x) + (C_1 + \|D^2 f\|_\infty)|x - z|^2 + C_1 C_2 h \end{aligned} \tag{4.6}$$

for all $x \in B_n(z, \delta_3)$ and $h \in (0, h_1)$ for some $h_1 > 0$ and $\delta_3 > 0$.

That is, (4.4) and (4.6) imply that for any $z \in \overline{\Omega}$ and any $\varepsilon > 0$, there are $C > 0$ and $h_0 > 0$ independent of ε and a $\delta = \delta(\varepsilon) > 0$ such that

$$[G_h f](x) \leq f(x) + Ch + \varepsilon \tag{4.7}$$

for all $x \in B_n(z, \delta)$ and $h \in (0, h_0)$. Since $\overline{\Omega}$ is compact, it may be covered by a finite number of such δ-neighborhoods and taking the largest of the C:s and the smallest of the h_0:s, (4.7) holds for all $x \in \overline{\Omega}$ and $h \in (0, h_0)$. Since ε is arbitrary, we get (4.3). □

LEMMA 4.5. (cf. [11, lemma 4.2]) *Let $f \in C(\overline{\Omega})$ in (4.1). Then $\overline{u}(x,0) = \underline{u}(x,0) = f(x)$ for all $x \in \overline{\Omega}$.*

Proof. Take a sequence $\{f_k\}$ in $C^2(\overline{\Omega})$ satisfying

$$\|f_k - f\|_\infty < \frac{1}{k}, \qquad \frac{\partial f_k}{\partial \hat{n}} > 0 \text{ on } \partial\Omega$$

and define u_k^m as in (4.1) with f_k instead of f.

Then, by lemma 4.4, there are constants C_k such that for each $k \in \mathbb{Z}_+$,

$$u_k^m(y,t) - f_k(y) \leq C_k t$$

for all $y \in \overline{\Omega}$ and $t \in [0, T)$. Because of proposition 3.2 (iv) and since $\|f_k - f\|_\infty < 1/k$, it follows that

$$|u^m(y,t) - u_k^m(y,t)| \leq \frac{1}{k}$$

and thus

$$u^m(y,t) - f(y) \leq C_k t + \frac{2}{k} \tag{4.8}$$

for all $y \in \overline{\Omega}$, $t \in [0, T)$ and $k \in \mathbb{Z}_+$.

For any fixed $x \in \overline{\Omega}$, we may now let $m \to \infty$, $y \to x$, $t \to 0$ and finally $k \to \infty$ to conclude by the definition (4.2) of $\overline{u}$ that

$$\overline{u}(x,0) \leq f(x).$$

Since $\underline{u}(x,0) \geq f(x)$ by a similar argument and $\underline{u}(x,0) \leq \overline{u}(x,0)$ by definition, the result follows. □

We are now ready to prove the main convergence theorem.

Proof. (of theorem 4.1) As already mentioned, the idea of the proof is to show that $\overline{u}(x,t)$ (defined by (4.2)) is a subsolution and $\underline{u}(x,t)$ a supersolution of the level set PDE (2.1). It then follows from lemma 4.5 and the comparison result by Sato [15] that $\overline{u}(x,t) \leq \underline{u}(x,t)$ for all $(x,t) \in \overline{\Omega} \times [0,T)$ and thus $u = \overline{u} = \underline{u}$ is a viscosity solution of (2.1). By remark 6.4 in Crandall, Ishii, Lions [4] it also follows that $u^m \to u$ locally uniformly.

We prove that $\overline{u}(x,t)$ is a subsolution and note that the proof that $\underline{u}(x,t)$ is a supersolution is similar.

So, we fix a function $\varphi \in C^2(\overline{\Omega} \times [0,T))$ and assume that $\overline{u} - \varphi$ has a strict maximum at (x_0, t_0). Since only the local behavior of φ is important, we may assume

that this maximum is global. Also, since all that matters is $\partial\varphi/\partial t$, $\nabla\varphi$ and $D^2\varphi$, we may choose φ on the form

$$\varphi(x,t) = \varphi_1(x) + \varphi_2(t)$$

for some functions φ_1, φ_2.

According to definition 4.2, in order to show that $\overline{u}$ is a subsolution, there are two distinct cases.

Case 1: $x_0 \in \Omega$ *and* $\nabla\varphi_1(x_0) \neq 0$, *or* $x_0 \in \partial\Omega$ *and* $\partial\varphi_1/\partial\hat{n}(x_0) > 0$.

Fix $\varepsilon > 0$. Then, setting $M \equiv |\nabla\varphi_1(x_0)|g(\mathrm{curv}(\varphi_1(x_0)))$, by proposition 3.3, there is a $\delta_1 > 0$ such that

$$[G_h\varphi_1](x) \leq \varphi_1(x) + Mh + \varepsilon h \tag{4.9}$$

for all $x \in B_n(x_0,\delta_1)$ and $h \in (0,\delta_1]$.

Since (x_0,t_0) is a strict global maximum point of $\overline{u} - \varphi$, by the definition of $\overline{u}$ there is an $m \in \mathbb{Z}_+$ such that

$$\sup_{B_{n+1}((x_0,t_0),\delta_1)} (u^m - \varphi)(x,t) > \sup_{(\overline{\Omega}\times(0,T))\setminus B_{n+1}((x_0,t_0),\delta_1)} (u^m - \varphi)(x,t)$$

and $h \equiv T/m < \delta_1$. Then, we can choose $(\xi,\tau) \in B_{n+1}((x_0,t_0),\delta_1)$ so that

$$(u^m - \varphi)(\xi,\tau) + \varepsilon h > (u^m - \varphi)(x,t) \tag{4.10}$$

for all $(x,t) \in \overline{\Omega}\times(0,T)$.

Now choose $l \in \mathbb{N}$ so that $lh \leq \tau < (l+1)h$. By the definition of u^m, we then have

$$u^m(x,\tau) = [G_{\tau-lh} \circ G_h u^m(\cdot,(l-1)h)](x)$$

for all $x \in \overline{\Omega}$. Also, from (4.10) follows that

$$u^m(x,(l-1)h) \leq \varphi_1(x) + \varphi_2((l-1)h) + \varepsilon h + (u^m - \varphi)(\xi,\tau),$$

which, using proposition 3.2 (i), (iii) and (4.9), gives us

$$\begin{aligned}[G_h u^m(\cdot,(l-1)h)](x) &\leq [G_h\varphi_1](x) + \varphi_2((l-1)h) + \varepsilon h + (u^m - \varphi)(\xi,\tau)\\ &\leq \varphi_1(x) + Mh + 2\varepsilon h + \varphi_2((l-1)h) + (u^m - \varphi)(\xi,\tau)\end{aligned} \tag{4.11}$$

for all $x \in B_n(x,\delta_1)$.

In order to get a similar inequality for each $x \in \overline{\Omega}$, we note that by proposition 3.2 (iv),

$$|u^m(x,t)| \leq \|f\| \tag{4.12}$$

for all $(x,t) \in \overline{\Omega}\times[0,T]$ and we choose $C > 0$ such that

$$2\|f\| + \|\varphi_1\| + 2\sup_{[0,T]}|\varphi_2| + |M|T \leq C, \tag{4.13}$$

$\delta_2 \in (0, \delta_1)$ and a function $\psi_1 \in C^2(\overline{\Omega})$ with $\psi_1 = \varphi_1$ in $B_n(x_0, \delta_2) \cap \overline{\Omega}$, $\psi_1 \geq \varphi_1$ in $\overline{\Omega}$ and $\psi_1 \geq C$ in $B_n(x_0, \delta_1)^c \cap \overline{\Omega}$. Then, by proposition 3.3, there is a $\delta_3 \in (0, \delta_2]$ such that

$$[G_h \psi_1](x) \leq \psi_1(x) + Mh + \varepsilon h \tag{4.14}$$

for all $x \in B_n(x_0, \delta_3)$ and $h \in (0, \delta_3)$. By (4.12) and (4.13), we note that

$$[G_h u^m(\cdot, (l-1)h)](x) - Mh - 2\varepsilon h - \varphi_2((l-1)h) - (u^m - \varphi)(\xi, \tau) \leq C,$$

and thus, using also (4.11) and the definition of ψ_1,

$$[G_h u^m(\cdot, (l-1)h)](x) - Mh - 2\varepsilon h - \varphi_2((l-1)h) - (u^m - \varphi)(\xi, \tau) \leq \psi_1(x) \tag{4.15}$$

for all $x \in \overline{\Omega}$.

Now, with m chosen so large that $h \leq \delta_3$, we apply $G_{\tau - lh}$ to both sides of (4.15) and use (4.14) to get

$$\begin{aligned} u^m(x, \tau) &= [G_{\tau - lh} \circ G_h u^m(\cdot, (l-1)h)](x) \\ &\leq \psi_1(x) + M(\tau - lh) + \varepsilon(\tau - lh) + Mh + 2\varepsilon h \\ &\quad + \varphi_2((l-1)h) + (u^m - \varphi)(\xi, \tau) \\ &\leq \varphi_1(x) + M(\tau - (l-1)h) + 2\varepsilon(\tau - (l-1)h) \\ &\quad + \varphi_2((l-1)h) + (u^m - \varphi)(\xi, \tau) \end{aligned}$$

for all $x \in B_n(x_0, \delta_3)$, since $h \leq \tau - (l-1)h$. Specifically for $x = \xi$, we get

$$u^m(\xi, \tau) \leq \varphi_1(\xi) + M(\tau - (l-1)h) + 2\varepsilon(\tau - (l-1)h) + \varphi_2((l-1)h) + (u^m - \varphi)(\xi, \tau),$$

that is, since $\varphi(\xi, \tau) = \varphi_1(\xi) + \varphi(\tau)$,

$$\varphi_2(\tau) - \varphi_2((l-1)h) \leq (M + 2\varepsilon)(\tau - (l-1)h).$$

Letting $\varepsilon \to 0$, so that $m \to \infty$, $h \to 0$ and $(\xi, \tau) \to (x_0, t_0)$, we get

$$\varphi_t'(x_0, t_0) \leq M = |\nabla \varphi(x_0, t_0)| g(\mathrm{curv}(\varphi(x_0, t_0))),$$

which is what we want according to definition 4.2.

Case 2: $x_0 \in \Omega$, $\nabla \varphi_1(x_0) = 0$ *and* $D^2 \varphi_1(x_0) = 0$.

We need to prove that $\varphi_t'(x_0, t_0) \leq 0$. This follows exactly as in [10, theorem 2.1], using lemma 4.3 and proposition 3.2. □

Finally, we show how to get around the assumption (2.3d) that g has bounded derivative from above and below. The next theorem shows that if g does not have bounded derivative it is enough to approximate g with functions g_ν, such that $g_\nu \to g$ uniformly as $\nu \to 0$. We remark that if g has unbounded derivative, the derivative of g_ν must obviously tend to infinity when $\nu \to 0$, which means that in order to fulfill the inequalities (1.6) restricting the choice of a_i and b_i, the radius r_2 of the support of the convolution kernel ρ_2 must also tend to infinity. In order to get a converging scheme, the choice of ν and the time-step h must be such that $r_2\sqrt{h}$, which is the radius of the support of the scaled kernel $\rho_2^{\sqrt{h}}$, is small. The exact choice of ν depends on the exact definition of the approximations g_ν and the choice of the kernel ρ_2.

THEOREM 4.6. *Let $g \in C(\mathbb{R})$ have the properties* (2.3a)–(2.3c) *and let $\{g_\nu\}_{\nu>0}$ be a family of functions in $C(\mathbb{R})$ that fulfill the assumptions* (2.3a)–(2.3d), *such that $g_\nu(\kappa) \to g(\kappa)$ as $\nu \to 0$, uniformly for $\kappa \in \mathbb{R}$. Define*

$$F_\nu(N_1, N_2) = \frac{1}{\sqrt{h}} \frac{b_2(\nu)N_1 - b_1(\nu)N_2}{d(\nu)} - g_\nu\left(\frac{1}{\sqrt{h}} \frac{a_2(\nu)N_1 - a_1(\nu)N_2}{d(\nu)}\right),$$

$$[G_h^\nu \varphi](x) = \sup\{\lambda \in \mathbb{R} \mid F_\nu(N_1(\lambda), N_2(\lambda)) \geq 0\},$$

with N_i, a_i, b_i and d as in (1.5). *Finally, for $f \in C(\overline{\Omega})$ define*

$$u_\nu^m(x,t) = [G_{t-lh}^\nu \circ (G_h^\nu)^l f](x),$$

with $h \equiv T/m$ and $lh \leq t < (l+1)h$.

Then there is a sequence $\{\nu_m\}_{m=1}^\infty$ with $\nu_m \to 0$ as $m \to \infty$, such that

$$u_{\nu_m}^m(x,t) \to u(x,t) \text{ as } m \to \infty,$$

locally uniformly for $(x,t) \in \overline{\Omega} \times (0,T)$, where $u(x,t)$ is the unique viscosity solution of the level set PDE (2.1).

Proof. It follows from remarks 6.3 and 6.4 in [4], that if v_n is a subsolution of an equation $\mathcal{F}_n(\nabla v, D^2 v) = 0$ in a set $\mathcal{O}$ for $n = 1, 2, \ldots$, and we define

$$\overline{v}(x) = \limsup_{n\to\infty}{}^* v_n(x) = \lim_{j\to\infty} \sup\{v_n(y) \mid n \geq j,\ y \in \mathcal{O},\ |y - x| < j^{-1}\}, \tag{4.16}$$

then $\overline{v}$ is a subsolution of the equation $\mathcal{F}(\nabla v, D^2 v) = 0$, where

$$\begin{aligned}\mathcal{F}(p, X) &= \liminf_{n\to\infty}{}_* \mathcal{F}_n(p, X) \\ &= \lim_{j\to\infty} \inf\{\mathcal{F}_n(q, Y) \mid n \geq j,\ (q, Y) \in W,\ \|p - q\| + \|X - Y\| \leq j^{-1}\},\end{aligned}$$

with W dense in $\mathbb{R}^n \times \mathcal{S}(n)$, where $\mathcal{S}(n)$ is the space of symmetric $n \times n$-matrices. Furthermore, $v_n \to \overline{v}$ locally uniformly. The equivalent conclusion holds for supersolutions (with $\liminf_*$ and $\limsup^*$ interchanged). There may also be boundary conditions included in $\mathcal{F}_n$ and $\mathcal{F}$ (see also section 7.A in [4]).

For the present case, this means that viscosity solutions u_{ν_m} of the equations $u_t - |\nabla u| g_{\nu_m}(\mathrm{curv}(u)) = 0$, with Neumann boundary conditions, tend locally uniformly to the viscosity solution u of the equation $u_t - |\nabla u| g(\mathrm{curv}(u)) = 0$ with Neumann boundary conditions. This is because

$$\liminf_{m\to\infty}{}_* |p| g_{\nu_m}(\mathrm{curv}(p, X)) = \limsup_{m\to\infty}{}^* |p| g_{\nu_m}(\mathrm{curv}(p, X)) = |p| g(\mathrm{curv}(p, X)),$$

since $g_{\nu_m} \to g$ uniformly. Here, $\mathrm{curv}(p, X)$ is defined as $\mathrm{curv}(u)$ (see (2.2)) with $p = \nabla u$ and $X = D^2 u$, but this notation is chosen to emphasize that the *-limits are with respect to p and X, and not u.

Furthermore, theorem 4.1 shows that for each fixed ν, the approximations u_ν^m tend to u_ν locally uniformly. Therefore, we can find a sequence $\{\nu_m\}_{m=1}^\infty$ such that for each compact $K \subset \overline{\Omega} \times [0, T)$, it holds that for each $\varepsilon > 0$, there is an $M \in \mathbb{Z}_+$ such that

$$|u_{\nu_m}^m(x,t) - u(x,t)| \leq |u_{\nu_m}^m(x,t) - u_{\nu_m}(x,t)| + |u_{\nu_m}(x,t) - u(x,t)| < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

for all $m \geq M$ and $(x,t) \in K$. That is, $u_{\nu_m}^m \to u$ locally uniformly as $m \to \infty$. □

5. Examples. We conclude with a few examples showing the output of the scheme. The evolutions have been computed by Alexei Heintz using the code developed by Richards Grzhibovskis for the case without boundary conditions and adapted to this new setting. The images in figure 5.1 were generated using the VRweb software [16].

The first example is shown in figure 5.1 and shows the generalized mean curvature flow, with $v = \kappa^3$, of a cylinder that has been placed slightly off center inside an ellipsoid. The intersections with the ellipsoid are to the left and right. First, we see that the surface becomes bent near the domain interface in order to fulfill the boundary condition. Then, evolution is faster in the middle and the neck becomes thinner, until it finally breaks off into two separate surfaces which then approach the boundary and disappear. Note that the time intervals between sequential images are not equal.

In figure 5.2, we see the evolution of the same cylinder using the three different velocities $v = \kappa$, $v = \kappa^3$ and $v = \kappa|\kappa|^{-1/2}$. This time, we show cross sections of the surfaces in the symmetry plane, for selected time steps. The differences between the evolutions for the three velocities are clearly visible, the most remarkable being that it is only when $v = \kappa^3$ that the surface splits into two. For the other two cases, the surface does not move fast enough in the middle to split there, but instead releases first from the boundary. There is also a clear difference between the cases $v = \kappa$ and $v = \kappa|\kappa|^{-1/2}$.

Acknowledgments. I wish to thank Alexei Heintz for introducing me to curvature flows and for suggesting this problem to me, as well as for generating the images in the examples section. Furthermore, I wish to thank Richards Grzhibovskis for taking time to adapt his computer program and for allowing us to use it.

REFERENCES

[1] G. Barles and C. Georgelin. A simple proof of convergence for an approximation scheme for computing motions by mean curvature. *SIAM J. Numer. Anal.*, 32(2):484–500, 1995.

[2] J. Bence, B. Merriman, and S. Osher. Diffusion generated motion by mean curvature. In *Computational Crystal Growers Workshop*, pages 73–83. American Mathematical Society, 1992.

[3] K. A. Brakke. *The Motion of a Surface by its Mean Curvature.* Princeton University Press, Princeton, NJ, 1978.

[4] M. G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.*, 27(1):1–67, 1992.

[5] K. Ecker. *Regularity Theory for Mean Curvature Flow.* Birkhäuser Verlag, 2004.

[6] L. C. Evans. Convergence of an algorithm for mean curvature motion. *Indiana Univ. Math. J.*, 42:533–556, 1993.

[7] L. C. Evans and J. Spruck. Motion of level sets by mean curvature, I. *J. Differential Geom.*, 33:635–681, 1991.

[8] Y. Giga and M.-H. Sato. Neumann problem for singular degenerate parabolic equations. *Differential and Integral Equations*, 6(6):1217–1230, 1993.

[9] R. Grzhibovskis and A. Heintz. A convolution-thresholding approximation of generalized curvature flows. *SIAM J. Numer Anal.*, 42(6):2652–2670, 2005.

[10] H. Ishii. A generalization of the Bence, Merriman and Osher algorithm for motion by mean curvature. In *Curvature flows and related topics (Levico, 1994)*, volume 5 of *GAKUTO Internat. Ser. Math. Sci. Appl.*, pages 111–127. Gakkōtosho, Tokyo, 1995.

[11] H. Ishii and K. Ishii. An approximation scheme for motion by mean curvature with right-angle boundary condition. *SIAM J. Math. Anal.*, 33(2):369–389, 2001.

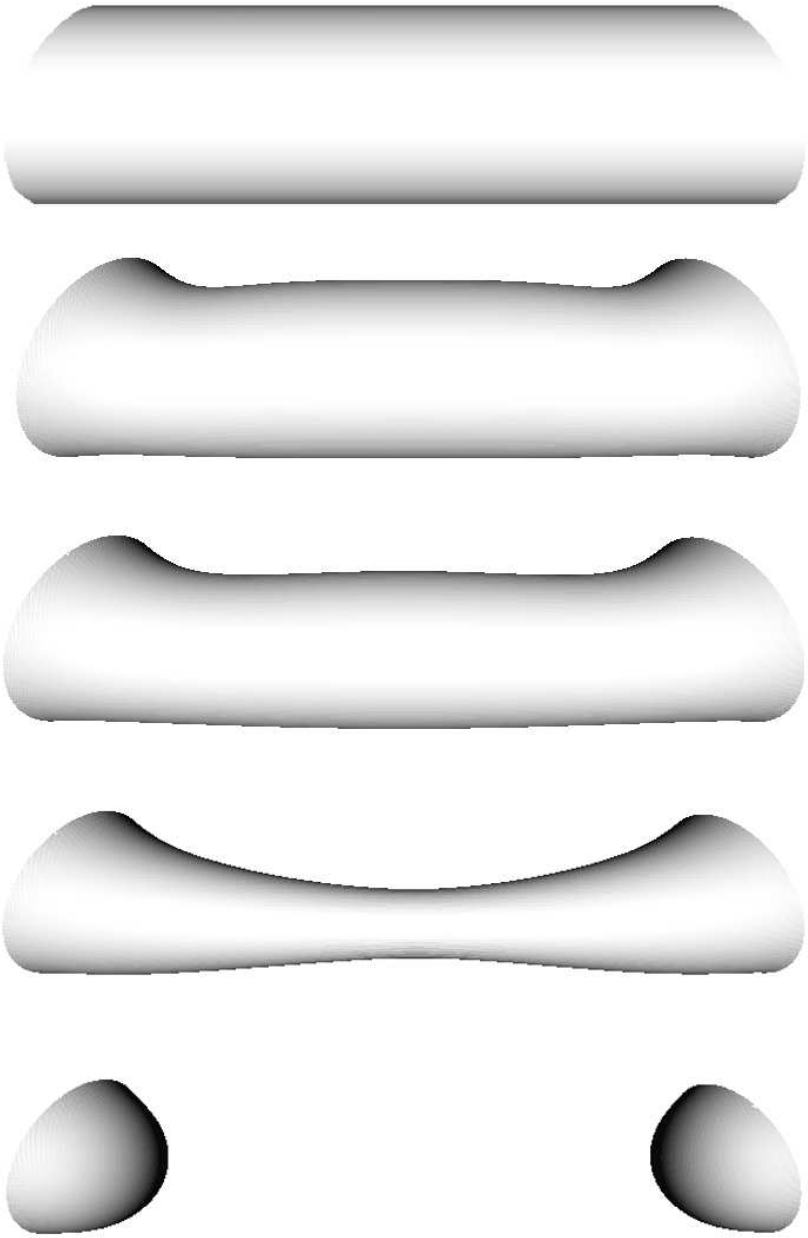

FIG. 5.1. *The evolution of a cylinder inside an ellipsoid, for $v = \kappa^3$. The cylinder is placed above the center of the ellipsoid and intersects the ellipsoid to the left and right in the images. The images shown are after 0, 20, 60, 100 and 110 time steps.*

[12] H. Ishii, G. E. Pires, and P. E. Souganidis. Threshold dynamics type approximation schemes for propagating fronts. *J. Math. Soc. Japan*, 51(2):267–308, 1999.

[13] S. Osher and J. A. Sethian. Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.*, 79:12–49, 1988.

[14] S. J. Ruuth. Efficient algorithms for diffusion-generated motion by mean curvature. *J. Comput. Phys.*, 144(2):603–625, 1998.

[15] M.-H. Sato. Interface evolution with Neumann boundary condition. *Adv. Math. Sci. Appl.*, 4(1):249–264, 1994.

[16] VRweb VRML browser, `http://www2.iicm.edu/vrweb`.

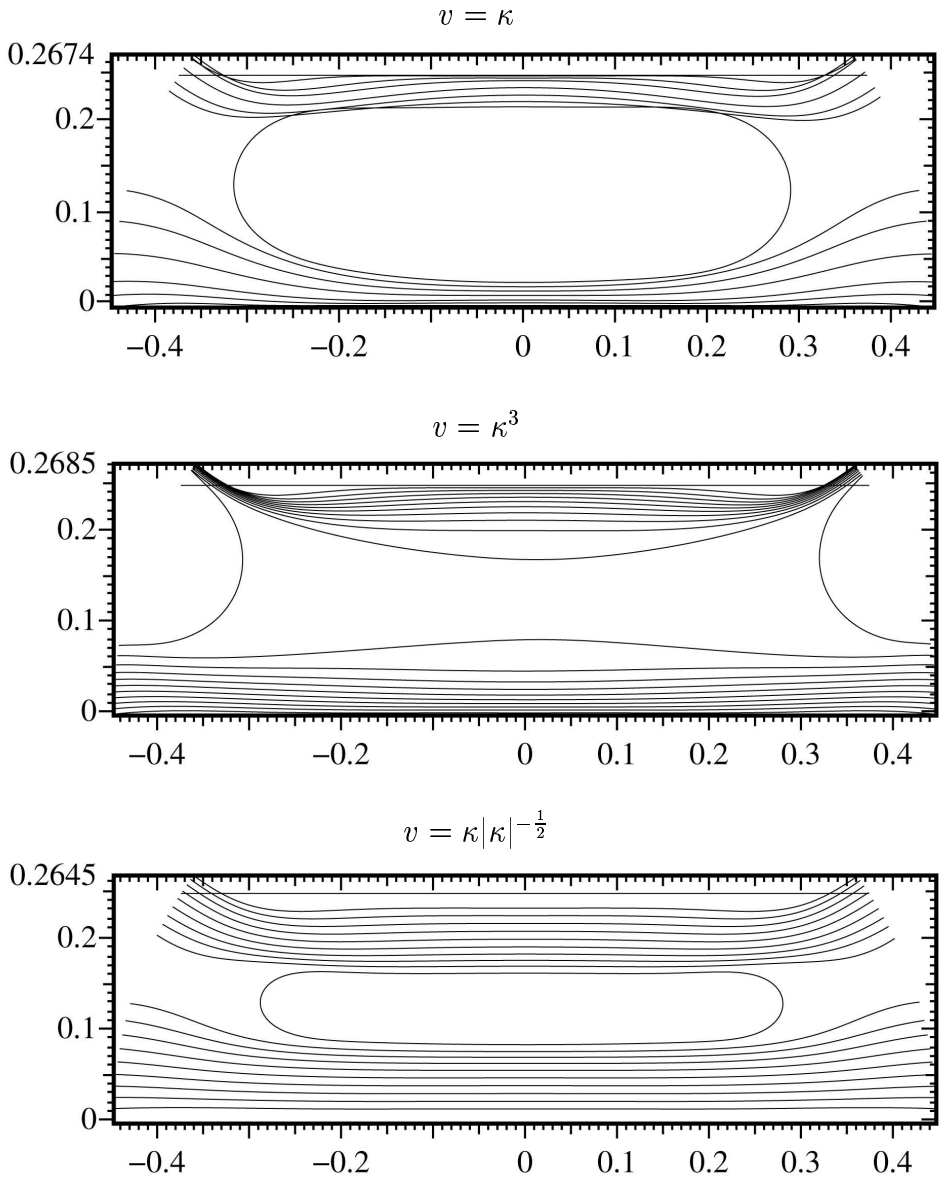

FIG. 5.2. *Cross sections in the symmetry plane for the evolution of a cylinder using three different velocity assignments. Note that the time steps are not equal in the three images.*