Diplomarbeit

Bewertung von hypotheken-basierten Wertpapieren unter Verwendung finiter Elemente

(Valuation of Mortgage-Backed Securities Using Finite Elements)

Angefertigt am

Institut für Numerische Simulation

Vorgelegt der

Mathematisch-Naturwissenschaftlichen Fakultät der

Rheinischen Friedrich-Wilhelms-Universität Bonn

im August 2007

von

Sebastian Paik

aus

Düren

Diploma Thesis

Valuation of Mortgage-Backed Securities Using Finite Elements

Developed at

Institute for Numerical Simulation

Presented to

The Faculty of Mathematics and Natural Sciences

of the Rheinische Friedrich-Wilhelms-Universität Bonn, Germany

August 2007

by

Sebastian Paik

From

Düren

Contents

Deutsche Zusammenfassung

Viele Probleme der Naturwissenschaften, Ökonomie und Finanzwissenschaften erfordern sehr komplexe mathematische Modelle, um sie zufriedenstellend lösen zu können. Insbesondere in letzterem Bereich sind stochastische Prozesse ein sehr bedeutsames Modellierungsmittel, weil sie einer stochastischen und nicht einer deterministischen Entwicklung der Finanzmärkte Rechnung tragen.
Im Falle regelmäßig an den Märkten gehandelter Derivate bestehen solche Modelle typischerweise aus stochastischen Prozessen, die das Verhalten der unterliegenden Variablen beschreiben, beispielsweise Aktienkurse. Nach Angabe eines bestimmten Modells besteht der nächste Schritt oft in der Herleitung einer partiellen Differentialgleichung für den Wert des Derivats. Dies bedeutet, daß Lösungen des Modells in Form einer partiellen Differentialgleichung dargestellt sind, die es zu lösen gilt, um den gesuchten Preis zu erhalten. Das vermutlich berühmteste Beispiel ist die Black-Scholes Gleichung, die Fischer Black, Myron Scholes und Robert C. Merton 1973 (vgl. [8] und [76]) für den Preis von Aktienoptionen veröffentlichten.
Als Teil der Komplexität eines Modells kann es passieren (was auch sehr häufig geschieht), daß keine analytische Lösung in geschlossener Form vorliegt. Deshalb muß man einen numerischen Ansatz bemühen, der die exakte Lösung hinreichend gut approximiert. Um diese Genauigkeit sicherzustellen sind sehr effiziente numerische Verfahren erforderlich, die speziell an die Eigenschaften des vorliegenden Problems angepaßt sind. Aufgrund der Komplexität der Modelle, die durch den Wunsch entsteht, die Wirklichkeit bestmöglich zu modellieren, handelt es sich dabei um eine anspruchsvolle Aufgabe, die gründliches Untersuchen der zugrundeliegenden mathematischen Aspekte erfordert.
Das Ziel des wissenschaftlichen Rechnens besteht neben diesen Modellierungsaspekten aus zwei Teilen, dem Diskretisieren und Lösen. Ersteres bezeichnet die Approximation der exakten Lösung durch Ersetzen der PDE durch ein lineares Gleichungssystem und letzteres im Bereitstellen effizienter Lösungsverfahren dazu.

Problemstellung dieser Arbeit

In dieser Diplomarbeit werden wir uns mit Mortgage-Backed Securities (Hypothekenverbriefungen, MBSs) beschäftigen, wobei es sich um Zinsderivate mit eingebetteten Optionen handelt. MBSs zahlen ihren Haltern gewöhnlich einen monatlichen Kupon, der durch Hypothekenzahlungen von zumeist privaten Krediten entsteht. Die Optionen rühren vom Recht des Kreditnehmers her, den Vertrag vorzeitig abzuzahlen oder durch Insolvenz die Zahlung einzustellen. In einem solchen Fall enden auch die Zahlungen an den Besitzer des MBS. Folglich hängt der Wert des Derivats von zwei Variablen ab, nämlich von Zinsen und Hauspreisen. MBSs sind von besonderem Interesse für die meisten Marktteilnehmer, weil sie eine Möglichkeit darstellen, Kreditrisiken zu verbriefen und weiterzuverkaufen. Im Sommer 2007 stieg ihre Bedeutung im Zuge der US-amerikanischen Hypothekenkrise sogar noch weiter, als viele Verträge von Kreditnehmern schlechter Bonität ausfielen. Der Grad der Krise und die Tatsache, daß weite Teile der amerikanischen Wirtschaft kreditfinanziert sind (z. B. die Mehrzahl der Fusionen und

Übernahmen), machen sie sehr anfällig gegenüber steigenden Risikoprämien, was sogar weltweit die Aktienmärkte in Aufruhr versetzte.
Eine Schlüsselrolle in dieser Diplomarbeit kommt der stochastischen Modellierung von Zinsen und Hauspreisen zu. Daraus werden wir eine partielle Differentialgleichung in Zinsen, Hauspreisen und der Zeit für den Wert eines solchen MBS herleiten. Auf die PDE werden wir dann effiziente numerische Verfahren anwenden, um sie zu lösen, womit unser Vorgehen bespielsweise dem von Black-Scholes, der Bewertung von Optionen mit stochastischer Volatilität oder aber der Bewertung amerikanischer Optionen sehr ähnlich wird (vgl. [50] oder [79]).

Modellierung

In dieser Arbeit benutzen wir die Ergebnisse von Downing, Stanton und Wallace (vgl. [28]) als Ausgangspunkt. Wir verwenden das Cox-Ingersoll-Ross Modell (vgl. [20]) als Zinsprozeß und den in [28] und [29] vorgeschlagenen für die Hauspreise. Letzterer ist bis auf die separat geschätzten Parameter fast identisch mit demjenigen für Aktienkurse des Black-Scholes Modells (vgl. [51]). Dies bedeutet, daß sowohl Drift als auch Volatilität proportional zu den Hauspreisen sind. Ferner gehen wir auf alternative Zinsmodelle ein, wie sie z. B. in [60] und [78] vorgestellt werden, und diskutieren Eigenschaften des Hauspreisprozesses. Darüber hinaus geben wir grundlegende Definitionen von Anleihen und Optionen und beschreiben Annuitäten. Darunter findet sich u. a. eine Herleitung der Standardannuitäten-Formel und wir erklären, warum sie auf festverzinsliche Kreditverträge angewendet werden kann.
Darüber hinaus leiten wir eine partielle Differentialgleichung für den Wert eines MBS her, indem wir die berühmte Itô-Formel (vgl. [51], [58] oder [78]) auf obige stochastische Prozesse anwenden. Dabei stellt sich heraus, daß wir es mit einer Gleichung zweiter Ordnung in Zinsen und Hauspreisen ohne gemischte Ableitung (der Grund dafür liegt in der Unkorreliertheit der Braunschen Bewegungen, die den Prozessen zugrundeliegen, was wir analog zu [28] annehmen) und erster Ordnung in der Zeit zu tun haben. Die rechte Seite ist konstant, nämlich gerade der monatliche Kupon. Folglich handelt es sich um eine parabolische Gleichung, die wir auf Regularität hin untersuchen und wobei wir auf das Verhältnis von Diffusion zu Konvektion achten. Nahe des Randes findet sich Konvektionsdominanz, die wir mit Hilfe von speziellen Transformationen und Stabilisierungstechniken bei der Diskretisierung zu behandeln suchen.
Bezüglich der Regularität geben wir sehr starke Hinweise auf schlechtere Glattheit als H^2 (zweiter Sobolevraum) selbst nach Transformation. Unsere Argumentation stützt sich darauf, daß der Differentialoperator nicht gleichmäßig parabolisch ist, wie es in [2], [11], [37] und [69] gefordert wird. Darüber hinaus vergleichen wir die Ergebnisse in [69] mit denen in [27] und wenden dies auf unser Problem an.
Neben der PDE entwickeln wir ein Modell für das vorzeitige Beenden des Kontrakts durch Vorabzahlung oder Insolvenz. Entscheidend dabei ist, daß es sowohl finanziell optimales Beenden als auch suboptimales Beenden berücksichtigt. Dies geschieht mit Hilfe von Funktionen, die den jeweiligen Abbruchwegen Wahrscheinlichkeiten zuordnen (vgl. [28] und [87]). Ferner zieht eine Beendigung Transaktionskosten nach sich. Diese wiederum werden wie in [28] als für jeden Kreditnehmer unterschiedlich angenommen und gehen in die Bewertung ein, ob ein Abbruch sinnvoll ist. Das wiederum arbeiten wir in die PDE-Lösung ein, indem wir diese mit den Werten bei Vorabzahlung oder Insolvenz gewichten. Allerdings werden wir auf [51] basierend argumentieren, daß es ausreicht, diese Gewichtung lediglich jeden Monat auszuführen und nicht in jedem Zeitschritt.

Diskretisierung

Die Ortsdiskretisierung der transformierten PDE erfolgt dann mit Hilfe der Finite-Elemente-Methode auf einem nicht-uniformen Gitter (vgl. [57], [70] oder [90]). Dabei plazieren wir mehr Freiheitsgrade in den Gebieten, die wir als problematisch erachten in bezug auf Regularität und

vorzeitiges Beenden. Einer der Hauptvorteile von finiten Elementen gegenüber finiten Differenzen ist die Verwendbarkeit von streamline upwind Petrov Galerkin (vgl. [19], [63] oder [96]), welches eine Veränderung der Testfunktionen benutzt, um Stabilisierung zu erreichen. Darüber hinaus leiten wir eine schwache Formulierung unserer PDE her, diskutieren die Standardfehlerabschätzungen für elliptische Probleme gemäß [11] und fassen semi-diskrete Fehlerabschätzung für parabolische Probleme, wie sie z. B. in [64] dargestellt sind, zusammen.
In der Zeit benutzen wir die sogenannten BDF-Formeln (vgl. [80]) innerhalb eines ordnungs- und schrittweitenadaptiven Zeitschrittverfahrens. Die genannten Kriterien entnahmen wir [40] und [86]. Wir versuchen also stets die Schrittweite so lang und die Ordnung zu hoch wie möglich (höchstens 5) zu wählen und das Ablehnen von Schritten zu vermeiden. Darüber hinaus untersuchen wir den Konsistenzfehler und geben basierend auf [80] an, unter welchen Bedingungen das Verfahren konvergiert.
Weiterhin zeigen wir, daß die wahrscheinlichkeitsgewichtete Lösung in L^2 konvergiert, falls die ungewichtete dies tut. Ferner merken wir an, daß dies das Funktionieren des gesamten Lösungsverfahrens sicherstellt, welches aus dem Lösen der PDE und dem Wahrscheinlichkeitsgewichten für alle 360 Monate der 30-jährigen Hypothek besteht.

Lösung des Modells

Beim Lösen des diskretisierten Problems versuchen wir abermals sehr effizient vorzugehen, indem wir ein Mehrgitter-Verfahren mit vier Vergröberungsebenen in einem V-Zyklus verwenden (vgl. [11], [45] und [64]). Dieses wiederum ist durch den iterativen GMRES-Löser (vgl. [82]) vorkonditioniert (selbst durch unvollständige LU-Zerlegung (vgl. [42] und [83]) vorgeglättet), während als Nachglätter das SORU-Vefahren benutzt wird. Der Vorteil dieses Lösers ist ein Verfahren in optimaler Komplexität, d.h. es ist von Ordnung $\mathcal{O}(N)$, wenn N die Anzahl der Unbekannten bezeichnet. Dies ist sehr gut, wenn man bedenkt, daß die klassische Gauß-Elimination von Ordnung $\mathcal{O}(N^3)$ ist (vgl. [47], [48] oder [89]). Nichtsdestotrotz verwenden wir für die Grobgitterkorrektur ein modifiziertes Gauß-Verfahren, nämlich den UMFPACK-Löser, der [21], [22], [23] und [24] entnommen ist.
Ferner führen wir unsere numerischen Betrachtungen innerhalb des kommerziellen Pakets Comsol Multiphysics aus, das eine Vielzahl verschiedener Löser, Einstellungen und Glättern bereitstellt (siehe z. B. [96], [97] und [98]). Darüber hinaus können wir davon ausgehen, daß die Implementierung schnell und zuverlässig ist.

Numerische Ergebnisse

Weiterhin testen wir unsere Diskretisierung, indem wir Konvergenzraten sowohl im Ort als auch der Zeit bestimmen. Im Ort finden wir eine schlechtere Rate als aufgrund der Standardfehlerabschätzen erwartet. Der Grund dafür liegt darin, daß unsere Lösung die Regularitätsvoraussetzungen verletzt. Darüber hinaus stellen sich die Einschränkungen des implementierten SUPG in Comsol Multiphysics als zu gravierend heraus (nur Elemente erster Ordnung, nur konstante PDE-Koeffizienten), so daß wir sogar ohne SUPG bessere Resultate bekommen als mit. In der Zeit erzielen wir die gute Konvergenzrate von 5.
Dieselben Tests führen wir für die wahrscheinlichkeitsgewichtete Lösung nach einem Monat aus, wobei die Raten erhalten bleiben. Dies geschieht jeweils für realistische monatliche Kupons, die sich aus einem Vertragszinssatz von plausiblen 5% ergeben.
Nachdem wir uns der Konvergenzraten versichert haben, widmen wir uns Rechnungen über die gesamte Laufzeit von 30 Jahren für verschiedene Vertragszinssätze (5.0%, 8.5% und 10.0%) und drei verschiedene Größenordnungen von Transaktionskosten (1.4%, 8.6% und 18.5%). Dabei führen wir genau Buch darüber, wie oft es optimal ist, den Vertrag vorzeitig zu beenden und vergleichen die verschiedenen Szenarien. Dies wird untermalt durch etliche graphische Aufbereitungen. Darüber hinaus bestimmen wir die Schwellen, an denen Insolvenz bzw. Vorabzahlung

optimal wird, und vergleichen dies mit den übrigen Daten. Ferner betrachten wir den Unterschied zwischen dem Aktivpreis (asset) und dem Passivpreis (liability).
Die Ergebnisse werden quantifiziert, indem wir für verschiedene Zeiten die Normen der MBS-Preise und Anzahl der optimal Vertragsbeendigungen ausrechnen. Dabei handelt es sich nicht nur um Plausibilitätsüberprüfungen, sondern bereits um die Bewertung einer verbrieften Hypothek mit unterliegendem Kontrakt, der gerade diesen vertraglichen Zinssatz hat und dessen Kreditnehmer sich den entsprechenden Transaktionskosten ausgesetzt sieht.

Eigene Beiträge

Diese Arbeit enthält die folgenden zahlreichen eigenen Beiträge zur Bewertung von Mortgage-Backed Securities:

- Regularitätsuntersuchung für das in [28] vorgestellte Modell
- erstmalige Verwendung finiter Elemente, um die ursprünglich in [28] vorgestellte Differentialgleichung zu diskretisieren, anstatt klassischer weitverbreiteter finiter Differenzen
- eine logarithmische Transformation, die Probleme in Hauspreisrichtung behebt, anstelle derjenigen aus [28]
- eine exponentielle Transformation, die zu schwächeren Singularitäten in Zinsrichtung führt als die in [28] vorgeschlagene
- die Benutzung nicht-uniformer Gitter, um mehr Freiheitsgrade dort zu plazieren, wo wir unter mathematischen Gesichtspunkten Probleme erwarten und die in bezug auf unser Modell interessant sind, im Gegensatz zu den uniformen Gittern, wie sie in [28], [29] und [87] benutzt wurden
- Anwendung des Mehrgitterverfahrens in jedem Zeitschritt, um die auftretenden linearen Gleichungssysteme zu lösen
- ein modernes ordnungs- und schrittweitenadaptives Verfahren in der Zeit anstelle des Hopscotch-Verfahrens (vgl. [44]), das in [28] verwendet wurde
- eine Bewertung des Pakets Comsol Multiphysics unter mathematischen, ökonomischen und finanzwissenschaftlichen Gesichtspunkten durch Anwendung auf relevante Probleme aus diesen Bereichen
- eine theoretische und numerische Untersuchung des Konvergenzverhaltens nach Gewichtung mit einer Wahrscheinlichkeitsverteilung, wie es in [28] vorgeschlagen wurde
- ein numerischer Vergleich der asset- und liability-Werte und damit einhergehend eine vollständige MBS Bewertung für realistische Kombinationen aus Transaktionskosten und vertraglichen Zinssätzen mit Hilfe der neuen Verfahren

Introduction

There are many problems in natural sciences, economics and finance which require very elaborate mathematical models in order to be able to cope with them. Particularly in the latter fields stochastic processes are a very important means of such modeling which account for a probability driven development rather than deterministic knowledge of financial markets and their future evolution.
In the case of derivative assets regularly traded at financial markets such models typically consist of stochastic processes governing the behavior of underlying variables, for instance stock prices. Having stated a certain model, the next step is very often the derivation and deduction of a partial differential equation for the value of the derivative. This means that solutions to the models are represented in the shape of a PDE and obtaining the desired price of the derivative then amounts to solving the PDE. Probably the most prominent example is the renowned Black-Scholes equation, which was proposed by Fischer Black, Myron Scholes and Robert C. Merton in 1973 (cf. [8] and [76]) for the price of stock options.
As part of the degree of a model's elaborateness it might (and in fact very often does happen) that an analytical solution to the problem does not exist in closed form. Therefore we have to rely on a numerical approach which approximates the analytical solution in a sufficiently accurate way. In order to secure this accuracy very efficient numerical methods are required which are designed to cope with the specific properties of the problem being dealt with. Due to the complexity of models brought in by the desire to describe reality as exactly as possible this is not an easy task and requires thorough study and knowledge of the underlying mathematical aspects.
The goal of scientific computing consists, besides the aforementioned modeling aspects, of two parts, discretization and solution. The former means approximation of the exact solution which revolves around replacing PDEs by linear systems of equations. The latter one is providing techniques to efficiently solve these linear systems.

Subject Matter of this Thesis

In this diploma thesis we shall focus on mortgage-backed securities (MBSs), which are interest rate derivatives with embedded options. MBSs are fixed-income instruments which pay (usually monthly) coupons to their holders. These cashflows are generated by mortgage payments on (mostly private) residential loans. The aforementioned embedded options arise from the borrower's right to default or prepay at any time during the life of the mortgage. In such a case the stream of monthly payments to the proprietor of a share in a mortgage-backed security is also ceased. In consequence, the value of an MBS depends on the development of two underlyings, namely interest rates and house prices. MBSs are of particular interest to most market participants (banks, financial institutions or other investors) because they constitute a way of securitizing credit risk and reselling it to other customers. In the summer of 2007 their importance rose even further with the spread of the US American mortgage crisis where many mortgage contracts of mortgagors with low credit worthiness were defaulted. The extent of this

predicament and the fact that large parts of US economy are loan financed (for example the majority of mergers and acquisitions) make the American economy very vulnerable to increasing risk premia which even rattled stock markets all across the world.
The key aspect of this diploma thesis therefore becomes to state a setting in which house prices (the relevant variable for possible default) and interest rates (the driving force behind prepayment) are modeled stochastically. From those stochastic processes we will derive a partial differential equation in house prices, interest rates and time for the value of a mortgage-backed security. To this PDE we then apply efficient numerical techniques in order to solve it. Thereby our pricing pattern is very similar to contingent-claim pricing as it is done, for example, in the Black-Scholes model (for references see above) or in its extensions, for instance to stochastic volatility or American-style options (cf. [50] or [79]).

Modeling

In this work we take the findings by Downing, Stanton and Wallace (cf. [28]) as a starting point. We prescribe the Cox-Ingersoll-Ross model (cf. [20]) as the stochastic process governing interest rate movements and the stochastic process proposed in [28] and [29] to account for house price development. The latter process happens to be practically identical to the one used in the Black-Scholes model for stock prices (cf. [51]) except for the parameters, which are estimated separately. This means that both drift (the coefficient of the deterministic part) and volatility (the coefficient of the Brownian motion driven part) are proportional to real estate prices. Besides we mention several alternative interest rate models taken from [60] and [78] and discuss properties of the house price process. Moreover basic definitions of assets like bonds or options are reiterated alongside examples of their purposes and we deal with fixed-income annuities. Among this we derive a formula for the present value of an annuity, the so-called standard annuity formula, and we point out that it can be applied to the fixed-rate mortgages we are considering.
In addition, we carry out the derivation of a PDE governing the price of a mortgage-backed security by application of the famous Itô-formula (see [51], [58] or [78]) to the aforementioned stochastic processes. This PDE turns out to be of second order in both house prices and interest rates without a mixed derivative (the cause for which is that the two Brownian motions appearing in the stochastic processes are assumed to be uncorrelated as proposed in [28]) and first order in time with a constant right-hand side, which is the fixed monthly coupon of the underlying mortgage. Hence we have to deal with a parabolic partial differential equation which is then further examined paying minute regards to regularity statements and the convection-to-diffusion ratio. We find convection dominance near the boundaries and we seek mending it through a combination of transformation to more favorable variables and special stabilization techniques when discretizing.
Concerning regularity we give strong evidence that a solution to our PDE is even after transformation not contained in the Sobolev space H^2 with respect to interest rates and house prices. The focal point around which our arguments revolve is that the differential operator is not uniformly parabolic as demanded in [2], [11], [37] or [69]. Moreover we make a comparison of the results in [69] to those in [27] and apply this our problem.
Apart from the partial differential equation we develop a model for premature mortgage termination either through prepayment or through default. A crucial feature of this model is that it accounts for both financially optimal and financially sub-optimal termination by virtue of so-called hazard functions which allow us to define prepayment and default probabilities (cf. [28] and [87]). Moreover transaction costs are incurred on any termination. Like in [28] we assume them to be different for any borrower. We then incorporate this model into the PDE solution process by adding a probability weighting of the PDE solution and the liability values (including

transaction costs) of either way of termination. In particular, the decision of whether or not termination is reasonable depends on frictions. Based on [51] we will argue that it suffices to perform this weighting only once every month rather than in any single time step.

Discretization

Discretization in space is then provided for the transformed PDE using the finite element method (cf. [57], [70] and [90]) on a non-uniform grid. In doing so we place more nodes and degrees of freedom to those areas we expect to be problematic with respect to regularity and premature termination. A major advantage of using finite elements rather than differences is that it enables us to apply a streamline upwind Petrov Galerkin approach (cf. [19], [63] or [96]), which uses a modification of the finite element testing functions, to our discretization in an attempt to ensure stabilization. Besides we provide a reformulation of the PDE in a weak form and discuss standard error estimates for finite element discretizations in the elliptic case based on [11] as well as summary of semi-discrete estimates for parabolic equations based on [64].
In time we make use of the so-called backward differentiation formulas (cf. [80]) and we conduct time-stepping adaptively in order (up to five) and step length. The illustrated criteria are based on those from [40] and [86]. Thus we seek to take time steps as large as possible and of as high an order as possible. Furthermore intended is avoiding the rejection of a step as often as possible, which is also taken care of by the above criteria. Apart from these aspects of choosing the steps we discuss the discretization error of these formulas, show local convergence (consistence) and state criteria under which overall convergence can be ensured as it is outlined in [80].
Moreover we prove that the probability-weighted solution converges in L^2 if the unweighted PDE solution converges in L^2. In addition, we point out that this ensures that the whole solution process, which consists of solving a PDE and weighting the solution afterwards for all 360 months of the 30-year mortgage, actually works.

Solving the Model

When it comes to solving the discretized problem we try to be very efficient, too, as we use a 4-level V-cycle multigrid iteration (cf. [11], [45] and [64]) which is pre-conditioned by the GMRES iterative solver (cf. [82]). In particular, the GMRES scheme is subjected to pre-smoothing itself using an incomplete LU-factorization ([42] and [83]) and the SORU iteration is used for post-smoothing purposes in the multigrid pattern. The benefit of this solver is a scheme which is optimal in the number of spatial unknowns, i. e. it is of order $\mathcal{O}(N)$ if N is the number of degrees of freedom in every time step. This is in fact very good taking into consideration that the classic and renowned Gaussian elimination scheme is of order $\mathcal{O}(N^3)$ (cf. for instance [47], [48] or [89]). Nonetheless we utilize a modified Gaussian elimination, namely the UMFPACK solver, as an exact coarse grid solver. The solver is taken from [21], [22], [23] and [24].
The numerical computations are conducted in the commercial package Comsol Multiphysics which provides a wide variety of different solvers, solver settings and smoothers to choose from (see, for example, [96], [97] and [98]). Moreover the implementation can be assumed to be fast and reliable.

Numerical Results

Furthermore we test the discretization by ascertaining the convergence rates of our scheme both in space and in time. In space we find a worse rate than expected on account of the discussed standard error estimates. The reason for this is that our solution violates the regularity prerequisites stated in the above standard error estimates. We additionally discern that the restrictions

in the implemented SUPG in Comsol Multiphysics (first order elements, constant PDE coefficients only) are too severe to overcome so that the non-SUPG solution actually happens to produce better results. In time we can attain good convergence rates of approximately 5.
In addition, the same tests are performed for the probability-weighted solution after one month and we find that our convergence rates are maintained. Both convergence examinations are considered for realistic monthly coupons (computed from the contractual interest rate which is sensibly chosen as 5% in our test setting).
Having established the above convergence rates we carry out computations over a complete mortgage lifespan of 30 years for varying contractual interest rates (5.0 %, 8.5 % and 10.0 %) and three levels of transaction costs (1.4 %, 8.6 % and 18.5 %). In doing so we keep track of optimal termination very closely and compare the termination data for those various combinations of costs and coupons. To stress the results many graphical illustrations are provided which underline the dependence on time, transaction costs and contractual interest rate. Furthermore threshold levels for optimal termination are fathomed out and evaluated with respect to the statistical data of how often termination is advisable. These considerations come along with a comparison of asset and liability prices.
The results are quantitatively stated by determining the norms of the MBS prices for several different times and the actual termination numbers. At the same time those results are not only plausibility checks but rather a valuation of a mortgage-backed security with an underlying mortgage contract of the respective contractual interest rate and a mortgagor facing the respective costs upon termination.

Own Contributions

There are various new contributions to the valuation of mortgage-backed securities proposed in this thesis. They are the following:

- a consideration of regularity of solutions of the model propounded in [28]
- a finite element approach to discretize the PDE originally stated in [28] instead of classic and widely used finite difference schemes
- a logarithmic transformation to remove problems in the house price variable direction rather than the one used in [28]
- an exponential transformation rather than the one proposed in [28] leading to less severe singularities in the interest rate variable
- the use of non-uniform grids in order to place more degrees of freedom in those areas which are mathematically problematic and more interesting in terms of the model in contrast to the uniform meshes used in papers [28] and [29] or in [87]
- the application of the multigrid iterative linear system solver to the discretized system in every time step
- the use of modern adaptive time-stepping in both order and stepsize rather than the Hopscotch method (cf. [44]) used in [28]
- an estimation of the Comsol Multiphysics package under mathematical, economical and finance aspects rather than engineering by applying it to relevant problems from those areas
- both a theoretical and numerical examination of how convergence aspects are maintained during probability weighting as suggested in [28]

- a numerical comparison of asset and liability values and thus a complete MBS valuation for several reasonable combinations of frictions and contractual interest rates using the improved numerical schemes

Outline of this Thesis

More detailedly this work is organized in five chapters as follows.
Chapter 1 is devoted to briefly introducing those financial assets we are going to deal with in this diploma thesis. This is attained by providing basic definitions alongside examples for their purposes. Moreover we explicitly discuss mortgages and derive several useful properties before finally getting acquainted with mortgage-backed securities. We expound what they are, to whom they are of importance and what their economic impact is.
In *Chapter 2* we in detail describe the financial and mathematical model which is the fundament of this work. We start with the consideration of different approaches to the MBS valuation problem and state our approach. Then the stochastic processes covering interest rates and house prices are described before outlining our approach to transaction costs for both a single mortgagor and a pool of borrowers. Moreover we point out that one has to meticulously distinguish between the asset and the liability price of an MBS. Next we derive our PDE and we impose boundary and initial conditions which are financially sensible and yield unique solvability of the initial boundary value problem. Afterwards we take care of transformations by discussing two transformations and their advantages. Eventually certain properties of our initial boundary value problem are outlined and we account for sub-optimal behavior by introducing a probability distribution for the behavior of mortgagors who might terminate the mortgage although they should not and who might not cease payments although they should.
The subject matter of *Chapter 3* is everything related to discretization. We start by reformulating the strong-form partial differential equation in a weak form. Next we derive a semi-discrete problem carrying out the replacement of the variational equation by a finite element approximation while the time dimension remains untouched until our time-discretization pattern is outlined and some of its convergence properties are shown. Moreover we provide estimates of the finite element approximation in the elliptic case, deal with the effect of numerical quadrature instead of exact integrals in the variational equation and we give error estimates for the semi-discrete problem. Then the convergence results are extended to probability weighting, which is the main theoretic result of this thesis. Finally we introduce numerical measures taken to change the infinite domain and cope with convection effects and we conclude the chapter by discussing the regularity of solutions.
In *Chapter 4* we deal with the solution process of our discrete initial boundary value problem. This starts by rewriting the semi-discrete problem in a differential algebraic system which is solved by a Newton method. We then explain our order and stepsize adaptiveness and the means by which both are determined. Concludingly the actual solver for the appearing linear systems, namely a multigrid method, is presented and pre- and post-smoothers as well as our coarse-grid solver are described.
In *Chapter 5* we finally deal with numerical results. First we explain how we measure both the spatial and time errors and which rate of convergence we can expect. We then move on justifying that our measurement procedures work by applying them to well known problems (namely the heat equation, for which even an analytical solution is examined, and a diffusion-convection equation with smooth enough solutions). We then turn our attention towards the actual MBS equation which is solved for the time of one month. Upon doing so we determine the convergence rates for both the SUPG case and the non-SUPG case. Moreover we deal with the time error. Afterwards convergence is tested and confirmed for the one-month probability-weighted solution. We do so for two different transaction costs and again both in space and time. Eventually we deal with an MBS over its entire lifespan of 30 years. This is done for three different transaction

costs and three different contractual interest rates. In the course of doing so we illustrate a lot of data about how often it is optimal to prematurely terminate and we distinguish between the asset and liability case.
Concludingly there is an unnumbered chapter providing summary of the results of this thesis and giving an outlook on possible future and forthcoming research in the area of pricing mortgage-backed securities.
In the appendices we summarize certain background information about the techniques from stochastic calculus (cf. Appendix A) and those from functional analysis (cf. Appendix B) which are used in this thesis. In Appendix A we distinguish between stochastic processes (cf. A.1) and stochastic differential equations (cf. A.2). In the latter section we show existence of solutions to the stochastic differential equations (SDEs) by which our stochastic processes are defined.

Acknowledgements

I would like to thank Dr. Matthias Reimer (formerly Deutsche Postbank AG, now Deutsche Bank AG / DWS) for introducing me to this problem during my internship at the former company in 2005 and creating a lot of interest in financial markets. Moreover I express my gratitude towards Prof. Dr. Michael Griebel for giving me the opportunity to realize this diploma thesis and support throughout its development. In addition, I received support from Dr. Thomas Gerstner for which I am also grateful. Furthermore I wish to thank Prof. Dr. Rolf Krause for counsel, support, co-supervision of this thesis and originally introducing me to numerics. Moreover my thanks go to Dominique Löbach for helping and advising me in questions of regularity.
And above all I wish to thank my parents for all their support.

Chapter 1

Mortgages, Options, Bonds and Securities

1.1 Basic Definitions of Options and Bonds

In this section I briefly define several financial instruments we are going to deal with in this work. First, I will give a definition of options.

Definition 1.1.1 (Option) *The holder of an option bears the right but not the obligation to buy (call option) or sell (put option) an asset at a pre-determined strike price K at a certain time (European style) or within a certain timespan (American style) in the future. Such options are often referred to as plain vanilla options.*

The underlying asset can be virtually everything. The most prominent examples are stocks or stock indices but the asset can as well be taken from commodities, such as gold or oil, or even from currencies or any other financial instrument. In this work, however, we are not going to face plain vanilla options but rather options embedded in mortgage contracts (details will be expounded in Section 2.1).
In the absence of transaction costs plain vanilla options yield the following profits P_{call} and P_{put} to their holders at maturity T:

$$\begin{aligned} P_{call}(S,T) &= \max(S(T) - K, 0) \\ P_{put}(S,T) &= \max(K - S(T), 0), \end{aligned}$$

where $S(T)$ denotes the price of the underlying asset at time T. These payoff functions are also illustrated graphically in Figure 1.1.
This reflects that the asset can be bought at a price K and sold at $S(T)$ in the case of a call or vice versa in the case of a put. Unless yielding a profit an option is not exercised. Hence the value of an option at maturity is known but the present value for a time $t < T$ is still unknown. In the renowned Black-Scholes model for option pricing a fair present value is established as the expectation (under a suitable probability measure) of the discounted profit, i. e. in the case of a call

$$\text{fair value} = E[e^{-rT} P_{call}].$$

In the 1970s Black, Scholes and Merton (cf. [8] and [76]) tackled the pricing problem by deriving a partial differential equation for the fair value, which can be solved analytically for European plain vanilla options. A different approach would be to directly deal with the expectation in a Monte Carlo setting.
So far we have described what options are and how a price can be obtained. What we have yet

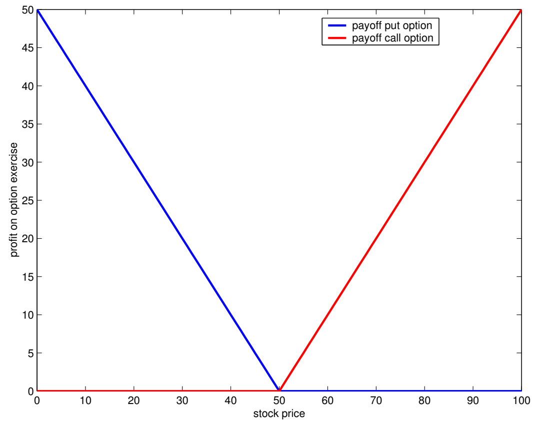

Figure 1.1: Payoff functions at maturity for both a call and a put with strike price $K = 50$

to elucidate is what purpose can be behind the use of options. There are three fields in which options are made use of, namely hedging, speculation and arbitrage. Speculation is a simple bet on rising or falling markets which can consist of various combinations of different kinds of options. Classic examples of such strategies include bull spreads (rising) or bear spreads (falling) which are besides others detailedly expounded in [51].
Hedging means protection against certain market movements and arbitrage means exploiting market inefficiencies to gain riskless profit. The following two examples for arbitrage and hedging are also taken from [51]:

Example 1.1.2 (Arbitrage) *If, for instance, a stock is traded at* $132 *in New York and at* €*100 in Frankfurt with an exchange rate of* $1.3500 *per Euro an arbitrageur could simultaneously purchase* 100 *shares of the stock in New York and sell them in Frankfurt. This would yield a risk-free profit of*

$$100\,(1.35\$\,100 - 132\$) = 300\$.$$

Example 1.1.3 (Hedging) *Concerning hedges we might think of an American company due to pay* €*1,000,000 to a European supplier in* 90 *days. It could now buy a call option to acquire* €*1,000,000 at an exchange rate of, for example,* $1.4000 *in* 90 *days. If the exchange rate happens to be above* 1.4000 *the company exercises the option and else it buys the money in the market and lets the option expire worthless. Thereby the company is safe against adverse movement while benefiting from favorable movement. But for this insurance the purchasing price of the options has to be paid.*

The other very important financial instrument we have to deal with is an interest rate derivative, namely bonds. In its simplest form a bond is defined in the following way:

Definition 1.1.4 (Bond) *A bond is a loan over a face value V by the issuer in order to borrow money for a certain time T. In exchange he or she pays a coupon C to the borrower at pre-determined intervals and repays the face value V at maturity T.*

Typically bond issuers are companies or governments, mostly countries or states but there are also municipal bonds. The coupon C to be paid for borrowing the money is determined by the issuer's credit-worthiness, i. e. the more likely he or she is to not being able to meet the repay obligation the more expensive the bond becomes. Concerning the lifespan no typical time can be stated since there are bonds far exceeding thirty years but also bonds of less than five years. The coupon C is usually determined by an interest payment, for instance 6% of the face value per year and it can be paid annually, every half year or at any other pre-determined times.

Compared to options obtaining a fair price for bonds is a relatively simple matter because according to [51] the present value, which is the sum of the discounted future payments, would be appropriate.

Example 1.1.5 (Present Value of a Bond) *Consider a five-year bond over* \$1000 *with 6% annual coupon payment. Moreover assume the spot rate* $r = 2.5\%$ *to be equal for any of the five years. Then the price of the bond is computed as follows:*

$$\frac{60\$}{1.025} + \frac{60\$}{1.025^2} + \ldots + \frac{60\$}{1.025^5} + \frac{1000\$}{1.025^5} = 1162.60\$.$$

According to [51] the spot rate is defined as follows:

Definition 1.1.6 (Spot Rate) *The n-year spot interest rate is the interest rate on an investment that is made for a period of time starting today and lasting for n years. The investment considered should be a pure n-year investment with no intermediate payments. This means that all the interest and the principal is repaid to the investor at the end of year n. The n-year spot rate is also referred to as the n-year zero-coupon yield. This is because it is, by definition, the yield on a bond that pays no coupons.*

In addition to spot rates there are forward interest rates for periods of time in the future.
The reason, why the above definitions were given is that we are going to price an interest rate derivative which can be interpreted as a combination of options and bonds.

1.2 Treating Mortgages as Annuities

In this section I propound and derive the standard annuity formula and elucidate why and how it can be applied to mortgages. Usually an annuity is a constant annual payment made to its holder.

1.2.1 The Standard Annuity Formula

The standard annuity formula is a formula for the present value of an annuity. In [28] the authors stated the formula without proof.

Lemma 1.2.1 (Standard Annuity Formula) *Assume $C > 0$ to be an annual payment made n times and further denote by $r > 0$ the risk-free interest rate. Then the present value B of the annuity satisfies:*

$$B = \frac{C}{r}\left[1 - \frac{1}{(1+r)^n}\right]. \tag{1.1}$$

Proof - Lemma 1.2.1:
For the present value of an annuity the following holds:

$$
\begin{aligned}
B &= \frac{C}{1+r} + \frac{C}{(1+r)^2} + \ldots + \frac{C}{(1+r)^n} \\
&= C + \frac{C}{1+r} + \frac{C}{(1+r)^2} + \ldots + \frac{C}{(1+r)^n} - C \\
&= C \underbrace{\frac{1-\left(\frac{1}{1+r}\right)^{n+1}}{1-\frac{1}{1+r}}}_{\frac{1+r}{r}\left(1-\left(\frac{1}{1+r}\right)^{n+1}\right)} - C \\
&= C\left[\frac{(1+r)(1-\left(\frac{1}{1+r}\right)^{n+1}) - r}{r}\right] \\
&= \frac{C}{r}\left[\left(1-\left(\frac{1}{1+r}\right)^{n+1}\right) + r\left(\left(1-\left(\frac{1}{1+r}\right)^{n+1}\right) - 1\right)\right] \\
&= \frac{C}{r}\left[1-\left(\frac{1}{1+r}\right)^{n+1} - r\left(\frac{1}{1+r}\right)^{n+1}\right] \\
&= \frac{C}{r}\left[1-(1+r)\left(\frac{1}{1+r}\right)^{n+1}\right] \\
&= \frac{C}{r}\left[1-\left(\frac{1}{1+r}\right)^{n}\right] \\
&= \frac{C}{r}\left[1-\frac{1}{(1+r)^n}\right].
\end{aligned}
$$

Altogether this establishes (1.1).

□

This formula can be slightly adjusted so that one can apply it to monthly payments. To do so we presuppose that the monthly interest rate $\tilde{r}$ and the monthly payment $\tilde{C}$ satisfy:

$$
\begin{aligned}
\tilde{C} &= C/12 \\
\tilde{r} &= r/12.
\end{aligned}
$$

When considering monthly payments there are twelve times as many as in the annual case so that we eventually arrive at:

$$
B = \frac{C}{r}\left[1-\left(\frac{1}{1+\left(\frac{r}{12}\right)^{12\,n}}\right)\right].
$$

In fact, the assumptions about $\tilde{r}$ and $\tilde{C}$ are a simplification because depositing money for one year at rate r and receiving interest at the end of the year yields a different benefit than depositing money for a year with monthly interest payments at a rate $r/12$.

1.2.2 Application to Fixed-Rate Mortgages

In this subsection I discuss why fixed-rate mortgages can be regarded as annuities. Intuitly this seems quite plausible because a fixed-rate mortgage consists of fixed monthly payments. But in

the following this statement shall be justified systematically.
First, we note a recursion formula for the remaining principal to be repaid on a mortgage. Assume a mortgage of N payments with a contactual interest rate c where at time n the remaining principal is R_n so that the following recursion formulae hold:

$$\begin{aligned} R_n &= (1+c)\, R_{n-1} - C \\ R_N &= 0 \qquad \text{which reflects amortization.} \end{aligned}$$

Lemma 1.2.2 *Under the assumptions about the mortgage outlined above we can state that the initial debt amounts to:*

$$R_0 = \frac{C}{1+c} + \ldots + \frac{C}{(1+c)^{N-1}} + \frac{C}{(1+c)^N}. \tag{1.2}$$

In addition, the same formula can be used to not only calculate the original debt but also the remaining principal at any given time $n < N$ by replacing N by $N - n$.

Proof - Lemma 1.2.2:
Let us start with the additional statement: At time $n < N$ there is an outstanding debt and one could sign in to a new loan over that sum which expires at the same time as the old loan. Then the new one has time $N - n$ to maturity and the same payments as the old loan so that the remaining principal on the old loan is the initial principal on the new one, i. e.

$$\begin{aligned} R_n &= \tilde{R}_0 \\ &= \frac{C}{1+c} + \ldots + \frac{C}{(1+c)^{N-n-1}} + \frac{C}{(1+c)^{N-n}}. \end{aligned}$$

The main part of the lemma is proved by mathematical induction.
For the basis, i. e. $N = 1$, we assert:

$$R_0 = \frac{C}{1+c}.$$

Since there is only one payment to be made and this payment is the amount of the loan plus interest the statement is apparently true.
Next we have to perform the induction step assuming assertion (1.2) to be true for N and then deducing the statement for $N+1$. Applying the recursion formula and the assumption to $N = 1$ we obtain

$$\begin{aligned} R_0\,(1+c) - C &= R_1 \\ &= \frac{C}{1+c} + \ldots + \frac{C}{(1+c)^{N-1}} + \frac{C}{(1+c)^N}. \end{aligned}$$

It follows:

$$R_0 = \frac{C}{(1+c)^2} + \ldots + \frac{C}{(1+c)^N} + \frac{C}{(1+c)^{N+1}} + \frac{C}{1+c},$$

which is exactly the desired statement.

□

Moreover we see that the formula in the lemma is the same formula as derived before for the present value of an annuity. Therefore we can state the following:

- the remaining principal on a fixed-rate mortgage can be computed using the standard annuity formula
- the remaining principal on a fixed-rate mortgage would be the present value of an annuity paying the same coupon if the current rate of risk-free interest was the contractual interest rate of the mortgage.

1.3 About Mortgage-Backed Securities

In this section I finally elucidate what mortgage-backed securities are. They are bond-like financial instruments and belong to the wider class of so-called asset-backed securities with the asset being mortgage payments.
What does that mean? An ordinary bond is basically a stream of future payments to the bond-holder. In the case of a mortgage-backed security these payments are provided or backed, as it is often called, by the usually monthly payments a mortgagor makes to the financial institution, from which he or she has borrowed money. These mortgage loans themselves can be backed by private homes or any commercial sites. Until now there is nothing special about them, but they have one very important additional feature: The mortgagor possesses the right to repay his or her outstanding debt prematurely at any time, which means that the duration of the cashflows to the holder of the mortgage-backed security is unknown. In addition, as it is the case with any bond, the holder faces the credit or default risk of the original mortgage loan. Therefore MBSs are more complex than plain vanilla bonds.
Mortgage-backed securities are created by financial institutions who intend to pass on or sell credit risk to customers. This is done by pooling several residential mortgages and dividing this pool into a certain number of shares then retailed. So by buying x % of the pool one acquires a claim upon x % of the payments (interest and principal) by the mortgagors. Therefore the holder of such a share is fully exposed to default risk. In the United States there are government institutions like the Federal National Mortgage Association (FNMA) or the Government National Mortgage Association (GNMA) who provide protection against credit risk in exchange for a servicing fee usually about 0.5 % of the cash flow. More information on that can, for instance, be found in [28], [29], [51], [87] or [88].
So far one might get the impression that mortgage-backed securities do not differ from regular government-issued fixed-income securities. But there is one major difference and that difference is the embedded prepayment option held by the borrower which according to [51] is a very typical feature in US mortgage contracts.
Altogether a secondary market to the primary mortgage market is created, in which credit risk is traded. As stated in [28] and [29] it turned out that this market is actually one of the fastest growing bond markets in the US. Therefore pricing mortgage contracts as well as possible is required and important.

1.3.1 Different Types of MBSs

Besides those standard passthrough mortgage-backed securities there are more elaborate ones like collateralized mortgage obligations (CMOs) or IOs and POs which are abbreviations for "interest only" and "principal only".
The latter ones belong to the class of stripped MBS which means that the cashflow of principal and interest is separated and channeled to different investors. In the case of POs the investor receives a certain amount of money and additionally benefits from high prepayment rates because the capital is returned early. The opposite is true for IOs because the total amount of cashflow is uncertain since interest payments are stopped if prepayment occurs. Hence the investor speculates on low prepayment rates.
In a CMO, for example, investors are divided into several classes, say A, B and C. First all class A investors are paid off, then class B and so forth. In consequence, all three classes differ in the risk to which they are exposed and hence a financial institution can match different investors' different intentions better. Further details can be found in [51].
In this work we shall focus on the classic passthrough MBSs, though.

1.3.2 Economic Importance

The next item I would like to stress is that MBSs are not only interesting by themselves as relatively new financial products. They also reflect important economic aspects from a wider perspective.
Mortgage-backed securities are closely linked to a major risk in US economy. This is because a large part of US economy is credit-financed by private consumers. And since these loans are mostly collateralized by houses, namely private homes, credit-worthiness of private consumers is particularly vulnerable by a change in house prices and interest rates. This is, for instance, pointed out in the articles [49] and [71] by Heaton and Leisinger respectively.
Since in recent years house prices in the United States tremendously rose many more and larger loans could be taken. This was fueled even further by historically low interest rates. How low interest rates actually were and still are is probably best illustrated in Figure 1.2 where the federal funds rates are plotted for roughly speaking the past fifty years.

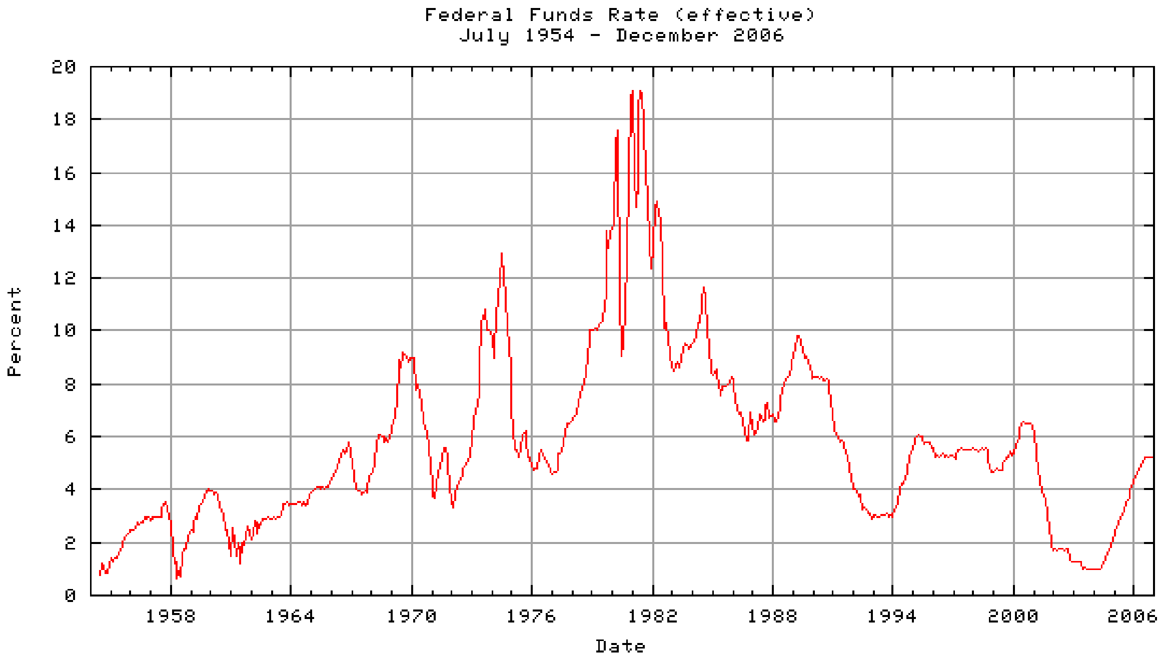

Figure 1.2: Historical Federal Fund Rates (taken from www.wikipedia.org)

In early 2007 the situation began to deteriorate and hamper US economy when house prices no longer increased and even declined. Moreover interest rates have risen over the past two years. The immediate effect was mostly in variable-interest-rate contracts from the so-called sub-prime sector, which is what the high-risk credit market is called. When interest rates rose those high-risk mortgagors could no longer match their increasing interest obligation and they were forced to default houses of less value than assumed before. Figure 1.3 gives an idea of the extent of this development. To better understand the economic implications it might be helpful to consider who the investors in mortgage-backed securities are. According to [49] these are not only financial investors like hedge funds, which one might have expected, but also less aggressive pension funds or even governments like China who seek diversification of their exposure by reducing currency reserves. The motivation behind investing in MBSs is the higher coupon compared to classic bonds. The reason for the higher coupon is the higher risk of an individual mortgage.

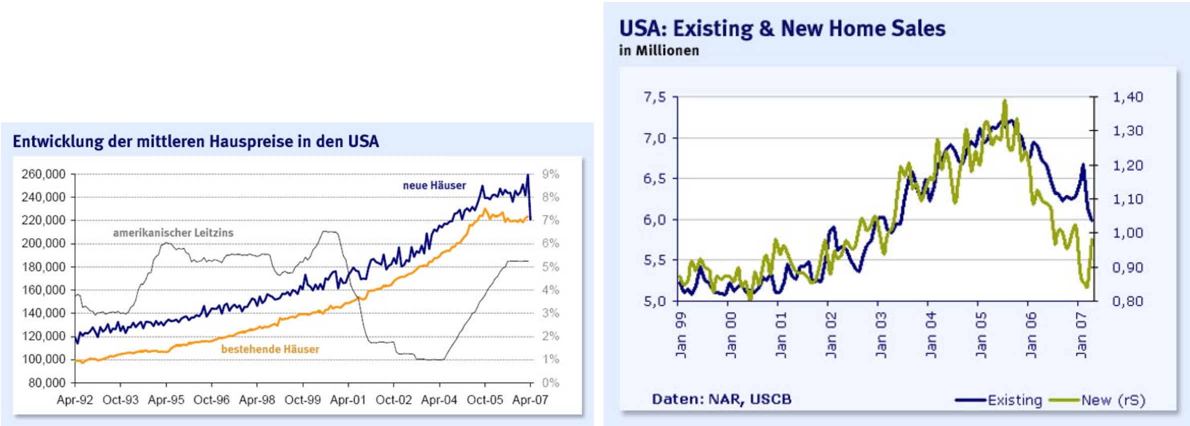

Figure 1.3: Left: development of US house prices and interest rates, right: development of US home sales (both taken from www.faz.net)

In the summer of 2007 it became apparent how large the market for trading of credit and mortgage risk is when a crisis broke out in the US. During this crisis several financial institutions ran into bankruptcy because of defaulted mortgages and hedge funds heavily invested in this market faced huge losses.
Moreover stock markets plummeted, too, because market participants were afraid that this quandary might slow down the development of the whole US economy. For example, the renowned Dow Jones index lost about $1,000$ points in comparison to its mid-July all-time high of over $14,000$ points. In addition, these problems spread all across the world with the German DAX, for instance, plunging from $8,150$ to about $7,300$ points.
The above and more information about that can be found in [99] and [100].

Chapter 2

The Model

2.1 Approach to the Problem

In the literature on pricing mortgage-backed securities two fundamentally different approaches have emerged, structural models and reduced-form models.
As outlined in [28] and [29], characteristic about structural models is the rational minimization of the liability value by the borrower in response to changes in the underlying state variables such as interest rates or house prices. As a result pricing becomes very similar to standard contingent-claims like American-style options. This approach was first carried out by Dunn and McConnell in 1981 (cf. [34] and [35]). But they did neither account for transaction costs nor for suboptimal option exercise. Though over the course of time these features were gradually added and prepayment and default were both considered at the same time. See, for example, [60], [61] and [62] or [87], [88]. The former ones all deal with prepayment and default but at least in part do not estimate results produced by their models empirically and the latter ones account for heteregeneous transaction costs.
In contrast to structural models reduced-form models do not explicitly simulate the underlying state variables. Instead termination is considered to be a function of a set of variables influencing the borrower's behavior, such as house prices and interest rates on the one hand but also personal factors like divorce or credit-worthiness on the other hand. More precisely at each time premature termination is considered to occur with a certain probability depending on the aforementioned factors. Examples for reduced-form models are [84], [25] and [26] in the case of mortgage valuation. According to [28] there are similar approaches for corporate risky debt, for instance [18], [30] or [53].
Among the advantages of reduced-form models is the relatively good ability to fit historical data about mortgage-termination. However, there is a lack of forecasting credibility because mortgage prices themselves or proxies for the borrower's options are used instead of directly valuing them.
In this diploma thesis, though, I shall focus on a structural approach.

In order to value mortgage-backed securities we have to define a mathematical model. Within this model we will derive and solve a partial differential equation for the desired value.

2.1.1 General Remarks

First we should mention what properties we expect our model to incorporate. As explained above in Section 1.3 an MBS consists of an underlying bond (the stream of cashflows) with additional prepayment and default options. The option of prepayment can be interpreted as a

call option on the underlying bond and the default option as a put on the house.
So to the holder of an MBS its value M_t is

$$M_t = B_t - V_p - V_d,$$

where B_t denotes the value of the underlying bond and V_p and V_d are the values of the options respectively. As part of the valuation of the MBS one has to find out when it is financially optimal to exercise one of the options, that is to prepay or default.
Heuristically speaking prepayment makes sense when interests rates drop low enough so that refinancing (picking up a new loan and using it to repay the other) is less expensive than continuing to make the original payments. Conversely defaulting is optimal when house prices are low enough, because you give away your house but save the remaining payments.
Our choice is a so-called two-factor model as proposed in [28] with the two factors being the rate of risk-free interest and the house prices. We consider this to be a sensible approach as it accounts for both prepayment and default risk.
To be more exact we choose two stochastic processes r_t and H_t for risk-free interests and house prices, which we will specify in Section 2.2. In order to understand the thresholds for optimal prepayment and default we do not need to know what the processes actually look like.
Moreover I would like to mention that prepaying and defaulting is associated with transaction costs. As I did with the interest rate and house price processes before, I will only explain in Section 2.3 how they are modelled. So for our current purposes we shall be content with knowing that there are proportional transaction costs X_d and X_p for default and prepayment respectively.

2.1.2 Optimality of Default

In the case of default the homeowner hands over his or her house to the financial institution and in exchange stops making the remaining scheduled payments on the mortgage. In consequence, it is optimal to exercise the default option if

$$M_t \geq H_t\,(1 + X_d). \tag{2.1}$$

Here we have to heed that for the house prices dropping below the value of the underlying bond does not suffice because the options also have values to their owners, which they forfeit by exercise. Moreover the two options are not independent of one another since exercising one precludes exercise of the other at a later time. So the house prices have to drop low enough to compensate for the forfeit of both options.

2.1.3 Optimality of Prepayment

The case of prepayment is slightly more complicated than default because it is not one of the underlyings that has to drop far enough but an expression depending on the interest rate. The idea of prepayment is that the mortgagor initiates a new loan for the value of the remaining principal to be paid. If the time to maturity remains the same he or she will have to pay a smaller monthly coupon $\tilde{C}$ due to the declined interest rates. More precisely a call option on the underlying bond is exercised, which is in-the-money because the lower interest rates are the higher is a bond's value. So we have to ponder under which circumstances doing so is financially optimal.
First we recall from Section 1.2 how to calculate the remaining principal F of a loan at time t, i. e. time $T - t$ to maturity, and deduce:

$$F(t) = \frac{C}{c}\left(1 - \frac{1}{(1 + \frac{c}{12})^{12(T-t)}}\right).$$

Optimal prepayment now means that the payment on refinancing, namely $(1 + X_p)F(t)$, is less than the values of the liability and the options, which again has to be taken account of since option exercise comes along with forfeiting possible future exercise of either option. So we procure the following condition:

$$M_t \geq (1 + X_p)F(t). \tag{2.2}$$

We should heed that the right-hand-side of the inequality is actually independent of interest rates and that a decrease of interest rates causes an increase in the value of M.
Furthermore I would like to point out that the optimality condition indeed has the desired property, that there are smaller payments to be made if the interest rates are small enough. But on the other hand it cannot be ensured that there is always some r such that prepayment is optimal because if transaction costs are high enough they will always exceed the effect of lower interest.
In addition, we make the facilitating assumption of not considering curtailments (partial prepayments). In [87] an argument is given why they can be virtually ruled out.

2.2 The Interest Rate and House Price Models

In this section I first describe the stochastic process we use for interest rates and then move on elucidating the one for house prices.
There are various models governing interest rate processes among which we have decided to use the one proposed by Cox, Ingersoll and Ross in 1985 (cf. [20]). It is probably the most prominent, renowned and utilized one in literature dealing with mortgages and mortgage derivatives. In this model interest rates are governed by the stochastic differential equation

$$dr_t = (\kappa(\theta_r - r_t) - \eta r_t)dt + \Phi_r \sqrt{r_t}\, dW_r. \tag{2.3}$$

We note that this is a mean-reverting process, where θ_r is the long-term mean and κ the rate of reversion to it; η is the so-called price of interest rate risk and Φ_r the proportional volatility in interest rates. The stochastic process W_r is assumed to be a standard Brownian motion.
Among the advantages of this process is the property that if the parameters satisfy certain conditions it is always non-negative if so initially, i. e. $r_0 > 0$. However, zero is not an absorbing barrier as interest rates can be positive at a time $t_2 > t_1$ even if $r_{t_1} = 0$ holds.
Besides the Cox-Ingersoll-Ross model there are several other famous one-factor models, i. e. those models which depend on merely one source of randomness. For example, the Vasicek model

$$dr_t = (a - br_t)dt + \sigma dW_t,$$

the Brennan and Schwartz model

$$dr_t = (a - br_t)dt + \sigma r_t dW_t$$

or Merton's model

$$\begin{aligned} r_t &= r_0 + at + \sigma W_t \\ dr_t &= a\, dt + \sigma\, dW_t. \end{aligned}$$

Details and references for those models can be found in [78]. A generalization of both the CIR-model and the Vasicek-model is formed by the Hull and White model

$$dr_t = a(b - cr_t)dt + \sigma r_t^\beta dW_t, \qquad 0 \leq \beta \leq 1.$$

One possible enhancement of the one-factor interest rate models would be to not only incorporate the spot rate r but also a long-term rate l, which makes the description of the actual term-structure more realistic since there are more degrees of freedom. The processes might look as follows:

$$\begin{aligned} dr &= [a_1 + b_1(l - r)]dt + \sigma_r \, dW_r \\ dl &= l(a_2 + b_2 r + c_2 l)dt + \sigma_l \, dW_l. \end{aligned}$$

Disadventageous is the consumption of more computing resources particularly in the presence of a third process governing the house prices because then one actually has to deal with a three-dimensional problem instead of a two-dimensional one. As pointed out in [60] and [87] it has nonetheless been used for instance by Schwartz and Torous in [84] but only when not considering default risk.

Having discussed our choice of the interest rate model I next wish to outline how the house prices are modelled. We assume H_t to evolve according to a geometric Brownian motion

$$dH_t = \theta_H H_t dt + \Phi_H H_t \, dW_H.$$

Here θ_H denotes the expected increase in house prices and Φ_H the corresponding volatility. To understand the next step it is helpful to briefly take notion of the fact that this process is practically the same as the one used for stock prices in the standard Black-Scholes model (cf. [8], [51], [76] or [78]) .
In the case of shares there is usually a cash flow, the dividend, to the proprietor of the stock. Considering house prices there is no dividend but a flow of rents which we denote by q_H. Moreover we consider a market price of risk λ_H for the house prices as well as the principle that the economy behaves risk-neutrally after adjusting, which means that the expected return equals the risk-free interest rate, i. e.

$$\theta_H - \lambda_H \Phi_H + q_H = r,$$

where $\theta_H - \lambda_H \Phi_H$ is the risk-adjusted appreciation rate of the real estate. Hence we obtain the following risk-adjusted stochastic process for the house prices:

$$dH_t = (r_t - q_H)H_t dt + \Phi_H H_t \, dW_H. \tag{2.4}$$

More details on this standard argument and the market price of risk can be found in [51] and [60].
Furthermore a very important feature of this process which will be exploited when imposing boundary data in Section 2.6 is that zero is an absorbing barrier. This can be seen by looking at the stochastic differential equation which tells us that if $H = 0$ for some t, dH also vanishes and thus H remains zero for $\bar{t} > t$.

Having discussed several alternative models we have decided to use the following ones in this diploma thesis:

$$\begin{aligned} dH_t &= \theta_H H_t dt + \Phi_H H_t dW_H \\ dr_t &= (\kappa(\theta_r - r_t) - \eta r_t)dt + \Phi_r \sqrt{r_t} \, dW_r. \end{aligned}$$

They are calibrated by the following set of parameters:

$$\begin{aligned} q_H &= 0.025 \\ \Phi_H &= 0.085 \\ \kappa &= 0.13131 \\ \theta_r &= 0.05740 \\ \Phi_r &= 0.06035 \\ \eta &= -0.07577. \end{aligned}$$

Concerning parameter estimation for these models as well as the parameters appearing in Section 2.3 on transaction costs we have to state that the results are taken from [28] and [29] and we did not carry out an estimation ourselves. In the two mentioned works a least squares approach was followed to fit historical market and termination data, thus leading to the problem of minimizing the non-linear functional

$$\chi(\Theta) = \sum_{i=1}^{N} \sum_{t=1}^{T_i} \left(\omega_{it} - \hat{\omega}(\Theta)\right)^2,$$

where Θ is the vector of all parameters to be estimated, N the number of mortgage pools observed and t the time of observation. It was solved by the Nelder-Mead downhill simplex algorithm.

2.3 Modeling of Transaction Costs

Concerning transaction costs there are various features we have not addressed yet. In this section I describe what we expect and how it is realized.
Until now we have merely considered mortgage holders who always act in a financially optimal way. But in reality this is certainly not the case as there are borrowers who repay their loans prematurely for non-monetary reasons such as relocation or divorce etc. and loaners who do not exercise their prepayment options although reason would suggest to do so. Moreover in a strictly rational world all mortgagors would optimally prepay once it becomes optimal for one borrower to do so. To amend these shortcomings we expect our model to at least satisfy the following conditions:

1. Borrowers face transaction costs when they opt to default or prepay
2. Different borrowers face different transaction costs
3. Prepayment and default occur depending on a certain probability rather than pure financial reasons.

Since we never consider single mortgages but always a pool of very similar mortgages (i. e. common time to maturity, common or just slightly different contractual interest rates etc.) we have to distinguish between two different probability aspects. The first is the likeliness of a single borrower to terminate his or her contract given his or her personal transaction costs and the second is the distribution of transaction costs across the pool of mortgages.
This is the same model as used in the papers [28], [29] and [87]. Moreover it allows us to deal with a pool of mortgages without knowing how many contracts are actually in the pool which is the case because we model the differences between the loaners through transaction costs rather than individual hazard functions (for an explanation see below). Justification can be attained by arguing that termination behavior translates to non-monetary transaction costs like the work to organize refinancing or the fact that some people are more reluctant to change contracts than

others even if it is optimal to do so. If transaction costs are interpreted as the sum of both monetary and individual non-monetary expenses this approach can be justified. Additionally it is a relatively convenient way of entering heterogeneity in the model since transaction costs are only one parameter whereas the below hazard functions have five parameters and a function covering time development we would have to meticulously estimate.

2.3.1 Premature Termination by a Single Borrower

First we introduce a random variable X defined as

$$X = \text{time until premature termination through default or prepayment},$$

which we furthermore assume to be exponentially distributed, i. e. the probability of early termination within time t conditional on not having occured prior is

$$P[X < t] = 1 - \exp(-\lambda\, t)$$

for some time- and state-dependent parameter λ. In order to obtain λ we make use of a so-called hazard function as elucidated in [29]. Such a function is supposed to describe the risk of premature termination and hence consists of parts referring to each source of risk. First there is a background hazard arising from the aforementioned suboptimal behavior and second the risk is increased if prepayment or default becomes optimal. The function looks as follows:

$$\begin{aligned} \lambda(t) &= \beta_0 + \beta_1 \arctan(t/\beta_2)\, P_t + \beta_3 \arctan(t/\beta_4)\, D_t \\ &=: \lambda_0 + \lambda_p + \lambda_d. \end{aligned}$$

The variables P_t and D_t are indicator variables which take the value one if prepayment or default respectively is optimal and zero otherwise.
At this point it is very important to heed that a probability computed in this way only governs the overall risk of early contract termination while we are interested in the single probabilities for prepayment and default. To resolve this problem we have to distinguish between four cases:

1. neither prepayment nor default is optimal
2. prepayment is optimal but default is not
3. default is optimal but prepayment is not
4. both prepayment and default are optimal.

In order to be able to account for suboptimal termination we presume that in the case of suboptimality prepayment occurs according to the background hazard rate. Doing so we are now able to deal with the four cases:

- case 1
  This is the simplest case because there is only the background hazard. Therefore we can obtain the desired probabilities as follows (for a justification see below):

$$\begin{aligned} P_d &= 0 \\ P_p &= 1/2\,(1 - \exp(-\lambda_0)). \end{aligned}$$

- case 2
  If prepayment is optimal and default is not the overall likeliness of premature termination and default probability have to be computed in the following way:

$$\begin{aligned} P_p &= 1 - \exp(-(\lambda_0 + \lambda_p)) \\ P_d &= 0. \end{aligned}$$

- case 3
  This case is slightly more complicated than the second one because it does not work analogously. The probabilities we prescribe are the following ones:

$$\begin{aligned} P_p &= 1/2\,(1-\exp(-\lambda_0)) \\ P_d &= 1/2\;[1-\exp(-(\lambda_0+\lambda_d))]\,. \end{aligned}$$

  Again, a justification is given below.

- case 4
  In this case, i. e. both prepayment and default are reasonable, we have to figure out the exercise of which option yields the higher benefit. More precisely we have to compare the payment incurred on default $H(1+X_d)$ to the one on prepayment, namely $(1+X_p)F$. The smaller payment then determines which way of termination is considered optimal. The other is not considered optimal so that we are either in case 2 or 3.

So the value of an MBS at time t is the probability weighted average of the solution to the partial differential equation (2.6) we are going to derive in Section 2.5 and the respective payments made on termination with the weights being the probabilties for the occurence of the respective event calculated in the above manner.
We are still due giving a justification for the above choices of the probabilities. If we blindly just added probabilities of the type

$$P = 1 - \exp(-\lambda)$$

we might face the problem of $P_p + P_d \geq 1$ which would be completely unreasonable. Therefore in the critical case we introduce the coefficient 1/2 which ensures the ratio between the two probabilities is maintained and none is favored over the other. We then obtain:

$$P_d + P_p = 1 - 1/2\ \underbrace{\exp(-\lambda_0)}_{\leq 1}[1 + \underbrace{\exp(\lambda_d)}_{\leq 1}].$$

Altogether this means $P_d + P_p \leq 1$ and thereby establishes the desired property.

2.3.2 Distribution of Transaction Costs

The subject matter of this section is the modeling of borrower heterogeneity, i. e. different mortgagors face different transaction costs on refinancing or defaulting. This is attained by assuming the frictions to be spread across a pool of mortgages according to a probability distribution. Our choice is the beta-distribution with the parameters $\alpha = 0.86703$ and $\beta = 5.09893$ (taken from [28] and [29]), the mean μ and variance σ^2 of which can be calculated as

$$\begin{aligned} \mu &= \frac{\alpha}{\alpha+\beta} \\ \sigma^2 &= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \end{aligned}$$

Since we lack any knowledge about the number of mortgages in a respective pool our strategy is to approximate a beta-distribution, or more precisely its cumulative distribution function, which is defined as

$$F(x) = P[X \leq x],$$

by finitely many values X_j and corresponding weights c_j. First I will state a lemma on the convergence of a certain approximation.

Lemma 2.3.1 *Let $m \in \mathbb{N}$ and let $F \in C^1((0,1))$. Consider the weights $c_j = 1/m$ and $X_j = F^{-1}(\frac{2j-1}{2m})$ for $j = 1, \ldots, m$ and the function*

$$\hat{F}^m(x) := \sum_{j=1}^{m} c_j I[X_j \leq x],$$

where I is an indicator function. Then $\hat{F}$ converges uniformly to F, i. e.

$$\sup_{x\in[0,1]} |F(x) - \hat{F}^m(x)| \overset{m\uparrow\infty}{\longrightarrow} 0.$$

Proof - Lemma 2.3.1:
First we note that the approximation is exact at the points $x = 0$ and $x = 1$ and at the points $x = F^{-1}(\frac{2j-1}{2m})$ the error is $\frac{1}{2m}$. Now choose an arbitrary $x \in (0,1)$. Then depending on x there is a j_0 such that $x \in [F^{-1}(\frac{2j_0-1}{2m}), F^{-1}(\frac{2(j_0+1)-1}{2m})] =: J$ and on J $\hat{F}^m$ is constant. So we obtain the estimate:

$$\begin{aligned}
|F(x) - \sum_{j=1}^{m} c_j I[X_j \leq x]| &\leq \sup_{x\in J}|F(x) - F(F^{-1}\left(\frac{2j_0-1}{2m}\right))| + \frac{1}{2m} \\
&\leq \frac{1}{2m} + \sup_{x\in J}|F'(x)|\,|x - F^{-1}\left(\frac{2j_0-1}{2m}\right)| \\
&\leq \frac{1}{2m} + \underbrace{\sup_{x\in[0,1]}|F'(x)|}_{<\infty}\,\underbrace{|x - F^{-1}\left(\frac{2j_0-1}{2m}\right)|}_{\longrightarrow 0} \\
&\overset{m\uparrow\infty}{\longrightarrow} 0
\end{aligned}$$

Since the above is valid for arbitrary x it also holds for the supremum as claimed.

□

This approximation of the cumulative distribution function was introduced in [87] and it is also used in the more recent papers [28] and [29]. In the former reference, however, no evidence of convergence was given but it was stressed that these choices of the quantiles c_j and costs X_j minimize

$$\sup_{x\in[0,1]} |F(x) - \hat{F}^m(x)|,$$

which motivates this approach.
Moreover we have to consider the proportions of the pool, for which prepayment and default respectively is optimal. As it is pointed out in [28] and [29] the optimality conditions (2.1) and (2.2) can be reinterpreted to saying that early termination is optimal if transaction costs are low enough, i. e. less or equal than thresholds X_p^* and X_d^*. Knowing the weights c_{jt} at time t one can write down the proportions by using indicator functions:

$$\begin{aligned}
P_{tp}^* &= \sum_{j=1}^{m} c_{jt}\, I[X_j \leq X_p^*] \\
P_{td}^* &= \sum_{j=1}^{m} c_{jt}\, I[X_j \leq X_d^*].
\end{aligned}$$

Using this we can compute the expectation for premature termination at time t:

$$\bar{\omega}_t^j = P_r^j\, P_{tp}^* + P_d^j\, P_{td}^* + P_e\,(1 - P_{tp}^* - P_{td}^*).$$

Here P_r^j is the probability for optimal prepayment, P_d^j the one for optimal default and P_e the likeliness for termination for exogenous reasons.
In addition, the proportions of the considered pool having transaction costs X_j have to be updated in each time step according to whether or not they have prepaid or defaulted the step before. Doing so one has to distinguish between four different cases:

$$c_{jt+1} = \begin{cases} \frac{c_{jt}(1-P_r)}{1-\bar{\omega}_t} & \text{if prepayment is optimal, but default is not} \\ \frac{c_{jt}(1-P_d)}{1-\bar{\omega}_t} & \text{if default is optimal, but prepayment is not} \\ \frac{c_{jt}(1-P_r-P_d)}{1-\bar{\omega}_t} & \text{if both prepayment and default is optimal} \\ \frac{c_{jt}(1-P_e)}{1-\bar{\omega}_t} & \text{if neither prepayment nor default is optimal.} \end{cases}$$

These updated weights can then be used to compute the value of a mortgage-backed security as the weighted average of the values of the single unpooled mortgages with transaction costs X_j, i. e.

$$M(t) = \sum_{j=1}^{m} c_{jt}\, M^j(t),$$

where for brevity other functional arguments than time have been omitted. The starting weights are $c_{j0} = \frac{1}{m}$ for $j = 1, \ldots, m$. So the valuation process for an MBS pool is the following:

1. Solving PDE (2.6) in every time step. According to Sections 2.5, 2.6 and 2.8, in which the initial boundary value problem is scrupulously stated this will be backwards in time.

2. Carrying out the weighting for $t = 0$

3. Working forward in time to procure weighted solutions for times $t > 0$ by adjusting the weights in the aforementioned way.

In particular, the last item means that one has to store the solution for at least every month throughout the entire computation instead of discarding them which is burdensome with respect to memory efficiency.

2.4 Asset or Liability

One important feature of mortgage-backed securities, which can very easily be overlooked, is the difference between the asset price and the liability price. This discrepancy arises because the asset holder does not receive all the payments the liability holder makes on premature termination since the latter one also has to pay for the transaction, i. e. the asset holder does not obtain the transaction costs.
Paying regard to this fact the prices are calculated in the following way:

$$\begin{aligned} M_l(r,H,t) &= P_d\, H_t(1+X_d) + P_p\, F_t(1+X_p) + (1-P_d-P_p)\, M_{PDE}(r,H,t) \\ M_a(r,H,t) &= P_d\, H_t + P_p\, F_t + (1-P_d-P_p)\, M_{PDE}(r,H,t). \end{aligned}$$

The subscripts a and l indicate the asset or liability case respectively. Moreover one has to heed that the asset price cannot be valued without pricing the liability simultaneously because whether or not premature termination is optimal depends on the liability price and not on the asset price.
Since the asset price differs from the liability value in the absence of transaction costs, the liability value is apparently always greater than the asset value.

2.5 Derivation of the PDE

In this section we are going to derive the partial differential equation for the value of a mortgage-backed security. It was stated in [28] but no proof or derivation was provided.
Given the two stochastic processes

$$dr_t = \underbrace{(\kappa(\theta_r - r_t) - \eta r_t)}_{=:\alpha_r} dt + \Phi_r \sqrt{r_t}\, dW_r$$

and

$$dH_t = (r_t - q_H) H_t\, dt + \Phi_H H_t\, dW_H$$

from Section 2.2 for the interest rates and house prices respectively, we aim to derive a partial differential equation for the value of the corresponding mortgage-backed security $M = M(r, H, t)$ paying a monthly coupon C. Moreover we assume the Brownian motions W_r and W_H to be uncorrelated, just as in [28].
The stretegy of the derivation is very similar to deriving the Black-Scholes equation as it is done in [51]. We start by constructing a riskless portfolio Π, but in our case we have to heed that the MBS is an interest rate derivative, which means that there is no underlying asset one can purchase to perform an intended hedge for the interest rate process. Therefore we have no choice but to have Π consist of two MBSs M_1 and M_2 differing in time to maturity and a certain amount of houses H. We balance it as follows:

$$\Pi := M_1 - \Delta_2\, M_2 - \Delta_1\, H. \tag{2.5}$$

Since our stochastic processes satisfy the assumptions of Theorem A.1.6 we can apply the two-dimensional Itô-formula and attain

$$\begin{aligned}
d\Pi &= dM_1 - \Delta_2 dM_2 - \Delta_1 dH - \Delta_1 q_H H dt + (1-\Delta_2)Cdt \\
&= \frac{\partial M_1}{\partial t}dt + Cdt + \frac{\partial M_1}{\partial H}dH + \frac{\partial M_1}{\partial r}dr + \frac{1}{2}\left(\Phi_r^2 r\frac{\partial^2 M_1}{\partial r^2} + \Phi_H^2 H^2 \frac{\partial^2 M_1}{\partial H^2}\right)dt - \\
&\quad \Delta_2\left[\frac{\partial M_2}{\partial t}dt + Cdt + \frac{\partial M_2}{\partial H}dH + \frac{\partial M_2}{\partial r}dr + \frac{1}{2}(\Phi_r^2 r\frac{\partial^2 M_2}{\partial r^2} + \Phi_H^2 H^2\frac{\partial^2 M_2}{\partial H^2})dt\right] - \\
&\quad \Delta_1\left[(r-q_H)H_t\,dt + \Phi_H H_t\,dW_H\right] - \Delta_1 q_H H dt \\
&= [\frac{\partial M_1}{\partial t} + C - \Delta_2\frac{\partial M_2}{\partial t} - \Delta_2 C - \Delta_1(r-q_H)H + \frac{1}{2}\Phi_r^2 r\frac{\partial^2 M_1}{\partial r^2} + \frac{1}{2}\Phi_H^2 H^2\frac{\partial^2 M_1}{\partial H^2} - \\
&\quad \Delta_2\frac{1}{2}\Phi_r^2 r\frac{\partial^2 M_2}{\partial r^2} - \Delta_2\frac{1}{2}\Phi_H^2 H^2\frac{\partial^2 M_2}{\partial H^2} - \Delta_1 q_H H]\,dt + \left[\frac{\partial M_1}{\partial r} - \Delta_2\frac{\partial M_2}{\partial r}\right]dr + \\
&\quad \left[\frac{\partial M_1}{\partial H} - \Delta_2\frac{\partial M_2}{\partial H}\right]dH - \Delta_1\Phi_H H dW_H \\
&= [\frac{\partial M_1}{\partial t} + C - \Delta_2\frac{\partial M_2}{\partial t} - \Delta_2 C - \Delta_1(r-q_H)H + \frac{1}{2}\Phi_r^2 r\frac{\partial^2 M_1}{\partial r^2} + \frac{1}{2}\Phi_H^2 H^2\frac{\partial^2 M_1}{\partial H^2} - \\
&\quad \Delta_2\frac{1}{2}\Phi_r^2 r\frac{\partial^2 M_2}{\partial r^2} - \Delta_2\frac{1}{2}\Phi_H^2 H^2\frac{\partial^2 M_2}{\partial H^2} - \Delta_1 q_H H]\,dt + \left[\frac{\partial M_1}{\partial r} - \Delta_2\frac{\partial M_2}{\partial r}\right](\alpha_r\,dt + \\
&\quad \Phi_r\sqrt{r}\,dW_r) + \left[\frac{\partial M_1}{\partial H} - \Delta_2\frac{\partial M_2}{\partial H}\right]((r-q_H)H_t\,dt + \Phi_H H_t\,dW_H) - \\
&\quad \Delta_1\Phi_H H dW_H \\
&= [\frac{\partial M_1}{\partial t} + C - \Delta_2\frac{\partial M_2}{\partial t} - \Delta_2 C - \Delta_1(r-q_H)H + \frac{1}{2}\Phi_r^2 r\frac{\partial^2 M_1}{\partial r^2} + \frac{1}{2}\Phi_H^2 H^2\frac{\partial^2 M_1}{\partial H^2} - \\
&\quad \Delta_2\frac{1}{2}\Phi_r^2 r\frac{\partial^2 M_2}{\partial r^2} - \Delta_2\frac{1}{2}\Phi_H^2 H^2\frac{\partial^2 M_2}{\partial H^2}] + \alpha_r[\frac{\partial M_1}{\partial r} - \Delta_2\frac{\partial M_2}{\partial r}] + \\
&\quad H(r-q_H)[\frac{\partial M_1}{\partial H} - \Delta_2\frac{\partial M_2}{\partial H}] \\
&\quad -\Delta_1 q_H H]dt + \Phi_r\sqrt{r}\left[\frac{\partial M_1}{\partial r} - \Delta_2\frac{\partial M_2}{\partial r}\right]dW_r + \Phi_H H\left[\left(\frac{\partial M_1}{\partial H} - \Delta_2\frac{\partial M_2}{\partial H}\right) - \Delta_1\right]dW_H \\
&\overset{!}{=} r\Pi\,dt
\end{aligned}$$

In the latter step we presume that the change in the value of our portfolio in the absence of risk, i. e. after eliminating the stochastic terms, should equal the profit gained on depositing the money at the riskless interest rate. To ensure the annihilation of the stochastic terms we have to make the following demands leading to certain balancing factors Δ_1 and Δ_2:

1. $\frac{\partial M_1}{\partial r} - \Delta_2\frac{\partial M_2}{\partial r} \overset{!}{=} 0 \quad \Rightarrow \quad \Delta_2 = \frac{\partial_r M_1}{\partial_r M_2}$

2. $[\frac{\partial M_1}{\partial H} - \frac{\partial_r M_1}{\partial_r M_2}\frac{\partial M_2}{\partial H}] - \Delta_1 \overset{!}{=} 0 \quad \Rightarrow \quad \Delta_1 = \partial_H M_1 - \frac{\partial_r M_1}{\partial_r M_2}\partial_H M_2.$

Inserting these choices into the equation will as desired eliminate the stochastic terms so that after using the definition of the portfolio (2.5) itself we obtain:

$$\begin{aligned}
r\Pi dt &= rM_1 dt - rM_2\frac{\partial_r M_1}{\partial_r M_2}dt - rH\left(\partial_H M_1 - \frac{\partial_r M_1}{\partial_r M_2}\,\partial_H M_2\right)dt \\
&= rM_1 dt - rM_2\frac{\partial_r M_1}{\partial_r M_2}dt - rH\,\partial_H M_1 dt + rH\frac{\partial_r M_1}{\partial_r M_2}\,\partial_H M_2 dt.
\end{aligned}$$

Combining both these calculations and rearranging in terms of M_1 and M_2 yields:

$$\begin{aligned} 0 &= rM_1 - (r-q_H)H\partial_H M_1 - \partial_t M_1 - C - \frac{1}{2}\left[\Phi_r^2 r\,\partial_r^2 M_1 + \Phi_H^2 H^2\,\partial_H^2 M_1\right] \\ &\quad -[\frac{\partial_r M_1}{\partial_r M_2}(rM_2 - (r-q_H)H\partial_H M_2 - \partial_t M_2 - C \\ &\quad -\frac{1}{2}\left[\Phi_r^2 r\,\partial_r^2 M_2 + \Phi_H^2 H^2\,\partial_H^2 M_2\right]]. \end{aligned}$$

Or equivalently after dividing by $\partial_r M_1$:

$$\begin{aligned} 0 &= \frac{1}{\partial_r M_1}(rM_1 - (r-q_H)H\,\partial_H M_1 - \partial_t M_1 - C - \frac{1}{2}[\Phi_r^2 r\,\partial_r^2 M_1 \\ &\quad + \Phi_H^2 H^2\,\partial_H^2 M_1]) - [\frac{1}{\partial_r M_2}(rM_2 - (r-q_H)H\,\partial_H M_2 \\ &\quad -\partial_t M_2 - C - \frac{1}{2}[\Phi_r^2 r\,\partial_r^2 M_2 + \Phi_H^2 H^2\,\partial_H^2 M_2])]. \end{aligned}$$

Since the first part of the equation depends only on a maturity T_1, whereas the second part depends merely on a maturity T_2, both parts can only equal one another if they are independent of maturity. Hence from here on one can omit the subscripts and write the equation in the following form:

$$a(r,t) = \frac{1}{\partial_r M}\left(rM - (r-q_H)H\,\partial_H M - \partial_t M - C - \frac{1}{2}[\Phi_r^2 r\,\partial_r^2 M + \Phi_H^2 H^2\,\partial_H^2 M]\right)$$

for some function $a(r,t)$. According to [4] a can additionally be written as:

$$a(r,t) = \alpha_r - \Phi_r\sqrt{r}\,\lambda(r,t),$$

where λ is the so-called market price of interest rate risk. Overall this yields the following partial differential equation for the value of a mortgage-backed security M after dividing by -1:

$$-rM + (r-q_H)H\partial_H M + \partial_t M + C + \frac{1}{2}[\Phi_r^2 r\partial_r^2 M + \Phi_H^2 H^2\partial_H^2 M] + (\alpha_r - \Phi_r\sqrt{r}\lambda(r,t))\,\partial_r M = 0.$$

If we set $\lambda = 0$, which we deem sensible, as the market price of risk has already been addressed in the model for the interest rates and house prices, we procure the following PDE for valuing mortgage-backed securities:

$$\frac{1}{2}\Phi_r^2 r M_{rr} + \frac{1}{2}\Phi_H^2 H^2 M_{HH} + [(\kappa(\theta_r - r) - \eta r)]\,M_r + (r-q_H)H\,M_H + M_t - rM + C = 0. \quad (2.6)$$

2.6 Boundary and Initial Values for the PDE

The previously derived partial differential equation (2.6) holds for any derivative depending on the variables r, H and t driven by the same stochastic processes. Therefore the characterics of an MBS are solely imposed by the boundary and initial values (or those at maturity respectively). Moreover as we know, for instance from [64] or [69], such PDEs do not have unique solutions without properly prescribing such boundary and initial data. Hence we have to specify those data both for mathematical and finance reasons.

To start with the initial/terminal data we heed that a mortgage-backed security has a finite lifespan. Generally speaking it has a value because of the scheduled future payments on the mortgage. Thus after expiration of the mortgage it is worthless because there are no more payments to be made. Hence the following terminal condition arises:

$$M(H,r,T) = 0,$$

which reflects the fact that at maturity all payments have been made because of amortization.

Concerning the boundary values there are four boundaries we have to deal with. These boundaries are infinitely large interest rates, infinitely large house prices, vanishing interest rates and vanishing house prices.
The simplest of the four is certainly the infinitely large interest rates. In this case the value becomes independent of house prices because the underlying bond, i. e. the present value of the scheduled future coupon payments, turns worthless. In consequence, the incorporated options are of no value to their holder either so that we arrive at the condition

$$\lim_{r\to\infty} M^l(H,r,t) = 0.$$

If on the other hand the house prices approach infinity default risk is practically absent because the default option is by far out-of-the-money so that within the mortgage-backed security only the interest rate risk remains. In other words, the MBS becomes a callable bond only depending on interest rates. Hence changes in the house prices do not affect the MBS. Therefore we impose a vanishing Neumann-condition on this boundary, thus

$$\lim_{H\to\infty} \frac{\partial M}{\partial H} = 0$$

is prescribed.
In the case of vanishing house prices immediate default occurs and consequently

$$M(0,r,t) = 0$$

has to hold. This becomes plain by interpreting vanishing house prices as the opportunity to exercise a call option on the underlying bond, and thus save the remaining payments, at a vanishing strike price. Moreover the situation cannot improve to a more beneficiary one to the mortgagor because as mentioned in Section 2.2 the underlying house price process has an absorbing barrier at zero which means that the price will forever remain at zero if it once is. And because of that most payment can be saved by immediate exercise.
Eventually we have to deal with vanishing interest rates which happens to be the most complicated case. The value we prescribe at this boundary is:

$$M(H,0,t) = \min(C\,(T-t),(1+X)F_t,(1+X)H_t).$$

This is actually a simplification because we assume the options to be exercised at once or never and then choose the least liability between prepayment, default and underlying bond without options. Nonetheless this simplification can be made palpable by considering the two options separately.
If it is not optimal to default now house prices are too high at the moment and have to drop in order to compensate for this deficiency in addition to the decline always needed because of the fewer number of payments to be saved on exercise so that altogether a relatively strong and therefore improbable movement is necessary. Much the same can be said about prepayment and interest rates except for the fact that they cannot drop any further than zero.
Due to those arguments we have decided to neglect the possibility of a strong movement and

live with the simplification instead of a free boundary. Moreover by our choices continuity of the boundary is ensured since for vanishing house prices and interest rates we have

$$\begin{aligned}\lim_{H\to 0} M(H,0,t) &= 0\\ \lim_{r\to 0} M(0,r,t) &= 0\end{aligned}$$

while on the other boundary

$$\lim_{H\to\infty} M(H,0,t) = \min(C\,(T-t),(1+X)F_t)$$

and thus

$$\partial_H\left(\lim_{H\to\infty} M(H,0,t)\right) = 0$$

holds. On the boundaries where we assume M to vanish identically those continuity checks are apparently satisfied, too.

2.7 Properties of the Initial Boundary Value Problem

In this section I am going to outline several very important properties of our initial boundary value problem derived in the previous two sections.
The first aspect one should take notion of is the infinite computational domain $(0,\infty)\times(0,\infty)$. This has to be addressed and altered for numerical solution but roughly speaking it poses only a relatively small problem. In the next section I will explain how we attempt to handle this by transformation and truncation.
The other very important and more complicated aspect is the differential operator itself. We are dealing with an equation of the following type:

$$\mathcal{L}u(r,H,t) - u_t(r,H,t) = f(r,H,t),$$

where $\mathcal{L}$ is defined as:

$$\mathcal{L}u(r,H,t) = \frac{1}{2}\Phi_r^2 r\,\partial_{rr}u + \frac{1}{2}\Phi_H^2 H^2\,\partial_{HH}u + [\kappa(\theta_r - r) - \eta r]\,\partial_r u + (r-q_H)H\partial_H u - r\,u.$$

We examine what happens to this operator at the boundary, particularly the boundaries $r=0$ and $H=0$. Therefore we have to consider the coefficient functions

$$\begin{aligned}k_1(r) &= \frac{1}{2}\Phi_r^2 r\\ k_2(H) &= \frac{1}{2}\Phi_H^2 H^2\\ b_1(r) &= \kappa(\theta_r - r) - \eta r\\ b_2(r,H) &= H(r-q_H)\\ a(r) &= -r.\end{aligned}$$

Here k_1 and k_2 represent diffusion, b_1 and b_2 convection and a is the reaction term. At the boundaries $r=0$ and $H=0$ the following happens:

$$\begin{aligned}k_1(0) &= 0\\ k_2(0) &= 0\\ b_1(0) &= \kappa\theta_r\\ b_2(0,H) &= -Hq_H\\ b_2(r,0) &= 0\\ a(0) &= 0.\end{aligned}$$

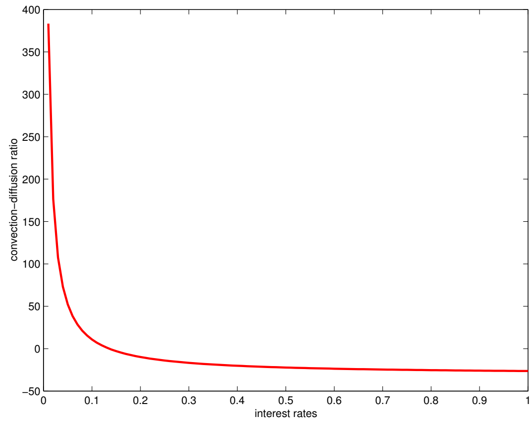

Figure 2.1: The plot shows the convection-diffusion ratio and and the convection dominance near the boundary $r = 0$

In particular, we note that the diffusion coefficients vanish so that $\mathcal{L}$ degenerates at the boundaries $r = 0$ and $H = 0$. Because of this we have to check how the convection coefficients behave in comparison to diffusion, i. e. we have to consider the following quotients:

$$\begin{aligned}
\lim_{r\to 0} \frac{b_1(r)}{k_1(r)} &= \lim_{r\to 0} \frac{\kappa(\theta_r - r) - \eta r}{\frac{1}{2}\Phi_r^2 r} \\
&= \infty \\
\lim_{H\to 0} \frac{b_2(r,H)}{k_2(r)} &= \lim_{H\to 0} \frac{H(r - q_H)}{\frac{1}{2}\Phi_H^2 H^2} \\
&= \infty \qquad \text{for fixed } r.
\end{aligned}$$

In Figure 2.1 this effect is also graphically illustrated for r. The above limites mean that for both small r and H the equation becomes convection dominated, so that classic numerical discretizations might turn instable. This is, for instance, outlined in [48] or [63]. In consequence, we have to find a transformation which resolves that problem or we have to deal with it separately when discretizing. In the next section it turns out that for H the former can be done whereas it remains a problem in r-direction.
Moreover we have to do the same checks for the other boundaries. Apparently all coefficients approach infinity for $r \to \infty$ and $H \to \infty$ respectively. Hence we consider the quotients:

$$\begin{aligned}
\lim_{r\to \infty} \frac{b_1(r)}{k_1(r)} &= \lim_{r\to 0} \frac{\kappa(\theta_r - r) - \eta r}{\frac{1}{2}\Phi_r^2 r} \\
&= -\frac{2(\eta + \kappa)}{\Phi_r^2} \\
&= -30.4987 \\
\lim_{H\to \infty} \frac{b_2(r,H)}{k_2(r)} &= \lim_{H\to \infty} \frac{H(r - q_H)}{\frac{1}{2}\Phi_H^2 H^2} \\
&= \lim_{H\to \infty} \frac{2(r - q_H)}{\Phi_H^2 H} \\
&= 0 \qquad \text{for fixed r.}
\end{aligned}$$

We notice that there are no problems on the far boundary. Moreover we have to heed that b_2 depends on both r and H so that we are in need of checking the following four simultaneous limites:

$$\begin{aligned}
\lim_{r,H\to 0} \frac{b_2(r,H)}{k_2(H)} &= \infty \\
\lim_{r\to\infty,H\to 0} \frac{b_2(r,H)}{k_2(H)} &= \infty \\
\lim_{r,H\to\infty} \frac{b_2(r,H)}{k_2(H)} &= \frac{2}{\Phi_H^2} \\
\lim_{r\to 0,H\to\infty} \frac{b_2(r,H)}{k_2(H)} &= 0.
\end{aligned}$$

We find that by the additional limes in r no further problems are raised.

2.8 Transformation of the PDE

We examine the partial differential equation (2.6) which reads as

$$\frac{1}{2}\Phi_r^2 r M_{rr} + \frac{1}{2}\Phi_H^2 H^2 M_{HH} + [(\kappa(\theta_r - r) - \eta r)]\, M_r^l + (r - q_H) H\, M_H + M_t - rM + C = 0$$

in the variables r, H and t subject to the boundary and terminal conditions from Section 2.6:

$$\begin{aligned}
M(0,r,t) &= 0 \\
M(H,r,T) &= 0 \\
\lim_{r\to\infty} M(H,r,t) &= 0 \\
\lim_{H\to\infty} M_H(H,r,t) &= 0 \\
M_r(H,0,t) &= \min(C\,(T-t), (1+X)F_t, (1+X)H_t)
\end{aligned}$$

for some transaction costs X. As outlined in the previous section we are facing the problems of a degenerating operator which causes at least in part dominant convection and of an infinite computational domain. In the following I will introduce two transformations which besides truncation will take care of that. However, we will find that that those problems cannot be completely removed so that we have to rely on further stabilization methods.
Transformation means defining a function U such that

$$U(z(H), y(r), s) := M(H,r,t).$$

Doing so we attain the following transformation formulae for the first and second derivatives in r, H and t:

$$\begin{aligned}
M_H &= U_z \frac{dz}{dH} \\
M_r &= U_y \frac{dy}{dr} \\
M_t &= U_s \frac{ds}{dt} \\
M_{HH} &= U_{zz} \left(\frac{dz}{dH}\right)^2 + U_z \left(\frac{d^2 z}{dH^2}\right) \\
M_{rr} &= U_{yy} \left(\frac{dy}{dr}\right)^2 + U_y \left(\frac{d^2 y}{dr^2}\right).
\end{aligned}$$

These transformation formulae hold for any transformation of the above type. Thus in a concrete example it suffices to state the explicit functional form of the transformation, compute the appearing derivatives and insert the results in the original PDE.
So I will first state such a transformation, then carry out the adjustments and afterwards I will examine which benefit the transformation yields.

2.8.1 Black-Scholes-Type Transformation

The aim of this transformation is to eliminate the variable H in the original PDE from the coefficients and to change the infinite domain in r. Moreover a time-reversal is used, which alters the terminal condition at maturity to an initial condition in the transformed problem. Precisely we choose the following:

$$\begin{aligned} x &= 1 - \exp(-\gamma_r r) \\ y &= A \ln(H) \\ s &= T - t \end{aligned}$$

The parameters A and γ_r of our choice are:

$$\begin{aligned} A &= 0.5 \\ \gamma_r &= \ln(20) \approx 2.9957. \end{aligned}$$

Along with this transformation comes a change in the domain on which the PDE is defined since the original domain $(0,\infty) \times (0,\infty)$ is mapped to $(0,1) \times (-\infty,+\infty)$ independently of the choices of γ_r and A. Why they are chosen in the above way is justified in Section 3.7 where a truncation strategy is proposed.
This leads to the following derivatives and inverse transformations which we will then be inserted into the transformation formulae and the original PDE:

$$\begin{aligned} \frac{dy}{dH} &= \frac{A}{H} \\ \frac{dx}{dr} &= \gamma_r(1-x) \\ \frac{d^2y}{dH^2} &= -\frac{A}{H^2} \\ \frac{d^2x}{dr^2} &= -\gamma_r^2(1-x) \\ H &= \exp(y/A) \\ r &= -\frac{1}{\gamma_r}\ln(1-x). \end{aligned}$$

So the r-, t- and H-derivative terms transform to

$$\begin{aligned} M_H &= U_y \frac{A}{H} \\ M_r &= U_x \gamma_r (1-x) \\ M_{HH} &= U_{yy}\frac{A^2}{H^2} - U_y \frac{A}{H^2} \\ M_{rr} &= U_{xx}\gamma_r^2(1-x)^2 - U_x \gamma_r^2 (1-x) \\ M_t &= -U_s. \end{aligned}$$

Altogether this results in the following transformed PDE:

$$\begin{aligned}
0 &= -\frac{1}{2}\Phi_r^2\gamma_r \ln(1-x)(1-x)^2\, U_{xx} + \frac{1}{2}\Phi_H^2 A^2\, U_{yy} \\
&\quad + \left(\alpha_r\gamma_r(1-x) + \frac{1}{2}\Phi_r^2\gamma_r(1-x)\ln(1-x)\right) U_x \\
&\quad - A\left(\frac{1}{2}\Phi_H^2 + \frac{\ln(1-x)}{\gamma_r} + q_H\right) U_y + \frac{\ln(1-x)}{\gamma_r} U - U_s + C.
\end{aligned} \tag{2.7}$$

Besides the PDE itself we have to transform its boundary conditions for which we obtain:

$$\begin{aligned}
\lim_{x\to\infty} u(x,y,s) &= \lim_{r\to\infty} M(r,H,t) \\
&= 0 \\
u(0,y(H),s) &= M(0,H,t) \\
&= \min(C\,(T-t),\, F_t(1+X),\, H(1+X)) \\
&= \min(C\,(T-t),\, F_t(1+X),\, \exp(y/A)(1+X)) \\
\lim_{y\to-\infty} u(x,y,s) &= M(r,0,t) \\
&= 0 \\
\lim_{y\to\infty} u_y(x,y,s) &= 0.
\end{aligned}$$

2.8.2 Benefit of the Transformation

This section shall be devoted to examining the properties of the transformed partial differential equation. As before we have to check the transformed differential operator for degeneration. Concerning H, which has become y, the problems were removed since the coefficients

$$\begin{aligned}
\tilde{b}_2(x,y) &= -A\left(\frac{1}{2}\Phi_H^2 + \frac{\ln(1-x)}{\gamma_r} + q_H\right) \\
\tilde{k}_2(y) &= \frac{1}{2}\Phi_H^2 A^2
\end{aligned}$$

are independent of y so that their quotient is constant. Heed that for $x \to 1$ convection becomes dominant. But this problem is solved by truncation as it is outlined in Section 3.7. Moreover y ranges from $-\infty$ to $+\infty$ which is roughly speaking even more infinite than before. Yet, as it is done in the case of the Black-Scholes equation this problem can also be solved by truncation (cf. Section 3.7).

Concerning x or r respectively things are more complicated and we have to closely look at the coefficient functions:

$$\begin{aligned}
\tilde{k}_1(x) &= -\frac{1}{2}\Phi_r^2\gamma_r(1-x)^2\ln(1-x) \\
\tilde{b}_1(x) &= \frac{1}{2}\Phi_r^2\gamma_r \ln(1-x)(1-x) + \kappa\theta_r\gamma_r(1-x) + (\kappa+\eta)(1-x)\ln(1-x)
\end{aligned}$$

and hence

$$\begin{aligned}
\tilde{b}_1(0) &= \kappa\,\theta_r\gamma_r \\
&= 2.22579e-2 \\
\tilde{k}_1(0) &= -\frac{1}{2}\Phi_r^2\gamma_r\ln(1) \\
&= 0 \\
\lim_{x\to 0}\frac{\tilde{b}_1(x)}{\tilde{k}_1(x)} &= \infty.
\end{aligned}$$

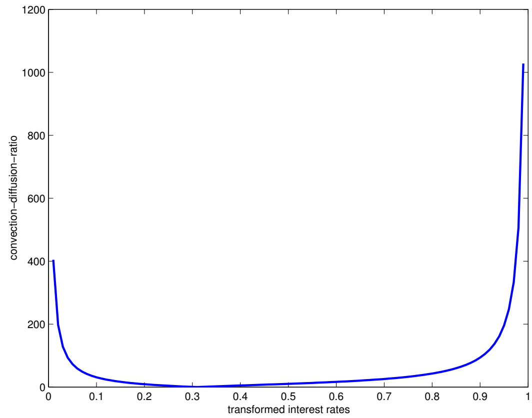

Figure 2.2: This plot shows the convection-diffusion ratio in x-direction of the transformed PDE (2.7) and the convection-dominance near both boundaries

We note that the operator still degenerates and near the boundary $x = 0$ and that convection is still dominant which is depicted in Figure 2.2. Seemingly the transformation went to no avail. However, a logarithmic singularity is less harmful than a linear one.
In addition, we have to look at the effects on the far boundary. Assuming the truncation results from Section 3.7 to be already known we compute the following:

$$\begin{aligned}
\lim_{x\to 1} \frac{\tilde{b}_1(x)}{\tilde{k}_1(x)} &= \lim_{x\to 1} \frac{1/2\Phi_r^2\gamma_r \ln(1-x)(1-x) + \kappa\theta_r\gamma_r(1-x) + \ln(1-x)(1-x)(\kappa+\eta)}{-1/2\Phi_r^2\gamma_r(1-x)^2\ln(1-x)} \\
&= \lim_{x\to 1} \left(\frac{-1}{1-x} - \frac{2\kappa\theta_r}{\Phi_r} \frac{1}{(1-x)\ln(1-x)} - \frac{2(\kappa+\eta)}{\Phi_r^2\gamma_r}\frac{1}{1-x} \right) \\
&= \infty \\
\tilde{k}_1(x_{max} = 0.95) &= -4.08574e-5 \\
\tilde{b}_1(x_{max} = 0.95) &= -8.00733.
\end{aligned}$$

Unfortunately we obtain a convection-to-diffusion ratio of $1.95984e5$, which shows that we might have to use stabilization techniques in x-direction.
Summing up those results one might say that the additional convection dominance at the far boundary was the price to be paid for a less severe singularity at the other boundary. This is acceptable because the near boundary is in some sense more important as it is the boundary where potential prepayment occurs.
Moreover it is worthwhile examining and interpreting the derivative of the transformation in x-direction, which is

$$x'(r) = \gamma_r \exp(-\gamma_r r).$$

Compared to the original domain in r the transformed domain is more thoroughly resolved if $x'(r) > 1$. This is the case if

$$r < \frac{\ln(\gamma_r)}{\gamma_r} = 0.3662.$$

The reason behind that is that any such r is mapped to an $x > r$. Hence small r, in which we are particularly interested, consume more space in x than in r so that after discretization of the x-space more degrees of freedom relate to small interest rates than it would have been the case if the r-space had been discretized. Altogether our transformation ensures that the possible prepayment region is more carefully handled.

2.8.3 Original Transformation

In this section I present the transformation of PDE (2.6) which was originally proposed in [28]. Moreover in the chapter on numerical results it will become plain that our transformation is far more advantageous than this one.
In contrast to our transformation the original domain $(0,\infty)\times(0,\infty)$ is, again independently of γ_r and γ_H, mapped to $(0,1)\times(0,1)$ by virtue of the following choices:

$$\begin{aligned} y &:= \frac{r}{\gamma_r + r} \\ z &:= \frac{H}{\gamma_H + H} \\ s &:= T - t. \end{aligned}$$

The two parameters are selected as in [28]:

$$\begin{aligned} \gamma_r &= 0.05740 \\ \gamma_H &= 1.25. \end{aligned}$$

For the derivatives of the new variables we obtain

$$\begin{aligned} \frac{dz}{dH} &= \frac{(1-z)^2}{\gamma_H} \\ \frac{dy}{dr} &= \frac{(1-y)^2}{\gamma_r} \\ \frac{ds}{dt} &= -1 \\ \frac{d^2z}{dH^2} &= \frac{-2(1-z)^3}{\gamma_H^2} \\ \frac{d^2y}{dr^2} &= \frac{-2(1-y)^3}{\gamma_r^2} \\ H &= \frac{\gamma_H z}{1-z} \\ r &= \frac{\gamma_r y}{1-y} \end{aligned}$$

and inserting them into the transformation formulas eventually yields:

$$\begin{aligned} M_t^l &= -U_s \\ M_H^l &= U_z \frac{(1-z)^2}{\gamma_H} \\ M_r^l &= U_y \frac{(1-y)^2}{\gamma_r} \\ M_{HH}^l &= U_{zz}\frac{(1-z)^4}{\gamma_H^2} - U_z \frac{2(1-z)^3}{\gamma_H^2} \\ M_{rr}^l &= U_{yy}\frac{(1-y)^4}{\gamma_r^2} - U_y \frac{2(1-y)^3}{\gamma_r^2}. \end{aligned}$$

These substitutions for the derivatives as well as those for r, H and t are now inserted into the original equation so that we have:

$$\begin{aligned}
0 &= \frac{1}{2}\frac{\Phi_r^2 r}{\gamma_r^2}\left[U_{yy}(1-y)^4 - U_y 2(1-y)^3\right] + \frac{1}{2}\frac{\Phi_H^2 H^2}{\gamma_H^2}\left[U_{zz}(1-z)^4 - U_z 2(1-z)^3\right] + \\
&\quad \left([\kappa(\theta_r - r) - \eta r - \Phi_r\sqrt{r}\lambda]\right) U_y \frac{(1-y)^2}{\gamma_r} + (r - q_H) H U_z \frac{(1-z)^2}{\gamma_H} - U_s - rU + C \\
&= \frac{\beta_r^2}{2} U_{yy}(1-y)^4 + \frac{\beta_H^2}{2} U_{zz}(1-z)^4 + \left[\frac{-\alpha_r}{\gamma_r}(1-y)^2 - \beta_r^2\,(1-y)^3\right] U_y + \\
&\quad U_z\left[\frac{\alpha_H}{\gamma_H}(1-z)^2 - \beta_H^2\,(1-z)^3\right] - U_s - \frac{\gamma_r y}{1-y} U + C
\end{aligned}$$

where the constants $\beta_r, \beta_H, \alpha_r$ and α_H have the following meanings:

$$\begin{aligned}
\beta_r &= \frac{\Phi_r}{\gamma_r}\sqrt{r} \\
\beta_H &= \frac{\Phi_H H}{\gamma_H} \\
\alpha_r &= \Phi_r\sqrt{r}\lambda[\kappa(\theta_r - r) - \eta r] \\
\alpha_H &= (r - q_H)H.
\end{aligned}$$

But unfortunately we have not yet fulfilled our aim of completely removing r and H from the equation and replacing them by y and z as they still appear in the constants. Therefore we rewrite the constants as follows:

$$\begin{aligned}
\beta_r &= \frac{\Phi_r}{\sqrt{\gamma_r}}\sqrt{\frac{y}{1-y}} \\
\beta_H &= \Phi_H \frac{z}{1-z} \\
\alpha_r &= -\Phi_r\sqrt{\frac{\gamma_r y}{1-y}}\lambda + \kappa\theta_r - (\kappa + \eta)\frac{\gamma_r y}{1-y} \\
\alpha_H &= \left(\frac{\gamma_r y}{1-y} - q_H\right)\frac{z}{1-z}.
\end{aligned}$$

Now they do no longer depend on r and H and can be inserted into the above equation, which yields:

$$\begin{aligned}
0 &= \frac{\Phi_r^2}{2\gamma_r} y(1-y)^3 U_{yy} + \frac{\Phi_H^2}{2} z^2(1-z)^2 U_{zz} + [\frac{\kappa\theta_r}{\gamma_r}(1-y)^2 - (\kappa+\eta)(1-y)y \\
&\quad - \Phi_r\lambda\sqrt{\frac{y}{\gamma_r(1-y)}}(1-y)^2 - \frac{\Phi_r^2}{\gamma_r}(1-y)^2 y]\, U_y \\
&\quad + \left[\left(\frac{\gamma_r y}{1-y} - q_H\right) z(1-z) - \Phi_H^2\,(1-z)z^2\right] U_z \\
&\quad + \gamma_r \frac{y}{1-y} U - U_s + C.
\end{aligned} \tag{2.8}$$

In the second step of the transformation process we have to heed that the boundary conditions also require transformation. Because of $z(0) = 0$ and $\lim_{r\to\infty} y(r) = 1$ we procure:

$$\begin{aligned}
U(0, y(r), s) &= 0 \\
U(z(H), y(r), 0) &= 0 \\
U(z(H), 1, s) &= 0 \\
U'(1, y(r), s) &= 0 \\
U(z(H), 0, s) &= \min(C\,(T-t), (1+X)\,F_t, (1+X)\,H).
\end{aligned}$$

2.9 Probability Weighting

As already mentioned in Section 2.3 a probabilty weighting of the solution of equation (2.6) has to be performed following every time step. But fortunately one can resort to only weighting the solution at the times of the coupon payments because exercise of either option is financially suboptimal at any other times. In this section I would like to justify why this is the case.
First we will consider a series of European call options on the underlying bond. The strike times of these options shall be the times of the coupon payments. We will then derive a lower bound for the value of such a European option, which is also a lower bound for both the corresponding American option and the American option which can be exercised during the whole lifespan of the underlying bond. The last statements hold because every right given to the holder of the European option is contained within the American options. Having these lower bounds we can justify why option exercise is only reasonable at the coupon payment dates. The arguments given are based on those expounded in [51].

Before starting we have to ensure that the results can actually be applied to mortgage-backed securities. Within the mortgage-backed security there is a joint call option on the underlying bond (prepayment or default) with a time-dependent strike. Since exercise of one option means forfeit of the other it suffices to merely consider one option with strike

$$X_t = (1 + X)\ \min(F_t, H_t).$$

This is an American-style option on a bond which can be exercised at any time during the life of the mortgage, so that the above strategy can actually be applied in our setting.

2.9.1 Lower Bound for European Options

Assume the times $t_1, \dots, t_{360}$ to be the times of the scheduled coupon payments and consider further European-style options with strike prices $X_{t_1}, \dots, X_{t_{360}}$ for those times.
We now fix an arbitrary path for H_t and consider the following portfolios:

- portfolio A:
  a European call-option c on the bond and an amount of cash equal to $X\ \exp(-\int_0^{t_i} r(t)dt)$
- portfolio B:
  the bond.

Concerning portfolio A we can state that upon investing the amount of cash at the riskfree interest rate its value increases to X_t at maturity of the option and if at that time $B_t > X_t$ the option is exercised and the value of the portfolio is $(B_t - X_t) + X_t = B_t$. If the reverse holds and $B_t \leq X_t$ the option expires worthless and the value of the portfolio becomes X_t. Overall we have:

$$\text{value of portfolio A} = \max(B_t, X_t).$$

The value of portfolio B is apparently always the value of the bond. Summing up yields that at the expiration date of the option portfolio A is always worth at least as much as portfolio B. Due to this certainty the same must be true for the present value of the portfolios which means:

$$c + X\ \exp\left(-\int_0^{t_i} r(t)dt\right) \geq B_t.$$

Since the value of an option is always non-negative we eventually obtain

$$c \geq \max(B - X\ \exp\left(-\int_0^{t_i} r(t)dt\right), 0).$$

2.9.2 Option Exercise at a Coupon Date

First we have to note the analogy between coupon payments on a bond and dividend payments on stocks. Both these cases are treated in a mathematically identical way and for a stock we know that once it goes ex-dividend, i. e. a dividend D is paid, its price S_t drops to $S_t - D$.
Returning to our bond we therefore have to consider what happens if the option is exercised at a coupon date. Due to the drop-off it is exercised only shortly prior to the payment and not afterwards. If so, the value is

$$B_{t_n} - X_{t_n}.$$

If the option is not exercise the value of the bond drops to $B_{t_n} - C$ and by applying the lower bound derived in the previous section we obtain that if exercise is not reasonable the following condition has to hold:

$$B_{t_n} - C - X_{t_{n+1}} \exp\left(-\int_0^{t_i} r(t)dt\right) \leq B_{t_n} - X_{t_n}.$$

Rewriting yields that it cannot be optimal to exercise if

$$C \leq X_{t_{n+1}} \exp\left(-\int_0^{t_i} r(t)dt\right) - X_{t_n}$$

or conversely exercise is only optimal if

$$C > X_{t_{n+1}} \exp\left(-\int_0^{t_i} r(t)dt\right) - X_{t_n}.$$

2.9.3 Option Exercise Between Coupon Dates

Having stated this condition for the optimality of option exercise we can ponder what happens if one exercises the option between two coupon payments instead of immediately prior to one. To do so we consider a time $t_i < \tau < t_{i+1}$. If exercise is optimal at least the following condition has to be satisfied

$$B_\tau - X_{t_{i+1}} \exp(-\int_0^{t_i} r(t)dt) \leq B_\tau - X_\tau$$

or equivalently

$$X_\tau \leq X_{t_{i+1}} \underbrace{\exp(-\int_0^{t_i} r(t)dt)}_{\leq 1}.$$

Hence this condition can only be fulfilled if $X_{t_{i+1}} \geq X_\tau$. Since F_t, the remaining principal to be repaid to the lender, is obviously monotone decreasing H_t has to increase significantly between τ and t_{i+1} in order to compensate for the exp-expression and the decrease in F_t, which makes such a development relatively improbable. Thus option exercise in between two coupon payment dates is highly unlikely.
Altogether this can be regarded as a justification why we can resort to considering option exercise, which is in fact done by the probability weighting described in Section 2.3, only at the coupon payment times.

Chapter 3

Discretization of the PDE

In this chapter I will in detail describe how the partial differential equation (2.6) for the valuation of MBSs is discretized. This process is at least in general independent of transformation and it consists of two steps, namely the approximation in the spatial coordinates and the time discretization. The aim of this process is to replace the PDE by a finite-dimensional linear system of equations we can solve numerically.
In addition, we should mention right at the beginning that it plays a significant role whether one first discretizes in space or in time. We choose to first discretize in space which is called the *method of lines*. In this case we will eventually obtain an ordinary differential equation in the finite element space.

3.1 Weak Formulation of the PDE

The first step of the discretization process is re-stating the partial differential equation in a weak formulation. Its derivation is independent of whether or not equation (2.6) is transformed. Therefore our strategy is the following: in Subsection 3.1.1 we first re-write a coefficient form PDE in a divergence-form PDE (for explanations see below), then we state the actual weak formulation in Subsection 3.1.2 and finally we mention the weak formulation in the concrete example of the transformed PDE (2.7).

3.1.1 Divergence Form

We are dealing with a partial differential equation of the following form, the so-called coefficient form:

$$k_1(x)\, u_{xx} + k_2(y)\, u_{yy} + b_1(x)\, u_x + b_2(x,y)\, u_y + c(x)\, u - u_t = f \quad \text{in } \Omega \times (0,T).$$

Transforming it to divergence form means that it can be written as follows

$$\nabla \cdot (-K(x,y)\nabla u) + \beta(x,y) \cdot \nabla u + c(x)\, u - u_t = f \quad \text{in } \Omega \times (0,T), \tag{3.1}$$

where $\beta \in (C^0(\Omega))^2$ and $K \in (C^1(\Omega))^{2\times 2}$ can be assumed to be diagonal since there are no mixed second derivatives. In the ensuing subsection it will become plain that the divergence form is more convenient for the purpose of integration by parts, a technique which will be used to obtain the weak formulation.
If $K_{11} = -k_1(x)$ and $K_{22} = -k_2(y)$ is chosen, the desired property is attained by selecting

$\beta_1 = b_1(x) - \partial_x k_1(x)$ and $\beta_2 = b_2(x,y) - \partial_y k_2(y)$ because

$$\begin{aligned}
f &= \nabla \cdot (-K(x,y)\nabla u) + \beta(x,y) \cdot \nabla u + c(x)\, u - u_t \\
&= -\nabla \cdot \begin{pmatrix} -k_1(x) u_x \\ -k_2(y) u_y \end{pmatrix} + \beta(x,y) \cdot \nabla u + c(x)\, u - u_t \\
&= (\partial_x(k_1(x)\, u_x) + \partial_y(k_2(y)\, u_y)) + (b_1(x) - \partial_x k_1(x))\, u_x \\
&\quad +(b_2(x,y) - \partial_y k_2(y))\, u_y + c(x)\, u - u_t \\
&= k_1(x)\, u_{xx} + k_2(y)\, u_{yy} + b_1(x)\, u_x + b_2(x,y)\, u_y + c(x)\, u - u_t.
\end{aligned}$$

We note that β has to be chosen in a way ensuring that the additional term from the derivation of the product $K(x,y)\nabla u$ is annihilated. Henceforth we can without loss of generality presuppose our PDE to be given in divergence form.

3.1.2 Weak Form

We assume the PDE (3.1) to be given on a domain $\Omega \times (0,T) \subset \mathbb{R}^2 \times (0,T)$ and for brevity we define a differential operator $\mathcal{L}$ in the following way:

$$\mathcal{L} := \nabla \cdot (-K(x,y)\nabla) + \beta(x,y) \cdot \nabla + c\, id \tag{3.2}$$

In addition to the equation itself we have to consider boundary values of the following type:

$$\begin{aligned}
u &= g_D \qquad \text{on } \Gamma_D \\
\nabla u \cdot \nu &= g_N \qquad \text{on } \Gamma_N.
\end{aligned}$$

Here $\partial\Omega = \Gamma = \Gamma_D \,\dot{\cup}\, \Gamma_N$ denotes the boundary of Ω and is decomposed into two disjoint parts denoted by Γ_D and Γ_N indicating the prescription of Dirichlet- or Neumann-values respectively. ν be the outward normal of Ω. Hence we are dealing with mixed boundary conditions instead of pure ones. Moreover they are not even homogeneous, i. e. g_D does not vanish identically on Γ_D.
Before we can continue with the actual weak formulation we have to address and resolve two problematic items:

- Reformulation of inhomogeneous boundary values to homogeneous boundary values and a constraint
- Choice of space from which the testing functions are taken

While the former item will be dealt with below we consider functions from the space $C^2(\Omega) \cap C^0(\bar{\Omega}) \cap H^1(\Omega)$ which vanish in a neighborhood of Γ_D concerning the second item. By V we denote the closure of this space under the H^1-norm (cf. Appendix B for a definition).

Multiplying the PDE with a testing function $v \in V$ and integrating over Ω yields:

$$\int_\Omega \nabla \cdot (-K(x,y)\nabla u)\, v + \int_\Omega (\beta(x,y) \cdot \nabla u)\; v + \int_\Omega c(x)\, u\, v = \int_\Omega f\, v + \int_\Omega u_t\, v \qquad \forall\, v \in V.$$

Applying integration by parts in the first integral we procure:

$$\begin{aligned}
\int_\Omega f\, v + \int_\Omega u_t\, v \;=\; & \int_\Omega (K(x,y)\nabla u) \cdot \nabla v + \int_{\Gamma_N} K(x,y) \underbrace{\nabla u \cdot \nu}_{=g_N=0}\, v + \int_\Omega (\beta(x,y) \cdot \nabla u)\; v \\
& + \int_\Omega c(x)\, u\, v \qquad \forall\, v \in V
\end{aligned}$$

The other boundary integral over Γ_D vanishes because of $v \in V$. Resolving the matrix and vector expressions in the first two integrals we obtain the weak formulation of equation (3.1):

$$\int_\Omega (k_1 u_x v_x + k_2 u_y v_y) + \int_\Omega (\beta_1 u_x v + \beta_2 u_y v) + \int_\Omega c(x) u\, v = \int_\Omega f\, v + \int_\Omega u_t\, v \qquad \forall\, v \in V \quad (3.3)$$

What we have yet to discuss is what happened to the inhomogeneous boundary values for which our testing space V does not account. To do so we introduce the following bilinear and linear forms, the claimed properties of which follow from the respective properties of the gradient operator and the Lebesgue integral:

$$\begin{aligned} A(u,v) &:= \int_\Omega (K(x,y)\nabla u)\cdot\nabla v + \int_\Omega (\beta(x,y)\cdot\nabla u)\, v + \int_\Omega c(x) u\, v \\ m(u,v) &:= \int_\Omega u\, v \\ l(v) &:= \int_\Omega f\, v. \end{aligned} \quad (3.4)$$

Then equation (3.3) can be written in the following variational form:

$$A(u,v) = l(v) + m(u_t, v) \qquad \forall\, v \in V. \quad (3.5)$$

Next, I will deduce that one can formulate the inhomogeneous equation as a homogeneous one with additional constraints. The argument is based on the one outlined in [11] for the elliptic problem.

In order to do so we introduce a sufficiently smooth function u_0 which be the extension of g_D to $\bar{\Omega}$. Note that using u_0 we can always write an inhomogeneous problem as a homogeneous one if we replace f by

$$\tilde{f} := f - \mathcal{L}u_0 + u_{0t}.$$

Moreover we can define a linear form $\tilde{l}$ in the following way:

$$\tilde{l}(v) := \int_\Omega \tilde{f}\, v.$$

Next we consider the function $w := u - u_0$ (u be a solution of the homogeneous problem) which vanishes on the Dirichlet-boundary by construction and satisfies:

$$A(w,v) = \tilde{l}(v) + m(u_t, v) \qquad \forall\, v \in V.$$

By definition we then have:

$$\begin{aligned} A(u,v) - A(u_0,v) &= A(w,v) \\ &= l(v) + \underbrace{m(\mathcal{L}u_0, v) - m(u_{0t}, v)}_{=A(u_0,v)} + m(u_t, v). \end{aligned}$$

So we eventually obtain the following variational equation with constraints:

$$\begin{aligned} A(u,v) &= l(v) + m(u_t, v) \qquad \forall\, v \in V \\ u - u_0 &\in V. \end{aligned} \quad (3.6)$$

This is the desired form with homogeneous data on Γ_D in the testing functions and an additional constraint which ensures that u actually takes on the value $g_D = u_0|_{\Gamma_D}$ on the respective part of the boundary.

3.1.3 The Weak Formulation of Our PDE

In this section the weak formulation of our partial differential equation covering mortgage-backed securities shall be explicitly stated. This is outlined for the PDE in its transformed face (2.7). Therefore we recall the coefficients from Section 2.8:

$$\begin{aligned}
k_1(x) &= -\frac{1}{2}\Phi_r^2\gamma_r\,(1-x)^2\,\ln(1-x) \\
k_2(y) &= \frac{1}{2}\Phi_H\,A^2 \\
b_1(x) &= \frac{1}{2}\Phi_r^2\gamma_r\,\ln(1-x)(1-x) + \kappa\,\theta_r\,\gamma_r\,(1-x) + (\kappa+\eta)\,\ln(1-x)(1-x) \\
b_2(x) &= -A\left(\frac{1}{2}\Phi_H^2 + \frac{\ln(1-x)}{\gamma_r} + q_H\right) \\
c(x) &= \frac{\ln(1-x)}{\gamma_r}.
\end{aligned}$$

According to Section 3.1.1 we have to state a matrix K and a vector-valued function β in the following way:

$$\begin{aligned}
K(x) &= \begin{pmatrix} -k_1(x) & 0 \\ 0 & -k_2 \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{2}\Phi_r^2\gamma_r\,(1-x)^2\,ln(1-x) & 0 \\ 0 & \frac{1}{2}\Phi_H\,A^2 \end{pmatrix} \\
\beta(x) &= \begin{pmatrix} b_1(x) - \partial_x k_1(x) \\ b_2 - \partial_y k_2 \end{pmatrix}.
\end{aligned}$$

Concerning β we exploit that k_2 is independent of y and that

$$\partial_x k_1(x) = \Phi_r^2\gamma_r(1-x)\,\ln(1-x) + \frac{1}{2}\Phi_r^2\gamma_r(1-x)$$

so that altogether we procure:

$$\beta(x) = \begin{pmatrix} -\frac{1}{2}\Phi_r^2\gamma_r\,\ln(1-x)(1-x) + \kappa\,\theta_r\,\gamma_r\,(1-x) + (\kappa+\eta)\,\ln(1-x)(1-x) - \frac{1}{2}\Phi_r^2\gamma_r(1-x) \\ -A\left(\frac{1}{2}\Phi_H^2 + \frac{\ln(1-x)}{\gamma_r} + q_H\right) \end{pmatrix}.$$

By this we have defined PDE (2.7) in its divergence form and by virtue of Section 3.1.2 in consequence also in its weak form.

3.2 Spatial Discretization

In this section I will detailedly explain how I use finite elements in order to provide a spatial discretization, which consists of deriving a linear system of equations based upon a given triangulation and a choice of approximating functions $V_h \subset V$. The starting point of this discretization is the constrained variational problem (3.6) we derived in Section 3.1.2:

$$\begin{aligned}
\int_\Omega k_1 u_x v_x + k_2 u_y v_y + \int_\Omega \beta_1 u_x v + \beta_2 u_y v + \int_\Omega c(x)\,u\,v &= \int_\Omega f v + \int_\Omega u_t v \qquad \forall\, v \in V \\
u - u_0 &\in V.
\end{aligned}$$

Since both the space of the testing functions V and the space in which we are seeking a solution to the PDE (it be denoted by W) have infinitely many dimensions we have to resort to approximating V and W by finite-dimensional subspaces $V_h \subset V$ and $W_h \subset W$.

Furthermore we make the simplifying assumption $W_h = V_h$ which according to [11] is called "conform". So we are seeking a finite-dimensional solution $u_h \in V_h$ which approximates u. In other words, u_h can be expressed in the basis of V_h and we have to derive a linear system of equations for the coefficients of u_h in this particular basis.

3.2.1 Linear System of Equations

In order to derive the linear system of equations corresponding to the partial differential equation we state the "discretized" version of (3.3):

$$\int_\Omega k_1\,(u_h)_x\,v_x + k_2\,(u_h)_y\,v_y + \int_\Omega \beta_1\,(u_h)_x\,v + \beta_2\,(u_h)_y\,v + \int_\Omega c(x)\,u\,v = \int_\Omega f\,v + \int_\Omega (u_h)_t\,v \qquad \forall v \in V_h$$

and rewrite it using the bilinear and linear forms A, M and l defined in (3.4). Hence the PDE becomes

$$A(u_h, v) = l(v) + m(u_t, v) \qquad \forall\, v \in\, V_h. \tag{3.7}$$

Suppose now that the dimension of V_h is $m \in \mathbb{N}$ and that there is a basis $\varphi_1, \ldots, \varphi_m$ of V_h. Then the solution u_h can be written as

$$u_h(x, y, t) = \sum_{i=1}^{m} U_i(t)\,\varphi_i(x, y).$$

Furthermore it suffices for the discrete variational equation (3.7) to be true that it holds for any of the basis functions so that by inserting u_h we procure

$$\sum_{i=1}^{m} U_i\,A(\varphi_i, \varphi_j) = l(\varphi_j) + \sum_{i=1}^{m} U_i'\,m(\varphi_i, \varphi_j) \qquad \text{for } j = 1, \ldots m.$$

Pending the time discretization we can now obtain a linear system of equations by defining matrices which shall also be denoted by A and M in an almost obvious way and a right-hand-side F as follows:

$$\begin{aligned} a_{ij} &= A(\varphi_i, \varphi_j) \qquad i, j = 1, \ldots, m \\ F_j &= l(\varphi_j) \qquad j = 1, \ldots, m \\ m_{ij} &= m(\varphi_i, \varphi_j) \qquad i, j = 1, \ldots, m, \end{aligned} \tag{3.8}$$

where $A = (a_{ij})_{i,j=1,\ldots,m}$, $M = (m_{ij})_{i,j=1,\ldots,m}$ and $F = (F_1, \ldots F_m)^T$ and thus we obtain the following linear system of equations

$$A\,U \;=\; F \;+\; MU'. \tag{3.9}$$

Having derived a linear system of equations representing the differential equation we now have to deal with the constraints. We demand that they be satisfied pointwise for the interpolation points $\zeta_i \in \Gamma_D$ (cf. Section 3.2.2 for an explanation) in the Dirichlet boundary Γ_D, i. e.

$$\begin{aligned} u_h(\zeta_i) &= \sum_{j=1}^{m} U_j(t)\varphi_j(\zeta_i) \\ &\overset{!}{=} g_D(\zeta_i) \qquad \forall\, i \,\text{with}\; \zeta_i \in \Gamma_D. \end{aligned}$$

By defining a matrix $B = (b_{ij})_{i,j=1,\ldots,m}$ and a vector $G = (g_i)_{i=1,\ldots,m}$ as

$$\begin{aligned} b_{ij} &= \varphi_j(\zeta_i) \\ g_i &= g_D(\zeta_i) \end{aligned}$$

one can also write these pointwise constraints in a linear system of equations:

$$BU = G. \tag{3.10}$$

But we have to keep in mind that we have yet to specify the approximation space V_h. Moreover in order to be able to solve this system we have to compute the entries of the matrices A and M and the appearing vectors which requires numerical integration and we have to discretize the time derivative.

3.2.2 Triangulation

This section on triangulating is supposed to provide some basic definitions concerning triangulations and corresponding vocabulary. Furthermore we attain some insight on what elementwise regularity means.
But first we should start with the basic definitions which are in their entirety taken from [11].

Definition 3.2.1 *A decomposition $\mathcal{T} = \{T_1, \ldots, T_M\}$ of Ω is called admissible if the following three properties are satisfied:*

1. $\bar{\Omega} = \bigcup_{i=1}^{M} T_i$
2. *If $T_i \cap T_j$ consists of one point only, it is a vertex of both T_i and T_j.*
3. *If $T_i \cap T_j$ consists of more than one point, $T_i \cap T_j$ is an edge of both T_i and T_j.*

Heuristically speaking this definition means that unless empty the intersection of two elements of the decomposition is a point (common vertex of both) or a common edge of both and the union of all should be the whole domain Ω (or more precisely its closure).
All triangulations used in this thesis satisfy the following important property:

Definition 3.2.2 (Shape-Regularity) *A family $\{\mathcal{T}_h\}_h$ of decompositions is called shape regular (quasi-uniform) if there is some $\kappa > 0$ such that any $T \in \mathcal{T}_h$ contains a disc of diameter $2\rho_T$ with $\rho_T \geq \frac{h_T}{\kappa}$ and h_T is half the diameter of T.*

Moreover we can state the following lemma on elementwise defined functions:

Lemma 3.2.3 *Let Ω be a bounded domain and $k \geq 0$. Then an arbitrarily often differentiable function $v : \bar{\Omega} \to \mathbb{R}$ is in $H^k(\Omega)$ if and only if it is in $C^{k-1}(\bar{\Omega})$.*

A proof of this lemma can be found in [11].
Furthermore I would like to stress that we are going to use non-uniform but still shape-regular grids which ensures that the more critical a region of the domain is the more degrees of freedom are placed there. Since this is done apriori based on the facts pointed out in the previous chapter this helps to considerably improve the solution but is still a step short of actual adaptivity based on aposteriori error estimators.
In Figure 3.1 we illustrate the shape of the grids we use in a schematic way, i. e. by plotting such a grid for far less unknowns than used during our computations.

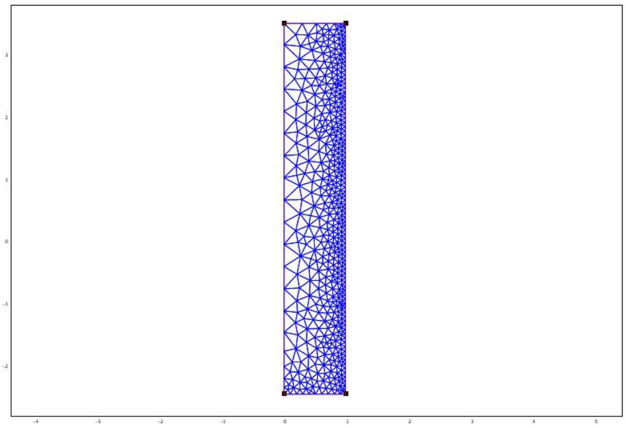

Figure 3.1: Schematic illustration of the non-uniformly refined meshes used in this thesis

3.2.3 Finite Element Spaces

Before starting with a formal definition it is my intention to point out several of the characteristics finite element spaces are supposed to have according to [11]:

- The computational domain is decomposed to polytopes, mostly triangles or rectangles in 2D and tetrahedrons or hexahedrons in 3D.
- Most of the time the restrictions to those entities are polynomials.
- The degree of the polynomial on a certain element is a local property.
- Overall the elements are supposed to be of regularity C^k.

First I will give a formal definition of what finite elements are.

Definition 3.2.4 (Finite Element) *A finite element is a triple (T, Π, Σ) with the following properties:*

1. *T is a polytope in $\mathbb{R}^d$*
2. *Π is a finite-dimensional subspace of $C(T)$. Its dimension is denoted by s and its elements are referred to as shape functions.*
3. *Σ is a set of s linearly independent functionals over Π. Any $p \in \Pi$ is uniquely characterized by the values of the s functionals from Σ.*

Next we will explain what an affine family is, which ensures that we can transform any element to a fixed reference element.

Definition 3.2.5 (Affine Family) *A family of finite element spaces S_h with decompositions $\mathcal{T}_h$ of Ω is called an affine family if there exists an element $(T_{ref}, \Pi_{ref}, \Sigma_{ref})$ with the following property: For any $T_j \in \mathcal{T}_h$ there is an affine mapping $F_j : T_{ref} \to \mathcal{T}_j$ such that for any $v \in S_h$ the restriction to T_j takes on the following shape:*

$$v(x) = p(F_j^{-1}x) \qquad \textit{for } p \in \Pi_{ref}.$$

In addition, any functional $l \in \Sigma$ looks like

$$l(v) = l_{ref}(p) \qquad \textit{with } p = v \circ F,\ l_{ref} \in \Sigma_{ref}.$$

Both these definitions were taken from [11].
Heuristically speaking the idea behind affine families is that it suffices to define the finite element functions on a reference element, while on other elements they are defined via an affine linear transformation to the reference element. This is extremely helpful in obtaining error estimates because the characteristics of the triangulation translate to the characteristics of the transformation matrix and can then be incorporated through the transformation formula of the Lebesgue integral. Moreover it enables us to give estimates on the reference element and by same the means extend them to the whole domain.
Now I want to introduce the most prominent examples of finite elements, the so-called Lagrange elements. They are actually polynomials of a given order and according to [11] and [96] look as follows:

- first order $\mathcal{M}_0^1$
  $u \in C^0(\Omega)$, $\Pi_{ref} = \mathcal{P}_1$, $\dim(\Pi_{ref}) = 3$
- second order $\mathcal{M}_0^2$
  $u \in C^0(\Omega)$, $\Pi_{ref} = \mathcal{P}_2$, $\dim(\Pi_{ref}) = 6$
- third order $\mathcal{M}_0^3$
  $u \in C^0(\Omega)$, $\Pi_{ref} = \mathcal{P}_3$, $\dim(\Pi_{ref}) = 10$
- higher orders analogously.

This means that in the case of second order elements, for instance, the finite element function is overall continuous on Ω but the restrictions to certain elements are second order polynomials, which look as follows in a two dimensional space:

$$p(x, y) = a_0 + a_1\, x + a_2\, y + a_3\, x^2 + a_4\, x\, y + a_5\, y^2.$$

Hence this leads to six degrees of freedom $a_0, \ldots, a_5$ as stated.
In order to make this more palpable I would like to state a remark on interpolation with such polynomials in the case of triangular elements, which is exactly what we are going to use when solving the equation:

Proposition 3.2.6 *Let $t \geq 0$. In the triangle T be given $s = 1+2+\ldots+(t+1)$ points $z_1, \ldots, z_s$ on $(t+1)$ lines. Then for any $f \in C(T)$ there is one and only one polynomial p of degree t satisfying:*

$$p(z_i) = f(z_i) \qquad \textit{for } i = 1, \ldots, s. \tag{3.11}$$

The following proof of this statement is taken from [11].
Proof - Proposition 3.2.6:
The proof will follow the strategy of mathematical induction. We note that there is nothing to be shown for $t = 0$ and assume the desired statement to already proven for $t - 1$.
Because of the invariance under affine-linear transformations we can assume an edge to coincide with the x-axis and furthermore this edge be the one with the points $z_1, \ldots, z_{t+1}$. Then there exists a polynomial $p_0 = p_0(x)$ such that

$$p_0(z_i) = f(z_i) \qquad \text{for } i = 1, \ldots, t+1.$$

By assumption there is also a polynomial $q = q(x, y)$ of degree $t - 1$ which satisfies:

$$q(z_i) = \frac{1}{y_i}\,(f(z_i) - p_0(z_i)) \qquad \text{for } i = t+2, \ldots, s.$$

Then $p(x, y) = p_0(x) + y\,q(x, y)$ satisfies (3.11).

□

This statement can be used to more formally define Lagrange-elements.

Definition 3.2.7 (Lagrange-Elements) *Concerning a finite element space there be a set of points such that the finite element functions are characterized by the values at those points. The set of functions which do not vanish for one and only one of these points constitute a nodal basis which is also referred to as Lagrange elements.*

3.2.4 Error Estimates

Eventually I revisit some of the basic error estimates for finite element discretizations of elliptic problems. Although we are dealing with a parabolic problem this is reasonable because semi-discrete and fully discrete estimates as described in Section 3.5 depend on the approximation properties of finite elements. More on that can be found in [64].
Elliptic problems can be written in the following variational form:

$$a(u, v) = (f, v) \qquad \forall\, v \in V.$$

Other theoretic results (cf. Appendix B) teach us that there is a unique solution to this problem if V is a Hilbert space and a is continuous and V-elliptic, which is defined as follows.

Definition 3.2.8 (Continuity, Ellipticity)

1. *A bilinear form $a : V \times V \to \mathbb{R}$ is called continuous if there is a constant $C < \infty$ such that*

$$|a(v, w)| \leq C\,||v||_V\,||w||_V \qquad \forall\, v, w \in V.$$

2. *A bilinear form $a : V \times V \to \mathbb{R}$ is called V-elliptic if there is some $\alpha > 0$ such that*

$$|a(w, w)| \geq \alpha\,||w||_V^2 \qquad \forall\, w \in V.$$

α is called the ellipticity of a.

In addition, we consider the discretized elliptical problem, for the solution of which we are seeking error estimates:

$$a(u_h, v) = (f, v) \qquad \forall\, v \in V_h.$$

The most fundamental error estimate, which is relatively easy to procure and which serves as the starting point for further considerations, is the Céa-Lemma.

Lemma 3.2.9 (Céa-Lemma) *Consider a V-elliptic bilinear form a and denote by u_h and u the respective solutions to the variational problems in V and V_h. Then the following estimate holds:*

$$||u - u_h||_V \leq \frac{C}{\alpha} \inf_{v_h \in V_h} ||u - v_h||_V.$$

The following proof and the above definitions are taken from [11].
Proof - Lemma 3.2.9:
By subtraction of the definitions of u and u_h we immediately obtain the so-called Galerkin-orthogonality:

$$a(u - u_h, v) = 0 \qquad \forall\, v \in V_h.$$

Consider now $v_h \in V_h$ and $v = u_h - v_h \in V_h$. Because of the properties of a this yields:

$$\begin{aligned}\alpha||u-u_h||_V^2 &\leq a(u-u_h, u-u_h)\\ &= a(u-u_h, u-v_h) + \underbrace{a(u-u_h, v_h-u_h)}_{=0}\\ &\leq C||u-u_h||_V||u-v_h||_V.\end{aligned}$$

Dividing by α and $||u-u_h||_V$ as well as taking into consideration that v_h was arbitrary establishes the desired inequality and completes the proof.

□

Having stated the Céa-lemma we recognize that in order to give an estimate for the approximate solution u_h we have to estimate the infimum. Therefore we will examine how well u is approximated by element-wise polynomials.
Before we are able to write down the estimate we have to explain the meaning of several expressions appearing in the below statements. To do so assume a decomposition $\mathcal{T}_h = \{T_1, \ldots, T_M\}$ of Ω and an $m \geq 1$ and define a norm in the following way:

$$||v||_{m,h} := \sqrt{\sum_{T_j \in \mathcal{T}_h} ||v||_{m,T_j}}.$$

Here $||\cdot||_m$ denotes the m-th Sobolev norm. If $m \geq 2$ the embedding theorems by Sobolev (cf. Appendix B) yield $H^m(\Omega) \subset C^0(\Omega)$ which strictly speaking means that any element $[v] \in H^m(\Omega)$ has a continuous representative v, to which in turn polynomial interpolation can be applied. In other words, there is a uniquely defined polynomial $I_h v \in V_h$ and we seek to estimate $||v - I_h v||_{m,h}$ through $||v||_{t,\Omega}$ for some $t \geq m$.
The main result on the quality of polynomial interpolation is the following:

Proposition 3.2.10 *Let $t \geq 2$ and $\mathcal{T}_h$ be a shape-regular (quasi-uniform) triangulation of Ω. Then interpolation by element-wise polynomials of degree $t-1$ satisfies the inequality*

$$||u - I_h u||_{m,h} \leq c(\Omega, \kappa, t)\, h^{t-m}\, |u|_{t,\Omega} \qquad \textit{for } u \in H^t(\Omega).$$

Concerning the validity of Proposition 3.2.10 we should strongly emphasize that $u \in H^t$ is required. If such a property is satisfied depends on the PDE. More precisely regularity theory would have to be carried out. Such is done in Section 3.9. Besides that Proposition 3.2.10 is the most important estimate concerning numerics because it tells us the convergence order we can expect and spot if our problem is smooth enough depending on the order of the approximating polynomials. A proof is given in [11].
But one can attain even more results by exploiting the given regularity of a problem. These further accomplishments are error estimates in both the energy- and the L^2-norm.

Theorem 3.2.11 (Energy norm) *By $\mathcal{T}_h$ a shape regular (quasi-uniform) family of decompositions of Ω be denoted. Then in the case of linear, quadratic or cubic triangle elements the following estimate holds for the finite element solution u_h:*

$$\begin{aligned}||u-u_h||_{H^1(\Omega)} &\leq c\,h\,||u||_{H^2(\Omega)}\\ &\leq c\,h\,||f||_{L^2(\Omega)}.\end{aligned}$$

Theorem 3.2.12 (L^2-norm) *Under the assumptions of Theorem 3.2.11 and suitable further presumptions we can state the following:*

1. *If $u \in H^1(\Omega)$ the following estimate is true:*

$$||u - u_h||_{L^2(\Omega)} \leq C\, c\, h\, ||u||_{H^1(\Omega)}.$$

2. *If furthermore $f \in L^2(\Omega)$ and hence $u \in H^2(\Omega)$ we even have*

$$||u - u_h||_{L^2(\Omega)} \leq C^2\, c^2\, h^2\, ||f||_{L^2(\Omega)}.$$

What is meant by suitable further presumptions can as well as proofs of both these theorems be found in [11].

3.3 Numerical Quadrature

This section shall be devoted to the effects of approximating the integrals in the bilinear and linear forms of Section 3.2 by quadrature formulas rather than computing them exactly.

3.3.1 Adjustments by Quadrature

As we know from (3.8) the entries of both the mass and the stiffness matrices M and A consist of integrals we, in general, cannot compute exactly. They look as follows:

$$\begin{aligned} a_{ij} &= \int_\Omega (K\nabla\varphi_j)\cdot\nabla\varphi_i + (\beta\cdot\nabla\varphi_j)\,\varphi_i + c\,\varphi_i\varphi_j \\ &=: \int v_{ij} \\ m_{ij} &= \int_\Omega \varphi_i\varphi_j. \end{aligned}$$

The assembling is performed elementwise by transformation to a certain reference element $\hat{T}$ where the integrals are of the following type and approximated in the following manner:

$$\int_{\hat{T}} \hat{v}(\hat{x})\, d\hat{x} \approx \sum_{i=1}^{R} \hat{\omega}_i\, \hat{v}(\hat{\xi}_i).$$

R is the number of quadrature points, $\xi_i, i = 1, \ldots, R$, are the quadrature points and $\omega_i, i = 1, \ldots, R$, are the respective weights. If we are dealing with an affine family there exists an affine-linear transformation $F_T(\hat{x}) = B\hat{x} + d$ for some B and d and one can compute the integral on the element T via the transformation formula:

$$\begin{aligned} \int_T v(x)dx &= \int_{\hat{T}} v(\hat{x})\, d\hat{x}\, |\det(B)| \\ &\approx \sum_{i=1}^{R} \omega_{i,T}\, v(\xi_{i,T}), \end{aligned}$$

where $\omega_{i,T} = \hat{\omega}_i\, |\det(B)|$ and $\xi_{i,T} = F(\hat{\xi}_i)$. Now the original stiffness matrix and the mass matrix, or to be more precise the bilinear forms and linear forms by which they are defined, are replaced by those forms one obtains when using quadrature. So we have:

$$a_h(v, w) := \sum_{T\in\mathcal{T}_h} \left[\sum_{l=1}^{R} \omega_{l,T}\, (((K\nabla v)\cdot\nabla w)(\xi_{l,T}) + (\beta\cdot\nabla v\, w)(\xi_{l,T}) + (c\, v\, w)(\xi_{l,T})) \right] \quad (3.12)$$

$$l_h(v) := \sum_{T\in\mathcal{T}_h} \sum_{l=1}^{R} \omega_{l,T} (f\, v)(\xi_{l,T}). \quad (3.13)$$

The above definitions and arguments are taken from [64].

3.3.2 Error Estimates

In this subsection I am going to write down several statements on the error caused by the replacements (3.12) and (3.13) as well as on the measures that might be taken to ensure that the error remains small.
First we have to introduce two variational problems, the first being the original problem and the second the approximated problem:

$$\begin{aligned} u \in V : a(u,v) &= l(v) \qquad \forall\, v \in V \\ u_h \in V_h : a_h(u_h,v) &= l(v) \qquad \forall\, v \in V_h, \end{aligned}$$

where $V_h \subset V$. The main estimate of the error is provided by the so-called first lemma of Strang:

Theorem 3.3.1 (First Lemma of Strang) *Presuppose a_h to be uniformly elliptic on V_h, i. e. there be an $\alpha > 0$ such that for any $h > 0$ and $v \in V_h$*

$$a_h(v,v) \geq \alpha\, ||v||^2.$$

Furthermore the bilinear form a be assumed to be continuous on $V \times V$. Then there is a constant C which is independent of V_h such that

$$||u - u_h|| \leq C \left(\inf_{v \in V_h} \left(||u - v|| + \sup_{w \in V_h} \frac{|a(v,w) - a_h(v,w)|}{||w||} \right) + \sup_{w \in V_h} \frac{|l(w) - l_h(w)|}{||w||} \right).$$

Interpreting this theorem we can make the following remarks:

- If a and a_h as well as l and l_h happen to equal one another, which is the case if the integrals are computed exactly, the statement is just the Céa-Lemma.
- One has to ensure uniform V_h-ellipticity.
- The consistence errors in the supremum expressions should vanish as h approaches zero.

According to [64] other lemmas and theorems state that the latter can be accomplished of order k if one uses a regular mesh, if u is sufficiently smooth, i. e. H^{k+1}, and if one uses quadrature formulas which are exact for polynomials of degree $2k - 2$. Also in [64] one can find a proof of this theorem.
Eventually I would like to state and prove a criterion for the ellipticity condition, also taken from [64]:

Lemma 3.3.2 *The bilinear form a be V-elliptic and there be a function $C(h)$ satisfying $C(h) \to 0$ for $h \to 0$ such that*

$$\begin{aligned} A_h(v) &\leq C(h)||v|| \\ A_h(v) &:= \sup_{v \in V_h} \frac{|a(v,w) - a_h(v,w)|}{||w||}. \end{aligned}$$

Then there is an $\bar{h} > 0$ such that a_h is uniformly V_h-elliptic for $h \leq \bar{h}$.

Proof - Lemma 3.3.2:
By assumption we have the following two estimates:

$$\begin{aligned} \alpha\, ||v||^2 &\leq a_h(v,v) + a(v,v) - a_h(v,v) \\ |a(v,v) - a_h(v,v)| &\leq A_h(v)\, ||v|| \\ &\leq C(h)\, ||v||^2. \end{aligned}$$

If we choose $\bar{h}$ in a way that $C(h) \leq \alpha/2$ for $h \leq \bar{h}$ it has the desired property:

$$\begin{aligned} a_h(v,v) &\geq \alpha\,||v||^2 - (a(v,v) - a_h(v,v)) \\ &\geq \alpha\,||v||^2 - |a(v,v) - a_h(v,v)| \\ &\geq \alpha\,||v||^2 - C(h)\,||v||^2 \\ &\geq \left(\alpha - \frac{\alpha}{2}\right)||v||^2. \end{aligned}$$

Hence for $h \leq \bar{h}$ a_h is uniformly elliptic with the constant $\frac{\alpha}{2}$.

□

3.4 Time Discretization

The subject matter of this section is a review on multistep-methods for ordinary differential equations which is of interest to us because after discretizing in space we obtain a system of ODEs. In Section 3.4.1 I will introduce us to the so-called backward-differentiation formulas, in Section 3.4.2 error estimates shall be provided and in Section 3.4.3 this is extended to systems of ODEs.
To do so we will always deal with the following problem:

Problem 3.4.1 *Let $I \subset \mathbb{R}$ be an interval with $t_0 \in I$. Find $y \in C^1(I)$ such that*

$$\begin{aligned} y'(t) &= f(t, y(t)) \qquad \text{for } t \in I \\ y(t_0) &= y_0, \end{aligned}$$

where $f : I \times (-\infty, +\infty) \to \mathbb{R}$ be simultaneously continuous in t and y.

3.4.1 Backward-Differentiation Formulas

In order to elucidate what the so-called backward differentiation formulas are we first consider a slightly more general case, the so-called q-step-methods.

Definition 3.4.2 (q-step-method) *For $q \geq 0$ a q-step-method is a method where for all $n \geq q-1$ u_{n+1} depends on u_{n+1-q} but not on u_k with $k < n+1-q$.*

Moreover we shall restrain ourselves to linear multi-step-methods, i. e. methods of the following type:

$$u_{n+1} = \sum_{j=0}^{p} a_j\, u_{n-j} + h\sum_{j=0}^{p} b_j f_{n-j} + hb_{-1}f_{n+1} \qquad \text{for } n = p, p+1, \ldots \tag{3.14}$$

These schemes are $(p+1)$-step-methods and they are implicit if $b_{-1} \neq 0$ and explicit else.
The backward-differentiation formulas, or BDF-schemes as they are abbreviated, are a special case of these methods, namely those where $b_j = 0$, $j = 0, \ldots, p$. Hence BDF-schemes look as follows:

$$u_{n+1} = \sum_{j=0}^{p} a_j\, u_{n-j} + hb_{-1}f_{n+1} \qquad \text{for } n = p, p+1, \ldots \tag{3.15}$$

As we will see in the next section these schemes are only stable for $p \leq 5$, so that it makes only sense to list the coefficients for those schemes. The list in Table 3.1 is taken from [80].

p	a_0	a_1	a_2	a_3	a_4	a_5	b_{-1}
0	1	0	0	0	0	0	1
1	$\frac{4}{3}$	$-\frac{1}{3}$	0	0	0	0	$\frac{2}{3}$
2	$\frac{18}{11}$	$-\frac{9}{11}$	$\frac{2}{11}$	0	0	0	$\frac{6}{11}$
3	$\frac{48}{25}$	$-\frac{36}{25}$	$\frac{16}{25}$	$-\frac{3}{25}$	0	0	$\frac{12}{25}$
4	$\frac{300}{137}$	$-\frac{300}{137}$	$\frac{200}{137}$	$-\frac{75}{137}$	$\frac{12}{137}$	0	$\frac{60}{137}$
5	$\frac{360}{147}$	$-\frac{450}{147}$	$\frac{400}{147}$	$-\frac{225}{147}$	$\frac{72}{147}$	$-\frac{10}{147}$	$\frac{60}{147}$

Table 3.1: Coefficients of the BDF-formulas for $p \leq 5$

Example 3.4.3 *We consider the special case $p = 1$ and show that in this case the BDF-formula happens to be the implicit-Euler time discretization. This is true because:*

$$\begin{aligned} u_{n+1} &= \sum_{j=0}^{p} a_j u_{n-j} + h\, b_{-1}\, f_{n+1} \\ &= u_n + h\, f_{n+1}. \end{aligned}$$

This is apparently the implicit Euler.

3.4.2 Error Estimates

In this section some results on convergence and stability of multi-step-methods shall be stated and for the most part also be proved. Since BDF-schemes are a special case of those the results remain true for them. The presented results and definitions are taken from [80].
First recall the definition of consistence:

Definition 3.4.4 *The multi-step-method (3.14) is said to be consistent if the local truncation error satisfies*

$$\tau(h) \to 0 \quad \text{for } h \to 0.$$

If, in addition, $\tau(h) = \mathcal{O}(h^q)$ holds for some $q \geq 1$ we say that the method is of order q.

In the previous definition we made use of the local truncation error which is defined as follows:

Definition 3.4.5 (Local Truncation Error) *The local truncation error (LTE) $\tau_{n+1}(h)$ of the multi-step-method (3.14) at t_{n+1} for $n \geq p$ is defined as*

$$h\,\tau_{n+1}(h) = y_{n+1} - \left[\sum_{j=0}^{p} a_j y_{n-j} + h \sum_{j=-1}^{p} b_j y'_{n-j} \right],$$

where $y_{n-j} = y(t_{n-j})$ and $y'_{n-j} = y'(t_{n-j})$ for $j = -1, \ldots, p$.

The first step on our way to convergence results is to know under which circumstances a scheme is at least consistent. This is provided by the following theorem, which states conditions the coefficients have to fulfil.

Theorem 3.4.6 *The multi-step-method (3.14) is consistent if and only if the following algebraic relations hold:*

$$\begin{aligned} \sum_{j=0}^{p} a_j &= 1 \\ -\sum_{j=0}^{p} j\, a_j + \sum_{j=-1}^{p} b_j &= 1. \end{aligned}$$

If furthermore $y \in C^{q+1}(I)$ for some $q \geq 1$, the scheme is of order q if and only if the above conditions hold and for $i = 2, \ldots, q$

$$\sum_{j=1}^{p} (-j)^i + i \sum_{j=-1}^{p} (-j)^{i-1} b_j = 1.$$

Proof - Theorem 3.4.6:
First y and f are expanded in a Taylor series which yields

$$\begin{aligned} y_{n-j} &= y_n - jh\, y_n' + \mathcal{O}(h^2) \\ f_{n-j} &= f_n + \mathcal{O}(h). \end{aligned}$$

Reinserting into the multi-step-scheme and neglecting terms of order higher than one leads to:

$$\begin{aligned} y_{n+1} &- \sum_{j=0}^{p} a_j\, y_{n-j} - h \sum_{j=-1}^{p} b_j\, f_{n-j} \\ &= y_{n+1} - \sum_{j=0}^{p} a_j\, y_n + h \sum_{j=0}^{p} j\, a_j\, y_n' - h \sum_{j=-1}^{p} b_j\, f_n - \mathcal{O}(h^2) \left(\sum_{j=0}^{p} a_j - \sum_{j=-1}^{p} b_j \right) \\ &= y_{n+1} - \sum_{j=0}^{p} a_j\, y_n - h\, y_n' \left(-\sum_{j=0}^{p} j a_j + \sum_{j=-1}^{p} b_j \right) - \mathcal{O}(h^2) \left(\sum_{j=0}^{p} a_j - \sum_{j=-1}^{p} b_j \right). \end{aligned}$$

In the last step we exploited the ODE and replaced f_n by y_n'. And by comparison with the definition of the local truncation error we procure:

$$h\, \tau_{n+1}(h) = y_{n+1} - \sum_{j=0}^{p} a_j y_n - h y_n' \left(-\sum_{j=0}^{p} j a_j + \sum_{j=-1}^{p} b_j \right) - \mathcal{O}(h^2) \left(\sum_{j=0}^{p} a_j - \sum_{j=-1}^{p} b_j \right).$$

And hence

$$\begin{aligned} \tau_{n+1}(h) &= \frac{y_{n+1} - y_n}{h} + \frac{y_n}{h} \left(1 - \sum_{j=0}^{p} a_j \right) + y_n' \left(-\sum_{j=0}^{p} j a_j + \sum_{j=-1}^{p} b_j \right) \\ &\quad - \mathcal{O}(h) \left(\sum_{j=0}^{p} a_j - \sum_{j=-1}^{p} b_j \right). \end{aligned}$$

Since $\frac{y_{n+1} - y_n}{h} \to y_n'$ for $h \to 0$ it follows that $\tau_{n+1}(h)$ converges to zero if and only if the algebraic relations claimed above are satisfied.
The rest of the proof, i. e. the statement concerning higher order can be obtained in exactly the same way by incorporating higher order terms in the initial Taylor series expansions.

□

This criterion for consistence does not suffice to yield convergence. Thus additional conditions have to be fulfilled by the scheme. The below considerations about this aspect are taken from [80].
Given a certain multi-step method we can consider the polynomial

$$\rho(r) = r^{p+1} - \sum_{j=0}^{p} a_j\, r^{p-j}$$

and we denote its roots by $r_0, \ldots, r_n$. Note that if our scheme is consistent 1 is a root. Now we can state the root-condition:

Definition 3.4.7 (Root Condition) *The multi-step-method (3.14) is said to satisfy the root-conditon if all the roots are situated within the unit circle of the complex plane and if those on the boundary are of multiplicity 1. Or equivalently:*

$$\begin{aligned} |r_j| &\leq 1 && \text{for } j = 0, \ldots, p \\ \rho'(r_j) &\neq 0 && \text{for those } j \text{ with } |r_j| = 1. \end{aligned}$$

Now we have all the ingredients necessary to state the main convergence result.

Theorem 3.4.8 *A consistent multi-step-method converges if and only if it satisfies the root-condition and the error in the initial data converges to zero for $h \to 0$.*
If furthermore the method is of order q and the initial data error behaves like $\mathcal{O}(h^q)$ for $h \to 0$, it also converges with order q.

Finally we have to note that for $p \leq 5$ the BDF-schemes satisfy the root-condition, whereas according to [80] for $p > 5$ this cannot be ensured generally.
Altogether this yields that for $p \leq 5$ the BDF-schemes converge. A proof of Theorem 3.4.8 can be found the aforementioned reference as well as further analysis and equivalence statements on the root-condition.

3.4.3 Systems of ODEs

So far we have focused solely on one-dimensional equations. But actually we have to deal with a system of potentially coupled ordinary differential equations. As long as we can find a way to decouple the equations that does not pose a problem because in this case the system could be treated as finitely many one-dimensional ones for which we already have convergence results.
The system we have to solve is equation (3.9) which resulted from the spatial discretization and which can also be written as

$$M^{-1}AU - U' = M^{-1}F.$$

If $M^{-1}A$ is diagonalizable there is matrix Q such that $\Lambda = Q^{-1}M^{-1}AQ$ is diagonal and we can transform the system in the variable U to an equivalent one in the variable $z' := Q^{-1}M^{-1}AU$. We obtain:

$$A^{-1}MQz' = Qz - M^{-1}F$$

or equivalently:

$$z' = \underbrace{Q^{-1}M^{-1}AQ}_{=:\Lambda} z + Q^{-1}M^{-1}AM^{-1}F.$$

In this way we have attained a diagonal system of equations:

$$z' = \Lambda z - \tilde{F},$$

where $\tilde{F} := Q^{-1}M^{-1}AM^{-1}F$ and to which the statements for one-dimensional ODEs can be applied.

3.5 Semi-Discrete Estimates

In this section I will state several so-called semi-discrete error estimates for parabolic problems. They estimate how the error caused by the spatial discretization spreads over time, while time is not yet discretized.

Unfortunately we are actually interested in an estimate for the fully discrete equation but nonetheless semi-discrete estimates can be helpful because one always has the following:

$$||u(t_k) - u_h^k|| \leq \underbrace{||u(t_k) - u_h(t_k)||}_{\text{semi-discrete error}} + \underbrace{||u_h(t_k) - u_h^k||}_{\text{error of the time discretization}} .$$

Concerning the notation: u denotes the exact solution to the variational equation (3.5), u_h the solution to the system of ODEs (3.9) resulting from the spatial discretization and u_h^k the fully discrete solution at time step k.
I will now give three estimates for the error of the semi-discrete equation where in contrast to our initial boundary value problem from Sections 2.5, 2.6 and 2.8 vanishing Dirichlet-boundary conditions are presumed. The first is in the L^2-norm, the second in the H^1- or energy norm and the third one for higher order finite elements. In all the following theorems it shall always be assumed that the solution is sufficiently smooth for the norms on the right-hand-side to be reasonably defined. Moreover v and v_h denote the respective initial data.
The estimates are taken from [64] where one can also find proofs.

Theorem 3.5.1 *If u and u_h are the solutions to the variational and semi-discrete variational equations (3.5) and (3.7) respectively for $t > 0$ we have the following estimate:*

$$||u_h(t) - u(t)||_{L^2} \leq ||v_h - v||_{L^2} + c\,h^2 \left(||v||_{L^2} + \int_0^t ||u_t||_{L^2} \right).$$

The next estimate deals with the H^1-semi-norm, i. e. it estimates the error made in the gradient:

Theorem 3.5.2 *Under the same assumptions as in the previous theorem one obtains for $t > 0$:*

$$|u_h(t) - u(t)|_{H^1} \leq |v_h - v|_{H^1} + C\,h \left\{ ||v||_{H^2} + ||u(t)||_{H^2} + \left(\int_0^t ||u_t||_{H^1}^2 \right)^{1/2} \right\}.$$

And eventually I state an estimate which yields a higher order approximation under suitable further assumptions:

Theorem 3.5.3 *If u and u_h are defined as previously and if furthermore the Ritz-projection satisfies the following estimate*

$$||R_h w - w|| \leq C\,h^r ||w||_{H^r} \tag{3.16}$$

we even have

$$||u_h(t) - u(t)||_{L^2} \leq Ch^r \left(||v||_{H^r} + \int_0^t ||u_t|| \right)$$

for $t \geq 0$.

The last theorem is indeed one for higher order finite elements because the additional condition (3.16) is only fulfilled if the finite element order is high enough. The Ritz-projection is, for instance, defined in [11]. Roughly speaking it is the a-orthogonal projection onto $V_h \subset V$.

3.6 The Effect of Probability-Weighting on Convergence

In this section I am going to show that a probability-weighted finite element solution still converges to the probability-weighted exact solution if the unweighted solution converges. This has to be examined because the above error estimates only account for the PDE solution process, whereas the process of probability weighting uses the finite element solution as its starting point.

Therefore we consider a prepayment probability P_p^h and a default probability P_d^h as defined in Section 2.3 which depend on h since in checking for optimality of either termination method the numerical PDE solution is used. We then compute

$$\tilde{U}_i^k = (1 - P_p^h - P_d^h)\, U_i + P_p^h\, (1 + X)\, F_t + P_d^h\, (1 + X)\, \exp(y/A)$$

for some transaction cost X in order to adjust the coefficients of the FEM solution at time step k. By doing so we indeed weight the whole finite element solution u_h since

$$\begin{aligned}
\tilde{u}_h^k &:= \sum_{i=1}^{DOF} \tilde{U}_i^k \varphi_i^h \\
&= (1 - P_p^h - P_d^h) \sum_{i=1}^{DOF} U_i^k \varphi_i^h + \sum_{i=1}^{DOF} (1 + X)\, P_p^h\, F_t\, \varphi_i^h + \sum_{i=1}^{DOF} (1 + X)\, P_d^h\, \exp(y/A)\, \varphi_i^h \\
&= (1 - P_p^h - P_d^h) \sum_{i=1}^{DOF} U_i^k \varphi_i^h + (1 + X)\, P_p^h\, F_t \underbrace{\sum_{i=1}^{DOF} \varphi_i^h}_{=1} + (1 + X)\, P_d^h\, \exp(y/A) \underbrace{\sum_{i=1}^{DOF} \varphi_i^h}_{=1} \\
&= (1 - P_p^h - P_d^h)\, u_h^k + (1 + X)\, P_p^h\, F_t + (1 + X)\, P_d^h\, \exp(y/A).
\end{aligned}$$

Accordingly we have a weighted exact solution $\tilde{u}$:

$$\tilde{u} := (1 - P_p - P_d)\, u + P_d\, (1 + X)\, \exp(y/A) + P_p\, (1 + X)\, F_t$$

Next, we prove that $\tilde{u}_h^k$ converges to $\tilde{u}(t_k)$ in the following sense:

Theorem 3.6.1 (Probability-weighted Convergence) *Assume u to be an exact solution of the initial boundary value problem (2.6) subject to the initial and boundary conditions stated in Section 2.6 and u_h^k the respective fully discrete finite element solution at the k-th time step. Moreover a convergence result*

$$||u_h^k - u||_{L^2} \overset{h \to 0}{\longrightarrow} 0 \tag{3.17}$$

be given. If $\tilde{u}_h^k$ and $\tilde{u}$ are defined as above we also have:

$$||\tilde{u}_h^k - \tilde{u}||_{L^2} \overset{h \to 0}{\longrightarrow} 0. \tag{3.18}$$

In addition, the same statement holds for the transformed initial boundary value problem (2.7) subject to the respective modified initial and boundary data in Section 2.8.

To be able to prove this statement we have to recall the Hölder-inequality, a proof of which can be found in [2].

Lemma 3.6.2 (Hölder-inequality) *Let $m \in \mathbb{N}$ and $f_i \in L^{p_i}(\mu)$ for $i = 1, \ldots, m$ and a measure μ with $p_i \in [1, \infty]$ and $p \in [1, \infty]$ satisfying $\sum_{i=1}^m \frac{1}{p_i} = \frac{1}{p}$. Then the product $f_1 f_2 \ldots f_m \in L^p(\mu)$ and the following estimate holds:*

$$||\prod_{i=1}^m f_i||_{L^p} \leq \prod_{i=1}^m ||f_i||_{L^{p_i}}.$$

Proof - Theorem 3.6.1:
The proof consists of two steps. First we have show that the probabilities P_d^h and P_p^h converge to P_d and P_p respectively in L^2. And exploiting that we then procede to show the desired

convergence result for the weighted solutions.
Let us begin with the second item:

$$\begin{aligned}
\|\tilde{u}_h^k - \tilde{u}\|_{L^2} &= \|(1 - P_p^h - P_d^h)\, u_h^k + (1+X)\, P_p^h\, F_t + (1+X)\, P_d^h\, \exp(y/A) \\
&\quad -((1 - P_p - P_d)\, u + P_d\, (1+X)\, \exp(y/A) + P_p\, (1+X)\, F_t)\|_{L^2} \\
&= \|(1 - P_p - P_d)\, u - (1 - P_p^h - P_d^h)\, u_h^k + (P_p - P_d^h)\, (1+X)\, \exp(y/A) \\
&\quad +(P_p - P_p^h)\, (1-X)\, F_t\|_{L^2} \\
&\leq \|(1 - P_p - P_d)\, u - (1 - P_p^h - P_d^h)\, u_h^k\|_{L^2} \\
&\quad +\|(P_d - P_d^h)\, (1+X)\, \exp(y/A)\|_{L^2} + \|(P_p - P_p^h)\, (1-X)\, F_t\|_{L^2} \\
&\overset{\text{Hölder}}{\leq} \|(1 - P_p - P_d)\, u - (1 - P_p^h - P_d^h)\, u_h^k\|_{L^2} \\
&\quad +(1+X)\, \underbrace{\|\exp(y/A)\|_{L^2}}_{<\infty}\, \underbrace{\|(P_d - P_d^h)\|_{L^2}}_{\overset{h\to 0}{\longrightarrow} 0} +(1+X)\, \underbrace{\|F_t\|_{L^2}}_{<\infty}\, \underbrace{\|(P_p - P_p^h)\|_{L^2}}_{\overset{h\to 0}{\longrightarrow} 0}.
\end{aligned}$$

In the last step we took advantage of the fact that the probabilities converge so that apparently we only have to estimate the first expression:

$$\begin{aligned}
\|(1 - P_p - P_d)\, u - (1 - P_p^h - P_d^h)\, u_h^k\|_{L^2} &\leq \|(1 - P_p - P_d)\, u - (1 - P_p - P_d)\, u_h^k\|_{L^2} \\
&\quad +\|(1 - P_p - P_d)\, u_h^k - (1 - P_p^h - P_d^h)\, u_h^k\|_{L^2} \\
&\leq \|1 - P_p - P_d\|_{L^2}\, \|u - u_h^k\|_{L^2} \\
&\quad + \|u_h^k\|_{L^2}\, \|(1 - P_p - P_d) - (1 - P_p^h - P_d^h)\|_{L^2} \\
&\leq \underbrace{\|1 - P_p - P_d\|_{L^2}}_{<\infty}\, \underbrace{\|u - u_h^k\|_{L^2}}_{\to 0 \text{ by } (3.17)} \\
&\quad + \underbrace{\|u_h^k\|_{L^2}}_{<\infty}\, \underbrace{\|(P_p^h - P_p) - (P_d - P_d^h)\|_{L^2}}_{\leq \|(P_p^h - P_p)\|_{L^2} + \|(P_d - P_d^h)\|_{L^2}} \\
&\overset{h\to 0}{\longrightarrow} 0.
\end{aligned}$$

What remains to be shown is the L^2-convergence for the prepayment and default probabilities. Because of the symmetry in their definitions it suffices to show the statement for one of them, say P_p^h. Consider the following difference between the statements checked for the optimality decision:

$$\|(1+X)\, F_t - u_h^k - ((1+X)\, F_t - u)\|_{L^2} \leq \|u - u_h^k\|_{L^2} \overset{h\to 0}{\longrightarrow} 0 \qquad \text{in } L^2.$$

Note that in the limes the optimality decision is the same no matter if accepted or rejected. In other words both indicator functions, one for optimality and one for sub-optimality, convergence in L^2, so that we obtain the following:

$$\begin{aligned}
|P_p - P_p^h| &= |(K_1 - \exp(-\lambda_1))\, 1_{\text{sub-optimal}} + (K_2 - \exp(-\lambda_1))\, 1_{\text{optimal}} - \\
&\quad (K_1 - \exp(-\lambda_1))\, 1_{\text{sub-optimal}}^h + (K_2 - \exp(-\lambda_2))\, 1_{\text{optimal}}^h) \\
&\leq (K_1 - \exp(-\lambda_1))\, |1_{\text{sub-optimal}}^h - 1_{\text{sub-optimal}}| + K_2 - \exp(-\lambda_2)\, |1_{\text{optimal}}^h - 1_{\text{optimal}}|.
\end{aligned}$$

Here we took into consideration that all probabilities from Section 2.3 can be written in the form:

$$P = K - \exp(-\lambda).$$

for some K and λ. By monotonicity of the Lebesgue integral and $||v||_{L^1} \leq ||v||_{L^2} \quad \forall\, v \in L^2$ we obtain:

$$\begin{aligned}
||P_p - P_p^h||_{L^2} \quad &\leq \quad ||(K_1 - \exp(-\lambda_1))\,|1^h_{\text{sub-optimal}} - 1_{\text{sub-optimal}}|\,||_{L^2} \\
&\quad\; + ||(K_2 - \exp(-\lambda_2))\,|1^h_{\text{optimal}} - 1_{\text{optimal}}|\,||_{L^2} \\
&\overset{\text{Hölder}}{\leq} \quad \underbrace{||K_1 - \exp(-\lambda_1)||_{L^2}}_{<\infty}\, \underbrace{||1^h_{\text{sub-optimal}} - 1_{\text{sub-optimal}}||_{L^2}}_{\overset{h\to 0}{\longrightarrow} 0} \\
&\quad\; + \underbrace{||K_2 - \exp(-\lambda_2)||_{L^2}}_{<\infty}\, \underbrace{||1^h_{\text{optimal}} - 1_{\text{optimal}}||_{L^2}}_{\overset{h\to 0}{\longrightarrow} 0}
\end{aligned}$$

Altogether this establishes the desired convergence result (3.18).

□

The just proved result is the central point for our numerical approach because it ensures that the convergence obtained for the PDE part is not hampered by the probability weighting. Furthermore I would like to point out that the same arguments can be applied to the problem obtained by prescribing the weighted solution as initial datum for the time interval of an ensuing month. Doing so we procure convergence for the second and later time steps, so that convergence is ensured for the whole lifespan of the underlying mortgage.

3.7 Truncation of the Domain

In this section I will discuss why the domain has to be truncated and how the boundary values are adjusted.

First we note once again that after transformation the partial differential equation (2.7) is defined on a domain with $\Omega = (0,1) \times (-\infty, +\infty)$, which is apparently infinite. Therefore we have to pick $y_{min} > -\infty, y_{max} < +\infty$ and replace the domain by $(0,1) \times (y_{min}, y_{max})$ whereby we attain a finite computational domain. But doing so we also recognize that our boundary values which have been defined for the limites $y \to \infty$ and $y \to -\infty$ can no longer be imposed without committing an error. However, we choose to tolerate it saying that the conditions are approximately satisfied if y_{min} is selected small enough and if y_{max} is sufficiently large.

The choices we make are:

$$\begin{aligned}
y_{max} &= 3.5 \\
y_{min} &= -2.3.
\end{aligned}$$

This choice can be justified by a scaling of the equation. More precisely we consider $100,000$ to be a typical house price and by assuming this to be the initial debt of a mortgage this would translate to a monthly coupon of $C = 320$ on assuming $c = 5\%$ contractual interest. This can be computed using the formulas from Section 1.2.

We now scale this coupon by a factor $a = 100,000$, thus changing the right-hand-side of the PDE to 1. So we have to alter the left-hand-side in the same way. But because of the linearity of the differential operator $\mathcal{L}$ defined in (3.2) the solution u changes by the same scaling factor, i. e.

$$\begin{aligned}
\mathcal{L}\left(\frac{u}{a}\right) - \partial_t\left(\frac{u}{a}\right) &= \frac{1}{a}\left(\mathcal{L}u - \partial_t u\right) \\
&= -\frac{1}{a}\,C.
\end{aligned}$$

Moreover according to Lemma 1.2.2 we know that for a fixed contractual interest rate c a change in the coupon C to C/a directly leads to a change of the initial debt R_0 to R_0/a and vice versa. Combining that with the aforementioned house prices can be interpreted as initial debt meaning that the upper and lower bounds y_{min} and y_{max} can be set relative to house price 1 rather than $100,000$. Nonetheless the above choice corresponds to

$$\begin{aligned} H_{max} &= 1.0966\,10^8 \\ H_{min} &= 1,005.184 \end{aligned}$$

in real house prices.
As already pointed out in Section 2.6 we have to ensure that the values on separate boundaries fit one another in their common points, the vertices. In our case we have forfeited the property

$$\lim_{y \to -\infty} u(0, y, t) = 0 \qquad \forall\, t > 0$$

by truncation because for $y \to y_{min}$ the boundary condition $\min(C(T-t), F_t(1+X), H_t(y)(1+X))$ remains positive and thus we obtain a jump at this boundary. In order to remove this jump we make use of interpolation.
The easiest way to interpolate would be linear interpolation between the value of the boundary condition and zero on a length $\delta > 0$. But unfortunately this is not only the easiest way but it also has a severe shortcoming in the fact that we would lose regularity as the resulting boundary datum would only be continuous but not in C^1. In order to overcome this we use the following function taken from [72] and [96]:

$$\chi(y, \delta) = I_{\{y > -\delta\}} I_{\{y < \delta\}} \left\{ \frac{1}{2} + \frac{y}{\delta} \left[\frac{15}{16} - \frac{5}{8} \left(\frac{y}{\delta} \right)^2 + \frac{3}{16} \left(\frac{y}{\delta} \right)^4 \right] \right\} + I_{\{y \geq \delta\}}.$$

Here I stands for indicator functions which take on the value one on the specified set and zero else. The advantage of this fifth-degree polynomial is that it smooths the jump from 0 to 1 with a continuous second derivative. Therefore the boundary condition can be mended in the following way:

$$g_{trunc}(y, t) := \chi(y - y_{min} - \delta, \delta)\, g(y, t),$$

where g denotes the original boundary condition. This choice ensures:

$$\begin{aligned} g_{trunc}(y_{min}, t) &= 0 \qquad \forall t > 0 \\ g_{trunc}(y, t) &= g(y, t) \qquad \forall y > y_{min} + 2\delta \quad \text{and} \quad \forall t > 0. \end{aligned}$$

In our computations we will use g_{trunc} instead of g as our boundary condition.

3.8 Stabilization

This section shall be devoted to the presentation of a stabilization technique. The necessity of stabilization can be inferred if the partial differential equation being solved contains a convective term, i. e. a term involving first spatial derivatives, and if it has the property of becoming large relative to the diffusive terms. Detailed examples of this phenomenon can be read and found in [48].
As outlined in Section 2.7 concerning our equation this is indeed the case. Therefore we are in need of applying a stabilization strategy and we choose the co-called streamline upwind Petrov-Galerkin technique (SUPG), which is conscientiously elucidated in [63]. The idea behind this is to add artificial diffusion and thus compensate for the convection dominance but it shall only be provided when needed. However, we have to make three important simplifications, namely

- we restrain ourselves to using first order finite elements only which yields vanishing second derivatives,
- whenever the derivative of a coefficient function has to be computed the dependence on x or y respectively is ignored and thus the coefficient function is treated as if it were constant,
- this technique is only applied in x-direction.

The first two of these simplifications are extremely severe and restrictive and we have to point out that they might really hamper the effectiveness of the whole process because our coefficients heavily depend on x. But at the same time I wish to emphasize that those restrictions are not voluntarily accepted but rather imposed by the limitations of the COMSOL MULTIPHYSICS package which we use in order to solve our initial boundary value problem. In [96], [97] and [98] it is said that the package is currently incapable of handling non-constant coefficients in connection with SUPG and trying nonetheless leads to undetermined outcome. So we should be aware and alert that our stabilization efforts might fall short and fail to be successful.
Stabilization is sought to be attained by abridging the weak formulation (3.3) of our equation. More precisely the testing function v is replaced by a different testing function $\tilde{v}$ defined as

$$\tilde{v} := v + \delta' 1_{\{\delta'>0\}}\, \beta_1 v_x,$$

where δ' is set in the following way:

$$\delta' = \frac{\delta\, h}{|\beta_1|} - \frac{k_1}{|\beta_1|^2}.$$

Here δ is a parameter and h the local mesh length. The other variables are taken from the differential operator $\mathcal{L}$ defined in (3.2). Hence after replacing the testing function and multiplying the equation with the new testing function the starting point of the derivation of the weak formulation looks as follows:

$$\int (\mathcal{L}u\, v - \partial_t u\, v) + \int (\delta'\beta_1 v_x)(\mathcal{L}u - \partial_t u) = \int -Cv - \int (\delta'\beta_1 v_x)C$$

The expressions appearing first on each side of the equation are exactly those treated when originally deriving the weak equation in Section 3.1. Therefore we can focus solely on the other integrals. Concerning the left-hand-side we obtain:

$$\begin{aligned}
\int (\delta'\beta_1 v_x)(\mathcal{L}u - \partial_t u) &= \int \delta'\beta_1 (v_x k_1 u_{xx} + v_x k_2 u_{yy} + \beta_1 v_x u_x + \beta_2 u_x u_y + au - \partial_t u) \\
&\overset{i.b.p.}{=} \int \delta'\beta_1 (\underbrace{v_{xx}}_{=0} k_1 u_x + \underbrace{v_{xy}}_{=0} k_2 u_y + \beta_1 v_x u_x + \beta_2 u_x u_y + au - \partial_t u) \\
&= \int \delta'\beta_1 \left(\beta_1 v_x u_x + \beta_2 v_x + au - \partial_t u v_x\right)
\end{aligned}$$

At this point one should heed that in the step where integration by parts was used we suppressed the dependence of δ', β and k_1 on the spatial variables x and y.
Hence incorporating SUPG with the above restrictions amounts to replacing k_1 and a by the expressions

$$\begin{aligned}
\tilde{k}_1 &= k_1 + \delta'\beta_1^2 \\
\tilde{a} &= a + a\delta'\beta_1
\end{aligned}$$

and adding the remaining expressions $\int \delta'\beta_1\beta_2\, u_y v_x$, $\int \delta'\beta_1 \partial_t u\, v_x$ on the left-hand-side and $-\int C\delta'\beta_1 v_x$ on the right-hand-side of the weak equation.

3.9 Regularity of Solutions

The subject matter of this section is to make plain and plausible that the solutions of our PDEs (2.6) and (2.7) subject to the respective boundary data from Sections 2.6 and 2.8 are of less regularity than the second Sobolev space. This is a very crucial aspect because in Proposition 3.2.10 it is pointed out that in the case of linear elements second order in required for the estimate to hold. Hence if the solution fails to be that smooth we cannot expect to discern a convergence rate which is that good and we have to be content with a worse rate.
Before being able to do so we have to introduce the concept of uniform parabolicity/ellipticity for a differential operator $\partial_t + \mathcal{L}$ with $\mathcal{L}$ being defined in (3.2). This is, for instance, elucidated in [37] or [69].

Definition 3.9.1 (Uniform Parabolicity) *We say that the partial differential operator $\partial_t + \mathcal{L}$ is uniformly parabolic if there exists a constant $\theta > 0$ such that*

$$\sum_{i,j=1}^{n} k_{ij}(x,t)\xi_i\,\xi_j \geq \theta\,|\xi|^2$$

for all $(x,t) \in U \times (0,T]$ and $\xi \in \mathbb{R}^n$, where $U \subset \mathbb{R}^n$ is the open set on which $\mathcal{L}$ is defined.

This definition is taken from [37].
The first thing we do now is argue that the condition of uniform parabolicity as claimed in the error estimates of Section 3.2 is violated by our differential operator in both the original and the transformed case.

Lemma 3.9.2 *The differential operators from the partial differential equations (2.6) and (2.7) are not uniformly parabolic.*

Proof - Lemma 3.9.2:
We consider the diffusion matrices from equations (2.6) and (2.7)

$$\begin{aligned} K_1(r,H) &= \begin{pmatrix} -1/2\,\Phi_r^2 r & 0 \\ 0 & -1/2\,\Phi_H^2 H^2 \end{pmatrix} \\ K_2(x,y) &= \begin{pmatrix} 1/2\,\Phi_r^2\gamma_r(1-x)^2\ln(1-x) & 0 \\ 0 & 1/2\,\Phi_H^2 A^2 \end{pmatrix} \end{aligned}$$

and for arbitrary but fixed $\xi \in \mathbb{R}^n$ we estimate:

$$\begin{aligned} \xi^T K_1 \xi &= -\frac{1}{2}\Phi_r^2\, r\xi_1^2 \;-\; \frac{1}{2}\,\Phi_H^2\, H^2\,\xi_2^2 \\ &\geq -\frac{1}{2}\,\underbrace{\min(\underbrace{\Phi_r^2 r}_{\overset{r}{\to}0\,0},\underbrace{\Phi_H^2 H^2}_{\overset{H}{\to}0\,0})}_{\to 0}\;\underbrace{(\xi_1^2+\xi_2^2)}_{=||\xi||_2^2} \\ &\to 0 \\ \xi^T K_2 \xi &= \frac{1}{2}\Phi_r^2\gamma_r\,(1-x)^2\,\ln(1-x)\,\xi_1^2 + \frac{1}{2}\,\Phi_H^2\,A^2\,\xi_2^2 \\ &\geq \frac{1}{2}\underbrace{\min(\underbrace{\Phi_r^2\gamma_r\,(1-x)^2\,\ln(1-x)}_{\overset{x}{\to}1\,0},\Phi_H^2\,A^2)}_{\to 0}\;\underbrace{(\xi_1^2+\xi_2^2)}_{=||\xi||_2^2} \\ &\to 0. \end{aligned}$$

Since in both cases the minimum estimate is optimal in the sense that $\min(a, b)$ is the smallest number z such $z \leq a$ and $z \leq b$, it is impossible to find a θ with the desired property which means that both times the differential operator fails to be uniformly parabolic. $\square$

In other words, the statement of the lemma means that $\theta = 0$ would be appropriate.
In Section 3.2 we saw that the Céa-Lemma (cf. Lemma 3.2.9) teaches us about approximate solutions:

$$\begin{aligned} ||u - u_h|| &\leq \frac{C}{\theta} \inf_{v_h \in V_h} ||u - v_h|| \\ &\to \infty \qquad \text{for } \theta \to 0. \end{aligned}$$

This along with the theorem of Lax-Milgram from Appendix B are strong hints that both regularity as well as the error estimates crucially depend on the modulus of ellipticity or parabolicity θ.
Next we state a theorem on uniqueness and existence in Sobolev spaces for equations of our type which is taken from [69]. Once again it shows how difficult it is to obtain H^2-regularity since parabolicity as well as a C^2-boundary are presumed (ours is only Lipschitz as it is rectangular). Definitions of smoothness of domains can be found in [2] and [69]. Moreover the authors deal with a pure Dirichlet problem rather than a mixed boundary value problem for which regularity of a certain level is harder to procure and they still barely show H^2.
Before being able to state the theorem we have to introduce some notation, i. e. several norms and spaces used in it. The definitions are also taken from [69].

Definition 3.9.3

1. *For $r, q \geq 1$ we define the norm*

$$||u||_{q,r,Q_T} = \left(\int_0^T \left(\int_\Omega |u(x,t)|^q \right)^{r/q} dt \right)^{1/r}$$

and the space

$$L_{q,r}(Q_T) = \{u \textit{ measurable} \mid \ ||u||_{q,r,Q_T} < \infty\}$$

and for the special case $r = q$ we set $||\cdot||_{q,q,Q_T} =: ||\cdot||_{q,Q_T}$ as well as $L_{q,q}(Q_T) =: L_q(Q_T)$.

2. *$W_q^{2l,l}(Q_T)$ for l integral ($q \geq 1$) is the Banach space consisting of the elements of $L_q(Q_T)$ having generalized derivatives of the form $D_t^r\, D_x^s$ with any r, s satisfying $2r + s \leq 2l$. Its norm is defined by*

$$||u||_{q,Q_T}^{(2l)} = \sum_{j=0}^{2l} \sum_{2r+s=j} ||D_t^r\, D_x^s u||_{q,Q_T},$$

where the summation is taken over all non-negative integers r, s satisfying $2r + s = j$.

3. *The local norm is defined as*

$$||f||_{r,Q_T}^{(loc)} = \sup_{q_T} ||f||_{r,q_T}, \qquad q_T = \omega \times (0,T), \omega = \Omega \cap cube.$$

In particular, I wish to emphasize that this definition of the parabolic Sobolev spaces $W_q^{2l,l}$ is a generalization of the definition of Sobolev spaces in Appendix B, where they are denoted by H,

to time-dependent functions. In the following W and H are used interchangeably if it is clear from the context which one is meant.
Now we have all the ingredients necessary to give the regularity and existence result on parabolic equations with Dirichlet boundary data which is taken from [69].

Theorem 3.9.4 (Regularity of Solutions of Parabolic Equations) *Let $q > 1$. Suppose that the coefficients c_{ij} of the operator $\mathcal{L}$ are bounded continuous functions in Q_T, while the coefficients b_i and a have finite norms $||b_i||^{(loc)}_{r,Q_T}$ and $||a||^{(loc)}_{s,Q_T}$ with*

$$r = \begin{cases} \max(q, n+2) & \text{for } q \neq n+2 \\ n+2+\epsilon & \text{for } q = n+2 \end{cases}$$

$$s = \begin{cases} \max(q, \frac{n+2}{2}) & \text{for } q \neq \frac{n+2}{2} \\ \frac{n+2}{2}+\epsilon & \text{for } q = \frac{n+2}{2} \end{cases}$$

and ϵ being an arbitrarily small positive number. Suppose, further, that the quantities $||b_i||^{(loc)}_{r,Q_t,t+\tau}$ and $||a||^{(loc)}_{s,Q_t,t+\tau}$ tend to zero for $\tau \to 0$. Let the boundary S be of class C^2. Then, for any $f \in L_q(Q_T)$, $\Phi \in W_q^{2-2/q}(\Omega)$ and $g \in W_q^{2-1/q,1-1/2q}(S_T)$, with $q \neq 3/2$, satisfying in the case $q > 3/2$ the compatibility condition

$$\Phi|_S = g|_{t=0}$$

the Dirichlet problem has a unique solution $u \in W_q^{2,1}(Q_T)$. It satisfies the estimate

$$||u||^{(2)}_{q,Q_T} \leq c\left(||f||_{q,Q_T} + ||\Phi||^{2-2/q}_{q,\Omega} + ||g||^{(2-1/q)}_{q,S_T}\right).$$

For any $f \in L_{3/2}(Q_T)$ and any $\Phi \in W^{2/3}_{3/2}(\Omega)$, $g \in W^{4/3,2/3}_{3/2}(S_T)$ such that the function

$$\Psi = \begin{cases} \Phi(x), & x \in \Omega \\ g(x,t), & x \in S \end{cases}$$

given on Γ_T, has the finite norm

$$\{\{\Psi\}\}^{(4/3)}_{3/2,\Gamma_T} := \text{cf. [69]}$$

the Dirichlet problem has a unique solution $u \in W^{2,1}_{3/2}(Q_T)$ which satisfies the inequality

$$||u||^{(2)}_{3/2,Q_T} \leq c\left(||f||_{3/2,Q_T} + \{\{\Psi\}\}^{(4/3)}_{3/2,\Gamma_T}\right).$$

A proof exceeds the scope of this diploma thesis but can nonetheless be found in [69].
This statement shows that even under this large array of assumptions we can only ensure W^2-regularity if the operator is uniformly parabolic which according to Lemma 3.9.2 is definitely violated in our case. Therefore we have to expect significantly worse properties than W^2 even if a way of relinguishing the severe restriction to C^2-boundaries and replacing it by a Lipschitz-boundary could be found. This is further stressed in [69] where it is pointed out that the uniform parabolicity is a very crucial presumption.
In contrast to this analysis of uniformly parabolic operators in Sobolev spaces in [27] degenerate equations are dealt with. These are equations which fail to be uniformly parabolic. However, in order to compensate for this shortcoming very strict growth conditions have to be imposed for the coefficient functions.
Unfortunately the author only deals with solvability in Hölder spaces so that we cannot compare

the results to ours in a Sobolev setting since there is no embedding result from Hölder to Sobolev spaces as stated in Appendix B for the reverse direction.
Altogether this section can be summed up saying that the solution to our PDE (either the original one (2.6) or the transformed one (2.7)) is of less regularity than H^2 so that we have to expect worse convergence rates than predicted by the standard error estimates in Section 3.2.

Chapter 4

Solving the Equation

This section deals with the solver we apply to the discretized initial boundary value problem (3.9) and (3.10).
The solver we choose is the commercial package COMSOL MULTIPHYSICS 3.2B which is designed to solve PDEs (also non-linear ones) by using the finite element method. Moreover it provides a variety of different linear system solvers and can be considered quite reliable.
Our strategy, which in fact copies the one implemented in COMSOL MULTIPHYSICS (cf. [96], [97] and [98])) is to rewrite the aforementioned problem as a differential algebraic equation and derive and use a Newton method to solve it.

4.1 Rewriting as a DAE

In Section 3.1 we have derived the following system of ordinary differential equations one procures after discretizing in the spatial variables (cf. (3.9) and (3.10)):

$$\begin{aligned} AU &= F + MU' \\ BU &= G. \end{aligned}$$

The subject matter of this subsection is now going to be the rewriting in a system of differential-algebraic equations which means that both the differential equation and the constraints appear in the same equation.
This is attained by enlarging the solution vector U to a vector $\tilde{U}$ not only containing the degrees of freedom situated in the interior $\bar{\Omega} - \partial\Omega$ of the computational domain Ω but also accounting for those on the Dirchlet-boundary Γ_D. In consequence, anything else also has to be abridged. We define the following:

$$\begin{aligned} \tilde{A} &:= \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \\ \tilde{M} &:= \begin{pmatrix} M & 0 \\ 0 & 0 \end{pmatrix} \\ \tilde{F} &:= \begin{pmatrix} F \\ G \end{pmatrix}. \end{aligned}$$

Having done so we obtain the following differential-algebraic equation:

$$\tilde{A}\tilde{U} = \tilde{F} + \tilde{M}\tilde{U}'. \tag{4.1}$$

This equation is solved using a Newton method I describe in the next section.

4.2 Newton-Iteration

The first step we have to take is writing the DAE (4.1) as $g(U, U', t) = 0$ where the function g is defined in the following way:

$$g(U, U', t) := F - AU + M\,U'.$$

Heed that for the sake of clarity $\tilde{U}, \tilde{A}, \tilde{F}$ and $\tilde{M}$ are now referred to as U, A, F and M assuming that we are dealing with a DAE as defined in Section 4.1. Moreover we take notion of the fact that only U and U' depend on t because the coefficient functions of the PDE are independent of time and that this g happens to be affine-linear in U and U'.
The following derivation is largely based on [13] and was only slightly abridged.

4.2.1 Derivation of the Newton-iteration

We now have to derive the Newton-iteration scheme. To do so discretizing the time-derivative by the backward-differentiation formula of order p (cf. (3.15)) as previously described is the first step. Therefore at time step $n \to n+1$ we assume to be given previous solutions $u_n, \ldots, u_{n-p}$ and we have

$$u_{n+1} = \sum_{j=0}^{p} a_j u_{n-j} + \tau_n b_{-1} f_{n+1},$$

where τ_n denotes the stepsize at step n. If interpreted as the discretization of a differential equation $u' = f$ one can write:

$$\begin{aligned} u'_{n+1} &= f_{n+1} \\ &= \frac{1}{\tau_n b_{-1}} \left(u_{n+1} - \sum_{j=0}^{p} a_j u_{n-j} \right) \\ &=: \alpha u_{m+1} + \beta, \end{aligned}$$

where

$$\begin{aligned} \alpha &= \frac{1}{\tau_n b_{-1}} \\ \beta &= -\sum_{j=0}^{p} \frac{a_j}{\tau_n b_{-1}} u_{n-j}. \end{aligned}$$

Applying this result we can write the discretized DAE at time t_n in the following way:

$$g(t_n, U_n, \alpha U_n + \beta) = 0.$$

The next step on our way to the iteration pattern is the linear approximation of g by a Taylor series expansion around a previous iterate (note that one has to distinguish between the time step $n+1$ and the $(m{+}1)$-st step of the Newton-iteration):

$$\begin{aligned} g(t_{n+1}, U^{m+1}, \alpha U^{m+1} + \beta) &\approx g(t_{n+1}, U^m, \alpha U^m + \beta) + (Dg)(U^{m+1} - U^m) \\ &\overset{!}{=} 0. \end{aligned}$$

In our case the Jacobian matrix Dg satisfies

$$Dg = \alpha\, M + A$$

so that we attain the Newton-iteration scheme:

$$U^{m+1} = U^m - (\alpha\, M + A)^{-1} g(t_{n+1}, U^m, \alpha U^m + \beta). \tag{4.2}$$

Thus performing a Newton-step amounts to solving the following linear system of equations:

$$Jx = b,$$

where J, x and b are defined as:

$$\begin{aligned} x &= U^{m+1} - U^m \\ J &= \alpha \tilde{M} + \tilde{A} \\ b &= -g(t_{n+1}, U^m, \alpha U^m + \beta). \end{aligned}$$

Summing up the aforementioned the Newton-iteration consists of:

1. evaluation of $g(t_{n+1}, U^m, \alpha U^m + \beta)$
2. solving the linear system $Jx = b$
3. updating the iterate: $U^{m+1} = U^m + x$

The starting iterate for the Newton-iteration is always zero. What we have yet to specify is when the iteration is terminated, which I will outline in the next section, how the appearing linear systems of equations are solved and how the time steps are chosen.

4.2.2 Controlling the Newton-Iteration

In this section I want to describe how the iteration is terminated. Of course, it should be terminated if an iterate is accurate enough. To determine if this is the case we make use of the rate of convergence ρ computed as:

$$\rho = \left(\frac{||U^{m+1} - U^m||}{||U^1 - U^0||} \right)^{1/m}.$$

If the condition

$$\frac{\rho}{1-\rho}\, ||U^{m+1} - U^m|| < 0.33$$

holds the iteration is terminated. If $\rho > 0.9$ or $m > 4$ the iteration is restarted with a smaller stepsize based on the criteria outlined in Section 4.3.
But we still have to explain which norm is taken in the above calculation. The norm is actually a user-defined one because it depends on absolute and relative tolerances, $ATOL$ and $RTOL$ respectively, prescribed by the user. Then weights w_i are defined as:

$$w_i := RTOL_i\, |U^i| + ATOL_i$$

and the norm is specified as

$$||x||_{WRMS} := \left[\frac{1}{N} \sum_{i=1}^{N} \left(\frac{x_i}{w_i} \right)^2 \right]^{1/2},$$

where WMRS stands for "weighted root-mean-square" and N is the number of degrees of freedom. By introducing a diagonal matrix $D := \sqrt{N}\, \mathrm{diag}(w_1, \ldots, w_N)$ this norm can be expressed in the Euclidean 2-norm:

$$||x||_{WRMS} = ||D^{-1}x||_2.$$

4.3 Time-Stepping

The subject matter of this section is the time-stepping algorithm, which is adaptive in both order and stepsize. By this one seeks to substantially reduce the computational work through choosing the stepsizes as large as possible and to enhance accuracy by choosing high-order formulas whenever possible. But out of this three main questions arise with which we have to deal in the following:

1. Under which circumstances do we wish to accept a step and when do we prefer rejecting it?

2. Based on which criteria do we choose the order for an ensuing step and how is the order selected?

3. Which stepsize do we use in the next step?

In order to answer these questions reasonably we have to use error estimators. So in the following I will outline which ones are applied and how they are used to determine the order and length of the next step.
This strategy is taken from [40] and [86] and whenever they are omitted for the sake of brevity and tractibility details can be found and read there.

4.3.1 Notation Used in the Following Discussion

Before really beginning with the explanation of the time-stepping algorithm I want to quickly list the definition of several expressions appearing below:

$$\begin{aligned}
\alpha_i(n+1) &= \frac{\tau_{n+1}}{t_{n+1}-t_{n+1-i}} \\
\alpha_s &= -\sum_{j=1}^{k} \frac{1}{j} \\
\alpha^0(n+1) &= -\sum_{j=1}^{k} \alpha_i(n+1) \\
[y_n] &= y_n \\
[y_n, y_{n-1}, \ldots, y_{n-k}] &= \frac{[y_n, y_{n-1}, \ldots, y_{n-k+1}] - [y_{n-1}, y_{n-2}, \ldots, y_{n-k}]}{t_n - t_{n-k}} \\
\psi_i(n+1) &= \tau_{n+1} + \tau_n + \ldots + \tau_{n+2-i} \\
&= t_{n+1} - t_{n+1-i} \\
\Phi_1(n) &= y_n \\
\Phi_i(n) &= \psi_1(n)\,\psi_2(n)\ldots\psi_{i-1}(n)[y_n, y_{n-1}, \ldots, y_{n-i+1}] \\
\tilde{\Phi} &= \Phi, \text{but evaluated at the correct solution } y(t) \\
\sigma_1(n+1) &= 1 \\
\sigma_i(n+1) &= \frac{\tau_{n+1}^i}{(i-1)!}\psi_1(n+1)\ldots\psi_i(n+1).
\end{aligned}$$

Further details about these expressions can be found in [40].

4.3.2 Accepting of a Step

As it is elucidated in [40] the derivative in a BDF-discretization and the local truncation error cannot merely be written as we did in Section 3.4.1 but also as:

$$\tau_{n+1}\, y'_{n+1} = \sum_{i=2}^{k+1} \alpha_{i-1}(n-1)\, \Phi_i(n+1) - (\alpha_s - \alpha^0(n+1))\, \Phi_{k+2}(n+1)$$

and the correct solution $y(t)$ satisfies:

$$\tau_{n+1}\, y'(t_{n+1}) = \sum_{i=2}^{k+1} \alpha_{i-2}(n+1)\, \tilde{\Phi}_i(n+1) - (\alpha_s - \alpha^0(n+1))\, \tilde{\Phi}_{k+2}(n+1) + T_{n+1},$$

where the local truncation error T_{n+1} is:

$$T_{n+1} = (\alpha_{k+1}(n+1) + \alpha_s - \alpha^0(n+1))\, \tilde{\Phi}_{k+2}(n+1) + \sum_{i=k+3}^{\infty} \alpha_{i-1}(n+1)\, \tilde{\Phi}_i(n+1).$$

If we now want to estimate the error made in a step we have to understand that there are two sources of error if we want to obtain a solution at a time $t_n \leq \hat{t} \leq t_{n+1}$:

1. the local truncation error which is the amount by which the solution to the DAE fails to satisfy the BDF-formula

2. the interpolation error for arbitrary times in between t_n and t_{n+1}.

Since the BDF-method itself can be interpreted as polynomial interpolation the principal term can be estimated by

$$(\alpha_{k+1}(n+1) + \alpha_s - \alpha^0(n+1))\, ||y_{n+1} - y_{n+1}^{(0)}||.$$

Concerning the second source of error it can be estimated by

$$\alpha_{k+1}(n+1)\, ||\Phi_{k+2}(n+1)||.$$

Combining all that we attain an error estimator for our step by which we can determine whether or not we wish to decline it. The step is accepted if the following condition holds:

$$\begin{aligned} ERR &= M\, ||y_{n+1} - y_{n+1}^{(0)}|| \\ &\leq 1.0, \end{aligned}$$

where the constant M is defined as:

$$M := \max(\alpha_{k+1}(n+1), |\alpha_{k+1}(n+1) + \alpha_s - \alpha^0(n+1)|).$$

As mentioned initially a more detailed justification for those estimators is omitted because it can be found in [40] and [86].

4.3.3 Order Selection

To be able to reasonably select an order for the next step one has to be able to estimate the error resulting from this. As one can read in [40] and [86] such estimates are provided by the following expressions:

$$\begin{aligned}
TERKM2 = ||(k-1)\,\sigma_{k-1}\,(n+1)\Phi_k(n+1)|| &\approx ||h^{k-1}\,y^{(k-1)}|| \\
TERKM1 = ||k\,\sigma_k\,(n+1)\,\Phi_{k+1}(n+1)|| &\approx ||h^k\,y^{(k)}|| \\
TERK = ||(k+1)\,\sigma_{k+1}\,(n+1)\,\Phi_{k+2}(n+1)|| &\approx ||h^{k+1}\,y^{(k+1)}|| \\
TERKP1 = ||(k+2)\,\sigma_{k+2}\,(n+1)\,\Phi_{k+3}|| &\approx ||h^{k+2}\,y^{(k+2)}|| \\
\sigma_{k+2} &= \frac{1}{k+2}.
\end{aligned}$$

They estimate the local error for a decrease of 2, a decrease of 1 etc. provided the preceding steps had been taken at equal length and order $k-1$, $k-2$ respectively and they require a successful completion of the current step. Roughly speaking those estimates are based on a Taylor series representation of the numerical solution computed with constant stepsize and they estimate the next-order term after termination. Details are outlined in [86].
Having said this, the question arises of how we utilize those estimates to determine the order of the next step. The philosophy behind the selection scheme is that a change of order has to show an immediate advantage and a tendency of improvement, i. e. $TERKM1 < TERK$ would indicate an advantage and $TERKM2 < TERKM1$ a tendency. Taking into account that such estimates are not always available the order is decreased by one if

$$\begin{aligned}
\max(TERKM1, TERKM2) &\leq TERK \qquad \text{for } k > 2 \text{ or} \\
TERKM1 &\leq 0.5\,TERK \qquad \text{for } k = 2
\end{aligned}$$

or if $TERKP1$ is available if

$$TERKM1 \leq \min(TERK, TERKP1).$$

The procedure for an increase of the order is quite similar. For $1 < k < maxorder$ the order is increased by one if

$$TERKP1 < TERK.$$

Given the results from the test for a decrease one automatically has

$$TERKP1 < TERK < \max(TERKM1, TERKM2).$$

In the case $k = 1$ the order is increased if

$$TERKP1 < 0.5\,TERK.$$

Except for the starting step and the procedure following a rejection of a step the order selection strategy is complete. Both exceptions shall be dealt with in the section on stepsize choosing.

4.3.4 Stepsize Selection

Given an order k for the step it is outlined in [86] that the estimated error after a step of length $r\,\tau$ is $r^{k+1}\,TERK$. Of course, it is our aim to choose r as large as possible, i. e. optimally the largest r such that

$$r^{k+1}\,TERK \leq \epsilon,$$

where ϵ is a user-prescribed error tolerance. But since in this case the estimated error would be approximately as large as the given error tolerance one is a little more careful and replaces ϵ by $0.5\,\epsilon$.
First it is checked whether or not the previous stepsize can be doubled, i. e. if

$$0.5\,\epsilon \geq 2^{k+1}\,TERK.$$

If this is indeed the case the stepsize is doubled but an enlargement with $r > 2$ is precluded at all times. Moreover if this is not possible the stepsize is not enlarged at all.
The second item to be discussed is whether one can keep the stepsize or if it has to be diminished which would be the case if

$$0.5\,\epsilon \geq TERK.$$

In case this condition holds we select r as claimed optimally above, namely

$$r = \left(\frac{0.5\,\epsilon}{TERK}\right)^{1/k+1},$$

but at the same time we restrain it to $r \leq 0.9$ and $r \geq 0.5$.
The philosophy behind this perhaps a little conservative strategy is to reduce the overall number of steps by avoiding as many rejected steps as possible.
If nonetheless a step is rejected the order is lowered by one. After three consecutive rejections it is set to one with the corresponding stepsize being the minimum of the optimal choice and half the previous length.
The last item we have to address is which procedure is an appropriate one at the beginning. Since there is no data available at this point we start using order one, which is adjusted as more data becomes available. In addition, the following is chosen for the initial step

$$\tau = \min\left(\tau_{input}, 1/4\sqrt{\frac{\epsilon}{|y'(t_0)|}}\right),$$

where τ_{input} is a user-defined maximal initial stepsize. An explanation why this is a sensible choice can be found in [86].
This concludes the section on the time-stepping procedure.

4.4 The Linear System Solver

In this chapter I intend to portray the multigrid method we apply in order to solve the linear systems of equations derived in the previous section. The multigrid method is an iterative solver and its presentation is based on [45] and [64].
For doing so we need a grid hierarchy $\mathcal{T}_0 \subset \ldots \subset \mathcal{T}_l$ with corresponding finite element spaces satisfying $V_k \subset V_{k+1}$ for $k = 0, \ldots, l-1$ and $V_l = V_h$ where V_h is the space in which we are seeking an approximate solution to our PDE. So in the following we assume such a nested sequence of spaces to be given.
Along with such a nested sequence of spaces come prolongation mappings $P_k : V_{k-1} \to V_k$ by which an element of V_{k-1} can be interpreted as one of V_k. Now let $\bar{x}_l$ be an iterate for the solution x_l to the linear system $J_l x_l = b_l$. Here the subscript l indicates that it belongs to the space V_l. We can then write down the error equation for the error $y_l := x_l - \bar{x}_l$:

$$J_l\, y_l = b_l - J_l \bar{x}_l.$$

One is interested in y_l because

$$\begin{aligned} J_l(\bar{x}_l + y_l) &= J_l x_l \\ &= b_l, \end{aligned}$$

which means that if one knows the error one knows the exact solution.
If y_l can be approximated well in V_{l-1} it seems reasonable to inexactly or iteratively solve the error equation in V_{l-1} in order to save computational work. Good enough approximation quality is ensured by so-called pre- and post-smoothing which I will explain below in Section 4.5. On the coarsest grid we can even afford to solve the problem exactly using a direct solver because of the small number of degrees of freedom. The step $k \to k+1$ of the multigrid iteration looks as follows:

Algorithm 4.4.1 (Multigrid Iteration) *Assume to be given the k-th iterate $x_l^{(k)}$ on level l.*

1. *pre-smoothing*
   Perform ν_1 pre-smoothing steps of an iterative pre-smoother S_l^{pre} and set
$$x_l^{(k+1/3)} = S_l^{\nu_1} x_l^{(k)}.$$

2. *coarse-grid correction*
   If $l = 1$ the following error equation be solved directly and exactly and if $l > 1$ μ steps of a multigrid iteration be taken:
$$J_{l-1} y_{l-1} = r_{l-1},$$
   where the residuals r_{l-1} and r_l are defined as
$$\begin{aligned} r_{l-1} &:= P_l^{-1} r_l \\ r_l &:= b_l - J_l x_l^{(k+1/3)}. \end{aligned}$$
   Then set
$$x_l^{(k+2/3)} = x_l^{(k+1/3)} + P_l y_{l-1}.$$

3. *post-smoothing*
   Perform ν_2 post-smoothing steps of an iterative post-smoother S_l^{post} and set
$$x_l^{(k+1)} = S_l^{\nu_2} x_l^{(k+2/3)}.$$

Moreover we have to state under which circumstances the iteration terminates. Our aim of course is to produce a solution, which is as good as possible or more precisely which satisfies the accuracy requirement
$$\rho\,|b - Jx| < TOL\,|b|.$$
Here TOL denotes a user-defined tolerance and ρ an additional factor, also to be specified by the user. But there is another way through which termination can be initiated namely a maximum number of iteration steps. If this number is reached the iteration process stops and returns the current iterate regardless of whether or not the accuracy requirement is matched.
Eventually we have to shed a few words on pre- and post-smoothing. Pre- or post-smoothers or pre-conditioners are iterative solvers for linear systems which can be used as solvers themselves. But when used as smoothers only very few steps are taken because it is not intended to obtain a reliable solution but rather to improve the quality of the current multigrid-iterate and have a better starting point for the next iteration step. In Section 4.5 I will outline the smoothers we used.

4.5 Pre- and Post-smoothers

In the preceeding sections I have already mentioned the use of pre- and post-smoothers. So in this section I intend to briefly summarize the smoothers we use and explain what such a smoother does.
Before it was our aim to solve a linear system of equations $Jx = b$ as efficiently as possible. Roughly speaking the biggest influence on efficiency comes from the condition of the linear system, which increases the number of iterations the larger it becomes (cf. [48] or [89]). Therefore it is a reasonable strategy to replace the linear system of equations by another equivalent but far more advantageous one

$$P^{-1}Jx = P^{-1}b.$$

Here P is the iteration matrix of an arbitrary iterative solver and the above is attained if it is better conditioned, i. e. if $\kappa(P^{-1}J) << \kappa(J)$. For a definition of the condition of a matrix the above references can be suggested.
In the following I am going to outline those concrete versions of P we utilized.

4.5.1 Incomplete LU-factorization

The description of the incomplete LU-factorization in this section is based on the articles [15] and [83].
The purpose of the incomplete LU factorization is to omit certain matrix entries during factorization and replace them by zero. If the matrix to be factorized is sparse this spares fill-in to a certain extent. More precisely it works in the following way:

Algorithm 4.5.1 (Incomplete LU) *For an n-by-n matrix A the factorization proceeds as follows:*

for $r = 1$ *until* $n-1$ *do*
  $d := 1/a_{rr}$
  for $i := (r+1)$ *until* n *do*
    if $(i, r) \in S$ *then*
      $e := d\, a_{ir},\ a_{ir} := e$
      for $j = r+1$ *until* n
        if $(i, j) \in S$ *and* $(r, j) \in S$ *then*
          $a_{ij} := a_{ij} - e\, a_{rj}$
        end if
      end for *(j-loop)*
    end if
  end for *(i-loop)*
end for *(r-loop)*

In this algorithm S is a subset of A (hence the name incomplete LU) and for the other entries we prescribe:

$$l_{ij} = 0 \quad \text{if } j > i \text{ or } (i, j) \notin S$$

and

$$u_{ij} = 0 \quad \text{if } i > j \text{ or } (i, j) \notin S.$$

As an example one might state that

$$S = \{(i, j) \,|\, a_{ij} \neq 0\}$$

would maintain the sparsity structure of A in the factors L and U, whereas

$$S = A$$

would simply replicate the classic Gaussian elimination.
Our choice for S which is taken from [96] is that an element is dropped if its absolute value is smaller than the Euclidean norm of the entire column times a drop tolerance, i. e. a_{ij} is dropped if

$$a_{ij} < TOL\,|A_j|,$$

where A_j denotes the j-th column of A. Our selection of the drop tolerance is

$$TOL = 0.01.$$

4.5.2 UMFPACK Coarse-Grid Solver

As pointed out earlier on the coarsest level of our multigrid hierarchy the linear systems shall be solved directly rather than iteratively up to a given tolerance. This is affordable since there are only relatively few variables on the coarsest level and on the other hand accuracy is improved. For instance, if we use four levels in 2D with uniform refinement in between there are $4^4 = 256$ times more unknowns on the finest level than on the coarsest. This means, if there are two and a half million degrees of freedom on the finest level there are only about ten thousand on the coarsest, which is not a particularly large system.
The direct solver of our choice is LU factorization. But since classic Gaussian elimination is of order N^3 (cf. for example [48] or [89]) it is not very efficient so that we make use of the UMFPACK-solver which is explicitly designed by Timothy Davis (for references see below) to solve systems $Ax = b$, where A is sparse and unsymmetric. More precisely it factorizes

$$PRAQ = LU.$$

Here P is a row re-ordering to maintain stability, Q is column re-ordering to limit and/or preserve an upper bound on the fill-in and R is a row re-scaling diagonal matrix. The fill-in is of interest and becomes an issue to deal with because if A is sparse, i. e. in our case it is a 5-striped matrix, the factorization LU does not have to be sparse. However, if it is that is very helpful because then during the process of actually solving the system $Ax = b$ many operations can be spared.
An overview over the solver's properties and interfaces can be found in [24]. In [21], [22] and [23] the implemented re-ordering strategies are described. The authors mostly rely on so-called unifrontal and multifrontal methods or combinations of both. Moreover those methods are extensively tested in the mentioned papers. According to [22] in unifrontal methods factorization proceeds as a sequence of partial factorizations and eliminations on dense submatrices (called frontal matrices), whereas in multifrontal methods there are several such submatrices. In addition, one has to distinguish between the symmetric and the unsymmetric case.
In pseudo-code the UMFPACK-algorithm for numerical factorization given in [23] looks as:

Algorithm 4.5.2 (UMFPACK)
initializations
$k = 0, i = 0$
for *each chain:*
    current frontal matrix is empty
    for *each frontal matrix in the chain*
        $i = i + 1$
        for $|C_i|$ *iterations:*

$k = k + 1$
find the k-th pivot row and column
apply pending updates to the k-th pivot column
if *too many zero entries in the frontal matrix (or new LU part)*
    apply all pending updates
    copy pivot rows and columns into LU data structure
end if
if *too many zero entries in the frontal matrix*
    create new contribution block and place on stack
    start new frontal matrix
else
    extend the frontal matrix
end if
assemble contribution blocks into current frontal matrix
scale pivot column
if $\sharp$ *pivots in current frontal matrix* $\geq n_B$
    apply all pending updates
    copy pivot rows and columns into LU data structure
end if
end for *($|C_i|$ iterations)*
end for *(each frontal matrix in the chain)*
apply all pending updates
copy pivot rows and columns into LU data structure
create new contribution block and place on stack
end for *(each chain)*

In the aforementioned R_i are the rows and C_i the candidate pivot columns. From my perspective giving further details or examples on the algorithm exceeds the scope of this section and amounts to reiterating past literature. So for the sake of clarity and brevity this shall be omitted.

4.5.3 Pre-Conditioned GMRES

In this section I am going to explain the main ideas behind the GMRES-algorithm based on the works referenced in [13] and [82]. This is of interest to us since we make use of this algorithm in its scaled, preconditioned and incomplete version. But first we will deal with the original unabridged one.

Once again the intention is to iteratively solve a linear system $Ax = b$ but terminate the iteration after only very few steps. The idea is to minimize the residual in a series of subspaces. This means if we think of x_0 as an initial guess and $x = x_0 + z$ we obtain an equivalent problem $Az = r_0$ where $r_0 = b - Ax_0$.

Considering the subspace

$$K_l = \text{span}\{r_0, Ar_0, \ldots, A^{l-1}r_0\}$$

a $z = z_l$ is uniquely determined by

$$\begin{aligned} ||b - Ax_l||_2 &= \min_{x \in x_0 + K_l} ||b - Ax||_2 \\ &= \min_{z \in K_l} ||r_0 - Az||_2. \end{aligned}$$

Given an orthonormal basis $\{v_1, \ldots, v_l\}$ of the Krylov subspace K_l one obtains matrices

$$\begin{aligned} V_l &= [v_1, \ldots, v_l] \ \in \mathbb{R}^{N\times l} \\ \mathbb{R}^{l\times l} \ni H_l &= V_l^T A V_l, \end{aligned}$$

where V_l satisfies $V_l^T V_l = \mathrm{id}_l$. As pointed out in [13] such a basis can be computed using the Arnoldi process or as it is done here by a modified Gram-Schmidt-procedure (cf. [38] for an explanation).

If the vectors $r_0, Ar_0, \ldots, A^l r_0$ are linearly independent we have $\dim(K_{l+1}) = l+1$ and one can define a matrix $\tilde{H}_l \in \mathbb{R}^{(l+1)\times l}$:

$$\begin{aligned} \tilde{H}_l &:= \begin{bmatrix} H_l \\ r^T \end{bmatrix} \\ r &:= (0, \ldots, 0, h_{l+1,l})^T \quad \in \mathbb{R}^l \end{aligned}$$

satisfying

$$AV_l = V_{l+1}\tilde{H}_l.$$

Furthermore, letting $z = V_l y$ we have

$$\begin{aligned} ||r_0 - Az||_2 &= ||r_0 - AV_l y||_2 \\ &= ||r_0 - V_{l+1}\tilde{H}_l y||_2 \\ &= ||e_1\beta - \tilde{H} y||_2 \end{aligned}$$

because V_{l+1} is orthogonal and where $\beta = ||r_0||_2$. The minimizing vector $y = y_l$ is then computed by a QR-factorization of $\tilde{H}_l$ using Givens-rotations. Denoting the $\sin\theta$ elements by s_j, $j = 1, \ldots, l$, one can even write

$$||b - Ax_l||_2 = \beta\, |s_1 \cdot \ldots \cdot s_l|$$

as is outlined in the aforementioned references. Altogether this results in the following basic GMRES-algorithm as listed in [13]:

Algorithm 4.5.3 (GMRES)

1. *Compute* $r_0 = b - Ax_0$ *and set* $v_1 = r_0/||r_0||_2$.
2. *For* $l = 1, \ldots, l_{max}$ *do*
   (a) *Form* Av_l *and orthogonalize it against* $v_1, \ldots, v_l$ *via*

$$\begin{aligned} w_{l+1} &= Av_l - \sum_{i=1}^{l} h_{il} v_i \\ h_{il} &= (Av_l, v_i) \\ h_{l+1,l} &= ||w_{l+1}||_2 \\ v_{l+1} &= w_{l+1}/h_{l+1,l}. \end{aligned}$$

   (b) *Update the QR-factorization of* $\tilde{H}_l$.
   (c) *Compute* $\rho_l = ||r_0||_2\, |s_1 \cdot \ldots \cdot s_l| = ||b - Ax_l||_2$.
   (d) *if* $\rho_l \leq \delta$, *go to step 3, otherwise return to a)*
3. *Compute* $x_l = x_0 + V_l y_l$.

We are now going to slightly modify this process in step 2a for the sake of significantly saving computational work. Step 2a is altered to not beginning the sum at $i = 1$ but rather at $i = i_0 = \max(1, l + 1 - p)$ for a $p \leq l + 1$.
Next we have to account for scaling and pre-conditioning. Scaling be done by the matrix D defined in Section 4.2 and pre-conditioning be performed by a matrix P. We can then restate the linear system $Ax = b$ as $\tilde{A}\tilde{x} = \tilde{b}$ where

$$\begin{aligned}\tilde{A} &= D^{-1}P^{-1}AD \\ \tilde{x} &= D^{-1}x \\ \tilde{b} &= D^{-1}P^{-1}b.\end{aligned}$$

Combining all of that leads to the Scaled Pre-conditioned Incomplete GMRES (again as listed in [13]):

Algorithm 4.5.4 (Scaled Pre-Conditioned Incomplete GMRES)

1. (a) $r_0 = b - Ax_0$, *stop if* $||r_0||_{WRMS} < \delta$
   (b) $\tilde{r}_0 = D^{-1}P^{-1}r_0$, *compute* $||\tilde{r}_0||_2 = ||P^{-1}r_0||_{WRMS}$, $\tilde{v}_1 = \tilde{r}_0/||\tilde{r}_0||_2$
2. *For* $l = 1, \ldots, l_{max}$ *do*
   (a) $\tilde{A}_l\tilde{v}_l = D^{-1}P^{-1}AD\tilde{v}_l$
   (b) $\tilde{h}_{il} = (\tilde{A}\tilde{v}_l,\, \tilde{v}_i), \tilde{w}_{l+1} = \tilde{A}_l\tilde{v}_l - \sum_{i=i_0}^{l} \tilde{h}_{il}\tilde{v}_i$
   (c) $\tilde{h}_{l+1,l} = ||\tilde{w}_{l+1}||_2$, $\tilde{v}_{l+1} = \tilde{w}_{l+1}/h_{l+1,l}$
   (d) *Update the QR-factorization of* $\tilde{H}_l = \tilde{h}_{ij} = Q_lR_l$ *(an* $(l + 1) \times l)$*-matrix*
   (e) *Compute the residual norm* ρ_l *indirectly using* $s_1, \ldots, s_l$
   (f) *if* $\rho_l < \delta$ *go to step 3, otherwise go to a)*
3. *Compute* $||\tilde{r}_0||Q_l^Te_1 = (\bar{g}_l, g)$, $\tilde{z} = \tilde{V}_l\tilde{R}_l^{-1}\bar{g}_l$, $x_l = x_0 + D\tilde{z}$

4.5.4 SOR/SORU Iteration

The SOR and SORU pre-conditioners are very basic modifications of the Gauss-Seidel iteration and can be stated rather quickly. They look as follows:

$$x^{k+1} = x^k + M^{-1}(b - Ax^k).$$

Apparently they only differ in the iteration matrix M. To explain the structure of M we assume the matrix A to be decomposed in the form $A = L+D+U$, where D is the diagonal of A and L and U are the strict lower and upper triangles respectively. After fixing a relaxation parameter $\omega \in (0, 2]$ the matrices are the following:

$$\begin{aligned}M_{SOR} &= \frac{L + D}{\omega} \\ M_{SORU} &= \frac{U + D}{\omega}.\end{aligned}$$

In the case $\omega = 1$ the Gauss-Seidel-algorithm is obtained.

4.5.5 Range of Application

Having outlined several pre-conditioners and a coarse grid solver we ought to mention where they are used.
In solving the linear systems in the Newton-iteration (4.2) we are going to use the multigrid solver, which shall be pre-conditioned by the GMRES-method desribed in Algorithm 4.5.4. This GMRES-algorithm will itself be pre-conditioned by the incomplete LU-factorization. As a post-smoother in the multigrid method we will use the SORU pattern.
When carrying out numerical experiments in Chapter 5 I will explicitly state the parameters used (like the number of multigrid levels) and the number of pre- and post-smoothing steps applied.

Chapter 5

Numerical Results

In this section I am going to outline numerical results of our valuation of mortgage-backed securities. But before doing so we will ascertain that our pattern works by applying it to two known problems, the heat equation and a convection-diffusion equation.
The first aspect we examine is the solution process of the PDE without any perturbations caused by the probability weighting, which means solving the PDE for the first of the 360 months only and we omit the ensuing probability adjustments.
Afterwards we consider the same for the probability-weighted one-month solution before dealing with full 30-year computations. The main goal of examining numerical results is procuring and confirming convergence rates as they are outlined in the error estimates of the sections on discretization.

5.1 The Procedure of Error Measurements

Confirming convergence rates of given error estimates amounts to measuring the numerical solutions or the error in the respective norms. And when it comes to measuring the numerical error we face the problem of lacking knowledge about an exact solution. This is crucial since the error estimates compare the numerical solution to the exact one but on the other hand this is a usual problem because if one knew the exact solution in a closed form the numerical approach would not be needed.
We therefore try to overcome this shortcoming by replacing the exact solution u by a solution $\bar{u}_h$ obtained on an extremely fine mesh. We deem that plausible because

$$||u - u_h|| \leq \underbrace{||u - u_{\bar{h}}||}_{\approx 0} + ||u_{\bar{h}} - u_h||.$$

If the mesh for $u_{\bar{h}}$ is chosen sufficiently fine the error $||u - u_{\bar{h}}||$ can be assumed to be approximately zero and we resort to measuring $||u_{\bar{h}} - u_h||$. Out of this two question arise:

1. Which norm is used and how is it computed?

2. How are the solutions from different grids compared?

5.1.1 Interpolation

Loosely speaking we use interpolation to represent a coarser solution u_h on the grid of the fine solution $u_{\bar{h}}$. More precisely, consider $\bar{h} << h$ and a corresponding interpolation operator $I_h : V_h \to V_{\bar{h}}$ where V_h and $V_{\bar{h}}$ respectively are the corresponding FEM-spaces for level $\bar{h}$ and

level h. We then compute $||u_{\bar{h}} - I_h(u_h)||$ and consider this the error of u_h.
I_h is defined in the following way:

$$I_h(u_h) = \sum_{i=1}^{DOF(V_{\bar{h}})} U_i \varphi_i^{\bar{h}},$$

where $U_i = u_h(x_i)$ and x_i is the i-th mesh point of the grid belonging to $V_{\bar{h}}$.
Heuristically this means the coarser solution is evaluated at the mesh points of the finer grid. In practice this evaluation is performed by the femlab built-in function postinterp which computes the value of an FEM-function at arbitrary coordinates.

5.1.2 Computation of the Norms

We consider two different norms, the L^2-norm and the discrete supremum norm, that is the L^∞-norm. The L^2-norm of an element $u_h \in V_h$ can be computed exactly using the mass matrix defined in (3.8) (the degrees of freedom of the FEM-space be denoted by N):

$$\begin{aligned} ||u_h||_{L^2}^2 &= \int_\Omega u_h^2 \\ &= \int_\Omega \left(\sum_{i=1}^N U_i \varphi_i \right)^2 \\ &= \sum_{i=1}^N \sum_{j=1}^N U_i U_j \underbrace{\int_\Omega \varphi_i \varphi_j}_{m_{ij}} \\ &= U^T M U. \end{aligned}$$

This is then applied to the error function $u_{\bar{h}} - I_h(u_h) \in V_{\bar{h}}$.
Secondly the maximum norm is computed in the following way:

$$||u_h||_{L^\infty} = \max_{i=1,\dots,N} |u_h(x_i)| = \max_{i=1,\dots,N} |U_i|.$$

In addition, it ought to be mentioned that all the appearing solutions are time-dependent and hence they are evaluated at the terminal time $t = 1/12 = 1$ month.

5.1.3 Expectations

Before plotting any results we should discuss what can be expected concerning the convergence rates. This is in turn based on the error estimates given in Section 3.2, which for instance yield third order convergence for quadratic finite elements.
But at this point one has to be careful and distinguish between the mesh length h used in the error estimates and the number of degrees of freedom (henceforth denoted by N) used for the plots. The latter one is a stricter measuring stick since h refers only to one spatial direction whereas we are dealing with two-dimensional problems. More generally the following connection holds between N, the dimension d and the mesh length h:

$$N = \left(\frac{1}{h}\right)^d$$

or equivalenty

$$h = N^{-1/d}.$$

Hence h^α-convergence corresponds to a rate $N^{-\alpha/d}$ or $-\frac{\alpha}{d}$ in doubly logarithmic plotting.

5.1.4 Measurement of the Time Error

The time error of the finite element solution is estimated similarly to the spatial error because once again we have to resort to using a fine reference solution but this time concerning the time evolution. Moreover we use a relatively fine grid in the spatial dimensions so that the spatial error can be neglected.
Since we can use the same grid for any time evolution examined we have the advantage that no interpolation is needed to compare solutions to the fine-evolution solution. Hence we measure the error in the following way:

$$||u - u_h^k|| \leq \underbrace{||u - u_{\bar{h}}^k||}_{\approx 0} + ||u_{\bar{h}}^k - u_h^k||.$$

Here $u_{\bar{h}}^k$ denotes the discrete solution obtained by using a very fine time evolution and u_h^k denotes the discrete solution the error of which we intend to measure. To avoid confusion it is my intention to point out that in both cases the index k is used to indicate that the same terminal time, say t_k, is used. But we have to heed that the evolution, i. e. the precise sequence of time steps taken to get there, is, in general, completely different as in the case of $u_{\bar{h}}^k$ many more steps are taken.
Because of the fact that we use the same spatial grid for all time evolutions $u_{\bar{h}}^k$ and u_h^k are defined on the same mesh so that their difference and there norm can be computed without interpolation. As before we are going to measure the error in both the L^2- and the L^∞-norm.
Concerning the expected convergence rate no precise statement can be made since the error estimates in Section 3.4.2 are only valid for a time evolution of constant order and stepsize. But in our case we can only compute a solution adaptively in both order and stepsize. Nonetheless the detected rates will be stated.

5.2 Justification for this Procedure

The purpose of this section is to give justification for the procedure of error measurement outlined in the previous section. This will be attained by applying it to two known problems. Such a justification process is absolutely mandatory because of the quite difficult differential operator which precludes knowledge of an exact solution. Therefore we examine our measurement strategy for several simplified operators where difficulties like degeneration are gradually added.
The first testing example is the two-dimensional heat equation for which we even know an exact solution and the second example is a diffusion-convection-reaction equation where all coefficients are set to one so that there is a sufficiently smooth solution as we know from [37] or [69]. Concerning initial and boundary data we take the same as in the MBS case.
In addition to that it is our aim to determine how much finer the reference solution has to be in comparison to the finest examined solution. That is, we are looking for a threshold γ such that

$$DOF(V_{\bar{h}}) \geq \gamma\, DOF(V_h) \qquad \text{for any examined } h.$$

This is achieved by approaching the reference solution ever and ever closer without recognizing a deterioration in the convergence rate. As the following examples show $\gamma = 1.5$ is an adequate choice.

5.2.1 The Heat Equation

We consider the open domain $\Omega = (0, \pi) \times (0, \pi)$ and suppose

$$u(x, y, t) = \sin(x)\, \sin(y)\, \exp(-t)$$

to be an exact solution to the heat equation

$$\triangle u - u_t = f \qquad \text{in}\, \Omega \times (0, 1/12).$$

In consequence, we have to prescibe

$$f(x, y, t) = -\sin(x)\,\sin(y)\,\exp(-t)$$

for the right-hand-side f which is obtained by applying the differential operator to u. Furthermore inserting the boundaries $x = 0, 2\pi$ and $y = 0, 2\pi$ yields $u|_{\partial\Omega} = 0$. As initial values we choose

$$u_0(x, y) := u(x, y, 0) = \sin(x)\,\sin(y).$$

We now have two different intentions, first to check convergence against the exact solution and second to examine convergence using a fine-mesh solution as a reference. Both times we measure the error for the terminal time $T = 1/12$ in the L^2- and L^∞-norms as described above. This means we have to compute the following:

$$\begin{aligned} ||u_h(\cdot, T) - u(\cdot, T)||_{L^2} &= \int_\Omega |u_h(x, y, T) - \sin(x)\,\sin(y)\,\exp(-T)|^2 dx\,dy \\ ||u_h(\cdot, T) - u(\cdot, T)||_{L^\infty} &= \max_{i=1,\ldots,DOF} |U_i - \sin(x_i)\sin(y_i)\exp(-T)|. \end{aligned}$$

The first integral is approximated using the femlab quadrature function *postint*. In both cases dividing by the norm of the exact solution yields the relative error. These norms can fortunately be computed exactly:

$$\begin{aligned} \sup_\Omega |u(\cdot, 1/12)| &= \sin(\pi/2)\,\sin(\pi/2)\,\exp(-1/12) \\ &= 0.920044415 \\ ||u||_{L^2(\Omega)} &= \int_0^\pi \int_0^\pi \sin^2(x)\,\sin^2(y)\,\exp(-1/6)\,dx\,dy \\ &= \exp(-1/6)\;[1/2\,x - 1/4\,\sin(2x)]_0^\pi\;\;[1/2\,y - 1/4\,\sin(2y)]_0^\pi \\ &= 2.088609939. \end{aligned}$$

In Figure 5.1 we have plotted the error versus the number of degrees of freedom in doubly logarithmic scaling for reference solutions of 539,905 and 2,157,569 DOFs. This allows us to take the convergence rate as the slope of the corresponding linear plot.
In this example we used the following set of parameters:

$$\begin{aligned} Atol &= 1e-12 \\ Rtol &= 1e-10 \\ maxorder &= 5 \\ mglevels &= 4 \\ presmooth &= GMRES \text{ (5 steps)} \\ postsmooth &= SORU \text{ (5 steps)} \\ FEM - order &= 2. \end{aligned}$$

Evaluating the data the expected rate of -1.5 is confirmed.
To conclude this section the above is verified by comparison to the convergence rate we observe when testing against the exact solution. This is crucial at this point because comparison against fine solutions is only an approximation of the truth and our actual interest. And the whole

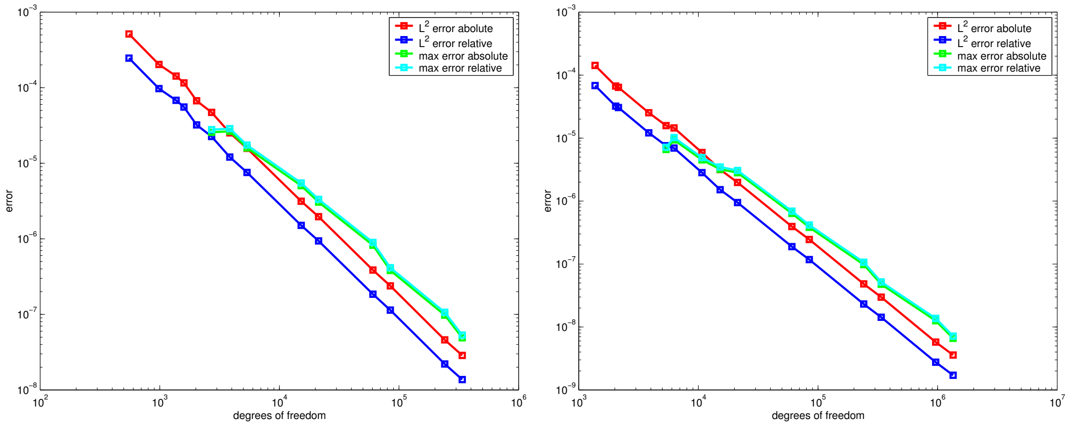

Figure 5.1: On the left: convergence plots for a reference solution with 539,905 DOFs and on the right: the same plots but with a reference solution of 2,157,569 DOFs

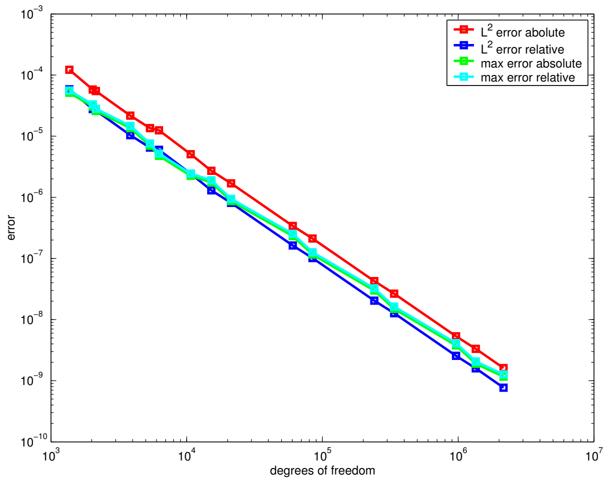

Figure 5.2: convergence plot for the exact solution of the heat equation

purpose of this section was justification of this procedure and to instill trustability for more elaborate differential operators.
Therefore in Figure 5.2 we plot the errors against the exact solution. Fortunately enough the in theory expected convergence rate of -1.5 is reproduced and it confirms the results of the previous measurements. Hence we can feel safe in deeming the above strategy reliable.

5.2.2 Second Example

We shall now examine a more advanced example. It still does not feature the same differential operator as the MBS equation but except for the coefficients every part is present. We consider the equation

$$u_{xx} + u_{yy} + u_x + u_y + u - u_t = 1 \qquad \text{in } (0, 2\pi)^2 \times (0, 1/12).$$

This differential operator is similar to the one of the MBS equation in the fact that only the dependence of the coefficients on the spatial coordinates is dropped and the coefficients are set to one. Moreover the right-hand-side equals one.
Next we have to state the boundary and initial data. The latter shall be identically zero exactly as in the MBS case and the boundary data only differ in the non-zero Dirichlet data. The vanishing Dirichlet and Neumann data are kept. On the respective fourth boundary we modify

the MBS data using $C = 1$, $c = 0.025$, $y_{min} = 0$ and no transaction costs.
This is an interesting example for our purposes because according to [37] and [69] it is smooth enough to satisfy the standard error estimates from Section 3.2 and just as for the MBS we lack knowledge of an exact solution.
As before we resort to measuring the error in the L^2-norm and the L^∞-norm by interpolation to two very fine reference grids and confirm the expected convergence rate of -1.5. The set of parameters used is:

$$\begin{aligned} Atol &= 1e-12 \\ Rtol &= 1e-10 \\ maxorder &= 5 \\ mglevels &= 4 \\ presmooth &= GMRES \text{ (5 steps)} \\ postsmooth &= SORU \text{ (5 steps)} \\ FEM-order &= 2. \end{aligned}$$

In Figure 5.3 the errors have been depicted.

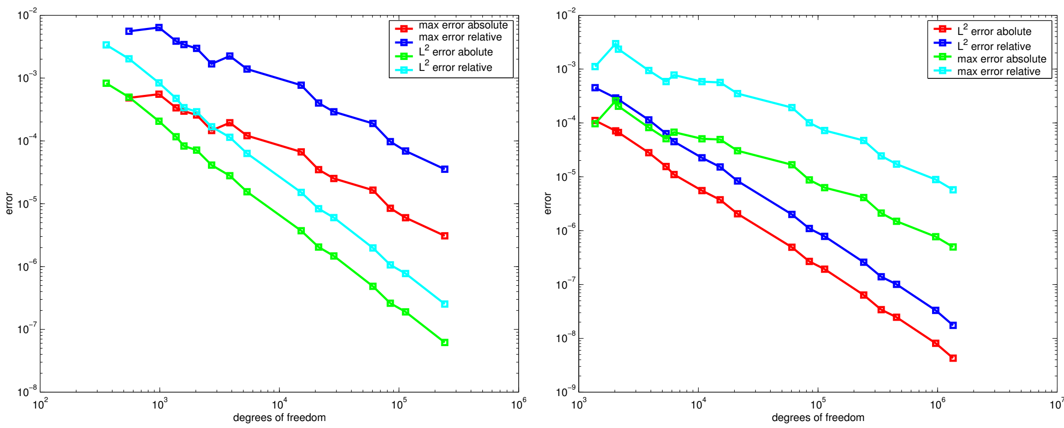

Figure 5.3: On the left: convergence plots for a reference solution with 539,905 DOFs and on the right: the same plots but with a reference solution of 2,157,569 DOFs

Moreover we test the behavior of time error measurement. To do so we choose a grid of $1,110,803$ degrees of freedom. Using the parameters

$$\begin{aligned} Atol &= 1e-15 \\ Rtol &= 1e-13 \end{aligned}$$

and same solver settings as above the solution process for the reference solution took 853 steps and $37,654.191$ seconds of CPU time. In addition, we computed the following norms for the reference solution:

$$\begin{aligned} ||u_h^k||_{L^2} &= 5.05810489567e-4 \\ ||u_h^k||_{L^\infty} &= 2.74549049071e-4. \end{aligned}$$

The results are plotted in doubly logarithmic scaling in Figure 5.4 and we spot a convergence rate of -4.63.

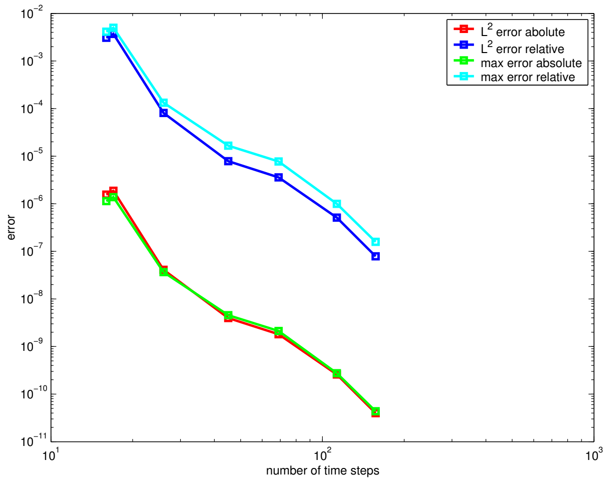

Figure 5.4: Convergence plot of the time error

5.3 The PDE Solution for One Month's Time

This section marks the starting point for the presentation of the actual results we are interested in. As it was pointed out before the solution process consists of solving a PDE for the time of one month, weighting the results by a certain probability distribution and prescribing the weighted result as the initial value for the next month's PDE solution process.
Because of this I have chosen to first examine the PDE result after one month without any influence by the probability weighting. Having said that in the following we will deal with the partial differential equation (2.7) as well as its mixed boundary values and the right-hand-side derived in Section 3.7. In doing so I will distinguish between the solution obtained without stabilization procedure and the one using the techniques outlined in Section 3.8.
For those computations the following set of parameters is used:

$$\begin{aligned} Atol &= 1e-14 \\ Rtol &= 1e-12 \\ maxorder &= 5 \\ mglevels &= 4 \\ presmooth &= GMRES \text{ (5 steps)} \\ postsmooth &= SORU \text{ (5 steps)} \\ FEM-order &= 1 \\ \delta &= 0.1. \end{aligned}$$

Of course, δ only comes into play when using streamline upwind Petrov-Galerkin. If not, it is redundant. As it was justified in the previous section we measure the error against a reference solution for both the L^2- and L^∞-norm. In Table 5.1 the results are shown when not using stabilization and Table 5.2 shows the same for use of SUPG. In both tables the relative error is listed for which to compute we needed the respective norms of the reference solutions. We obtained the following values for the non-SUPG case:

$$\begin{aligned} ||u_h^{ref}||_{L^2} &= 6.17814479641e-4 \\ ||u_h^{ref}||_{L^\infty} &= 2.66664495818e-4. \end{aligned}$$

The computation of this solution took 550 time steps and and lasted for 217,643.713 seconds.

DOFs	steps	error in $\|\|\cdot\|\|_{L^2}$		error in $\|\|\cdot\|\|_{L^\infty}$		CPU time [sec]
		absolute	relative	absolute	relative	
161,372	393	4.68406e-6	7.58166e-3	1.09129e-4	4.09237e-1	891.555
204,920	453	7.22186e-7	1.16894e-3	3.22098e-5	1.20788e-1	1,330.154
252,096	432	4.41437e-7	7.14514e-4	3.03360e-5	1.13761e-1	1,614.049
284,875	493	3.04206e-7	4.92391e-4	2.03981e-5	7.64935e-2	2,072.309
329,078	473	2.14785e-7	3.47653e-4	1.89088e-5	7.09086e-2	2,563.025
386,452	464	1.81798e-7	2.94260e-4	1.48693e-5	5.57603e-2	2,857.486
469,204	480	1.54652e-7	2.50321e-4	1.26802e-5	4.75511e-2	3,491.322
547,129	531	1.16906e-7	1.89225e-4	1.01577e-5	3.80917e-2	5,179.038
656,343	549	9.63076e-8	1.55884e-4	9.72044e-6	3.64519e-2	6,383.517
747,316	517	8.20302e-8	1.32775e-4	6.52038e-6	2.44516e-2	7,415.362
821,238	477	7.45132e-8	1.20608e-4	5.67078e-6	2.12656e-2	7,468.554
914,215	539	6.70462e-8	1.08522e-4	4.27394e-6	1.60274e-2	9,603.386
1,094,681	635	5.91295e-8	9.57075e-5	4.12644e-6	1.54743e-2	14,809.388
1,640,104	509	4.72004e-8	7.63990e-5	2.43771e-6	9.14149e-3	20,792.545
1,830,547	603	4.23001e-8	6.84673e-5	1.08416e-6	4.06563e-3	30,378.172
1,939,171	576	4.12556e-8	6.67767e-5	1.10839e-6	4.15650e-3	31,144.646
2,059,265	508	4.04153e-8	6.54166e-5	1.06231e-6	3.98369e-3	33,621.298
2,196,543	561	3.97445e-8	6.43308e-5	1.09757e-6	4.11592e-3	35,472.926

Table 5.1: Numerical error of the MBS PDE without SUPG and with a reference solution of 10,221,317 DOFs

In the case of using SUPG the solution process of the reference solution also took 550 steps and consumed 237,697.470 seconds of CPU time. Moreover we procured the following values for the norms of the reference solution:

$$\begin{aligned} ||u_h^{ref}||_{L^2} &= 6.17894489878e-4 \\ ||u_h^{ref}||_{L^\infty} &= 2.96059925054e-4. \end{aligned}$$

The data from Tables 5.1 and 5.2 is graphically illustrated in Figure 5.5. In the non-SUPG case we find a convergence rate of -0.89 and using SUPG we detect -0.52.

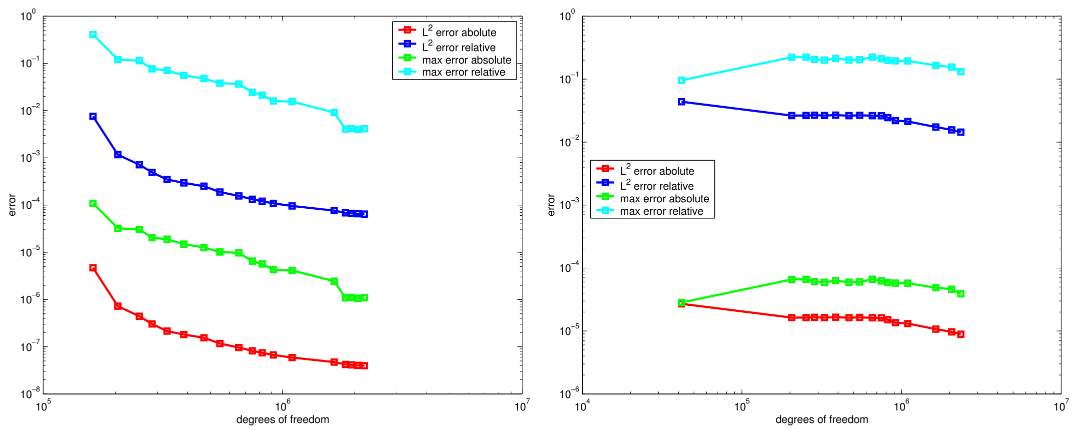

Figure 5.5: Left: convergence plot for the PDE without SUPG, right: same with use of SUPG

Next we have to measure the time error for both the use of streamline-upwind-Petrov-Galerkin and for omitting it. The results of the former test are shown in Table 5.4 and the latter in Table

DOFs	steps	error in $\|\|\cdot\|\|_{L^2}$		error in $\|\|\cdot\|\|_{L^\infty}$		CPU time [sec]
		absolute	relative	absolute	relative	
41,815	382	2.69997e-5	4.36963e-2	2.83690e-5	9.58218e-2	228.989
204,920	465	1.63056e-5	2.63890e-2	6.59344e-5	2.22706e-1	1,854.345
252,096	439	1.63495e-5	2.64600e-2	6.60645e-5	2.23146e-1	2,239.007
284,875	493	1.64816e-5	2.66738e-2	6.07134e-5	2.05071e-1	2,721.007
329,078	480	1.63181e-5	2.64092e-2	5.94534e-5	2.00815e-1	2,888.096
386,452	495	1.65643e-5	2.68077e-2	6.30727e-5	2.13040e-1	3,647.817
469,002	480	1.62561e-5	2.63089e-2	5.95074e-5	2.00998e-1	4,244.836
547,129	531	1.64470e-5	2.66178e-2	6.01774e-5	2.03261e-1	5,791.319
656,343	549	1.62064e-5	2.62284e-2	6.63193e-5	2.24006e-1	7,570.655
747,316	517	1.61829e-5	2.61904e-2	6.23900e-5	2.10734e-1	8,808.673
821,238	477	1.51003e-5	2.44383e-2	5.88359e-5	1.98730e-1	8,958.268
914,215	539	1.35739e-5	2.19680e-2	5.77285e-5	1.94989e-1	11,453.809
1,094,681	635	1.31166e-5	2.12279e-2	5.73943e-5	1.93860e-1	19,340.785
1,640,104	509	1.07264e-5	1.73596e-2	4.88645e-5	1.65049e-1	29,918.569
2,059,265	508	9.66687e-6	1.56449e-2	4.60790e-5	1.55641e-1	41,146.901
2,352,519	548	8.89161e-6	1.43902e-2	3.90031e-5	1.31741e-1	37,673.877

Table 5.2: Numerical error of the MBS PDE with SUPG and with a reference solution of 10,221,317 DOFs

5.3. Moreover those results are depicted in Figure 5.6. From this figure we take convergence rates of -5.16 and -4.71 for non-SUPG and SUPG respectively.

To obtain those results we used the same solver settings as we did during the measurement of the spatial error. In addition, we used a grid of $1,110,803$ degrees of freedom besides the following tolerance settings:

$$\begin{aligned} Atol &= 1e-15 \\ Rtol &= 1e-13 \end{aligned}$$

In the case without SUPG the time evolution took 761 steps, consumed $20,308.017$ seconds and

tolerances		steps	error in $\|\|\cdot\|\|_{L^2}$		error in $\|\|\cdot\|\|_{L^\infty}$		CPU time [sec]
Atol	Rtol		absolute	relative	absolute	relative	
1e-4	1e-2	16	2.67521e-7	4.33013e-4	8.60772e-7	3.22792e-3	1,724.001
1e-5	1e-3	16	2.66709e-7	4.31698e-4	8.56547e-7	3.21208e-3	1,863.851
1e-6	1e-4	20	4.28599e-7	6.93735e-5	1.12651e-7	4.22444e-4	2,144.488
1e-7	1e-5	33	6.37663e-10	1.03213e-6	4.66344e-9	1.74880e-5	2,246.203
1e-8	1e-6	51	8.18256e-11	1.32444e-7	8.40935e-10	3.15353e-6	2,583.257
1e-9	1e-7	78	1.93560e-11	3.13298e-8	1.89017e-10	7.08819e-7	3,195.378
1e-10	1e-8	112	5.85377e-12	9.47498e-9	5.67845e-11	2.12943e-7	3,567.380
1e-11	1e-9	162	4.38884e-13	7.10383e-10	4.22371e-12	1.58390e-8	5,086.678
1e-12	1e-10	231	6.72656e-14	1.08877e-10	6.44815e-13	2.41807e-9	6,572.675

Table 5.3: Time error of the MBS PDE without use of SUPG

tolerances		steps	error in $\|\|\cdot\|\|_{L^2}$		error in $\|\|\cdot\|\|_{L^\infty}$		CPU time [sec]
Atol	Rtol		absolute	relative	absolute	relative	
1e-4	1e-2	16	2.68886e-7	4.35173e-4	8.64584e-7	2.88732e-3	1,457.495
1e-5	1e-3	16	2.68067e-7	4.33848e-4	8.60326e-7	3.22625e-3	1,643.468
1e-6	1e-4	20	4.30058e-8	6.96018e-5	1.15347e-7	4.32555e-4	1,799.526
1e-7	1e-5	33	6.41927e-10	1.03891e-6	4.94267e-9	1.85352e-5	2,220.834
1e-8	1e-6	51	8.22375e-11	1.33096e-7	8.47483e-10	3.17809e-6	2,449.747
1e-9	1e-7	78	1.94743e-11	3.15178e-8	1.90060e-10	7.12730e-7	3,219.143
1e-10	1e-8	112	5.89055e-12	9.53344e-9	5.71124e-11	2.14169e-7	3,583.206
1e-11	1e-9	162	4.41917e-13	7.15211e-10	4.26097e-12	1.59788e-8	5,340.402

Table 5.4: Time error of the MBS PDE using SUPG

the following norms could be computed:

$$||u_h^k||_{L^2} = 6.1781351015e-4$$
$$||u_h^k||_{L^\infty} = 2.6666463561e-4.$$

Using stabilization 772 steps were taken during 19,749.949 seconds and we obtained the below norms:

$$||u_h^k||_{L^2} = 6.17883003096e-4$$
$$||u_h^k||_{L^\infty} = 2.94418347888e-4.$$

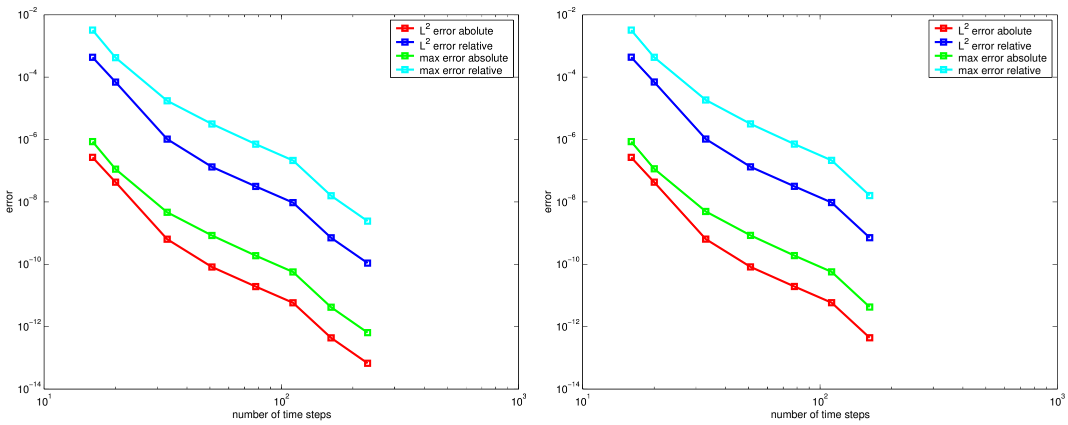

Figure 5.6: Left: time convergence plot for the PDE without SUPG, right: same with use of SUPG

The solutions for both the use of SUPG and not using it have been plotted in Figure 5.7 on a regular and uniform grid (200-by-200) in both spatial directions for the terminal time $T = 1/12$. The values plotted were obtained by interpolation.
Having stated the results of our spatial and time convergence tests in Tables 5.1, 5.2, 5.3 and 5.4 and graphically illustrated them in Figures 5.5 and 5.6 we have to interpret the results. Doing so we distinguish between convergence in time and spatial variables.
The time convergence plots of the MBS PDE show that the streamline-upwind Petrov-Galerkin scheme does not have too much of an effect on the convergence plot. This is the case because concerning time there are no smoothness and regularity problems which could be resolved when using SUPG. Moreover the use was intended to deal with problems in x-direction rather than

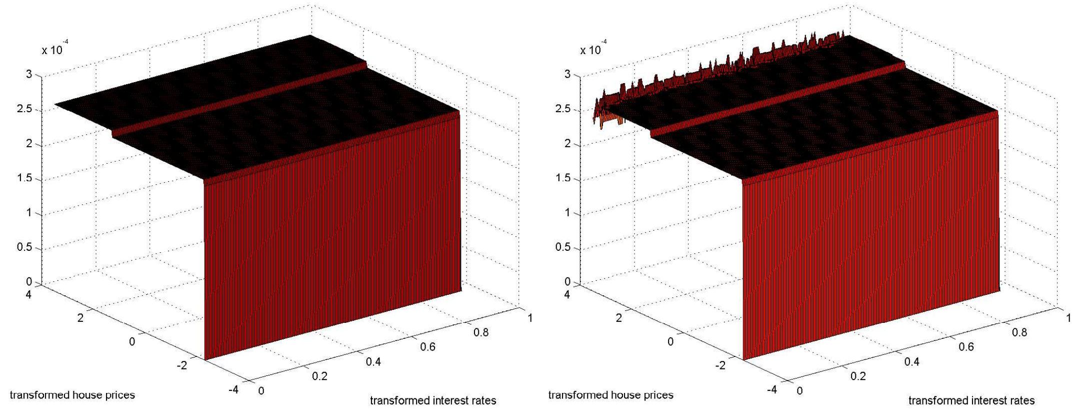

Figure 5.7: Left: solution plot for the PDE without SUPG, right: same with use of SUPG

in the time variable.
This discernment is underlined by the fact that the time evolution took an identical number of time steps for each set of parameters *Atol* and *Rtol*. Only in the errors measured very slight differences appear. Nonetheless there is a difference in the number of time steps taken to compute the reference solution as in the non-SUPG case the solver takes only 761 steps, whereas in the SUPG case 772 are necessary, roughly speaking a 1.5 % difference.
In addition, I would like to point out that very similar results for both cases could be expected since in the time direction there were not any changes at all to the discretization.
In the spatial direction things are completely different. Although for a given number of degrees of freedom mostly the same number of time steps was taken we can notice a considerably larger error in all computations when using streamline upwind Petrov-Galerkin which even exceeds three magnitudes. Moreover there is a significant drop in the observed convergence rate. Using SUPG we obtain -0.52 whereas without SUPG -0.89 is detected.
Altogether we have to come to the conclusion that the limitations of this procedure imposed by COMSOL MULTIPHYSICS and expounded in Section 3.8 are too severe simplifications to overcome. Therefore we cannot but admit that this stabilization procedure does not yield the desired effect so that we shall no longer use it in our computations.
However, examining the solution plots in Figure 5.7 we find that there are scarcely any problems at all without SUPG as one might have expected. Such problems could have been oscillations particularly where there are steep gradients as well as a worse convergence rate than expected which in fact appear when using SUPG as the solution plot shows. Despite the fact that there are indeed steep gradients near the vanishing Dirichlet boundaries there are no oscillations at all so that our solution pattern can be deemed stable. Moreover we do not spot a bad convergence rate because we cannot expect one better than one because of Section 3.9 where we pointed out that we have less regularity than H^2 which is the minimum requirement for a rate of one or better according to the estimates given in Section 3.2.

Before moving on to our consideration of the probability-weighted solution I will make a few remarks on the overall rate of convergence. This is the rate discerned when simultaneously refining in space and time and plotting the error against the number of degrees of freedom times the number of time steps. Yet, it is based on the results obtained for independently considering space and time.
The problem is that the number of time steps has to be adjusted in a way which exactly and appropriately matches the spatial convergence rate or vice versa. To be able to predict the overall rate consider rates α and β in space and time respectively for the actual number of unknowns (say N in space and M in time), i. e. our pattern is of order $\mathcal{O}(N^{-\alpha} + M^{-\beta})$.

Atol	Rtol	steps	a	b	predicted DOF
1e-5	1e-3	16	1.25	3.64637	1,000
1e-6	1e-4	20	1.65	18.23552	3,646
1e-7	1e-5	33	1.55	12.69101	66,494
1e-8	1e-6	51	1.53	11.77051	782,662
1e-9	1e-7	78	1.44	8.28222	9,212,329
1e-10	1e-8	112	1.45	8.62128	76,298,535
1e-11	1e-9	162			657,791,038

Table 5.5: Prediction of number of DOFs if refined in accordance to time convergence rate

From the computations already analysed in this section we have learned that the number of time steps is unchangeably imposed by the tolerance settings, so that we have to set the degrees of freedom according to the given refinements in time. More precisely, in such a refinement step there is given an a and we are looking for a b such that

$$\begin{aligned} M' &:= a\,M \\ N' &:= b\,N. \end{aligned}$$

Since our scheme is of order β in time the error changes by $a^{-\beta}$ and we seek to attain the same in the spatial direction. Hence $b = a^{\beta/\alpha}$ is the adequate choice because

$$\begin{aligned} (N')^{-\alpha} + (M')^{-\beta} &= (b\,N)^{-\alpha} + (a\,M)^{-\beta} \\ &= a^{-\beta}\left[N^{-\alpha} + M^{-\beta}\right]. \end{aligned}$$

If we take the previously determined $\alpha = 0.89$ and $\beta = 5.16$ (and hence $\beta/\alpha = 5.79775$) we have to choose a and b in the way stated in Table 5.5. In the same table we also predict the number of degrees of freedom imposed by those parameters if the coarsest grid consists of $1,000$ unknowns. Unfortunately it makes plain that such an examination is practically impossible. On the one hand the starting grid of $1,000$ nodes is too coarse and even in this case we would have to deal with several hundred millions of unknowns. So a more appropriate starting grid would only lead to a deterioration of the situation. On the other hand dropping those cases and considering only the ones with fewer unknowns would lead to too few data points to draw reasonable conclusions.

In addition to these numerical experiments we carried out a comparison between this transformation which was introduced in Section 2.8.1 and the one propounded in the original paper (cf. [28]) and reviewed by us in Section 2.8.3. Of course, we seek demonstration of the superiority of our suggestion.

To do so chose the following set of parameters for our computations:

$$\begin{aligned} Atol &= 1e-12 \\ Rtol &= 1e-10 \\ maxorder &= 5 \\ mglevels &= 4 \\ presmooth &= GMRES \text{ (5 steps)} \\ postsmooth &= SORU \text{ (5 steps)} \\ FEM-order &= 1 \end{aligned}$$

As a matter of fact it turned out that our transformation easily prevails since the convergence aspects are so bad because of the worse boundary singularities that our solution pattern could not even be completed for the computation of the reference solution. It was actally aborted after 49 time steps at a time of

$$t = 4.5108536914e - 5.$$

Until that point the duration of the computation had already consumed consumed $800,967.414$ seconds which amounts to about 9.5 days. From our point of view this warrants the conclusion that no reasonable results can be procured within our setting or within any acceptable time.

5.4 The Probability-Weighted One Month PDE Solution

In this section we are going to deal with the same problem as in the previous section. But this time we consider the errors after performing the probability weighting of the PDE solution as discussed in Section 2.9. In particular, we are interested in the following items:

- convergence rates for both the spatial error and the time error
- dependence of both the solution and the error on transaction costs.

Therefore we carry out the solution process as before and apply the probability weighting for two different transaction costs, namely

$$\begin{aligned} X_1 &= 0.001 \\ X_2 &= 0.03, \end{aligned}$$

where X_1 means that transaction costs are virtually absent (0.1%) and X_2 means 3% transaction costs. We have to point out that the minimizing effect of the probability weighting should be very hard to detect because of our time reversal, i. e. our one-month solution corresponds to "one month to maturity", which means that a benefit of early termination has to be huge to materialize over the course of one month. With respect to this examining a longer period of time would have been more sensible but the effect of probability weighting can be seen best when it is kind of isolated which is only the case after one month.
For our spatial convergence test we utilize the following set of solver parameters for both levels of transaction costs:

$$\begin{aligned} Atol &= 1e - 12 \\ Rtol &= 1e - 10 \\ maxorder &= 5 \\ mglevels &= 4 \\ presmooth &= GMRES \text{ (5 steps)} \\ postsmooth &= SORU \text{ (5 steps)} \\ FEM - order &= 2. \end{aligned}$$

In Table 5.6 the detailed results are given for X_1 and in Table 5.7 the same is done for X_2. As before, the appearing relative errors are computed using the norms of the probability-weighted solutions:

$$\begin{aligned} ||u_h^k||_{L^2}^{X_1} &= 6.17807099965e - 4 \\ ||u_h^k||_{L^\infty}^{X_1} &= 2.66587165019e - 4 \\ ||u_h^k||_{L^2}^{X_2} &= 6.17865138159e - 4 \\ ||u_h^k||_{L^\infty}^{X_2} &= 2.666788952219e - 4. \end{aligned}$$

DOFs	steps	error in $\|\|\cdot\|\|_{L^2}$		error in $\|\|\cdot\|\|_{L^\infty}$		CPU time [sec]
		abs	rel	abs	rel	
161,372	198	3.70670e-6	5.99977e-3	8.95529e-5	3.35924e-1	771.652
204,920	203	3.87099e-7	6.26569e-4	1.78894e-5	6.71053e-2	1,103.289
252,096	219	1.99856e-7	3.23493e-4	1.40728e-5	5.27887e-2	1,409.128
284,875	208	1.13059e-7	1.83000e-4	9.54968e-6	3.58220e-2	1,702.268
329,078	212	6.77227e-8	1.09618e-4	7.01506e-6	2.63143e-2	2,033.393
386,452	228	5.29531e-8	8.57114e-5	5.67727e-6	2.12961e-2	2,932.220
469,204	224	4.43424e-8	7.17739e-5	4.51410e-6	1.69329e-2	3,469.436
547,129	291	2.76113e-8	4.46924e-5	3.16423e-6	1.18694e-2	5,664.572
656,343	225	2.07524e-8	3.35904e-5	2.95162e-6	1.10719e-2	6,180.426
747,316	228	1.64914e-8	2.66934e-5	1.63728e-6	6.14163e-3	6,735.411
821,238	230	1.42835e-8	2.31197e-5	1.21703e-6	4.56522e-3	7,349.478
914,215	230	1.21078e-8	1.95980e-5	9.43110e-7	3.53772e-3	8,395.141
1,094,681	247	1.05122e-8	1.70153e-5	8.73149e-7	3.27529e-3	10,877.745
1,640,104	238	7.77897e-9	1.25913e-5	3.67943e-7	1.38020e-3	18,740.613

Table 5.6: Numerical error of the probability-weighted MBS PDE with a reference solution of 10,221,317 DOFs and transaction costs X_1

DOFs	steps	error in $\|\|\cdot\|\|_{L^2}$		error in $\|\|\cdot\|\|_{L^\infty}$		CPU time [sec]
		absolute	relative	absolute	relative	
161,372	198	3.70682e-6	5.99940e-3	8.95529e-5	3.35808e-1	822.327
204,920	203	3.87103e-7	6.26517e-4	1.78894e-5	6.70822e-2	1,172.371
252,096	219	1.99858e-7	3.23465e-4	1.40728e-5	5.27706e-2	1,551.371
284,875	208	1.13061e-7	1.82987e-4	9.54968e-6	3.58097e-2	1,646.890
329,078	212	6.77245e-8	1.09610e-4	7.01506e-6	2.63053e-2	2,094.708
386,452	228	5.29554e-8	8.57071e-5	5.67727e-6	2.12888e-2	2,592.235
469,204	224	4.43449e-8	7.17712e-5	4.51410e-6	1.69271e-2	3,255.139
547,129	291	2.76137e-8	4.46921e-5	3.16423e-6	1.18653e-2	5,012.824
656,343	225	2.07547e-8	3.35910e-5	2.95162e-6	1.10681e-2	4,853.303
747,316	228	1.64945e-8	2.66960e-5	1.63728e-6	6.13952e-3	5,711.476
821,238	230	1.42865e-8	2.31224e-5	1.21703e-6	4.56365e-3	6,517.478
914,215	230	1.21109e-8	1.96012e-5	9.43110e-7	3.53650e-3	7,777.320
1,094,681	247	1.05151e-8	1.70184e-5	8.73149e-7	3.27416e-3	9,332.201
1,640,104	238	7.78234e-9	1.25955e-5	3.67943e-7	1.37972e-3	17,128.482
1,830,547	239	6.33671e-9	1.02558e-5	2.79864e-7	1.04944e-3	18,268.549
1,939,171	239	6.17466e-9	9.99254e-5	2.52523e-7	9.46918e-4	20,463.217
2,059,265	241	6.11445e-9	9.89609e-5	2.14548e-7	8.04518e-4	21,194.452
2,351,857	253	5.76943e-9	9.33768e-5	1.99687e-7	7.48792e-4	22,703.448

Table 5.7: Numerical error of the probability-weighted MBS PDE with a reference solution of 10,221,317 DOFs and transaction costs X_2

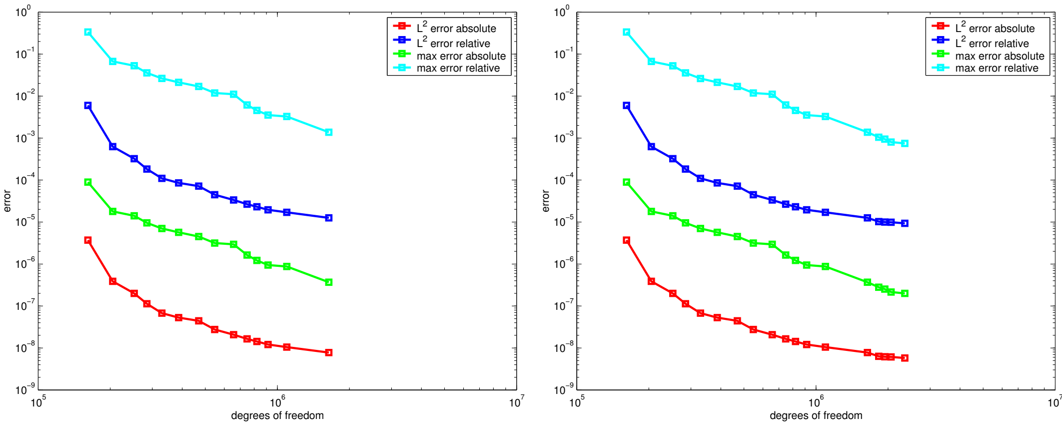

Figure 5.8: Left: spatial convergence plot for the probability-weighted PDE with transaction costs X_1, right: same with transaction costs X_2

tolerances		steps	error in $\|\| \cdot \|\|_{L^2}$		error in $\|\| \cdot \|\|_{L^\infty}$		CPU time [sec]
Atol	Rtol		absolute	relative	absolute	relative	
1e-4	1e-2	16	2.66891e-7	4.31998e-4	8.58947e-7	3.22201e-3	1,977.586
1e-5	1e-3	16	2.66081e-7	4.30687e-4	8.54731e-7	3.20619e-3	2,337.588
1e-6	1e-4	20	4.27539e-8	6.92028e-5	1.12412e-7	4.21670e-4	2,589.208
1e-7	1e-5	33	6.36187e-10	1.02975e-6	4.65355e-9	1.74560e-5	3,010.848
1e-8	1e-6	51	8.16460e-11	1.32155e-7	8.39152e-10	3.14776e-6	3,742.608
1e-9	1e-7	78	1.93129e-11	3.12605e-8	1.88616e-10	7.07520e-7	4,901.363
1e-10	1e-8	112	5.84066e-12	9.45388e-9	5.66641e-11	2.12554e-7	5,926.292
1e-11	1e-9	162	4.37882e-13	7.08770e-10	4.21476e-12	1.58100e-8	8,617.047
1e-12	1e-10	231	6.71106e-14	1.08627e-10	6.43449e-13	2.41365e-9	11,067.075

Table 5.8: Time error of the probability-weighted one-month PDE solution for transaction costs X_1

The results from Tables 5.6 and 5.7 are plotted in Figure 5.8. As it can be seen in the tables there are hardly any differences in the error so that it is almost impossible to see a difference in the graphs for X_1 and X_2.
Besides the error of the spatial discretization we have to examine the time discretization which is also done for both transaction costs X_1 and X_2 using the same strategies as before. Table 5.8 shows the results for X_1 and Table 5.9 does so for X_2. The relative errors were computed using the norms of the reference solution:

$$\begin{aligned} ||u_h^k||_{L^2}^{X_1} &= 6.1780592587e-4 \\ ||u_h^k||_{L^\infty}^{X_1} &= 2.6658743066e-4 \\ ||u_h^k||_{L^2}^{X_2} &= 6.17807099965e-4 \\ ||u_h^k||_{L^\infty}^{X_2} &= 2.66587165019e-4. \end{aligned}$$

Obtaining these reference solutions took 764 and 781 steps respectively and consumed $91,318.486$ and $30,391.620$ seconds. The results are graphically illustrated in the plots of Figure 5.9.
Apart from the convergence plots we have plotted the solutions of the probability-weighted PDE as before on a (200-by-200)-grid with the values being obtained via interpolation in Figure 5.10.

tolerances		steps	error in $\|\|\cdot\|\|_{L^2}$		error in $\|\|\cdot\|\|_{L^\infty}$		CPU time [sec]
Atol	Rtol		absolute	relative	absolute	relative	
1e-4	1e-2	16	2.66954e-7	4.03384e-4	8.58947e-7	3.22090e-3	2,316.794
1e-5	1e-3	16	2.66143e-7	4.02158e-4	8.54731e-7	3.20509e-3	2,116.144
1e-6	1e-4	20	4.27690e-8	6.46266e-5	1.12412e-7	4.21525e-4	2,427.245
1e-7	1e-5	33	6.36311e-10	9.61505e-6	4.65355e-9	1.74500e-5	3,185.550
1e-8	1e-6	51	8.16521e-11	1.23381e-7	8.39152e-10	3.14667e-6	3,671.460
1e-9	1e-7	78	1.93149e-11	2.91860e-8	1.88616e-10	7.07277e-7	4,871.417
1e-10	1e-8	112	5.84136e-12	8.82665e-9	5.66641e-11	2.12480e-7	5,819.222
1e-11	1e-9	162	4.37953e-13	6.61774e-10	4.21476e-12	1.58046e-8	8,448.239

Table 5.9: Time error of the probability-weighted one-month PDE solution for transaction costs X_2

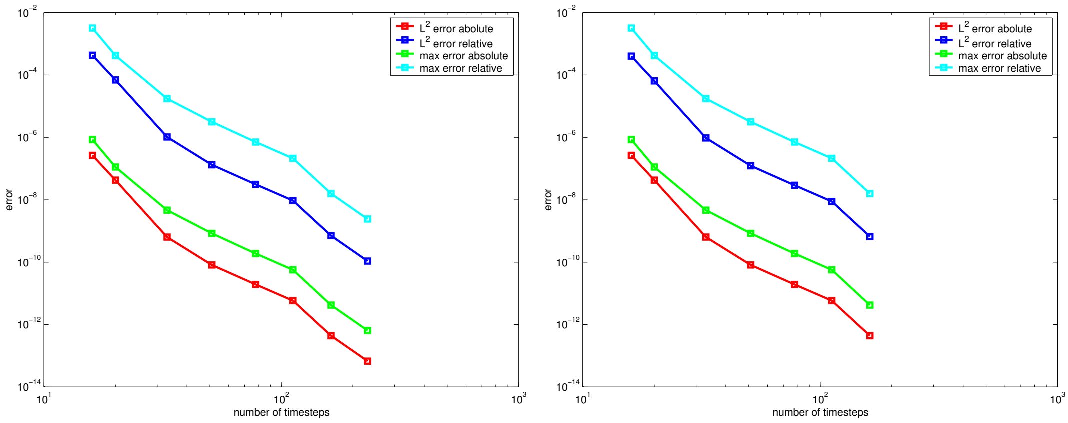

Figure 5.9: Left: time convergence plot for the probability-weighted PDE with transaction costs X_1, right: same with transaction costs X_2

In interpreting the probability-weighted results we distinguish, as before, between the time and spatial error.
Concerning the former one we see that there is hardly any difference in the errors between the solutions for transaction costs X_1 and X_2 just as we expected it. Moreover we observe convergence rates -4.71 and -4.58 respectively which are almost identical to the one for the non-weighted case.
Much the same can be said about the spatial error. Just as it was the case with the time error there are only very slight differences in the errors measured and we procure convergence rates of -1.15 and -1.07 for X_1 and X_2 respectively. In particular, we note that better rates are obtained than before, the reason for which is that we used second-order elements instead of first-order elements as we are no longer subjected to the limitations imposed by SUPG. But we are still short of the desired rate of -1.5 if our solution was smooth enough.
Altogether this can be considered as a confirmation of Theorem 3.6.1 on probability-weighted convergence and as indicative that our valuation pattern produces reasonable results since we predicted such a low effect of early termination in the aforementioned. Nonetheless we notice that in the case of higher transaction costs X_2 the overall liability measured in the L^2-norm is slightly larger compared to X_1 as it should be because termination is more expensive while moving on without altering anything remains the same. In addition, we take notion of the fact that both times probability-weighting causes a diminishing of the liability compared to not

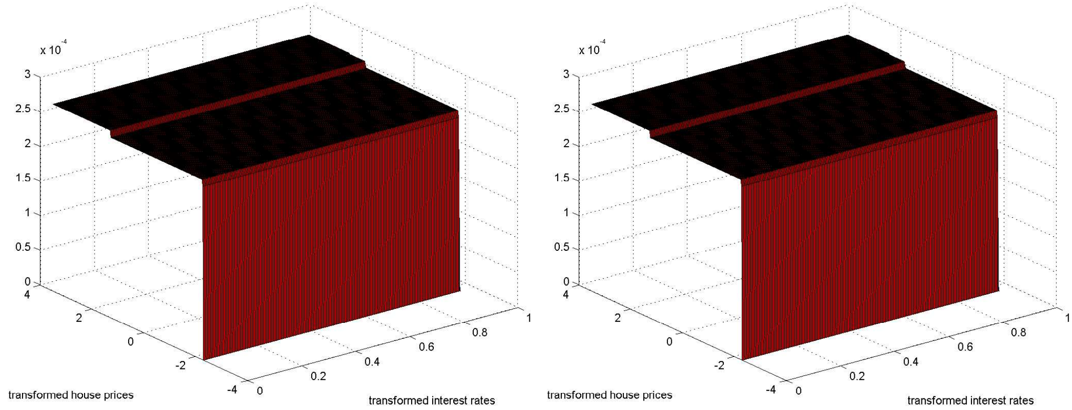

Figure 5.10: Left: solution plot of the probability-weighted PDE with transaction costs X_1, right: same with transaction costs X_2

considering possible option exercise such that

$$U \geq U_{X_2} \geq U_{X_1}$$

which is quite plausible because the higher transaction costs are the higher the liability should be. Moreover closely looking at the solution plots in Figure 5.10 confirmes that there is scarcely any difference resulting from the disparate transaction costs with only one month remaining on the mortgage contract.

5.5 Long-Term Computations

Having exhaustively analysed the convergence aspects of both the partial differential equation and the probability-weighted solution to the equation for one month's time we now move on examining what happens at larger time scales. Doing so we face five questions of major interest:

1. find an appropriate grid and time evolution in order to yield good enough accuracy while maintaining an acceptable computational time
2. determine how often early termination occurs
3. examine the dependence of early termination on both time and transaction costs and distinguish between prepayment and default in doing so
4. find the thresholds in both house prices and interest rates at which refinancing becomes advisable depending on both time and transaction costs
5. account for mortgage pools in which each mortgagor faces different transaction costs.

In the following answers to those questions will be procured by carrying out all the necessary computations.

5.5.1 Selection of Grids and Time Evolution

Over the course of our convergence analyses in the previous two sections (cf. 5.3 and 5.4) it became apparent that even for the time of one month an enormous amount of CPU time can be required. But since the valuation of an MBS backed by a 30-year mortgage amounts to considering 360 months this poses a huge problem.

DOF	Atol	Rtol	time	30y (approximately)
656,343	1e-6	1e-4	474.304	48 hours
747,316	1e-6	1e-4	567.308	57 hours
821,238	1e-6	1e-4	703.477	70 hours
914,215	1e-6	1e-4	767.763	77 hours

Table 5.10: Computational times for several grids varying in the number of degrees of freedom

There are two ways to reduce the overall work, the first being coarser grids in every time step and the second being a smaller number of time steps. However, both measures result in lower accuracy which is nonetheless acceptable because for convergence purposes we were seeking a level of accuracy far better than needed for "normal" computations. Therefore we have to find a compromise or more precisely as coarse a grid as possible and as few time steps as affordable while maintaining a certain level of accuracy. The effect is illustrated in the following example:

Example 5.5.1 (CPU Time) *Consider three different grids and time evolutions which take about* 700*,* 1,000 *and* 1,300 *seconds for the one-month solution process. This leads to estimated times of*

$$\begin{aligned} 70 \textit{ hours} &\approx 3 \textit{ days} \\ 100 \textit{ hours} &\approx 4 \textit{ days} \\ 130 \textit{ hours} &\approx 5.5 \textit{ days} \end{aligned}$$

if we assume that in every month there should be approximately the same number of time steps required but without considering that the probability weighting also takes its time in every month. (The stated values were obtained by taking 360 *times the duration for one month's time.)*

Altogether this example shows that it is worthwhile focusing on economic computations.
Our judgment is based on the results from Section 5.3 in which a time error of about $5e-7$ measured in the L^2-norm was determined for $Atol = 1e-6$ and $Rtol = 1e-4$. The spatial grid of our choice is taken from the examples in Table 5.10 where we carried out the one-month computation on several differently fine grids and projected the duration for a 360-month computation.
We choose the grid of about 650,000 unknowns because of the relatively short duration of the computations. In Figure 5.11 the unweighted one-month-PDE-solution is plotted for two different meshes in order to illustrate that this choice seems accurate enough.

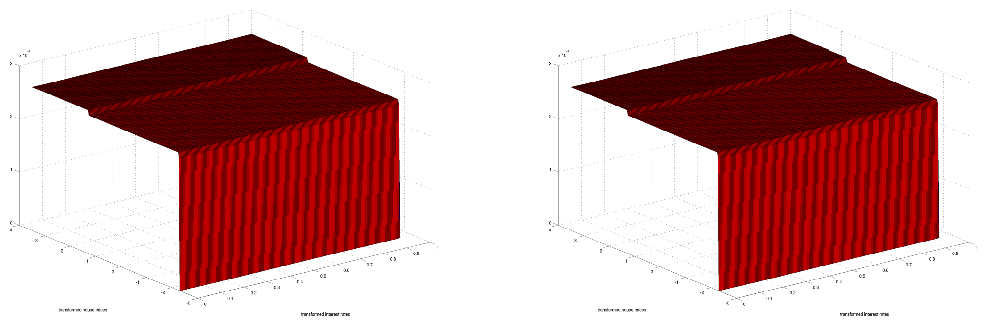

Figure 5.11: Left: one-month-PDE solution plot for 656,343 DOFs, right: same with 747,316 DOFs

5.5.2 Full 30-Year-Computation

The subject matter of this section is going to be the valuation of a mortgage-backed security over its entire lifespan of 360 months by the means outlined in thesis. To do so we use the grid determined in the previous subsection and the following set of parameters for the PDE solution process:

$$\begin{aligned}
Atol &= 1e-6\\
Rtol &= 1e-4\\
maxorder &= 5\\
mglevels &= 4\\
presmooth &= GMRES \text{ (5 steps)}\\
postsmooth &= SORU \text{ (5 steps)}\\
FEM-order &= 2.
\end{aligned}$$

As described in Section 2.3 we are interested in a pool across which transaction costs are spread according to a β-distribution. To attain this we pick a number $m \in \mathbb{N}$ of different transaction costs and calculate $X_1, \ldots, X_m$ according to Lemma 2.3.1 which means that we have to use the inverse cumulative distribution function F^{-1} by virtue of

$$X_j = F^{-1}\left(\frac{2j-1}{2m}\right), \qquad j = 1, \ldots, m.$$

We choose $m = 15$ which is a relatively small number but taking into account that each computation takes at least three days we have no choice but keeping it small.
However, we shall not carry out the actual mortgage pool computation because of the tremendous amount of computational time it would consume. Instead we resort to a comparison of the solution processes for three different levels of transaction costs (low-, high- and medium-level costs). If we wanted to price an MBS backed by whole pool we would merely take the average of the 15 solutions as outlined in Section 2.3. Presumably this would not lead to much further insight.
We first determine the 15 costs in the aforementioned way to obtain:

$$\begin{aligned}
X_1 &= 0.003739\\
X_2 &= 0.013564\\
X_3 &= 0.025073\\
X_4 &= 0.038022\\
X_5 &= 0.052428\\
X_6 &= 0.068422\\
X_7 &= 0.086225\\
X_8 &= 0.106174\\
X_9 &= 0.128759\\
X_{10} &= 0.154715\\
X_{11} &= 0.185203\\
X_{12} &= 0.222206\\
X_{13} &= 0.269578\\
X_{14} &= 0.336672\\
X_{15} &= 0.462117.
\end{aligned}$$

Among those we opt to closely examine the solution process for X_2, X_7 and X_{11}. Doing so we measure the duration of solving the PDE and weighting its result for every of the 360 months and during probability weighting we count for how many degrees of freedom prepayment is optimal, default is advisable and neither of the two is the case. In addition, we seek approximation of the threshold levels x^* and y^* with the property that they are the largest numbers such that

- default is optimal for every $y \leq y^*$
- prepayment is optimal for every $x \leq x^*$

Problematic about this is to actually maximize the thresholds because if early termination is optimal for some x or y choosing $x^* = x_{min}$ and $y^* = y_{min}$ will always satisfy these two conditions. The quality of this approximation is given by the grid we use because the finer the grid is the more accurately the thresholds can be resolved. We measure them in the following way:

Algorithm 5.5.2 (Computation of Thresholds)
The computation of the thresholds works in the following way:
Set $x^* := x_{min}$ *and* $y^* := y_{min}$
For $i = 1, \ldots, DOF$
  if default is optimal for (x_i, y_i) *and* $y_i > y^*$ *set* $y^* = y_i$
  if prepayment is optimal for (x_i, y_i) *and* $x_i > x^*$ *set* $x^* = x_i$
end for

In particular, we have to point out that there can be nodes for which both ways of termination are advisable (i. e. better than keeping the loan without altering anything). In such a case the more beneficiary way is determined in the sense that a minimization of the liability is intended, so that the other one is not considered optimal any longer. Therefore our measurement is kind of very strict because it accounts for optimality in both directions simultaneously instead of considering them independently. Moreover x^* and y^* are time-dependent, so that they have to be computed in every time step.

In the three plots of Figure 5.12 we have illustrated how often early termination occurs. Concerning the total number we find that the more time to maturity there is remaining the more often mortgages should be terminated. This behavior is qualitatively independent of transaction costs, which means that for higher transaction costs there are still more terminations advisable with more time remaining on the contracts.
However, transaction costs have an influence quantitatively. Intuitly one would immediately predict that for higher transaction costs there should be fewer terminations if the remaining time is held constant because the optimality criteria (2.1) and (2.2) are harder to be satisfied. We actually find our intuition confirmed as there are more terminations for X_2 than for X_7 and even fewer for X_{11}. Moreover the aforementioned difficulties in attaining optimality translate to more time remaining before the termination is optimal for the first time. In the plot we read that for X_2 it first happens in month 6 while for X_7 and X_{11} this is the case in months 48 and 68 respectively.
So far we have only analyzed the overall behavior but assiduously distinguishing between default and prepayment yields further insight. We see that generally there has to be more time remaining on the contract for default to become optimal than for prepayment. But transaction costs have a larger effect on prepayment. The latter statement means that prepayment occurs less often for higher transaction costs with X_{11} being large enough to rule out prepayment at all. In the case of default such a preclusion cannot be observed for X_{11}. For X_2 and X_7 there

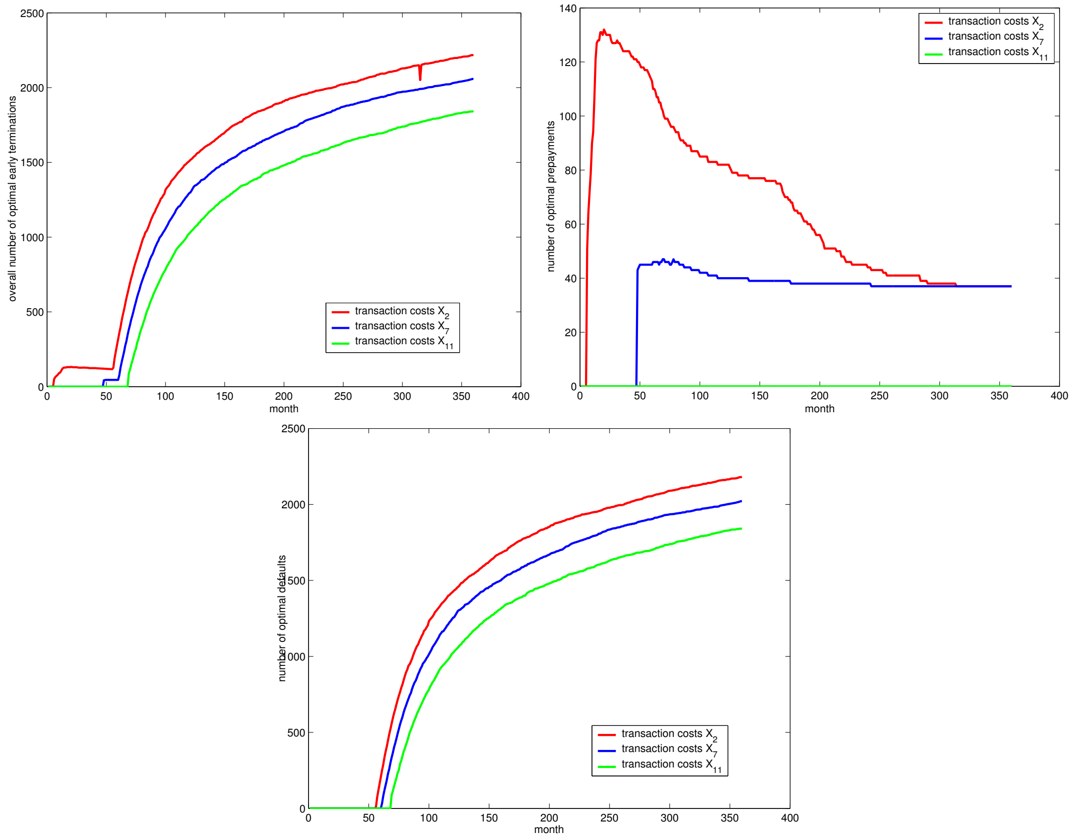

Figure 5.12: Number of degrees of freedom for which termination is financially optimal

is a strong increase in the number of prepayments for very few months remaining which is followed by a slow decrease once default starts to become optimal. Concerning X_2 prepayment is first optimal in month 6 while default only becomes advisable in month 56. In the case of X_7 the respective numbers are month 48 and 61. The development of default is in contrast to prepayment monotone in remaining time to maturity.

In Figure 5.13 we have plotted the threshold levels r^* and H^* of real interest rates and house prices at which taking action to terminate the mortgage becomes financially optimal. Their behavior is kind of parallel to the number of terminations as the more terminations occur the higher these levels are. This is absolutely plausible because more terminations mean that optimality is more easily attainable which by definition directly translates to higher thresholds. In addition, we notice that our interest rate thresholds are always smaller than the contractual interest rate c which has to be the case as they have to compensate for both transaction costs and forfeit of possible future option exercise.
Furthermore we perceive far more optimal default situations than there are for prepayment. Part of the reason for this lies in the design of the optimality checks. When it comes to defaulting we directly check against real estate prices while in the case of prepayment we check the remaining principal against a function of the interest rate, namely the PDE solution. By virtue of the boundary datum for $r = 0$ it cannot become arbitrarily small in contrast to the house prices. The second contribution we ought to mention is that we assumed a relatively small contractual interest rate of 5% which realistic but nonetheless small with respect to historical data (cf. Figure 1.2 in Section 1.3).
Another argument we can give why there are far more nodes at which termination through

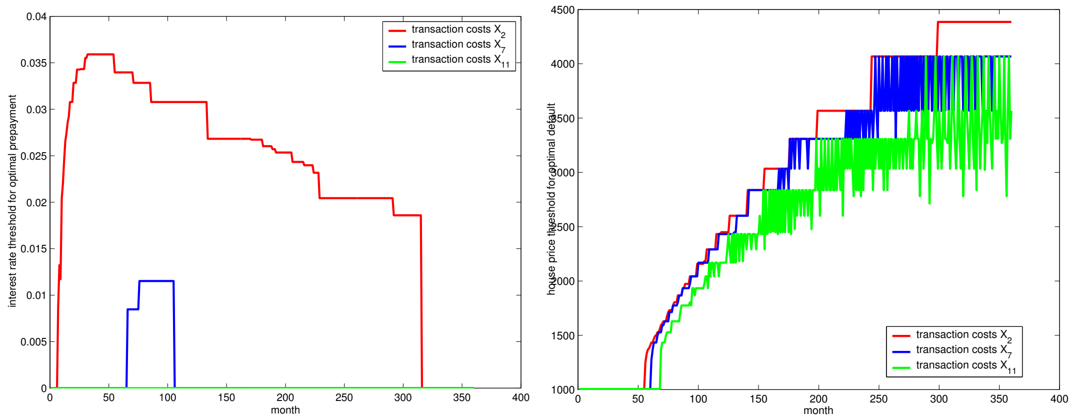

Figure 5.13: left: Thresholds for which prepayment becomes optimal, right: Same for default

default is optimal than there are for prepayment is that the mesh has more unknowns in the default region than in the prepayment region. The respective regions are those areas in which the respective termination method is optimal. Since those nodes are close to the boundaries $r = 0$ and $H = 0$ the validity of this argument becomes apparent from Figure 3.1 in which our mesh is plotted qualitatively.

Furthermore we are interested in gathering knowledge about the time development of the MBS which is why we not only consider the solution at the terminal time $T = 30$ but also at intermediate times. Of course, it does not make sense to plot every single one of the 360 months so that we resort to focusing on seven intermediate times, namely

$$\begin{aligned} t_1 &= 1 \\ t_2 &= 2 \\ t_3 &= 10 \\ t_4 &= 15 \\ t_5 &= 20 \\ t_6 &= 25 \\ t_7 &= 30, \end{aligned}$$

and in order to make plain that higher transaction costs in fact lead to a higher liability value we measure it in the L^2-norm in the same way as we did before during error measurement. The results are listed in Table 5.11. In the same table we have also stated the asset values of the MBSs and find that compared to the respective liability values they are always smaller. This confirms our observation from Section 2.4.

The aforementioned solution plots can be found in Figures 5.14, 5.15, 5.16, 5.17, 5.18, 5.19 and 5.20.

5.5.3 Comparison of Different Contractual Interest Rates

In this section I will examine the influence of a change in the contractual interest rates on the number of early terminations and the thresholds at which premature termination becomes optimal.

In the same way as in Section 5.5.2 we count the number of prepayments and defaults and

cost	X_2		X_7		X_{11}	
time	liability	asset	liability	asset	liability	asset
1 year	6.62877e-3	6.62605e-3	6.63627e-3	6.62817e-3	6.64567e-3	6.62817e-3
5 years	2.44426e-2	2.43691e-2	2.46859e-2	2.45444e-2	2.48466e-2	2.45569e-2
10 years	3.83243e-2	3.81723e-2	3.90278e-2	3.85580e-2	3.96007e-2	3.86869e-2
15 years	4.76407e-2	4.74761e-2	4.86272e-2	4.78544e-2	4.95832e-2	4.80261e-2
20 years	5.42244e-2	5.40383e-2	5.55344e-2	5.43221e-2	5.65995e-2	5.44649e-2
25 years	5.89139e-2	5.87022e-2	6.01331e-2	5.88877e-2	6.16194e-2	5.89981e-2
30 years	6.22804e-2	6.20407e-2	6.36147e-2	6.21933e-2	6.52767e-2	6.22389e-2

Table 5.11: Asset and liability values in L^2-norm depending on time and transaction costs

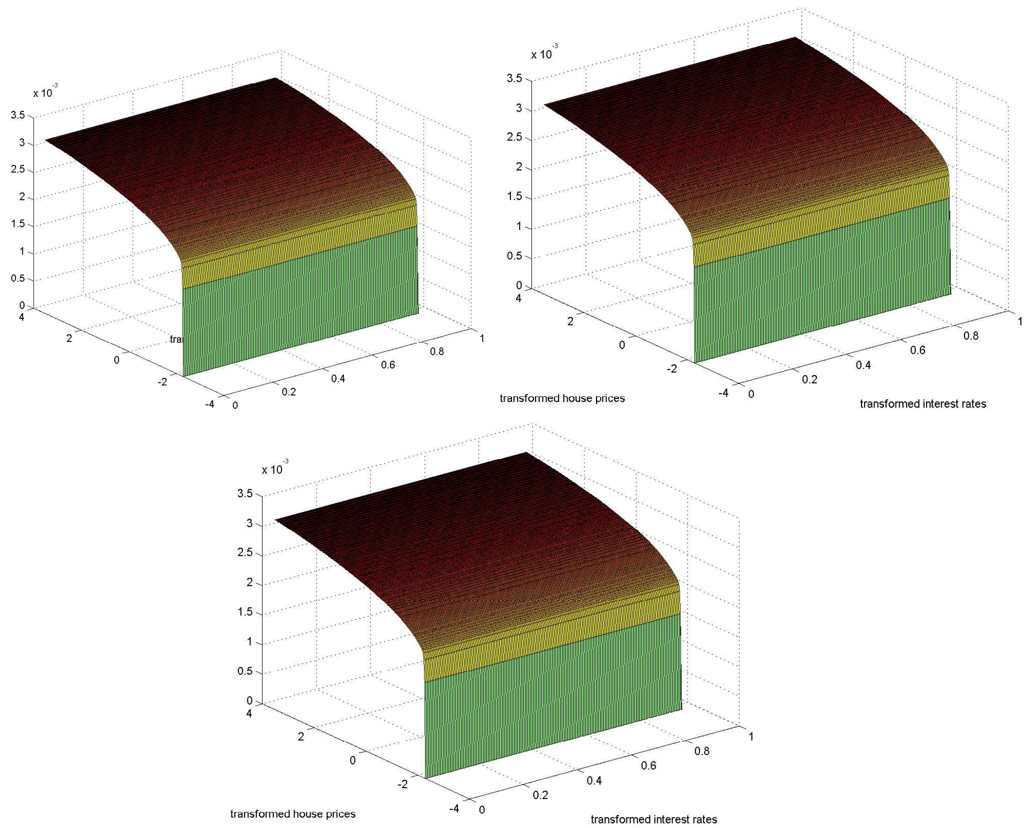

Figure 5.14: MBS value with one year remaining for transaction costs X_2, X_7 and X_{11}

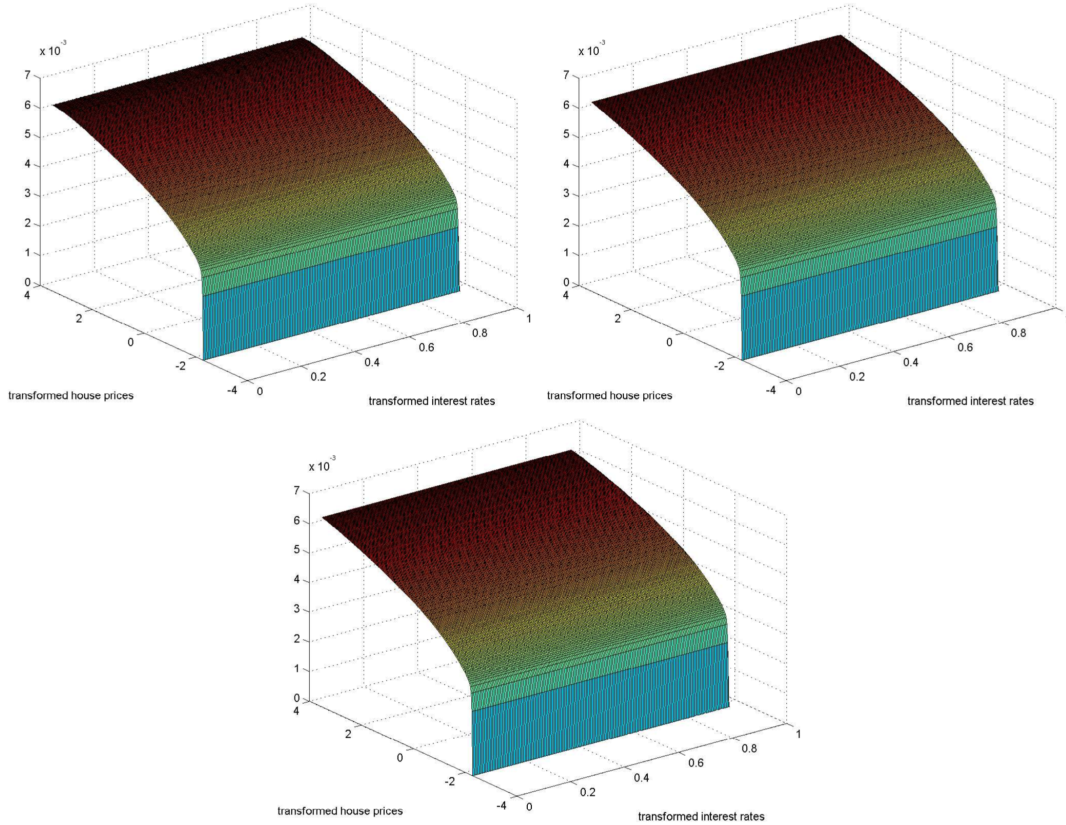

Figure 5.15: MBS value with two years remaining for transaction costs X_2, X_7 and X_{11}

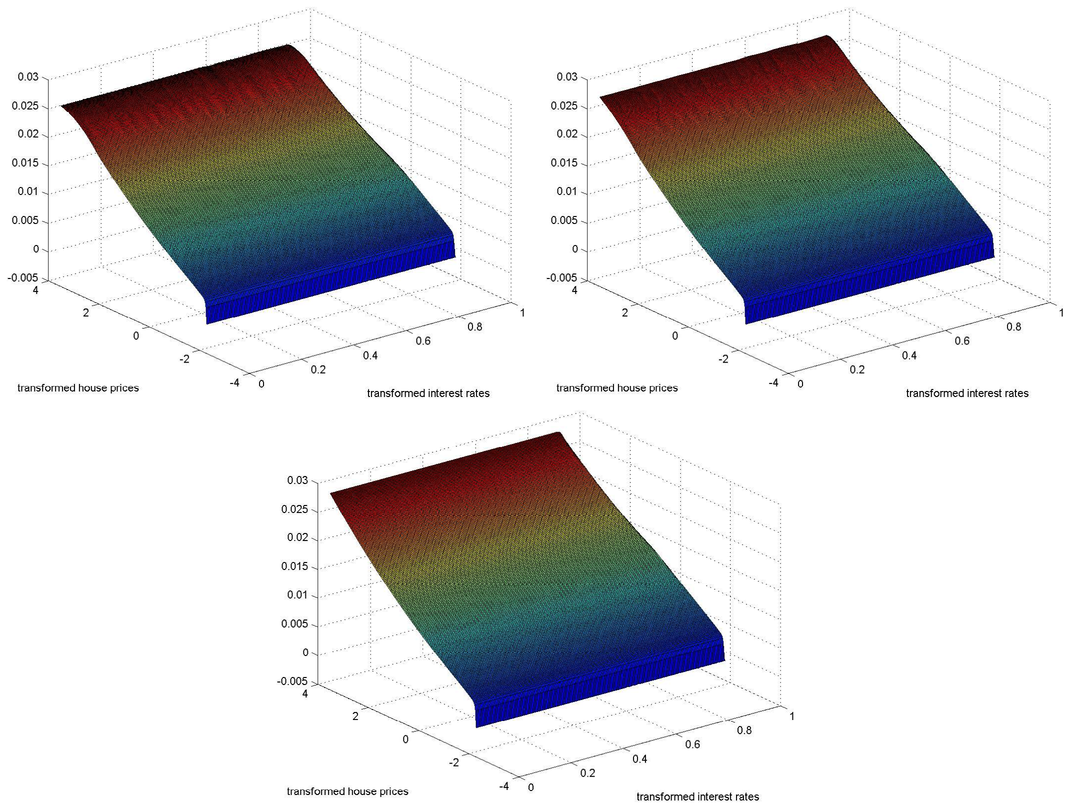

Figure 5.16: MBS value with ten years remaining for transaction costs X_2, X_7 and X_{11}

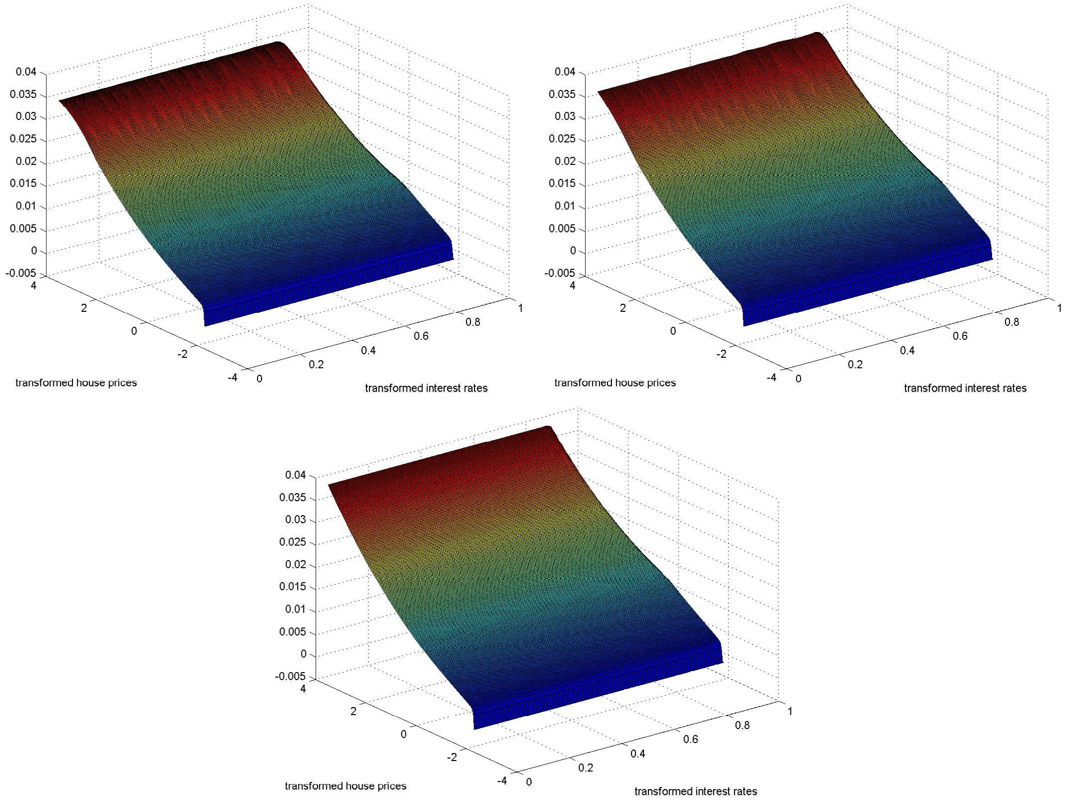

Figure 5.17: MBS value with fifteen years remaining for transaction costs X_2, X_7 and X_{11}

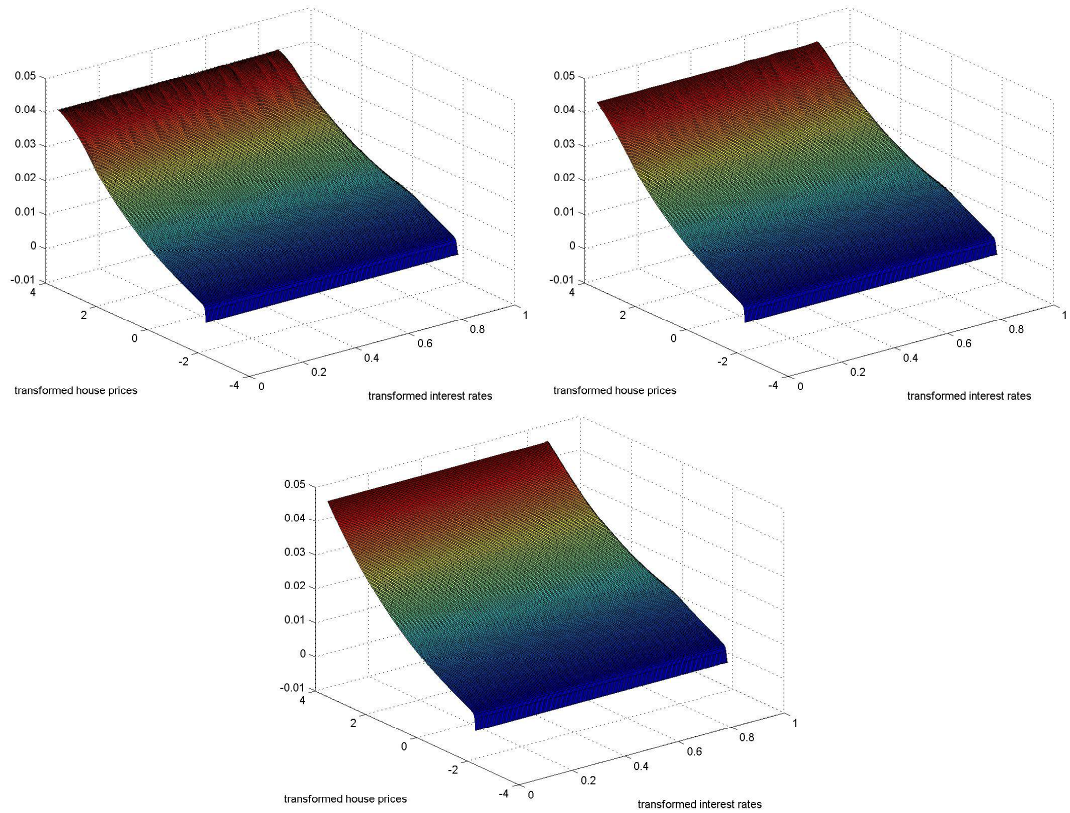

Figure 5.18: MBS value with twenty years remaining for transaction costs X_2, X_7 and X_{11}

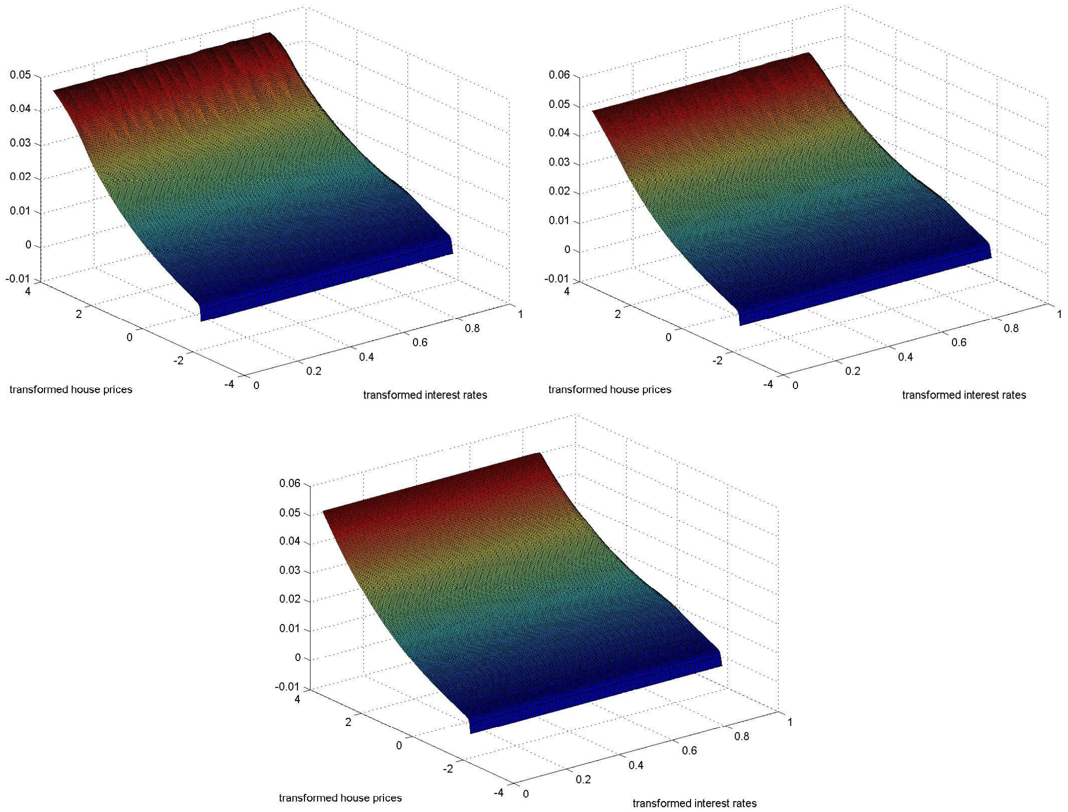

Figure 5.19: MBS value with twenty-five years remaining for transaction costs X_2, X_7 and X_{11}

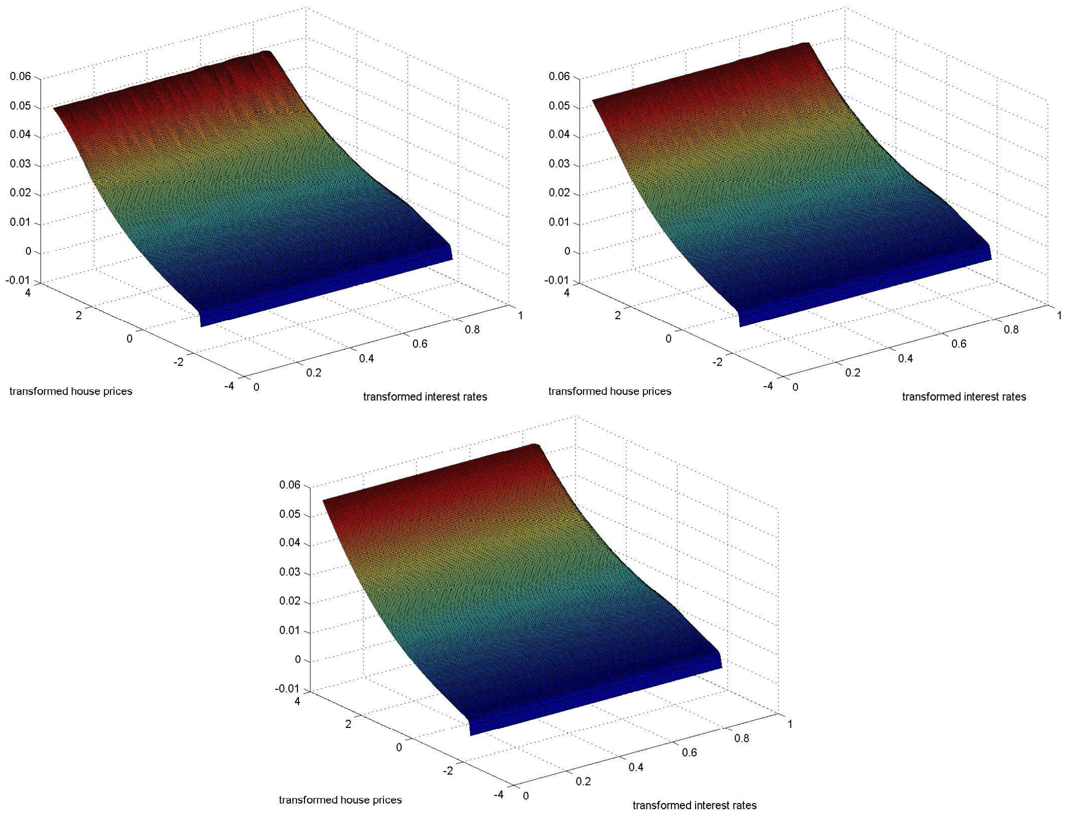

Figure 5.20: MBS value with thirty years remaining for transaction costs X_2, X_7 and X_{11}

determine the threshold levels according to Algorithm 5.5.2. We do so for contractual rates of

$$\begin{aligned} c_2 &= 0.085 \\ c_3 &= 0.1, \end{aligned}$$

which are also realistic levels if we take into account that many of the mortgages resold as MBSs were first initiated some ten or fifteen years ago. And at those times eight to ten percent interest obligation was quite common.

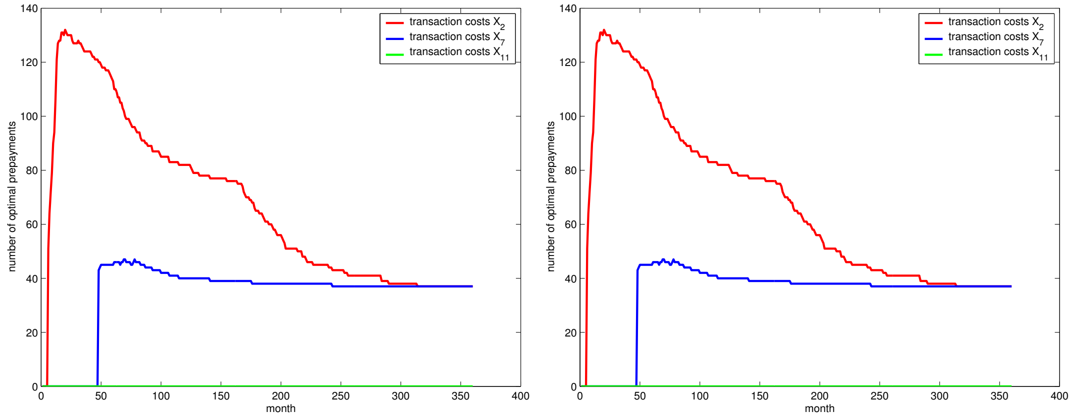

Figure 5.21: Left: Number of nodes for which prepayment is optimal in the case of c_2, Right: same for c_3

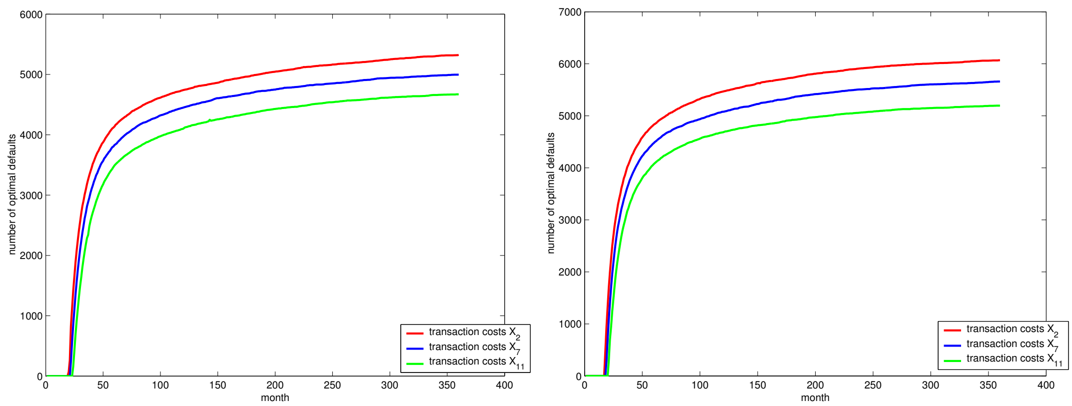

Figure 5.22: Left: Number of nodes for which default is optimal in the case of c_2, Right: same for c_3

For the sake of comparison the occurence numbers of prepayment are plotted in Figure 5.21, while in Figure 5.22 the figures for default are illustrated and Figure 5.23 shows the data for the overall numbers of optimal early termination through either way.
In order to minutely evaluate the data we should first ponder what to expect for different levels of contractual monthly coupons. The higher the scheduled monthly payment is the easier it should become to benefit from market movements. Since in the case of higher coupons the overall liability is also higher both interest rates and house prices do not have to drop as low for termination to materialize as a liability minimization. In consequence, we expect to spot a higher number of terminations in all three categories than before in Section 5.5.2 for $c_1 = 0.05$. Moreover for c_3 they should be even higher than for c_2. However, qualitatively the plots can be

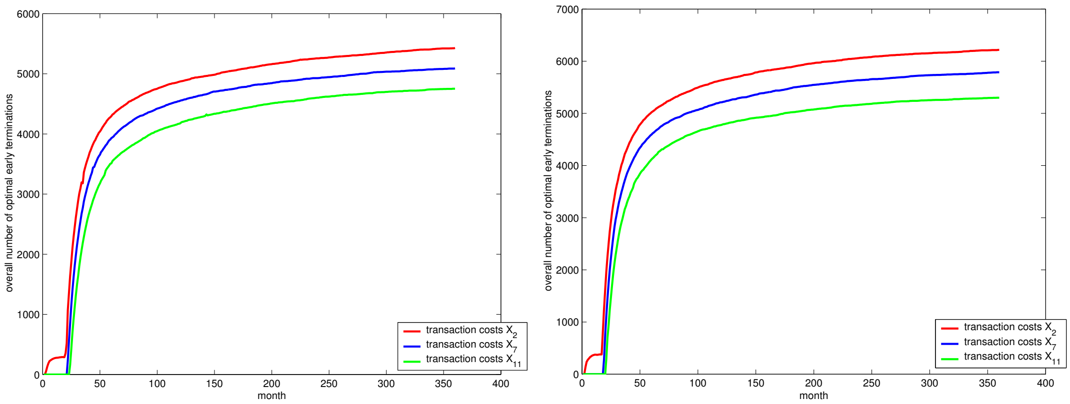

Figure 5.23: Left: Number of nodes for which either termination way is optimal in the case of c_2, Right: same for c_3

expected to look very much the same as for c_1. In the dependence on transaction costs we do not expect any qualitative altering either compared to what we have found in Section 5.5.2. Indeed, our expectations are confirmed. In all three categories we see that the higher transaction costs are the more unlikely premature termination is or the harder the optimality conditions (2.1) and (2.2) are to be satisfied. In addition, we take notion of the fact that the same rise in optimal defaults we already discerned in Figure 5.12 is seen in Figure 5.22 while concerning prepayment (cf. Figure 5.21) there once again is an increase followed by a slow decrease. Figure 5.23 looks very similar to Figure 5.22 because of the higher number of degrees of freedom for which default is optimal than it is the case for prepayment. The reasons for this are the same as in Section 5.5.2.
Yet, the actual numbers are slightly different. In the case of c_3 we detect more optimal early termination situations than in the case of c_2 for each level of transaction costs. For instance, there are $6,202$ optimal premature termination environments for X_2 and c_3 while only $5,409$ can be identified for X_2 and c_2. Both numbers are those for 340 months remaining on the contract. Moreover we find that the first time at which default is optimal depends on the contractual interest rate. In the case of X_7 and c_3, for example, month 18 marks the first time while for X_7 and c_2 month 22 is spotted. Once again this stresses the fact that the less convenient the contractual situation for a mortgagor is the more often termination becomes financially reasonable. Furthermore I would like to emphasize and point out that in contrast to the previous section where $c_1 = 0.05$ is dealt with, X_{11} is neither in the case of c_2 nor in the case of c_3 a precluding level for optimal prepayment to occur. However, there is a tendency that for c_2 the optimality condition for prepayment is more difficult to be satisfied than for c_3 translating to later months at which it first becomes optimal. We observe 3-to-3, 25-to-20 and 55-to-45 for X_2, X_7 and X_{11} respectively. In particular, the difference increases with increasing transaction costs.

Besides those absolute numbers of early termination we have also plotted the thresholds r^* and H^* at which prepaying or defaulting turns optimal in Figures 5.24 and 5.25. As it was the case in Section 5.5.2 the levels reflect the data, i. e. the more often either termination method is financially optimal the higher the respective threshold is. Qualitatively the plots look similar to those obtained for c_1 but the thresholds are higher because termination occurs more often for c_2 and c_3 and for c_3 even more often than for c_2. Except for those discernments one can state the same explanations for the shape of the respective graphs as in Section 5.5.2.

Moreover we examine the liability values for c_2 and c_3 and we compare them to the respective

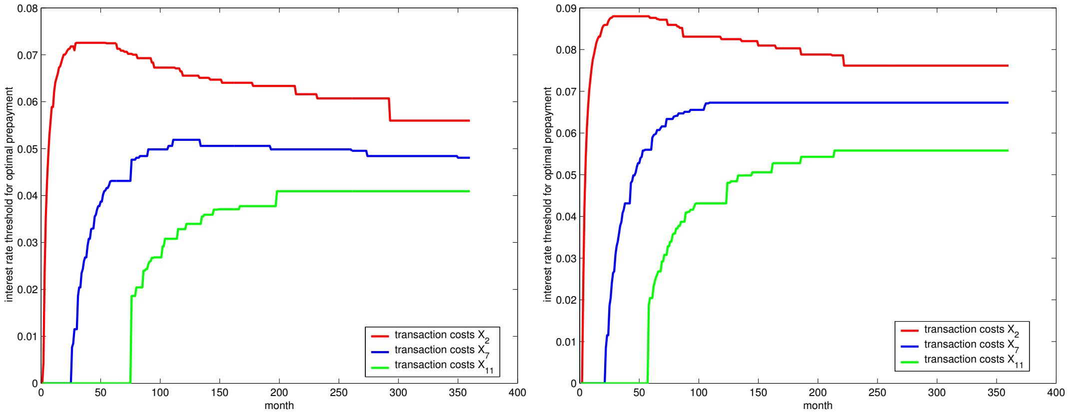

Figure 5.24: Left: Interest rate thresholds at which prepayment becomes optimal for c_2, Right: same for c_3

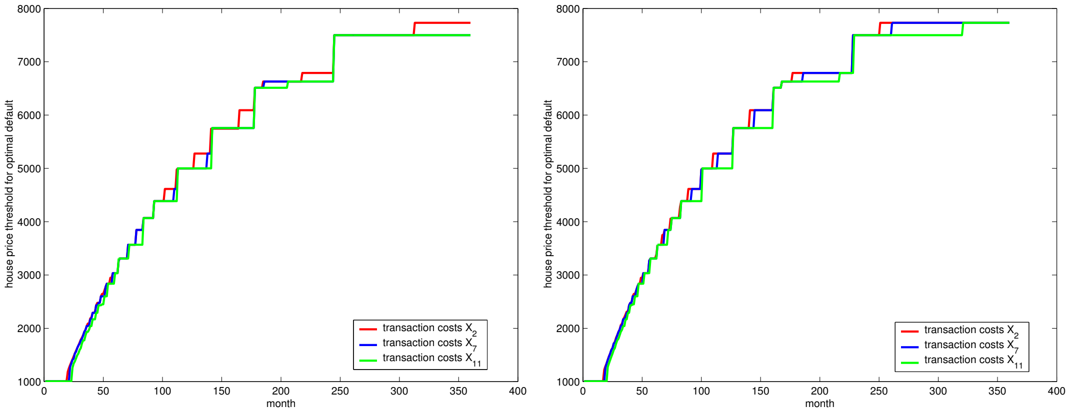

Figure 5.25: Left: house price thresholds at which default becomes optimal for c_2, Right: same for c_3

asset values in Tables 5.12 and 5.13 as well as the values for c_1 as stated in Table 5.11. Once again we measure them in the L^2-norm and observe that the asset value is always smaller than the liability value which is another confirmation of our argument from Section 2.4. In addition, we again find the expected dependence on transaction costs and time, i. e. the more time there is remaining on the contract the higher the liablity is and the lower transaction costs are the lower the liability is. Comparing the values for c_2 and c_3 to those for c_1 we see that a higher monthly coupon (or equivalently a higher contractual interest rate) translates to a higher liability at all times which is absolutely plausible from intuition.

cost	X_2		X_7		X_{11}	
time	liability	asset	liability	asset	liability	asset
1 year	1.60372e-2	1.60188e-2	1.60687e-2	1.60498e-2	1.60906e-2	1.60498e-2
5 years	5.73626e-2	5.71172e-2	5.85584e-2	5.85176e-2	5.93783e-2	5.85176e-2
10 years	8.58730e-2	8.54032e-2	8.88438e-2	8.91741e-2	9.14806e-2	8.91741e-2
15 years	1.02734e-1	1.02106e-1	1.06230e-1	1.07195e-1	1.10718e-1	1.07195e-1
20 years	1.13193e-1	1.12448e-1	1.17315e-1	1.18433e-1	1.22745e-1	1.18433e-1
25 years	1.19808e-1	1.18976e-1	1.24353e-1	1.25565e-1	1.30414e-1	1.25565e-1
30 years	1.24033e-1	1.23143e-1	1.28864e-1	1.30114e-1	1.35332e-1	1.30114e-1

Table 5.12: Asset and liability values for contractual interest rates c_2 and c_3

cost	X_2		X_7		X_{11}	
time	liability	asset	liability	asset	liability	asset
1 year	1.82803e-2	1.82536e-2	1.83270e-2	1.83055e-2	1.83515e-2	1.83055e-2
5 years	6.45409e-2	6.42249e-2	6.60773e-2	6.50785e-2	6.73573e-2	6.63387e-2
10 years	9.49820e-2	9.43676e-2	9.82345e-2	9.59403e-2	1.01834e-1	9.88995e-2
15 years	1.11968e-1	1.11162e-1	1.16301e-1	1.13166e-1	1.21333e-1	1.17045e-1
20 years	1.21889e-1	1.20976e-1	1.26886e-1	1.23216e-1	1.32836e-1	1.27704e-1
25 years	1.27996e-1	1.26813e-1	1.33168e-1	1.29159e-1	1.39676e-1	1.34017e-1
30 years	1.31317e-1	1.30319e-1	1.36936e-1	1.32719e-1	1.43786e-1	1.37811e-1

Table 5.13: Asset and liability values for a contractual interest rate c_3

Summary and Outlook

To conclude this diploma thesis it is my intention to summarize the main results and how they were attained. Moreover I wish to give suggestions and an outlook on possible further research.

Main Results

In this diploma thesis we have presented a strategy for the valuation of mortgage-backed securities by taking a given approach as outlined in [28] and enhancing it with the latest numerical techniques. Thereby we were able to deal with a complex financial asset from the field of economics in a mathematically very efficient way and thus to combine these two subjects of research. We are the first to use finite elements rather than differences in a structural model of mortgage-backed securities and the first to apply backward-differentiation formulas and a multigrid iteration in this field. Also a new transformation is proposed which significantly reduces problems encountered in [28].

There were three main problems in realizing these acchievements. The first was using a finite element discretization for the partial differential equation which ensures convergence even after weighting the numerical solution with the liabilities on prepayment or default. We found and proved that this can be attained even by standard Lagrange elements provided only that the PDE solution converges. The second problem was coping with a lack of regularity of the PDE solution. Since it is worse than H^2 the standard error estimates for finite elements cannot be applied and in consequence we perceive worse convergence rates than predicted in these estimates. Our attempt to improve matters was to use the streamline upwind Petrov Galerkin method which fell short because of the shortcomings of the solver COMSOL MULTIPHYSICS which only allows for first order elements and constant PDE coefficients. However, we were able to introduce a transformation which removes the violation of uniform parabolicity (the cause for the missing regularity) at least in house price direction so that after transformation all difficulties stemmed from the interest rate direction.

Concerning time discretization we were able to use backward-differentiation formulas rather than the inefficient and 30-year-old Hopscotch scheme (cf. [44]), which is used in [28], and a strategy to adaptively select order and length of a step and thus minimizing the number of steps taken.

Having considerably improved discretization we solved the equations in optimal complexity which we achieved with a multigrid solver which is pre-conditioned by the GMRES algorithm.

Eventually we tested the solver package by carrying out numerous computations for different transaction costs and contractual interest rates upon distinguishing between the asset and liability value. Thus we were able to show the robustness of our pattern. Moreover we demonstrated that our transformations are more favorable than the ones proposed in [28].

Possible Enhancements Within the Model

Among the shortcomings of our scheme is the fact that streamline upwind Petrov Galerkin did not work as intended because of the limitations of the implemented version in COMSOL

Multiphysics. In addition, we did not have a scheme which is fully adaptive in space in every time step and we rather imposed a non-uniform grid which remains the same over all time steps. Moreover in an entirely addaptive pattern time steps could be chosen optimally according to the given mesh.This full adaptivity is absolutely desirable because on the one hand it improves the quality of the solution at any given time step with respect to the spatial error since additional degrees of freedom are placed exactly where problems occur and not where we expect them to be based on apriori judgment. On the other hand this extent of adaptiveness saves computational work because no more variables than necessary are used to attain a given level of accuracy. And using such improved solutions one might even spare several time steps.
However, such a method is very hard to implement since this should coincide with a change from the *method of lines* to *Rothe's method*. The latter means to first discretize in time and then solve an elliptic problem in every time step. Doing so one faces the problem of having to map solutions from previous time steps (if multi-step methods are used) to a possibly completely different grid at the current time step. So an interpolation operator has to be provided and one has to store at least two complete grids in every step.
Moreover it might be interesting to obtain further theoretic results, particularly about regularity. This could include a rigorous proof that the solution of our PDE (2.6) is less regular than H^2 or a proof that the solution is contained in a weighted Sobolev space (see for instance [67] and [77]). But such research would better fit the field of functional analysis.
In addition, one might try to find very elaborate finite elements which ensure better convergence properties than standard Lagrange elements.
These suggested enhancements have in common that they could be done within the existing model and setting as they either seek improvement of the discretization or of the solution process. In addition, one can also conduct research outside the scope of our model.

Further Applications

Apart from the numerics for this approach or further theory another starting point for future research (and possibly huge numerical progress) might be modifying the model and interpreting the MBS valuation as a free boundary value problem (with the free boundaries being the points where early termination is optimal) as it is proposed in [29]. But following this approach would require completely different numerical methods such as linear complementarity. This technique is, for instance, used in the valuation of American-style options, which can also be interpreted as a free boundary problem (cf. [50]).
Besides that our setting cannot only be used to price mortgage-backed securities (either in a pool or as a single mortgage) but also for credit risk considerations by banks or other lenders. This is the case because unless a mortgage contract is securitized and resold as an MBS the lender of the original money bears and holds all the risk, i. e. in other words he or she owns a mortgage-backed security although the mortgage is not officially declared as such. And hence a bank might use this model to value its mortgage exposure.
Moreover our computations constitute the starting point for considering whole pools of mortgages as we would only need to carry out more computations for additional transaction costs (which would have to be chosen appropriately in accordance with a certain probability distribution) and take the average.

Appendix A

Stochastic Calculus

In this appendix I want to review the basic aspects of stochastic calculus needed for our purposes in this diploma thesis. This is largely based on the book on stochastic calculus by Karatzas and Shreve (cf. [58]).
The appendix is organized in two sections the first dealing with stochastic processes in general and the second is about stochastic differential equations by which the stochastic processes appearing in this work are defined.

A.1 Stochastic Processes

Since all the stochastic processes appearing are driven by Brownian motions we commence by stating the definition of a one-dimensional Brownian motion for which we need to know what an adapted stochastic process is.

Definition A.1.1 *The stochastic process X is adapted to the filtration $\mathcal{F}_t$ if for each $t \geq 0$ X_t is an $\mathcal{F}_t$-measurable random variable.*

And secondly the actual definition of a one-dimensional Brownian motion is:

Definition A.1.2 (Brownian Motion) *A standard one-dimensional Brownian motion is a continuous, adapted process $B = \{B_t, \mathcal{F}_t; 0 < t < \infty\}$ defined on some probability space $(\Omega, \mathcal{F}, P)$ with the following properties:*

1. *$B_0 = 0$ almost surely*
2. *For $0 \leq s < t$ the increment $B_t - B_s$ is independent of $\mathcal{F}_s$ and is normally distributed with mean 0 and variance $t - s$.*

Next I am going to state the multi-dimensional version of the Itô-formula. It describes the behavior of a stochastic process which is the function of finitely many other stochastic processes. But we need the following three definitions:

Definition A.1.3 *Let $X = \{X_t, \mathcal{F}_t; 0 \leq t\infty\}$ be a right-continuous martingale. We say that X is square integrable if $E[X_t^2] < \infty$ for every $t \geq 0$. If, in addition, $X_0 = 0$ a.s. we write $X \in \mathcal{M}_2$ or $X \in \mathcal{M}_2^c$ if X is also continuous.*

Definition A.1.4 *For $X \in \mathcal{M}_2$ we define the quadratic variation of X to be the unique (up to indistinguishability), adapted, natural, increasing process $< X >$ for which $< X >_0 = 0$ a.s. and $X^2 - < X >$ is a martingale.*

In [58] it is pointed out that $< X >_t$ is the process from the Doob-Meyer decomposition of X^2. Moreover evidence of existence is given for $X \in \mathcal{M}_2$.

Definition A.1.5 *For any two martingales $X, Y \in \mathcal{M}_2$ we define their cross-variation process $< X, Y >$ by*

$$< X, Y >_t := \frac{1}{4}\left[< X + Y >_t - < X - Y >_t\right]$$

for $0 \leq t < \infty$ and observe that $XY - < X, Y >$ is a martingale.

The reason for the latter claim is, according to [58], that for any two elements $X, Y \in \mathcal{M}_2$ both $(X+Y)^2 - < X + Y >$ and $(X-Y)^2 - < X - Y >$ are martingales and therefore so is their difference $4\,X\,Y - [< X + Y > - < X - Y >]$. Furthermore the authors elucidate that $< X, Y >$ is, up to indistinguishability, the only process of the form $A = A^{(1)} - A^{(2)}$ with $A^{(j)}$ adapted and natural increasing $(j = 1, 2)$ such that $XY - A$ is a martingale. But for our purposes we shall be content with merely stating these facts without proof. In particular, $< X, X > = < X >$ holds.
Having stated all these definitions and properties we are finally able to assert the renowned Itô-formula which is one of the most utilized techniques in standard contingent-claim pricing and hence extremely important for financial mathematics. In this thesis it is applied during the derivation of equation (2.6) in Section 2.5.

Theorem A.1.6 (Itô-Formula) *Let $\{M_t := (M_t^1, \ldots, M_t^d), \mathcal{F}_t; 0 \leq t < \infty\}$ be a vector of local martingales in $M^{c,loc}$ and $\{B_t := (B_t^1, \ldots, B_t^d), \mathcal{F}_t; 0 \leq t < \infty\}$ be a vector of adapted processes of bounded variation with $B_0 = 0$ and set*

$$X_t = X_0 + M_t + B_t,$$

where X_0 is an $\mathcal{F}_0$-measurable random vector in $\mathbb{R}^d$. Let $f(t, x) : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}$ be of class $C^{1,2}$. Then $P - a.s.$ for $0 \leq t < \infty$ the following holds:

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t \partial_t f(s, X_s)\, ds + \sum_{i=1}^d \int_0^t \partial_i f(s, X_s)\, dB_s^i + \sum_{i=1}^d \int_0^t \partial_i f(s, X_s)\, dM_s^i \\ &\quad + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \int_0^t \partial_i \partial_j f(s, X_s)\, d < M^i, M^j >_s . \end{aligned}$$

A proof of the Itô-formula is given in [58]. Due to its length and complicatedness we have opted to omit it at this point.

A.2 Stochastic Differential Equations

This section is devoted to strong solutions of stochastic differential equations with respect to Brownian motion. That is, we consider equations of the following type:

$$dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dW_t. \tag{A.1}$$

Here we suppose $b_i(t, x), \sigma_{ij}(t, x), 1 \leq i \leq d, 1 \leq j \leq m$ to be Borel-measurable mappings from $[0, \infty) \times \mathbb{R}^d$ into $\mathbb{R}$. We define the drift vector $b(x, t) = \{b_i(t, x))\}_{1 \leq i \leq d}$ and the dispersion matrix $\sigma(x, t) = \{\sigma_{ij}(t, x)\}_{1 \leq i \leq d, 1 \leq j \leq m}$ and read the above equation componentwise. This as well as the following statements and proofs are taken from [58].
This is of interest to us because we are dealing with such equations for interest rates and house

prices in the one-dimensional case. But before we can state the definition of a strong solution we have to make several other presumptions. We consider a probability space $(\Omega, \mathcal{F}, P)$ and an m-dimensional Brownian motion $W = \{W_t, \mathcal{F}_t^W; 0 \leq t < \infty\}$. In addition, we suppose the space to accomodate a vector $\xi \in \mathbb{R}^d$, independent of $\mathcal{F}_\infty^W$, with a given distribution

$$\mu(\Gamma) = P[\xi \in \Gamma], \quad \Gamma \in \mathcal{B}(\mathbb{R}^d).$$

Moreover we consider the left-continuous filtration

$$\mathcal{G}_t := \sigma(\xi) \vee \mathcal{F}_t^W = \sigma(\xi, W_s; 0 \leq s \leq t) \quad 0 \leq t < \infty$$

as well as the collection of null sets

$$\mathcal{N} := \{N \subseteq \Omega \mid \exists G \in \mathcal{G}_\infty \text{ with } N \subseteq G \text{ and } P[G] = 0\}$$

in order to create the augmented filtration

$$\begin{aligned} \mathcal{F}_t &:= \sigma(\mathcal{G}_t \cup \mathcal{N}), 0 \leq t < \infty && \text{(A.2)} \\ \mathcal{F}_\infty &:= \sigma(\bigcup_{t \geq 0} \mathcal{F}_t). && \text{(A.3)} \end{aligned}$$

In [58] it is shown that both $\{W_t, \mathcal{F}_t; 0 \leq t < \infty\}$ and $\{W_t, \mathcal{G}_t; 0 \leq t < \infty\}$ are Brownian motions. Now we have all the ingredients necessary to give the definition of a strong solution:

Definition A.2.1 (Strong Solution) *A strong solution of the stochastic differential equation (A.1) on the probability space $(\Omega, \mathcal{F}, P)$ with respect to the fixed Browian motion W and initial condition ξ is a process $X = \{X_t, 0 \leq t < \infty\}$ with continuous sample paths and with the following properties:*

1. *X is adapted to the filtration $\{\mathcal{F}_t\}$*
2. *$P[X_0 = \xi] = 1$*
3. *$P[\int_0^t \{|b_i(s, X_s)| + \sigma_{ij}^2(s, X_s)\}\, ds < \infty] = 1$ holds for every $1 \leq i \leq d$*
4. *the integral version*

$$X_t = X_0 + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dW_s, \qquad 0 \leq t < \infty$$

of (A.1) holds almost surely.

Next we have to define a concept of uniqueness:

Definition A.2.2 (Strong Uniqueness) *Let the drift vector $b(t, x)$ and dispersion matrix $\sigma(t, x)$ be given. Suppose that whenever W is an r-dimensional Brownian motion on some $(\Omega, \mathcal{F}, P)$, ξ is an independent, d-dimensional random vector, $\{\mathcal{F}_t\}$ is given by (A.2) and (A.3) and X, $\tilde{X}$ are two strong solutions of (A.1) relative to W with initial condition ξ, then*

$$P[X_t = \tilde{X}_t, 0 \leq t < \infty] = 1.$$

Under those conditions we say that strong uniqueness holds for the pair (b, σ).

Having stated definitions of strong solutions and uniqueness of the stochastic differential equation (A.1) we desire a criterion under which circumstances unique strong solutions exist.

Theorem A.2.3 (Existence of Unique Strong Solutions) *Suppose that the coefficients $b(t,x)$, $\sigma(t,x)$ are locally Lipschitz-continuous in the space variable, i. e. for every integer $n \geq 1$ there exists a constant $K_n > 0$ such that for every $t \geq 0$, $||x|| \leq n$ and $||y|| \leq n$*

$$||b(t,x) - b(t,y)|| + ||\sigma(t,x) - \sigma(t,y)|| \leq K_n \, ||x - y||.$$

Then strong uniqueness holds for (A.1).

A proof of this statement can be found in [58].
Next it is my intention to apply this result to our stochastic differential equations (2.3) and (2.4) governing the interest rate and house price processes:

$$\begin{aligned} dr &= (\kappa\theta_r - (\kappa + \eta)\, r)\, dt + \Phi_r \sqrt{r}\, dW_r \\ dH &= (r - q_H)\, H\, dt + \Phi_H H\, dW_H. \end{aligned}$$

Lemma A.2.4 *Assume the two stochastic differential equations (2.3) and (2.4) to be given. Then there are unique strong solutions on $(0, \infty)$ in the case of r and on $[0, \infty)$ in the case of H.*

Proof - Lemma A.2.4:
Concerning H the criterion is obviously satisfied for fixed $r \in [0, \infty)$ because both the drift and dispersion coefficients are linear in H, independently of t, and linear functions are even globally Lipschitz-continuous. Hence we choose:

$$K_n = \max(r - q_H, \Phi_H).$$

In the case of r things are slightly more complicated because the drift vector is linear again whereas the dispersion matrix is not. Therefore we consider some $\epsilon > 0$ and by application of the mean value theorem we estimate for $r_1, r_2 > \epsilon$:

$$\begin{aligned} |\sqrt{r_1} - \sqrt{r_2}| &\leq ||\frac{1}{2\sqrt{r}}||_\infty \, |r_1 - r_2| \\ &= \frac{1}{2\,\epsilon}\, |r_1 - r_2|. \end{aligned}$$

The latter claim holds because of the monotonicity of $\sqrt{\cdot}$. Since ϵ was chosen arbitrarily this altogether yields existence of unique strong solutions for r and H on $(0, \infty)$ and $[0, \infty)$ respectively.

$\square$

Appendix B

Functional Analysis

In this appendix I briefly summarize those statements and definitions from functional analysis which yield existence and uniqueness of solutions to variational equations in the elliptic case and which are helpful in procuring regularity statements. This is sensible because after reformulating our PDE in a weak form in (3.3) variational equations are what we have to deal with. Moreover the semi-discrete estimates from Section 3.5 require existence and regularity results for elliptic problems.
Therefore we start by stating the representation theorem of Riesz and the theorem of Lax and Milgram.

Proposition B.0.1 (Riesz) *Let X be a Hilbert-space, denote by X' its dual space and by $(\cdot,\cdot)_X$ its scalar product. Then*

$$J(x)(y) := (y,x)_X \qquad \forall x, y \in X$$

defines an isometric linear isomorphism $J : X \to X'$.

Using this statement one can obtain the following statement:

Proposition B.0.2 (Lax-Milgram) *Let X be a Hilbert space and $a : X \times X \to \mathbb{R}$ be bilinear, elliptic and continuous with constants c_0 and C_0 respectively. Then there is a uniquely defined mapping $A : X \to X$ satisfying*

$$a(y,x) = (y, Ax)_X \qquad \forall x, y \in X.$$

In addition, $A \in \mathcal{L}(X)$ is invertible and fulfils

$$\begin{aligned} ||A|| &\leq C_0 \\ ||A^{-1}|| &\leq \frac{1}{c_0}. \end{aligned}$$

Concerning proofs of both those statements we point out that they can be found in [2].
The desired existence and uniqueness results for elliptic problems can now be obtained by applying those results to the Sobolev space $H_0^{1,2}(\Omega)$, the bilinear form a and linear form l defined in Section 3.1.2.
Of course, we still have to define what Sobolev spaces are. They are spaces such that their elements satisfy integrability conditions and weak differentiablity. More precisely according to [2] this means:

Definition B.0.3 (Sobolev Spaces) *Let $m \geq 0$, $m \in \mathbb{Z}$, $1 \leq p \leq \infty$, $\Omega \subset \mathbb{R}^n$. Then we define the Sobolev space of order m with exponent p in the following way:*

$$\begin{aligned} H^{m,p}(\Omega) \;=\; & \{f \in L^p(\Omega) | \textit{ for any multi-index } |s| \leq m \textit{ there is } f^{(s)} \in L^p(\Omega) \textit{ with} \\ & f^{(0)} = f \textit{ and condition (B.1)}\}. \end{aligned}$$

Condition (B.1) is the following:

$$\int_\Omega \partial^s \zeta \, f^{(0)} = (-1)^{|s|} \int_\Omega \zeta \, f^{(s)} \qquad \forall \zeta \in C_0^\infty(\Omega). \tag{B.1}$$

Moreover the Sobolev space shall be equipped with the following norm:

$$||f||_{H^{m,p}(\Omega)} := \sum_{|s| \leq m} ||f^{(s)}||_{L^p(\Omega)}.$$

We would like to point out that in the case $p = 2$ $H^{m,2}(\Omega)$ is often referred to as $H^m(\Omega)$. The expressions $f^{(s)}$ are called weak derivatives because if f is differentiable one can apply the rule of integration by parts and $f^{(s)}$ happens to coincide with $\partial^s f$.
In order to be able to define the spaces $H_0^{m,p}$ we need the following statement which further elucidates the Sobolev spaces.

Proposition B.0.4 *For $f \in H^{m,p}(\Omega)$ with $1 \leq p < \infty$ there is a sequence $\{f_j\}_{j \in \mathbb{N}}$, $f_j \in H^{m,p}(\Omega) \cap C^\infty(\Omega)$, such that*

$$||f - f_j||_{H^{m,p}(\Omega)} \longrightarrow 0 \qquad \textit{for } j \to \infty.$$

A proof of this theorem can be found in [2]. Now we are able to state the following definition:

Definition B.0.5 ($H_0^{m,p}(\Omega)$-spaces) *Let $\Omega \subset \mathbb{R}^n$ open, $1 \leq p < \infty$ and $m \geq 0$ integral. Then we define:*

$$H_0^{m,p}(\Omega) := \{f \in H^{m,p}(\Omega) | \textit{ there is a sequence } \{f_k\}_{k \in \mathbb{N}} \textit{ satisfying condition (B.2)}\}.$$

Condition (B.2) is the following:

$$||f - f_k||_{H^{m,p}(\Omega)} \longrightarrow 0 \qquad k \to \infty. \tag{B.2}$$

These Sobolev functions can be further characterized, i. e. one can give criteria under which such a Sobolev function is contained in another Sobolev space. This is expressed in the following embedding theorem.

Theorem B.0.6 (Embedding in Sobolev Spaces) *Let $\Omega \subset \mathbb{R}^n$ be open and bounded with Lipschitz boundary. Further assume $m_1, m_2 \geq 0$ integral and $1 \leq p_1 < \infty$, $1 \leq p_2 < \infty$. Then the following assertions hold:*

1. *If $m_1 - \frac{n}{p_1} \geq m_2 - \frac{n}{p_2}$ and $m_1 \geq m_2$ there exists a continuous embedding $id : H^{m_1,p_1}(\Omega) \to H^{m_2,p_2}(\Omega)$, i. e. for a constant $C = C(n, \Omega, m_1, m_2, p_1, p_2)$ we have an estimate:*

$$||u||_{H^{m_2,p_2}(\Omega)} \leq C \, ||u||_{H^{m_1,p_1}(\Omega)}.$$

2. *If $m_1 - \frac{n}{p_1} \geq m_2 - \frac{n}{p_2}$ and $m_1 > m_2$ there is an embedding $id : H^{m_1,p_1}(\Omega) \to H^{m_2,p_2}(\Omega)$ and it is continuous and compact.*

3. *For arbitrary open and bounded domains $\Omega \subset \mathbb{R}^n$ the first two assertions hold for $H_0^{m_i,p_i}(\Omega)$ instead of $H^{m_i,p_i}(\Omega)$.*

Proofs of these assertions can once again be found in [2].
Interpreting the results we see that the higher the dimension n is the easier the embedding condition can be fulfilled and the higher the integration order p_2 is the harder it becomes. In particular, we note that for $p_1 = p_2$ $H^{m_1,p_1} \subseteq H^{m_2,p_2}$ which seems very reasonable and plausible from a common sense perspective.
Besides this embedding theorem there is another one which deals with the question under which circumstances Sobolev functions can be continuous. More precisely the embedding theorem is set in Hölder spaces:

Theorem B.0.7 (Embedding from Sobolev Spaces into Hölder Spaces) *Let $\Omega \subset \mathbb{R}^n$ be open and bounded with Lipschitz boundary. Moreover we assume an integer $m \geq 0$ and $1 \leq p < \infty$ as well as an integer $k \geq 1$ and $0 \leq \alpha \leq 1$. Then the following assertions hold:*

1. *If $m - \frac{n}{p} = k + \alpha$ there is a continuous embedding $id : H^{m,p}(\Omega) \to C^{k,\alpha}(\bar{\Omega})$. More precisely, for $u \in H^{m,p}(\Omega)$ there is one and only one continuous function which almost everywhere coincides with u (and it be also denoted by u) such that*

$$||u||_{C^{k,\alpha}(\bar{\Omega})} \leq C(\Omega, n, m, p, k, \alpha)\ ||u||_{H^{m,p}(\Omega)}.$$

2. *If $m - \frac{n}{p} > k + \alpha$ there is a continuous and compact embedding $id : H^{m,p}(\Omega) \to C^{k,\alpha}(\bar{\Omega})$.*

3. *For arbitrary open and bounded sets $\Omega \subset \mathbb{R}^n$ the first two statements hold for $H_0^{m,p}(\Omega)$ instead of $H^{m,p}(\Omega)$.*

Having stated this embedding result we have yet to define what Hölder spaces are.

Definition B.0.8 (Hölder Spaces) *For $\Omega \subset \mathbb{R}^n$ open, bounded and $m \geq 0$ as well as a Banach space Y with norm $|\cdot|$ and for $0 < \alpha < 1$ the Hölder spaces are defined in the following way:*

$$C^{m,\alpha}(\bar{\Omega}; Y) := \{f \in C^m(\bar{\Omega}; Y) | H\ddot{o}l_\alpha(\partial^s f, \bar{\Omega}) < \infty \text{ for } |s| = m\},$$

where the Hölder constant $H\ddot{o}l_\alpha(f, \bar{\Omega})$ is defined as

$$H\ddot{o}l_\alpha(f, \bar{\Omega}) := \sup \left\{ \frac{|f(x) - f(y)|}{|x - y|^\alpha} \ | \quad x, y \in \bar{\Omega}, x \neq y \right\} \in [0, \infty].$$

Together with the norm

$$||f||_{C^{m,\alpha}(\bar{\Omega})} := \sum_{|s| \leq m} ||\partial^s f||_{C^0(\bar{\Omega})} + \sum_{|s| = m} H\ddot{o}l_\alpha(\partial^s f, \bar{\Omega})$$

they form Banach spaces.

This definition is taken from [2] where, in addition, both a proof of the latter claim on the Banach space property and a proof of the embedding statements can be found.

Bibliography

[1] R. Adams: Sobolev Spaces, Academic Press, 1975

[2] Hans-Wilhelm Alt: Lineare Funktionalanalysis, Springer Verlag, Berlin, Heidelberg 2002

[3] R. Anderson, S. Sundaresan: A Comparative Study of Structural Models of Corporate Bond Yields: An Exploratory Investigation, Journal of Banking and Finance 24, 2000, pp. 255-269

[4] Kristina Andersson: Stochastic Volatility, U.U.D.M Project Report 2003:18, Department of Mathematics, Uppsala University, November 2003

[5] Clifford A. Ball, Antonio Roma: Stochastic Volatility Option Pricing, Journal of Financial and Quantitative Analysis 29 No. 4, 1994, pp. 589-607

[6] Heinz Bauer: Wahrscheinlichkeitstheorie, de Gruyter Verlag, Berlin New York 2002

[7] S. Becketti, C. S. Morris: The prepayment experience of FNMA mortgage-backed securities, Working paper, New York University Salomon Center, 1990

[8] Fischer Black, Myron Scholes: The Pricing of Options and Corporate Liabilities, Journal of Political Economy 81, May-June 1973, pp. 637-659

[9] Fischer Black: The Pricing of Commodity Contracts, Journal Financial Economics 3, pp. 167-179, 1976

[10] J. Boudoukh, M. Richardson, R. Stanton, R. F. Whitelaw: A New Strategy for Dynamically Hedging Mortgage-Backed Securities, Journal of Derivatives 2, 1995, pp. 60-77

[11] Dietrich Braess: Finite Elemente, Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie, Springer Verlag, Berlin, Heidelberg 2003

[12] K. Brenan, S. Campbell, Linda Petzold: Numerical Solution of Initial Value Problems in Differential-Algebraic Equations, Elsevier, New York, 1989

[13] Peter N. Brown, Alan C. Hindmarsh, Linda R. Petzold: Using Krylov Methods in the Solution of Large-Scale Differential-Algebraic Systems, SIAM J. Sci. Comput., Vol. 15, No. 6, pp. 1467-1488, 1994

[14] Peter N. Brown, Alan C. Hindmarsh, Linda R. Petzold: Consistent Initial Condition Calculations for Differential Algebraic Systems, SIAM J. Sci. Comput., Vol. 19, No. 5, pp. 1495-1512, September 1998

[15] Tony F. Chan, Henk A. Van Der Vorst: Approximate and Incomplete Factorizations, in: David E. Keyes, Ahmed Samed, V. Venkatkrishnan, editors, Parallel Numerical Algorithms, ICASE/LaRC Interdisciplinary Series in Science and Engineering, pp. 167-202, Kluver Academic, Dordecht, 1997

[16] Ren-Raw Chen, Louis Scott: Multi-Factor Cox-Ingersoll-Ross Models of the Term Structure: Estimates and Tests From a Kalman Filter Model, The Journal of Real Estate Finance and Economics, pp. 143-172, 2003

[17] Nigel Clarke, Kevin Parrott: Multigrid for American Option Pricing With Stochastic Volatility, Appl. Math. Finance, 6: 177-197, 1999

[18] P. Collin-Dufresne, B. Solnik: On the Term Structure of Default Premia in the Swap and LIBOR Markets, Journal of Finance 56, pp. 1095-1115, 2001

[19] S. Scott Collis, Matthias Heinkenschloss, Analysis of the Streamline Upwind/Petrov Galerkin Method Applied to the Solution of Optimal Control Problem, March 2002

[20] J.C. Cox, J.E. Ingersoll, S.A. Ross: A Theory of the Term Structure of Interest Rates, Econometrica 53, 781-793, 1985

[21] Timothy A. Davis, Iain S. Duff: An Unsymmetric-Pattern Multifrontal Method for Sparse LU Factorization, SIAM J. Matrix Anal. Appl., Vol. 18, No. 1, pp. 140-158, January 1997

[22] Timothy A. Davis, Iain S. Duff: A Combined Unifrontal/Multifrontal Method for Unsymmetric Sparse Matrices, ACM Transactions on Mathematical Software, Vol. 25, No. 1, March 1999, Pages 1-20

[23] Timothy A. Davis: A Column Pre-Ordering Strategy for the Unsymmetric-Pattern Multifrontal Method, ACM Transactions on Mathematical Software, Vol. 30, No. 2, June 2004, Pages 165-195

[24] Timothy A. Davis: Algorithm 832: UMFPACK V4.3 - An Unsymmetric Pattern Multifrontal Method, ACM Transactions on Mathematical Software, Vol. 30, No. 2, June 2004, Pages 196-199

[25] Yongheng Deng, John M. Quigley: Woodhead Behavior and the Pricing of Residential Mortgages, Working Paper, University of Southern California, 2002

[26] Yongheng Deng, John M. Quigley, Robert Van Order: Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options, Econometrica 68, 275-307, 2000

[27] Emmanuele DiBenedetto: Degenerate Parabolic Equations, Springer Verlag, New York 1993

[28] Chris Downing, Richard Stanton, Nancy Wallace: An Empirical Test of a Two-Factor Mortgage Valuation Model: How Much Do House Prices Matter?, Working Paper, University of California, Berkeley, 2003

[29] Chris Downing, Richard Stanton, Nancy Wallace: An Empirical Test of a Two-Factor Mortgage Valuation Model: How Much Do House Prices Matter?, Real Estate Economics 33, pp. 681-710, 2005

[30] G. R. Duffee: The Relation Between Treasury Yields and Corporate Bond Yield Spreads, Journal of Finance 53, pp. 2225-2242, 1998

[31] Daniel J. Duffy: Finite Difference Methods in Financial Engineering, A Partial Differential Equation Approach, Wiley, The Atrium, Southern Gate, Chichester, West Sussex, England 2006

[32] Daniel Duffy, D. Lando: The Term Structure of Credit Spreads with Incomplete Accounting Information, Econometrica 69, pp. 633-664, 2001

[33] Daniel Duffy, K. J. Singleton: Modeling Term Structures of Defaultable Bonds, Review of Financial Studies 12, pp. 687-720, 1999

[34] K.B. Dunn, J.J. McConnell, A Comparison of Alternative Models for Pricing GNMA Mortgage-Backed Securities, Journal of Finance 36, pp. 471-483, 1981

[35] K.B. Dunn, J.J. McConnell, Valuation of Mortgage-Backed Securities, Journal of Finance 36, pp. 599-617, 1981

[36] K. B. Dunn, C. S. Spatt: The Effect of Refinancing Costs and Market Imperfections on the Optimal Call Strategy and the Pricing of Debt Contracts, Working Paper, Carnegie-Mellon University, 1986

[37] Lawrence C. Evans: Partial Differential Equations, Graduate Studies in Mathematics Volume 19, American Mathematical Society, Providence, Rhode Island 1998

[38] Gerd Fischer: Lineare Algebra, Fiedr. Vieweg & Sohn Verlagsgesellschaft mbH, Braunschweig / Wiesbaden 2000

[39] Mark B. Garman: A General Theory of Asset Valuation Under Diffusion State Processes, Working Paper No. 50, Graduate School of Business Administration, University of California, Berkeley, July 1976

[40] C.W. Gear: Simultaneous Numerical Solution of Differential/Algebraic Equations, IEEE Trans. Circuit Theory, CT-18, 1 (1971), pp. 89-95

[41] Thomas Gerstner, Michael Griebel: Praktikumsausarbeitung Computational Finance, Institut für Angewandte Mathematik, Rheinische Friedrich-Wilhelms-Universität Bonn, 2003

[42] John R. Gilbert, Sivan Toledo: An Assessment of Incomplete LU-Preconditioners for Non-Symmetric Linear Systems, Informatica 24, pp. 409-425, 2000

[43] A.R. Gourlay: Hopscotch: a Fast Second-Order Partial Differential Equation Solver, J. Inst. Maths. Applics. 6, pp. 375-390, 1970

[44] A.R. Gourlay, S. McKee: The Construction of Hopscotch Methods for Parabolic and Elliptic Equations in Two Space Dimensions with a Mixed Derivative, Journal of Computational and Applied Mathematics 3, pp. 201-206, 1977

[45] W. Hackbusch: Multigrid Methods and Applications, Springer Verlag, Berlin, Heidelberg, New York 1985

[46] E. Hairer, G. Wanner: Solving Ordinary Differential Equations 2, Stiff and Differential-Algebraic Problems, Springer Verlag, 1991

[47] G. Hämmerlin, K. H. Hoffmann: Numerische Mathematik, Springer Verlag, 1994

[48] Martin Hanke-Bourgeois: Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens, Teubner Verlag, Stuttgart, Leipzig, Wiesbaden 2002

[49] Cris Shelfo Heaton: US Housing Market Is Different this Time - It's Worse, www.moneyweek.com, 2006

[50] Markus Holtz: Konstruktion B-Spline basierter monotoner Mehrgitterverfahren zur Bewertung amerikanischer Optionen, Rheinische Friedrich-Wilhelms-Universitt Bonn, Diplomarbeit, Mai 2004

[51] John C. Hull: Options, Futures and Other Derivatives, Prentice Hall, Upper Saddle River, New Jersey 2006

[52] John C. Hull, Alan White: The Pricing of Options on Assets with Stochastic Volatility, The Journal of Finance 42 No. 2, June 1987

[53] R. A. Jarrow, S. M. Turnbull: Pricing Options on Financial Securities Subject to Default Risk, Journal of Finance 50, pp. 53-86, 1995

[54] W. Jay, K. Anderson: The Role of Mortgage Hedging in U.S. Fixed Income Markets, Mimeograph, Lehman Brothers and Blackrock, 2002

[55] Herb Johnson, David Shanno: Option Pricing when the Variance Is Changing, Journal of Financial and Quatitative Analysis 22 No. 2, June 1987

[56] E. Johnston, L. Van Drunen, Pricing Mortgage Pools with Heterogeneous Mortgagors: Emperical Evidence, Working Paper, University of Utah, 1988

[57] Ansgar Jüngel: Das kleine Finite-Elemente-Skript, Vorlesungsskript, Johannes-Gutenberg-Universität Mainz, 2001

[58] Ioannis Karatzas, Stephen E. Shreve: Brownian Motion and Stochastic Calculus, Springer Verlag, New York 1988

[59] Takeaki Kariya, Fumiaki Ushiyama, Stanley R. Pliska: A 3-Factor Valuation Model for Mortgage-Backed Securities (MBS), April 2002

[60] James B. Kau, Donald C. Keenan: An Overview of the Option-Theoretic Pricing of Mortgages, Journal of Housing Research 6, pp. 217-244, 1995

[61] J. B. Kau, D. C. Keenan, W. J. Muller, J. F. Epperson: A Generalized Valuation Model for Fixed-Rate Residential Mortgages, Journal of Money, Credit and Banking 24, pp. 279-299, 1992

[62] J. B. Kau, D. C. Keenan, W. J. Muller, J. F. Epperson: The Valuation at Origination of Fixed-Rate Mortgages with Default and Prepayment, Journal of Real Estate Finance and Economics 11, pp. 5-39, 1995

[63] Frank Kiefer: Multiskalen-Verfahren für Konvektions-Diffusionsprobleme, Dissertation, Rheinische Friedrich-Wilhelms-Universität Bonn, 2001

[64] Peter Knabner, Lutz Angermann: Numerik partieller Differentialgleichungen, Eine anwendungsorientierte Einfhrung, Springer Verlag, Berlin, Heidelberg 2000

[65] Helmut Kopka: Latex, Band 1: Eine Einführung, Addison-Wesley Verlag, München 2002

[66] Helmut Kopka: Latex, Band 2: Ergänzungen, Addison-Wesley Verlag, München 2002

[67] Alois Kufner: Weighted Sobolev Spaces, Teubner-Texte zur Mathematik Band 31, Teubner Verlag, Leipzig 1980

[68] Gerd Kunert: Aposteriori Error Estimation for Convection Dominated Problems on Anisotropic Meshes, Preprint SFB 393, Chemnitz 2002

[69] O.A. Ladyzenskaja, V.A, Solonnikov, N.N. Ural'ceva: Linear and Quasi-Linear Equations of Parabolic Type, Translations Of Mathematical Monographs Volume 23, American Mathematical Society, Providence, Rhode Island, USA 1968

[70] Stig Larsson, Vidar Thomee: Partielle Differentialgleichungen und numerische Methoden, Springer Verlag, Berlin, Heidelberg 2005

[71] Christof Leisinger: Der amerikanische Häusermarkt bleibt schwach, FAZ.NET, 2007

[72] Z. Liu, J.G. Korvink, R. Huang: Structure Topology Optimization: Fully Coupled Level Set Method via FEMLAB, Struct. Multidisc. Optim. 29, pp. 407-417, 2005

[73] F. A. Longstaff, E. S. Schwartz: A Simple Approach to Valuing Risky Fixed and Floating Rate Debt, Journal of Finance 50, pp. 789-819, 1995

[74] J. Mattey, N. Wallace: Housing Price Cycles and Prepayment Rates of U.S. Mortgage Pools, Journal of Real Estate Finance and Economics 23, pp. 161-184, 2001

[75] C. Mayer, D. Genesove: Equity and Time to Sale in the Real Estate Market, American Economic Review, pp. 255-269, June 1997

[76] R.C. Merton: Theory of Rational Option Pricing, Bell Journal of Economics and Management Science 4, pp. 141-183, Spring 1973

[77] Charles B. Morrey: Multiple Integrals In the Calculus of Variations, Die Grundlagen der mathematischen Wissenschaften, Springer Verlag, Berlin Heidelberg New York 1966

[78] Marek Musiela, Marek Rutkowski: Martingale Methods in Financial Modelling, Springer Verlag, Berlin, Heidelberg 2005

[79] C.W. Oosterlee: On Multigrid for Linear Complementarity Problems with Application to American-Style Options, Electron. Trans. Numer. Anal. 15, pp. 165-185, 2000

[80] Alfio Quarteroni, Riccardo Sacco, Fausto Saleri, Numerische Mathematik 2, Springer Verlag, Berlin Heidelberg 2002

[81] Christoph Reisinger: Numerische Methoden für hochdimensionale parabolische Gleichungen am Beispiel von Optionspreisaufgaben, Inaugural Dissertation, Ruprecht-Karls-Universität, Heidelberg, Januar 2004

[82] Yousef Saad, Martin H. Schultz: GMRES: A Generalized Minimal Residual Algorithm For Solving Nonsymmetric Linear Systems, SIAM J. Sci. Stat. Comput., Vol. 7, No. 3, July 1986

[83] Yousef Saad: ILUT: A Dual Threshold Incomplete LU Factorization, Report umsi-92-38, Computer Science Department, University of Minnesota

[84] E.S. Schwartz, W.N. Torous: Prepayment and the Valuation of Mortgage-Backed Securities, Journal Of Finance 44, pp. 375-392, 1989

[85] R. Seydel: Tools for Computational Finance, Springer Verlag, 2002

[86] C.F. Shampine, M.K. Gordon: Computer-solution of Ordinary Differential Equations, WH Freeman & Co., San Francisco, California, 1975

[87] Richard Stanton: Rational Prepayment and the Valuation of Mortgage-Backed Securities, The Review of Financial Studies 8, pp. 677-708, 1995

[88] R. H. Stanton: Unobservable Heterogneity and Rational Learning: Pool-specific Versus Generic Mortgage-Backed Security Prices, Journal of Real Estate Finance and Economics 12, pp. 243-263, 1996

[89] Josef Stoer: Numerische Mathematik 1, Springer Verlag, Berlin Heidelberg 1999

[90] Vidar Thomée: Galerkin Finite Element Methods for Parabolic Problems, Springer Verlag 1997

[91] Jürgen Topper: Finite Element Modeling of Exotic Options, Discussion Paper No. 26, ISSN 0949-9962

[92] R. Verfürht: A Review of Aposteriori Error Estimation and Addaptive Mesh-Refinement Techniques, Teubner Verlag and J. Wiley, Stuttgart 1996

[93] Dirk Werner: Funktionalanalysis, Springer Verlag, Berlin Heidelberg 1997

[94] Paul Wilmott, Sam Howison, Jeff Dewynne: The Mathematics of Financial Derivatives, A Student Introduction, Cambridge University Press, Cambridge 1995

[95] Peter G. Zhang: Exotic Options - A Guide to Second Generation Options, World Scientific Publishing, Singapore 1998

[96] Comsol Multiphysics User's Guide, Version 3.2, 2005

[97] Comsol Multiphysics Scripting Guide, Version 3.2, 2005

[98] Comsol Multiphysics Command Reference, Version 3.2, 2005

[99] www.faz.net

[100] www.handelsblatt.com